MINIMALIST INVESTIGATIONS IN LINGUISTIC THEORY

Howard Lasnik

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

MINIMALIST INVESTIGATIONS IN LINGUISTIC THEORY

During the past decade the minimalist approach has been highly influential in syntactic theorizing. The articles collected in this book represent empirical and technical explorations of syntactic phenomena that have implications for this minimalist approach to syntax. Taken together they display some of the changes and developments in the minimalist program during the past decade.

This book is a considerable contribution to the development of the approach in its exploration and refinement of theoretical concepts, comparison of competing implementations and resurrection of much older lines of analysis. Among the central topics considered are verbal morphology, ellipsis, anaphora and the determination of anaphoric connection, and the blending of minimalist devices with features of earlier generative accounts.

Howard Lasnik is Professor of Linguistics at the University of Maryland. For thirty years he has played a prominent role in syntactic theorizing in the Chomskian framework from the Extended Standard Theory and Government-Binding Theory to Minimalism.

ROUTLEDGE LEADING LINGUISTS
Series editor Carlos P. Otero

1 ESSAYS ON SYNTAX AND SEMANTICS
James Higginbotham

2 PARTITIONS AND ATOMS OF CLAUSE STRUCTURE
Subjects, agreement, case and clitics
Dominique Sportiche

3 THE SYNTAX OF SPECIFIERS AND HEADS
Collected essays of Hilda J. Koopman
Hilda J. Koopman

4 CONFIGURATIONS OF SENTENTIAL COMPLEMENTATION
Perspectives from romance languages
Johan Rooryck

5 ESSAYS IN SYNTACTIC THEORY
Samuel David Epstein

6 ON SYNTAX AND SEMANTICS
Richard K. Larson

7 COMPARATIVE SYNTAX AND LANGUAGE ACQUISITION
Luigi Rizzi

8 MINIMALIST INVESTIGATIONS IN LINGUISTIC THEORY
Howard Lasnik

9 DERIVATIONS
Exploring the dynamics of syntax
Juan Uriagereka

MINIMALIST INVESTIGATIONS IN LINGUISTIC THEORY

Howard Lasnik

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

First published 2003
by Routledge
11 New Fetter Lane, London EC4P 4EE

Simultaneously published in the USA and Canada
by Routledge
29 West 35th Street, New York, NY 10001

Routledge is an imprint of the Taylor & Francis Group

This edition published in the Taylor & Francis e-Library, 2005.

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging in Publication Data
Lasnik, Howard.
Minimalist investigations in linguistic theory/Howard Lasnik.
p. cm.
Includes bibliographical references and index.
1. Minimalist theory (Linguistics) 2. Grammar, Comparative and general – Syntax.
3. Generative grammar. I. Title
P158.28 .L38 2002
415–dc21 2002068749

ISBN 0-203-98720-9 Master e-book ISBN

ISBN 0-415-18194-1 (Print Edition)

CONTENTS

List of figures vii
Preface ix
Acknowledgments xi

1 Introduction 1

2 Patterns of verb raising with auxiliary "be" 6

3 Last resort and attract F 22

4 Levels of representation and the elements of anaphora 42

5 Pseudogapping puzzles 55

6 On feature strength: three minimalist approaches to overt movement 83

7 A gap in an ellipsis paradigm: some theoretical implications 103

8 On a scope reconstruction paradox 118

9 Some reconstruction riddles 125

10 Chains of arguments 139

Notes 158
References 177
Index 187

FIGURES

2.1	Pollock's split Infl hypothesis	9
5.1	The relevant stage where the accusative NP has raised out of the "small clause" but the V has not yet raised	58
5.2	Koizumi's split VP hypothesis	61
5.3	Portion of the structure that would cause a PF crash is deleted	62
5.4	Double object structure	63
5.5	Underlying structure of a simple transitive	65
5.6	Pseudogapping structure for double object construction with two remnants; lowest VP to be deleted	67
5.7	Verb *give* has overtly raised to its correct ultimate destination, passing the intermediate shell V in transit	68
5.8	Tense head governs the VP that is to be deleted	70
5.9	Deletion of lower VP when the licensing head, Past, is more remote, with two maximal projections intervening	71
5.10	Bad Pseudogapping with two maximal projections intervening between Tense and the deleted VP	72
5.11	A potential source for HNPS and Pseudogapping	74
5.12	*A lot of advice* undergoes HNPS to a position outside VP_1 adjoined to TP, VP, or $AgrP_1$ and finally VP_1 deletes	75
6.1	Koizumi's split VP hypothesis	90
6.2	Source for simple Pseudogapping	90
6.3	The strong feature not overtly checked causes a PF crash	91
6.4	Source for matrix Sluicing	93
6.5	Structure for matrix Sluicing with strong feature indicated	94
6.6	A PF crash will be avoided if either pied-piping or deletion of a category containing V takes place	100
6.7	A PF crash will be avoided if either pied-piping or deletion of a category containing Infl takes place	101

7.1 Split VP structure with the strong feature of the shell V attracting a matching feature of the lower V 110
7.2 Strong feature of C attracting a matching feature of Infl 112
10.1 The relevant portion of an ECM structure with ECM subject and matrix V both raising 152

PREFACE

The articles collected in this book represent empirical and technical explorations of syntactic phenomena that have implications for the minimalist approach to syntax. Taken together, they display some of the changes, developments, and fits and starts in the minimalist program, from the mid-1990s through the late 1990s. Alongside detailed investigation of a number of syntactic phenomena, I also compare alternative instantiations some of the leading ideas of Minimalism. It is clear that there remain many possible ways of pursuing the basic approach, and it is still very far from clear which ways are correct. I hope that this book provides at least a hint of what some of the paths might be.

The articles appear in their original form, with the following exceptions: Typographical errors have been corrected; occasionally, a clearer example has been substituted for the original; minor changes have been made to bring the articles more into line with each other in terms of style and exposition. Also all bibliographic references have been updated, cross-references to chapters have been added, and a unified bibliography provided. Finally, I have attached a brief introductory note to each chapter putting the material in context and flagging changes in the analyses and theories.

I would like to take this opportunity to once again acknowledge my vast intellectual debt to Noam Chomsky. His work has been a constant source of stimulation and inspiration for me (as for so many others), and I have also been very fortunate to have had the benefit of his advice and guidance over the last thirty-three years. Others who have contributed important ideas, suggestions, and comments are Cedric Boeckx, Bob Freidin, Roger Martin, Masao Ochi, and, especially, Željko Bošković. The majority of this material was developed in my annual syntax seminar at the University of Connecticut. All of the participants made important contributions to the development and clarification of the ideas.

Finally, I would like to thank Carlos Otero for his encouragement in this project, Arthur Stepanov for indispensable help with the preparation of the manuscript and Cedric Boeckx and Tomohiro Fujii for proofreading and indexing. Last, but far from least, I thank my wife Roberta for her unfailing encouragement and support for thirty years.

Howard Lasnik
April, 2002

ACKNOWLEDGMENTS

Chapter 2 is the written version of my presentation at the 1995 University of Massachusetts, Amherst, Conference on African-American English. Chapter 3 is the written version of my presentation at the 1995 meeting of the Formal Linguistics Society of Mid-America. It first appeared in the Volume 1 of the Proceedings of that conference, pp. 62–81. Chapter 4 is the written version of my presentation at the 1996 Workshop on Atomism and Binding, HIL/Leiden University. It first appeared in H. Bennis, P. Pica, and J. Rooryck (eds) *Atomism and Binding*. Foris, pp. 251–68. It is reproduced here with the kind permission of Holland Academic Graphics. Chapter 5 is the written version of my presentation at the 1996 Workshop on Fragments: Studies in Ellipsis, Conjunction and Gapping, SOAS, London. It first appeared in S. Lappin and E. Benmamoun (eds) *Fragments: Studies in Ellipsis and Gapping*. Oxford University Press, pp. 141–74. It is reproduced here with the kind permission of Oxford University Press. Chapter 6 first appeared in the journal *Linguistic Inquiry* 1999, 30: 197–217. It is reproduced here with the kind permission of MIT Press Journals. Chapter 7 first appeared in the journal *Linguistic Analysis* 1997: 166–85. It is reproduced here with the kind permission of that journal. Chapter 8 first appeared on the 1998 web page Celebration: An Electronic Festschrift in Honor of Noam Chomsky's 70th Birthday. Chapter 9 is the written version of my presentation at the 1998 Penn Linguistics Colloquium. It first appeared in *University of Pennsylvania Working Papers in Linguistics*, Vol 5.1, pp. 83–98. Chapter 10 first appeared in S. Epstein and N. Hornstein (eds) *Working Minimalism*. MIT Press, pp. 189–215. It is reprinted here with the kind permission of MIT Press.

1

INTRODUCTION

The articles collected here, like the large majority of my work, are syntactic investigations in the Chomskian framework, especially Minimalism, its most recent development. I was first attracted to Chomskian generative grammar by its quest for "hidden causes." A hallmark of Chomsky's work (and that of many of his students, associates, and followers) is the attempt to explain the apparent chaos of language by appeal to the interaction of simple underlying principles, structures, and operations. A related impulse has been to show that languages that are superficially very different are actually far more similar just a bit below the surface.

Hidden causes at the morpho-syntactic level have been especially interesting to me. One classic analysis in such terms is presented in Chomsky (1955), and given implicit justification there, and explicit justification in Chomsky (1957). The analysis centers around Chomsky's phenomenally successful attempt to make sense of the apparent chaos of English verbal morphology, with the very simplest kinds of basic sentences seeming to exhibit the most unpredictably idiosyncratic alternations under negation and interrogation:

(1)	John left	John didn't leave
	John should leave	John shouldn't leave
	John has left	John hasn't left
	John is leaving	John isn't leaving
(2)	John left	Did John leave
	John should leave	Should John leave
	John has left	Has John left
	John is leaving	Is John leaving

Chomsky's breakthrough was the insight that the tense/agreement morpheme in English is syntactically an autonomous entity even though it is invariably realized as a bound morpheme. It is available to transformational manipulation just as much as, say, a modal auxiliary is.

The beauty of this analysis had a profound effect on me, though, interestingly, not the very first time I saw it. In early 1968, I read through *Syntactic Structures* and did not really "get it." Having a degree in mathematics, I had no technical difficulty with the automata discussion or the phrase structure and transformational formalism. Rather, I failed to grasp what it was all in aid of. That was rectified in the fall of that year when I took "baby syntax," the introductory graduate syntax course at MIT. This was the first linguistics course I ever took, and one of the very best. It was team taught by two inspirational teachers, Morris Halle and John Robert ("Haj") Ross, who easily held the rapt attention of the class for the three two-hour meetings each week. Halle taught the first half of the course, emphasizing general issues of linguistics, especially what kind of a thing a human language is, what its special properties are (structure, infinity, …), and what a theory of language should be: an account of what speakers know and how that knowledge arises. When Halle then began to present a version of the *Syntactic Structures* account of the English verbal system (followed up later by Ross's detailed presentation of virtually the entire syntax of English, as then understood), I finally "got it," and was overwhelmed by its beauty. My reaction is still the same. In the Introduction to Lasnik (2000a), I call Chomsky's early treatment of verbal morphology "the best set of analyses in the history of our field." I continue to use it as a model for my own theorizing, sometimes explicitly, as in my work on verbal morphology (Lasnik 1981, 1995e) and Chapter 2 of this volume), but always at least implicitly. What particularly struck me was how seemingly different sentence types, which obviously *had* to be brought together somehow, were, indeed, brought together, and by a system that was simpler than any description that kept them apart.

Chomsky's analysis crucially embodies a hidden cause, an element that has no overt realization of its own, but has striking indirect effects. Chomsky posits an *S morpheme* for the present tense of verbs with (third person) singular subjects, and a zero morpheme Ø for other present tense forms. Given that the latter forms are virtually always phonologically indistinguishable from the bare citation form, Chomsky observes that "An alternative we did not consider was to eliminate the zero morpheme and to state simply that *no* affix occurs if the subject is not third person singular." Chomsky (1957: 64). The reason for rejecting that alternative out of hand was that it would have substantially complicated the system with no concomitant benefit. In Lasnik (1981), I present a completely parallel argument for the existence of a zero imperative morpheme in English, and in Chapter 2, a similar one for a zero habitual morpheme in African-American English.

A closely related Chomskian theme that has greatly influenced my work is that "if some phenomenon is observed overtly in certain languages, then it probably applies covertly (i.e. without overt expression at PF) in all languages in some manner; that is, the overt expression is probably a consequence of requirements of UG, which must be satisfied at S-structure and LF, even if not overtly observed at the PF level" Chomsky (1987: 68–9). Chomsky specifically mentions Case,

which is richly expressed in such languages as Latin, Finnish, and Sanskrit, hence "probably … required in all languages, even in languages such as English in which there is only a very marginal residue at PF …." Case theory, as developed especially by Chomsky out of an original proposal by Jean-Roger Vergnaud in a personal letter to Chomsky and me (commenting on a draft of Chomsky and Lasnik (1977)) led the way to insightful analyses of many phenomena in numerous languages. In much of my work from the early 1980s to the present, I have relied on Vergnaud's and Chomsky's basic insight about the distribution of nominal expressions, and I have explored various ways of developing and extending Case theory. Examples are Lasnik (1992a, 1993, 1995a,b, 2001), Lasnik and Freidin (1981), Lasnik and Saito (1991), and Chapters 3 and 10. One early question I still worry about is exactly what class of expressions is subject to the Case Filter. Early on, Bob Freidin and I began worrying about Chomsky's original formulation in which phonological realization was key. In Lasnik and Freidin (1981), we showed that the trace of *wh*-movement must have Case, even when the operator that has moved is phonetically null. This demonstration was, in part, responsible for the move to the "visibility" approach to Case suggested by Chomsky (1981), under which all and only arguments (whether phonologically realized or not) obey the Case Filter. Over two decades later, the ramifications of that move are still being worked out. This, along with a number of questions about the role of Case, is at the core of much minimalist theorizing, including my own.

Chomsky also mentions *wh*-movement in the context of deep similarities between languages hidden by superficial differences, suggesting that "If languages such as English exhibit overt *wh*-movement in interrogatives, then probably languages such as Chinese and Japanese, which do not, nevertheless have LF-movement of *wh*-phrases …." In the very early 1980s, when Huang (1981/82) was first laying the empirical and theoretical groundwork for a theory of covert *wh*-movement, Mamoru Saito, then a graduate student at MIT, and I simultaneously became intrigued by some of the technical questions that arose in Chomsky's initial presentation of a theory based on Huang's. Our collaboration on this topic resulted in Lasnik and Saito (1984) and, several years later, Lasnik and Saito (1992). A major concern of both works was the specification of the Empty Category Principle (ECP), a locality constraint on (the traces of) overt and covert *wh*-movement. To this day, I regard Huang's initial argument that *wh* in situ obeys the ECP as still the single strongest argument for covert movement. The general question of how overtness of movement is forced is the theme of Chapter 6.

At the more foundational level, Chomsky (1965, chapter 1), was a major influence on my thinking about language. Parallel to my experience with Chomsky (1957), I initially had difficulty understanding exactly what Chomsky was getting at. But when I came to understand Chomsky's notion of "explanatory adequacy," I made that a goal of virtually all of my work. It was the move towards increased emphasis on explanatory adequacy that made the advent of the "extended standard theory" so appealing to me, and the "principles and parameters" theory even more so. Lasnik (1990) collected nine of my articles, in the principles and

parameters framework, on the explanatory themes of restrictiveness and learnability. I see Minimalism as a natural development of these themes.

One major Minimalist line of research with evident roots in the earliest work in transformational grammar (e.g. Chomsky (1955)) concerns levels of representation, particularly the question of exactly what levels there are. (For a historical survey see Lasnik (In press)). In the late 1960s and early 1970s, questions were raised about the existence of deep structure as a level, most directly by Postal (1972). Chomsky in his minimalist writings raises this same question (even if from a somewhat different perspective), and extends the question to surface structure. Anaphora has provided a major testing ground for the hypothesis that surface structure does not exist, since so many anaphoric connections had been thought to crucially rely on surface structure configuration. In Chapters 4 and 8–10, I discuss the issue, showing how apparent surface structure effects can be reanalyzed as dependent on LF instead (by positing more overt movement than is usually assumed), but I also continue to explore the question raised in Lasnik (1993) of why assumed covert movement does not generally create new configurations of anaphora.

Thinking about my research and my teaching, I discern a rather robust strain of what can be characterized as conservatism. I often find myself trying to resurrect old analyses or maintain current analyses that are being supplanted, feeling that the rejected accounts were rejected prematurely, or for debatable reasons. I talk about this trend specifically in my teaching in the Introduction to Lasnik (2000a):

> At the beginning, I want to say a few words about why I am going to base a large part of the discussion on something that seems so old and outmoded by the standards of current syntactic theorizing. I have three reasons. First, many of the terms, concepts, and analyses in recent work are much easier to understand against a backdrop of their ancestors of a few decades ago. Second, our field is a relatively young one without a very large number of good arguments and analyses to use as models. We can't yet afford to ignore some of the good arguments and analyses, even if we conclude that they are ultimately incorrect. Finally, and probably most importantly, I believe that many of the analyses of the 1950's are actually correct, fundamentally and (sometimes) even in detail.

In addition to my investigations of verbal morphology, some of my other recent "conservative," or even "reactionary" work explores Exceptional Case Marking (ECM), ellipsis, island violation "repair," and clause-mate conditions on syntactic relations. Starting with Lasnik and Saito (1991), I have been arguing that the accusative subject of (certain) infinitive constructions raises into the higher clause, an approach to the construction that was standard in the late 1960s and early 1970s (as in Rosenbaum (1967) and Postal (1974)) but which was largely abandoned following the arguments of Chomsky (1973). Some of my most recent

discussions of (an updated "minimalist" version) of this old approach appear in Chapters 3 and 10 and in Lasnik (2001). Classically, the issue of "raising to object" has been intimately connected with the formulation of the locality condition on anaphoric relations, among others. The traditional view, elaborated in great detail in Postal (1974), was that the relevant locality obtains when the two items to be related are in the same clause – clause-mates. Work since the early 1970s has taken alternative approaches, for example, in terms of the Tensed Sentence Condition of Chomsky (1973) or the Governing Category of Chomsky (1981). My work on ECM constructions implicitly argues for the classic view. I make that argument explicit in Lasnik (2002). It is an interesting question whether the classic approach is more or less in accord than its successors with minimalist ideals. My gut instinct is that, if anything, clause-mate is actually a more minimalist concept, since the notion "clause" is quite plausibly a primitive of syntactic description.

Some of my current work investigates ellipsis, one of the most persistent topics in generative grammar. In this realm, too, I have assumed, and sometimes argued for, a traditional analysis, one based on deletion, rather LF copying, the latter being the far more standard view in recent syntactic theories. Chapters 5–7 analyze some properties of ellipsis in terms of deletion. Chapter 6 argues that a certain complementary between normally obligatory movement and ellipsis receives a natural account in these terms. With respect to ellipsis, again, one might ask whether Minimalism suggests one approach over another. The answer is far from clear, but a strict interpretation of the Inclusiveness condition of Chomsky (1995a) might suggest a deletion approach, since on that approach, it is clear that no new elements are added in the course of a derivation.

As noted above, I see Minimalism, with its extreme simplification of the model, as a natural development of the drive towards explanatory adequacy. I think we still have quite a way to go to achieve explanatory adequacy, but is not unreasonable to begin to raise the possibility suggested by Chomsky (2001) that we can eventually move beyond it.

2

PATTERNS OF VERB RAISING WITH AUXILIARY "BE"

In this chapter, I reprise the "hybrid" analysis of verbal morphology that I presented in Lasnik (1995e), where Infl can be either a bundle of syntactic features, as in Chomsky (1993) (in which case it attracts a verb), or an affix, as in Chomsky (1955, 1957) (in which case a low-level process of Affix Hopping associates it with a verb). In terms of this theory, I suggest an account for the patterning of habitual *be* in African-American English (AAE). Based on the work of Green (1993), I propose an AAE habitual morpheme "Hab" and suggest that its morpho-syntactic behavior is parallel to that of the imperative morpheme (a zero affixal morpheme). In the course of the presentation, I offer a speculation about the nature of *do* support, suggesting that it is merely the phonetic manifestation of a finite Infl (or Imp or Hab) that has not been able to undergo Affix Hopping. I also bring new data to bear on the old puzzle that adverbs do not block Affix Hopping, a process normally requiring adjacency. The new data suggest that, contrary to initial appearance, adjacency might actually obtain in these cases.

One of the major breakthroughs in the history of generative transformational grammar was the discovery by Chomsky (1955, 1957) of the regularities underlying English verbal morphology. Much of the apparent chaos of this central portion of Standard American English (SAE) morpho-syntax was rendered systematic by the fundamental insight that the tense agreement inflectional morpheme is syntactically independent, even though always a bound morpheme superficially. The analysis was brilliantly successful, and paved the way for numerous refinements and extensions over the past forty years, the large majority of them sharing the same fundamental insight. Labov in a number of publications (1969, 1972) demonstrated that African-American English (AAE) falls under many of the same fundamental generalizations, the apparent striking differences often being the result of rather low-level operations. More recently, Green (1993) has explored especially the aspectual system of AAE, showing (unsurprisingly to

linguists) that the same general syntactic principles are operative, though there are significant parametric differences. In this chapter, I will consider some of the revisions of Chomsky's original proposal, ultimately presenting an analysis (Lasnik (1995e)) that constitutes something of a return to that original proposal. Finally, I will show how a modified version of Green's analysis of the AAE habitual construction fits naturally into the approach to verbal morphology I advocate.

2.1. Earlier accounts

Many of the successive refinements of the Chomsky (1957) theory can be viewed as attempts to maintain the leading ideas of the analysis but to reconcile them with the growing concern for explanatory adequacy. For example, Lasnik (1981) was particularly concerned with the stipulated rule ordering and the arbitrary marking of particular transformations as obligatory or optional in Chomsky's early system, and proposed that these problematic language particular formal mechanisms can be eliminated in favor of the general filter in (1):

(1) The "stranded affix" filter: A morphologically realized affix must be a syntactic dependent of a morphologically realized category, at surface structure (Lasnik (1981)).

Notice that this filter crucially assumes, along with Chomsky (1957) and many succeeding analyses, that the inflectional material on a verb is a morphological affix, even though it begins its syntactic existence as an autonomous entity. Given this assumption, and given (1) and the restrictive theory of transformations it presupposes, a typical analysis of the English verb system of the early 1980s looks something like (2):

(2)(a) S is the maximal projection of the inflectional morpheme Infl (= C of Chomsky 1957).
(b) Infl takes VP as its complement.
(c) When the head of VP is *have* or *be* it raises to Infl, the next head up.
(d) Otherwise Infl lowers to V: Affix Hopping.
(e) Otherwise *do* adjoins to Infl.

Such a system is descriptively comparable to that of Chomsky (1957), handling the familiar paradigms below, and is superior in terms of explanatory adequacy,[1] for the reasons already alluded to.

(3)

John left	*John leftn't
John has left	John hasn't left
John is leaving	John isn't leaving

(4)

John left	*Left John
John has left	Has John left
John is leaving	Is John leaving

Emonds (1978), based on a similar model, insightfully explored certain differences between SAE and French. Taking *pas* to be the analog of *not*, he was concerned with the fact that while only auxiliary (finite) verbs precede negation in SAE, any (finite) verb does in French:

(5)(a) *John likes not Mary
(b) Jean (n')aime pas Marie

Emonds proposed that the basic difference between SAE and French is that in the latter language, verb raising is not limited to auxiliaries. Then, given the priorities in (2), Affix Hopping will never be necessary in French.

Pollock (1989) developed Emonds' idea further, offering an explanation of the verb raising difference between SAE and French. First, he argued that Infl should be split:

(6) "Infl" is not one head; it consists of (at least) Tense and Agr(eement), each heading its own projection. Raising to tense proceeds via Agr.

Given (6), the difference between SAE and French is accounted for by (7):

(7)(a) SAE Agr, because not morphologically rich, is OPAQUE to θ-role transmission. Thus, if a verb with θ-roles to assign were to raise, it would be unable to assign them, resulting in a violation of the θ-criterion.
(b) French Agr, because morphologically rich, is TRANSPARENT to θ-role transmission, so any sort of verb can raise.

Chomsky (1991), building on Pollock's analysis, offers the following economy-based explanation of why raising takes place whenever it can:

(8) Raising is preferred to lowering, because lowering will leave an unbound trace that will have to be remedied by re-raising in LF.

Notice that (8) assumes re-raising is, in general, possible, when not blocked by a more economical derivation. If this were not so, even (9) would not be possible:

(9) John likes Mary

The next question, then, is why (10), with overt lowering and LF re-raising, is not the SAE version of (5b).[2]

(10) *John not likes Mary

Chomsky's answer to this question is stated in terms of his more articulated version of Pollock's split Infl hypothesis. Note that Chomsky follows Pollock in taking negation to be a head (Figure 2.1).[3]

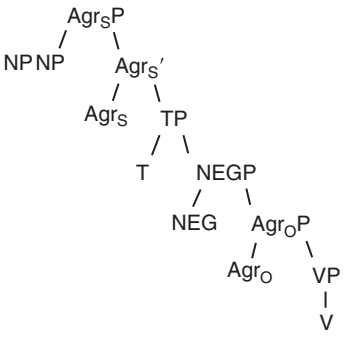

Figure 2.1 Pollock's split Infl hypothesis.

According to Chomsky, the Head Movement Constraint (reduced to an Empty Category Principle (ECP) antecedent government requirement) prevents the LF re-raising needed in the derivation of (10).[4] The intervening head NEG cannot be crossed. On the face of it, the overt raising across negation in French, and that in SAE with *have* and *be*, would seem to run afoul of the same requirement, Chomsky accounts for the difference in the following way, where (11) lists the relevant principles, and (12) and (13) sketch the French and SAE derivations, respectively:

(11)(a) If Agr moves, its trace can be deleted, since it plays no role in LF.
(b) If V moves, its trace cannot be deleted.
(c) Deletion of an element leaves a category lacking features, [*e*].
(d) Adjunction to [*e*] is not permitted.

(12)(a) When V OVERTLY raises (French), (5b), it first adjoins to Agr_O, creating $[_{Agr_O}$ V $Agr_O]$;
(b) Next, Agr_O raises to T, crossing NEG, thus leaving a trace that is marked [−γ], indicating a violation of the ECP. That trace is an Agr;
(c) In accord with (11a), the [−γ] trace is deleted, so there is no ECP violation (where ECP is, as in Lasnik and Saito (1984, 1992), an LF filter: *[−γ]).

(13)(a) When V vainly attempts to COVERTLY (re-)raise in LF (English), (10), Agr_s has already lowered overtly to T, leaving an Agr trace (which deletes, leaving [*e*]), and creating a complex T.
(b) Which has already lowered overtly to Agr_O, leaving a T trace and creating a still more complex Agr.
(c) Which has already lowered overtly to V, leaving an Agr trace (which deletes, leaving [*e*]), and creating a complex V. (Evidently, the deletion is obligatory, which follows from Full Interpretation, and necessarily *immediate*, which does not seem to follow from anything.)

(d) This complex V raises to the [*e*] left by the deletion of the Agr_O trace, a movement that is, by (11d), necessarily substitution, thus turning [*e*] into V.

(e) This element now raises across NEG to (the trace of) T, leaving behind a [−γ] trace which is, crucially, a V trace, hence non-deletable. The resulting LF is in violation of the ECP.

There is a potential technical problem with this account of French, in that (11a), (12c) seem to be inconsistent with a central economy condition of Chomsky (1991): Deletion is only permitted to turn an ill-formed LF object into a well-formed LF object, where the relevant well-formed objects are operator–variable pairs and UNIFORM CHAINS (chains all of whose members are X^0s; are in A-positions; or are in A′-positions.) This is precisely to prevent making a short licit head-, A-, or adjunct-movement, followed by a long illicit movement, with subsequent deletion of the offending trace of the latter movement.[5] But just such a short head-movement followed by a long one is crucially being allowed here.

A related problem is that generally, a long movement (i.e. one in violation of relativized minimality) results in some degradation (e.g. Subjacency effects, as discussed by Chomsky and Lasnik (1993)), even if the offending trace is eventually eliminated. But the long overt V-movement at issue here is fully grammatical. I will not pursue these technical problems further, since the minimalist framework of Chomsky (1993), which I turn to now, rejects the account for other reasons, and provides another perspective on verbal morphology.

2.2. A lexicalist approach

Chomsky (1993) departs in an important respect from his earlier treatments of verbal morphology by adopting a strictly lexicalist view, under which verbs are taken from the lexicon already fully inflected. They still must associate syntactically with the appropriate functional heads, but only in order for their inflectional properties to be CHECKED against abstract features of the functional heads (rather than acquired as affixes). This checking approach mirrors Chomsky's checking view of Case, which holds that Case features are already associated with (the heads of) DPs (Determiner Phrases) as they are first inserted into syntactic structures. These DPs must wind up in positions where the Case they already have can be suitably licensed.[6] Note that on this view, there is no obvious need for Affix Hopping. The fact that verbs overtly appear with their inflectional morphology even in English is no longer a relevant consideration in determining exactly how the derivation proceeded.

Intrinsic to this checking theory is that the features of verbs and functional heads must be checked against each other, but that this checking can, in principle take place anywhere in a derivation on the path to LF. Chomsky also proposes, as a matter of execution, that once a feature of Agr has done its checking work, it disappears.

From this point of view, the difference between French and SAE is not verb raising vs affix lowering. Rather, it is whether verb raising takes place in overt syntax (French) or in the LF component (SAE). Further, since Chomsky argues that LF and PF are the only levels of linguistic representation, this difference cannot be attributed, as it might have been in previous theories, to any S-structure property. Chomsky thus proposes (14) as the core difference between French and SAE. The relevant notions are explicated in (15):

(14)(a) In French, the V-features of Agr (i.e. those that check features of a V) are strong.
(b) In SAE, the V-features of Agr are weak.

(15)(a) V-features are not legitimate PF objects.
(b) Strong features are visible at PF; weak features are not. Surviving strong features cause the derivation to CRASH at PF.

In French, since the V-features of Agr are strong, if V raises to Agr overtly, the V-features of Agr check the features of the V in overt syntax and disappear. Both LF and PF are thus well-formed. If, on the other hand, V were to delay raising until LF, the V-features of Agr would survive into PF, causing the derivation to crash at that level, even though LF requirements would be satisfied. This correctly forces overt V-raising in French.

In SAE, delaying V-raising until LF does not result in an ill-formed PF object, so such a derivation is POSSIBLE. What makes it NECESSARY is:

(16) PROCRASTINATE: Delay an operation until LF whenever possible, that is, whenever delaying would not cause the derivation to crash.

(16) thus plays a central role in excluding (5a), repeated as (17):

(17) *John likes not Mary

But, as already discussed, *have* and *be* do raise overtly. Chomsky proposes that this happens because *have* and *be* are semantically vacuous, hence not visible to LF operations. Thus, if they have not raised overtly, they will not be able to raise at all. Their unchecked features will cause the LF to crash.

This proposal raises certain questions. First, it is not clear that *be* is always semantically vacuous, yet the syntactic behavior of *be* in finite clauses is always the same. For example, it is reasonable to assume that in (18), *is* has the meaning of *exists*. Yet, as seen in (19), it raises overtly nonetheless:

(18) There is a solution
(19)(a) There is not a solution
(b) Is there a solution

Second, even apart from the empirical considerations just mentioned, there is the conceptual question of whether syntactic operations, even those in the LF

component, should be sensitive to purely semantic properties. LF is, after all, a syntactic, rather than a semantic, component. Finally, there is reason to believe that even instances of *have* and *be* that are vacuous in Chomsky's sense can undergo LF raising. For example, as noted by Wilder and Čavar (1994) and Wexler (1994), in Swedish, auxiliary verbs pattern exactly with main verbs in remaining in situ in embedded clauses, even though they are undoubtedly inflected:[7]

(20)(a) …, om hon inte ofte **har** sett honom
whether she not often has seen him
(b) * om hon **har** inte ofte sett honom
(c) * om hon inte **har** ofta sett honom

Even putting these questions aside, it is important to note that Chomsky (1993) provides no account of the crucial (10), repeated as (21):

(21) *John not likes Mary

The analysis of Chomsky (1991), summarized in (13) above, does not carry over to this framework, nor can it be made to, since it relies so centrally on properties of the traces left by affix lowering, an operation that is eliminated in Chomsky (1993).[8] This much is clear: it must be ruled out, but its derivation must not crash. If it crashed, it could not block (22), since Procrastinate only chooses among convergent derivations.

(22) *John likes not Mary

The analytic options are severely limited, particularly under the proposal of Chomsky (1994), apparently contra Chomsky (1991), that a [−γ] trace causes a derivation to crash. Below, I will suggest a new (actually very old) perspective on these facts that avoids this particular problem.

2.3. A hybrid theory

We have seen that Chomsky's lexicalist–minimalist account of verbal morphology demands that Agr and T are just abstract features that check against features of fully inflected verbs which raise to them. The earlier accounts treated such Infl items exclusively as bound morphemes that had to become affixes on otherwise bare verbs. We have seen that each approach has substantial problems. I will argue that the most important of these problems can be overcome under a hybrid approach that allows both mechanisms to coexist. (23) sketches such a possibility, where the fundamental difference between French and SAE (and between SAE auxiliary and main verbs) is with respect to choice of mechanism, that is, with respect to lexical representation:

(23)(a) French verbs are fully inflected in the lexicon (possibly correlating with the fact that there are no bare forms; even the infinitive has an ending).

(b) *Have* and *be* are fully inflected in the lexicon (possibly correlating with the fact that they are highly suppletive, but see below).
(c) All other SAE verbs are bare in the lexicon.

With the lexical properties of verbs outlined in (23), no further stipulations are needed for Infl, at least for the core phenomena. As I will show momentarily, (24), the null hypothesis under the theory I advocate, suffices for French and SAE finite clauses:

(24) Infl is freely an affix or a set of abstract features.

Given that SAE *have* and *be* behave just like French verbs, and given that SAE main verbs are not lexically represented with inflectional features (23c), the Infl feature strength difference posited by Chomsky ((14) and (15)) above) becomes superfluous. Instead, we have (25):

(25) Finite featural Infl is strong in both French and SAE.

The final necessary mechanism is, for all intents and purposes, the original one: Affix Hopping. Further, as conjectured by Lasnik (1981), and developed further by Halle and Marantz (1993) and Bobaljik (1994), the rule is a morphophonemic one rather than a syntactic one:

(26) Affixal Infl must merge with a V, a PF process (distinct from syntactic head-movement) demanding adjacency.

Consider now the various combinations made available by this theory. First, suppose that we select a verb with inflectional features (notated here as +F), and a featural (as opposed to affixal) Infl:

(27) … Infl … V …
+F +F

This configuration is, of course, well-formed. V raises (overtly) to Infl, and all relevant features are checked. This is the situation with all French verbs, as well as with SAE *have* and *be*.

Next, consider the case of a bare verb and an affixal (as opposed to featural) Infl:

(28) … Infl V …
Af bare

This is the situation with SAE main verbs. In this configuration, PF merger takes place, as long as adjacency obtains, and the PF affixal requirement of Infl is satisfied.

Given (24), two other configurations could potentially arise, but, since both of them will ultimately crash, there is no need to replace (24) with a stipulation. The first such mis-matched configuration is shown in (29):

(29) … Infl … V …
+F bare

Here, the features of Infl will not be checked, so the derivation crashes at LF. And under the assumption that the features are strong, there is a PF crash as well.

Finally, consider (30), the reverse of (29):

(30) … Infl … V …
Af +F

This time, the features of V will fail to be checked, causing an LF crash. Additionally, if affixal Infl cannot attach to an already inflected verb, this failure leads to a PF crash.

(31) summarizes the immediately preceding discussion:

(31)(a)	… Infl … V … +F +F	OK. V will overtly raise.
(b)	… Infl … V … Af bare	OK. PF merger.
(c)	… Infl … V … +F bare	* at LF. +F of I won't be checked; * at PF as well, since +F is strong.
(d)	… Infl … V … Af +F	* at LF. +F of V won't be checked. * at PF also, if merger fails.

Thus, it follows automatically from the lexical properties of French verbs that French finite Infl will always have to be featural, just as it follows from the lexical properties of *be* and auxiliary *have* that SAE finite Infl will always have to be featural, when the verb is *be* or *have*. The parallelism in behavior between French verbs and SAE auxiliaries has a unified account in this theory, following from a parallelism in morphological properties. With a main verb in SAE, on the other hand, English finite Infl will always have to be affixal, and this too follows automatically.

Consider now the problematic ill-formed negative sentences in English. (32) involves an apparently inflected verb in situ in overt syntax:

(32) *John not likes Mary

Recall that Chomsky had assumed that such a configuration is universally ungrammatical, invariably leading to an ECP violation. However, we have seen that the theory leading to that conclusion for Chomsky, in which LF verb-raising is preceded by overt affix lowering, has been rejected in favor of a lexicalist approach. And in the latter, the result does not obtain. Further, there is empirical reason for doubting the conclusion in the first place, as seen, for example, in the Swedish example (20) above.

My account of (32) is simply that of Chomsky (1957).[9] *Walked* is not in the lexicon of SAE: all main verbs are "bare." Hence, (32) must arise from the merger of affixal Infl with *walk*. But *not* intervenes between Infl and *walk* so the former cannot merge with the latter. Crucially, then, the Swedish example (20a) must not involve merger. Rather, it must involve the covert analog of the overt raising seen in French or with SAE *have* and *be*. That is, the verb is pulled from the lexicon fully inflected. Infl, then, is necessarily featural, so the verb must raise to Infl for the matching features to be checked. The different property of Swedish is that the V-features of Infl are weak, while those in French and SAE are strong. Procrastinate dictates that the Swedish verb will remain in situ in overt syntax.

(33) is fundamentally similar:

(33) *John likes not Mary

Likes is not in the lexicon, so even though featural Infl exists in English, and even though its V-features are strong, (33) could never be generated by raising.[10]

We have seen that raising across negation overtly, as in (34), (35), and covertly, as in (36), is available:

(34) Jean (n′)aime pas Marie
(35) John has not left
(36) …, om hon inte ofte **har** sett honom

We now must ask why this is possible. One possibility, along the lines of Roberts (1993, 1994), is that NEG and V are heads of different sorts (A′ vs A), and that relativized minimality is even more relativized than in the original proposal of Rizzi (1990). If a head only blocks movement of a head of the same type, NEG would then not block movement of V. A second possibility is that NEG is not a head, but a modifier. Note that, at least for Chomsky, its central (perhaps only) role as a head had been to block (32), via the ECP. But under the present approach, the ECP is irrelevant to the issue. I will not attempt to choose between these two approaches here.

One remaining question for an adjacency account of examples like (32) is raised by Pollock (1989). Pollock notes that unlike *not*, adverbs do not seem to block adjacency. For an example like (37), there is no particular difficulty.

(37) John never left

The adverb *never* is relatively free in its distribution, occurring after or before an auxiliary in Infl:

(38) John will never leave
(39) John never will leave

hence presumably after or before Infl more generally, even when Infl is an affix. However, as Pollock notes, there are adverbs that seem to occur only after the auxiliary verbs in a sentence:

(40) John will completely lose his mind
(41) *John completely will lose his mind

It would seem that these adverbs necessarily occur between Infl and V, perhaps as left adjunct of VP. But even these adverbs are seemingly invisible for Affix Hopping:

(42) John completely lost his mind

Based on this fact, Bobaljik (1994) suggests that adverbs (or, more generally, adjuncts) are not relevant to PF adjacency, while heads and specifiers are relevant. While this approach deals with the problem Pollock raised, it seems somewhat stipulative. Another direction to explore is suggested by the following observation: under certain circumstances, VP ellipsis being one, even such "low" adverbs as *completely* can appear before an element in Infl, in fact, even before the pure morphological realization of Infl:

(43) John partially lost his mind, and Bill completely did

What makes this possible is, at present, mysterious.[11] But the fact that it is possible suggests that even in (42), the adverb might precede Infl, and thus might not interfere with the adjacency required between Infl and V.

2.4. Further evidence: VP-ellipsis and imperatives

I have argued that there is a fundamental morphological difference between French verbs and SAE main verbs, and that this difference is mirrored internal to SAE by one between auxiliary and main verbs. Certain surprising facts about VP ellipsis first discussed by Warner (1986) provide interesting additional evidence supporting the English internal claim.[12]

It has long been known that VP ellipsis can ignore certain inflectional differences between antecedent verb and elided verb. For example, Quirk *et al.* (1972), reported by Sag (1976), observe that a finite form of a verb can antecede the deletion[13] of the bare form that follows a modal, as in the following example:

(44) John slept, and Mary will too
(45)(a) *John *slept*, and Mary will *slept* too
(b) John *slept*, and Mary will *sleep* too

In (44), the past tense form *slept* serves as antecedent for the deletion of the bare form *sleep*. The present tense form can also antecede the bare form:

(46) John sleeps (every afternoon), and Mary should too

(47)(a) *John *sleeps*, and Mary should *sleeps* too
(b) John *sleeps*, and Mary should *sleep* too

But, as Warner notes, there are certain exceptions to this general pattern. (48) is seemingly parallel to (44), but, surprisingly, it is unacceptable:

(48) *John was here, and Mary will too
(49)(a) *John was here and Mary will was here too
(b) John was here and Mary will be here too

Evidently *was* cannot antecede *be*; nor can *is* antecede *be*:

(50) *John is here, and Mary will too

There is no general prohibition on VP ellipsis of a VP headed by *be* following a modal. (51) is virtually perfect, and far better than (48) and (50):

(51) John will be here, and Mary will too

Similar effects obtain with auxiliary *have*. Ellipsis is markedly better in (52), with identical forms of *have*, than in (53) with distinct ones:

(52)(a) John should have left, but Mary shouldn't have left
(b) ?John should have left, but Mary shouldn't ~~have left~~

(53)(a) John has left, but Mary shouldn't have left
(b) *John has left, but Mary shouldn't ~~have left~~

Note that, as might be expected, the ellipsis site in (53b) is fine when interpreted as *leave*. That is roughly the situation we have seen before, with one form of a main verb anteceding a distinct form of that verb (in this case, perfect *left* anteceding bare *leave*). Note too that the identity of form demanded for ellipsis of auxiliary *have* is somewhat abstract, making reference to morphological features, and not just phonetic ones. (54) is no better than (53b):

(54) *The men have left, but the women shouldn't ~~have left~~

The present plural of auxiliary *have* cannot antecede the bare form, even though they are both superficially *have*. Note too that main verb *have* patterns with other main verbs, and not with auxiliary *have*:

(55) John has a driver's license, but Mary shouldn't

The descriptive generalization is stated in (56):

(56) The bare form of a verb V other than *be* or AUXILIARY *have* can be DELETED UNDER IDENTITY with any other form of V. *Be* or auxiliary *have* can only be deleted under identity with the very same form.

As Warner observes, this difference does not follow directly from (degree of) suppletion. The paradigm of *go* is highly suppletive, yet that verb patterns with all other main verbs:

(57) John went, and now Mary will ~~go~~

Thus, the relevant difference seems to be between MAIN verbs and AUXILIARIES, where the latter category includes *be* and certain instances of *have*.

Sag (1976) briefly discusses the main verb phenomena, taking them to be representative. He observes that these cases could be accounted for by ordering verb phrase deletion before Affix Hopping. Note that on the strictly lexicalist view discussed above, there is no such point in a derivation. However, on the analysis of Chomsky (1957), adopted in its essentials here, there is indeed such a point. I have departed from Chomsky (1957) in just one major respect: for him *all* verbs are introduced into syntactic structures bare, and achieve their inflectional form via Affix Hopping. On the other hand, I have argued that auxiliaries are pulled from the lexicon fully inflected.[14] This difference between SAE main and auxiliary verbs was part of my explanation of the verb-raising asymmetries. Strikingly, the very same difference can explain the ellipsis asymmetries, along essentially the lines suggested by Sag:

(58) A form of a verb V can only be deleted under identity with the very same form. Forms of *be* and auxiliary *have* are introduced into syntactic structures already fully inflected. Forms of "main" verbs are created out of lexically introduced bare forms and independent affixes.

Given (58), deletion under apparent incomplete identity is actually deletion under full identity, but at a point in the derivation before the bare stem has associated with the inflectional affix. This is schematically illustrated in (59), a structure for *John slept, and Mary will too*:

(59) John Infl sleep, and Mary will ~~sleep~~ too

The conclusions about the negation and ellipsis phenomena reviewed thus far potentially provide a microscope for the examination of additional inflectional forms of verbs. Consider simple imperatives in SAE:

(60) Leave

What is the morphological analysis of such a sentence? Lasnik (1981) argues that there is an imperative affix (occupying the position normally occupied by Tense) that must associate with the bare stem, based on the ungrammaticality of (61):

(61) *Not leave

The ungrammaticality is due to the stranded affix filter, since lack of adjacency blocks the merger of Imp and *leave*. The analysis fits completely into the framework I have outlined here. Two alternatives are excluded. It cannot be that there is no Imp morpheme at all, since that would leave (61) unexplained. Nor could there be a featural Imp to which already inflected *leave* would raise, since if the hypothesized feature were weak, (61) would be good, and if it were strong, (62) would:

(62) *Leave not

Thus far, Imp is behaving just like finite Infl. The parallelism extends still further: (61) is salvaged (however that is to be captured in the theory) by *do* support:

(63) Do not leave

The parallelism breaks down with respect to auxiliary verbs, however. Not even *be* can raise:

(64) *Be not foolish

This suggests (65), on the present account, as a lexical property of English:

(65) The Imperative morpheme is strictly affixal, hence there will never be raising to it (just merger with it).

In this particular construction, then, *be* is pulled from the lexicon bare, just as main verbs are. This predicts that imperative of *be* should parallel imperative of main verbs in ellipsis behavior. (66) shows that the prediction is confirmed:[15]

(66)(a) Leave. I don't want to.
I won't.
(b) Be quiet. I don't want to.
I won't.

(66) is in direct contrast with (67), the properties of the latter following from the fact that *is* never arises via affixation.

(67)(a) Mary left. I don't want to.
(b) Mary is quiet. *I don't want to.

2.5. An analysis of *be* in AAE

With all of this as background, I turn now to the syntax of habitual *be* in AAE, as outlined by Green (1993). Green shows that habitual *be* fails to undergo raising:

(68)(a) Bob be writing his assignments
(b) *Bob ben't writing his assignments

(69) *Be you happy when you talk to your sister?

Further, like imperatives, and unlike SAE subjunctives, habituals do display *do* support.

(70) Bob don't be writing his assignments
(71) Do you be happy when you talk to your sister?

Compare subjunctives:

(72) I require that you (*do) not be late for class

Green correctly rejects a Pollock-type account, on the basis that habitual *be* has no θ-roles to assign, so should not be prevented from raising. She ultimately proposes that this AAE habitual construction involves an AspP and, above it, an Aux position containing *do*. Then, under certain circumstances, the *do* is deleted. This analysis is not unreasonable, but, based on the striking parallelism between the AAE habitual construction and the SAE (and, I assume, AAE) imperative construction, I will suggest an analysis of the former parallel to the one I offered for the latter. In both, we seem to begin with the bare form of the verb (*be* in the habitual, all verbs, including *be*, in the imperative). In neither construction does the verb raise. Finally, in both constructions, supportive *do* appears.[16]

The parallel analysis would proceed as follows. In place of Tense or Imp, there is a habitual morpheme (call it "Hab"). This habitual morpheme is, like the imperative morpheme, a phonetically null affix. Under the hypothesis that the morpheme is strictly affixal, lacking the formal features that would attract the raising of a verb, (68b) and (69) are straightforwardly excluded. Hab takes as its complement *be*, and merges with it under adjacency. When adjacency does not obtain, either because Hab has been fronted (71) or because negation intervenes (70), *do* support applies. The SAE subjunctive presumably differs from imperatives and habituals in just this respect: there is no affix, hence *do* support is neither required nor permitted, as seen in (72) above.[17]

2.6. Summary

In summary, I have argued that a version of the theory of Chomsky (1957), with affixal functional morphemes, provides a principled account of major properties of

the SAE verbal system. Further, the affixal hypothesis naturally extends to cover two constructions not analyzed by Chomsky: the SAE imperative and the AAE habitual. Whether the striking parallelism seen between the latter two constructions is merely fortuitous, and more generally, whether principles or lexical idiosyncrasy determine affixal vs featural properties, awaits further investigation.

3

LAST RESORT AND ATTRACT F

The nature of the "last resort" restriction on movement is the concern of this chapter. I argue that Chomsky's (1993) Greed, which permits movement only in order to satisfy requirements of the moving item, is too strong; there are derivations where movement is solely to satisfy the needs of the position to which movement is taking place. There are two principal cases considered: movement to "Extended Projection Principle (EPP)" positions, and, in much more detail (covert) movement of the "associate" of the expletive *there*. With respect to the latter operation, following Chomsky (1995a), I argue that the relevant movement is just of the formal features of the associate, and to satisfy the agreement needs of the Agr head of the clause, *there* lacking agreement features of its own. This approach, as Chomsky notes, solves the scope problem of existential constructions – that the associate always takes scope in situ. Departing from Chomsky, I also argue that for all c-command purposes (binding, negative polarity item licensing, pronouns as bound variables), not just scope, the associate displays low behavior. And given this line of reasoning, I argue, following Koizumi (1993, 1995) that there is much more overt movement in English that is standardly assumed, with, in particular, nominal complements and Exceptional Case Marking (ECM) subjects raising overtly to [Spec, Agr_O], with the verb raising overtly to a still higher V position. Further consequences and arguments are examined in Chapters 4 and 5.

3.1. The "Greed" constraint

Since the earliest work in "Government-Binding" syntax, a major concern in investigations of syntactic movement has been the search for the "driving force" for movement. In fact, the roots of this line of investigation go back still further, as seen in this passage from Chomsky (1965): "… it has been shown that many of the optional singulary transformations of Chomsky (1955, 1957, 1962) must be reformulated as obligatory transformations, whose applicability to a string is determined by presence or absence of a certain marker in the string." In modern

work, A-movement, as in (1), is rendered obligatory by the need for the nominative Case of the moving NP to be checked.

(1) He is likely to win

In this particular situation, there might seem to be an additional driving force for the movement, the feature responsible for the Extended Projection Principle (EPP). This is seen in (2), which is bad with unfilled subject position.

(2)(a) It is likely he will win
(b) *Is likely he will win

However, particularly since Chomsky (1986b), the apparent "last resort" nature of A-movement has been emphasized. If an NP already is in a Case-marked position, it cannot move to another, even if such movement would result in satisfaction of the EPP:

(3) *He is likely will win

Chomsky (1993, 1994), therefore, has proposed that the movement of an item α is driven exclusively by requirements of α itself, even if failure to move results in a "crashed" derivation with an unchecked EPP feature, as in (2b). Chomsky calls this version of the last resort constraint "Greed."

Chomsky (1993) discusses another example with the abstract properties of (3), indicating that Greed is necessary to achieve the correct results:

(4) _seems to [$_{\alpha}$ a strange man] [that it is raining outside]

Chomsky argues that raising of α must be blocked, both in LF ((4) is not a well-formed sentence) and in overt syntax:[1]

(5) *A strange man seems to *t* that it is raining outside

Chomsky's proposal is that α has no reasons of its own for moving, either overtly or covertly, since its Case is licensed by *to*. Again, according to Chomsky, Greed is implicated, since even though movement would help the derivation satisfy the EPP, it still is blocked. However, on closer inspection it becomes evident that the examples considered so far provide no firm evidence for Greed. In fact, under independently plausible assumptions, they argue against Greed, since that constraint turns out to be redundant for such examples.

3.2. Some arguments against Greed

First, it is generally assumed that the Case feature of an NP must be checked, as must the matching Case feature of the corresponding Case licensing head. A simple

technical instantiation of this, consistent with Chomsky's analyses, is as follows: If the Case feature survives until the LF interface level, the derivation crashes. Establishing a checking configuration avoids this crash if the feature disappears when it is checked. But then, if the derived subjects in (3) and (5) have already had their Case checked before they move to subject position, the nominative Case feature of Tense will never be checked, and *that* will cause the derivation to crash. Greed is superfluous in explaining the ungrammaticality, redundant with other principles.[2]

Next, note that for (5), there is even further redundancy, since a single NP could not possibly check both the nominative feature of Tense and the accusative or oblique feature of *to*. One or the other would necessarily remain unchecked, even apart from the considerations of the preceding paragraph.

Consider now a related example discussed by Chomsky (1994) with respect to Greed:

(6) *It is believed [a man to seem to *t* that S]

On the face of it, this forms the basis for a much stronger argument for Greed. There is exactly one Case feature to be discharged (that of *to*), and exactly one NP to check that feature (*a man*). The movement is *solely* to satisfy the EPP, thus is altruistic, in violation of Greed. However, here, too, there is an alternative account, one made available by Chomsky's (1994) analysis of (7).

(7) *There is likely [someone to be [*t* here]]

In Chomsky's theory of phrase structure, at the outset of a derivation, all lexical items to be used are selected, constituting the "numeration." The derivation includes generalized transformations, which cyclically combine these lexical items into phrasal units, and ultimately into one structure. Consider then (8), one stage in the derivation of (7).

(8) [$_{\gamma}$ to be [$_{\beta}$ someone here]]

At stage (8), there is a choice: it is possible to fill [Spec, γ] by selecting *there* from the numeration and inserting it, or by raising *someone*. Chomsky argues that the latter move would violate Procrastinate.

(9) Procrastinate: LF movement is preferred to overt movement.

Then when γ is further embedded, *there* will be raised to the higher subject position. This movement is permitted by Procrastinate since satisfaction of the EPP can no longer be delayed. There will be no need to move *someone* at all in the overt syntax, and, in particular, no need to move it to the position vacated by *there*, as

the EPP has already been satisfied. Procrastinate (along with general economy considerations) thus suffices to rule out (7), in favor of (10).

(10) There is likely to be someone here

With that in mind, reconsider (6), repeated as (11).

(11) *It is believed [a man to seem to *t* that S]

One stage in the derivation of (11) would be (12).

(12) [$_{\gamma}$ to seem to a man that S]

At this point in the derivation [Spec, γ] must be filled. The choices are the raising of *a man*, or the insertion of *it*. But just as in Chomsky's discussion of (7), Procrastinate favors the latter over the former, blocking (11) in favor of (13):

(13) It is believed [*t* to seem to a man that S]

Here too, Greed is irrelevant to the choice.

The next phenomenon to consider, also from Chomsky (1994), is the apparent non-existence of simple transitive verbs with θ-roles for subject and object, but no objective Case feature. For example, *HIT* in (14) is like *hit*, except that it has no Case feature.

(14) *John$_i$ Infl [$_{VP}$ t_i [$_{V'}$ HIT t_i]]

John has originated in complement position, picking up the object θ-role of the verb, then moved to [Spec, V], picking up the subject role, on its way to [Spec, I]. Thus, if (14) were grammatical it would mean "John hit himself." Chomsky notes that on one set of assumptions, Greed is involved in the explanation of the ill-formedness of this derivation, thus explaining why *HIT* does not exist. The Greed account of (14) raises a question. Elsewhere, Chomsky indicates that Greed has a sort of "global" character. In particular, a specific instance of movement need not immediately result in the checking of a feature of the moved item. Rather, it need only be a necessary step in a sequence of movements that ultimately result in the required checking configuration. For Chomsky, the intermediate steps in a successive cyclic movement derivation are licensed in this way. So if [Spec, V] is a necessary landing site on the way to [Spec, I], movement through that position would, in fact, be permitted and the subject θ-role of *HIT* could be discharged. The issue is far from clear-cut (see Bošković (1994a) developing ideas of Saito and Murasugi (1993)) but the economy condition "shortest move" might demand, hence license, that step of movement. Thus, the badness of (14) does not constitute a clear argument for Greed, since Greed would not suffice to exclude it.

One final consideration is suggested by Željko Bošković, who observes, following Chomsky, that the non-existence of the specific lexical item *HIT* is not of concern. This might simply be a lexical gap. Rather, the central empirical claim is that there can be no verbs at all with the Case and thematic properties of *HIT*. While this claim is, to the best of my knowledge, very generally true, it might not be entirely true. As Alan Munn points out, examples such as those in (15) receive a natural description if the verbs are regarded as *WASH, DRESS*, and *SHAVE*, with the surface subject having originated in deep thematic object position and having passed through deep thematic subject position.

(15)(a) John washed (= John washed himself)
(b) John shaved (= John shaved himself)
(c) John dressed (= John dressed himself)

I have already argued (contra Chomsky) that it is by no means certain that Greed does exclude verbs of the *HIT* type, so there is no clear argument for Greed in that direction. Now we see from (15) that if Greed does have such an effect, that might actually count as an argument *against* that constraint. Pending further investigation of the class of verbs in (15), I will leave the matter in this unsettled stated.[3]

3.3. Enlightened self-interest

The precise nature of the last resort condition is of fundamental importance for the treatment of simple existential constructions. In a series of articles, beginning with Chomsky (1986b), Chomsky argued that the "associate" of *there* moves to *there* in LF in such a sentence as (16).

(16) There is a man here

This provides the basis for an account of the familiar superficially bizarre agreement paradigms displayed by these constructions, with the verb agreeing with something that is not its formal subject:

(17)(a) There is/*are a man here
(b) There are/*is men here

Chomsky (1986b) proposed that the LF movement of the subject is substitution, with the associate replacing *there*. There are, however, difficulties with substitution. For example, it would result in identical LFs for (18) and (19).

(18) A man is likely to be here
(19) There is likely to be a man here

But the interpretive possibilities diverge. In (18), *a man* can evidently have wide or narrow scope with respect to *likely*, while (19) allows only narrow scope for *a man*. Partly for this reason, Chomsky (1991) modified his (1986b) substitution analysis, proposing instead that the associate adjoins to *there*, the latter being a sort of LF clitic.[4] As for the driving force for the movement, Chomsky claims that this movement is A-movement and, as in many other instances of A-movement, it is motivated by the Case requirements of the moved item. Chomsky thus maintains that the Case of the argument in the constructions under consideration is not licensed without movement to subject position, but is licensed in that position. This is in accord with Greed.

Note, though, that Case checking in this affixation configuration raises a conceptual question. Chomsky (1993) proposes that the "crucial properties and relations [of the grammar] will be stated in the simple and elementary terms of X-bar Theory." In particular, he argues that such notions as government by a head, clearly not a core X-bar relation, must be dispensed with. All structural Case is thus recast in terms of the Spec–head relation. However, just as "government" was an extension of the core head–complement relation, Chomsky extends the Spec–head relation to "checking domain." This notion includes not just the Spec of a head, but also items adjoined to the Spec. It is this latter configuration that is relevant to Case licensing in existential constructions, according to Chomsky. Given Chomsky's argument against government, there is, thus, reason to suspect that Case does not in fact provide the driving force for the movement. Belletti (1988), Lasnik (1992a, 1995a,b) provide further arguments for this conclusion, and propose that *be* (and unaccusatives) are Case licensers (licensing a Case that Belletti calls "partitive"), so that the Case of the associate is licensed without movement to the expletive. Further, if *there*, like other NPs in A-positions, must bear a Case of its own, "partitive" will be the only Case available for the associate of *there*.[5]

Chomsky (1994) in passing indicates acceptance of the conclusion that *be* licenses Case, but indicates that Greed can still be maintained, suggesting that it is the *agreement* features of the associate that need to be licensed by the movement. This seems unlikely, however. First, even if the ϕ-features of the associate have to be checked, they would already be checked against the Agr projection that Chomsky assumes constitutes the small clause in which the associate originates. In an existential construction such as (20)

(20) There is someone available

this agreement would be manifested on the adjective (as it is overtly in languages with richer morphology than English). Further, there are reasons to doubt that NPs do have to be so checked. First, in a coordinate NP, it is not clear how the individual conjuncts could be checked, particularly when they disagree in ϕ-features ("Mary and the boys"). Additionally, in a language like English with "natural" (as opposed to purely formal) agreement features on NPs, these features are semantically

relevant so presumably must survive to the LF level, hence cannot be checked in Chomsky's sense, since that would entail deletion.[6]

We have arrived at the conclusion that there are no features of the associate that need to be checked by the movement, thus, that Greed is too strong a constraint. Now if any version of last resort is correct, the movement must satisfy *some* formal requirement of some item. Martin (1992a) and Groat (1993) make the suggestion that *there* lacks ϕ-features. The raising of the associate in well-formed existential constructions is then driven by the need for Agr to discharge its ϕ-features. This analysis is inconsistent with Greed, unless we assume, with Groat, that the ϕ-features of an NP also have to be checked, a possibility that Chomsky (1994: 36) also seems to favor (in accord with his apparent acceptance of the partitive Case analysis). However, as already discussed, it is unlikely that NPs have to be so checked. Thus, even under the Martin/Groat proposal, it would still follow that Greed is too strong a constraint: the relevant movement is altruistic. Consider then a slightly weakened version of Greed, which I will call "enlightened self interest." These two versions of last resort are stated in (21).

(21)(a) Greed: Movement of α to β is for the satisfaction of formal requirements of α.

(b) "Enlightened self interest": Movement of α to β is for the satisfaction of formal requirements of α or β.

Even apart from Case and agreement properties of existential constructions, there are reasons for favoring (21b) over (21a). Chomsky (1993) proposes that raising of a *wh*-operator to [Spec, C] is driven by the need for a morphological Q-feature to be checked. In a simple interrogative clause, C has this feature, as does the operator that raises to it. Further, if the Q-feature of C is strong, the raising will be overt, as in English. Thus, it seems that the operator raises to check its own feature, and in so doing, it satisfies the feature of the head it raises to. So far so good. But what of multiple interrogation? Chomsky (1993) argues, exactly along the lines of Chomsky (1973), that the *wh*-phrases that are in situ overtly, remain in situ at LF, and are interpreted in the appropriate Comp without movement to that Comp at any level of representation. As Martin (1996b) notes, for the one *wh* that actually does move, we must identify a driving force, in particular (under Greed) a morphological feature of that *wh* that must be checked. Further, that feature must distinguish the *wh* that moves from the ones that do not, because if all had the feature, the unmoved ones would cause the derivation to crash. Or, just as bad as far as Chomsky is concerned, they should all move. Alternatively, if the feature is simply freely assigned to any *wh*-phrase, then there is no description of standard Superiority effects, as in (22).

(22) *****What did who buy

This is so since we could have freely assigned Q to *What* and not to *who*, with the result that *What* would be the highest *wh* capable of moving. These problems

disappear once Greed is relaxed to enlightened self-interest. All that needs to be said is that interrogative C has the Q feature, that the feature is strong, and that it can be checked by any *wh*-operator, and that the Q feature of the operator need *not* be checked (it can survive to the interface level).[7] None of this goes beyond what is explicitly or tacitly assumed in most discussions of the phenomenon. Enlightened self-interest, but not Greed, allows the movement of an operator to [Spec, C] to be entirely for the benefit of the target.

ECM constructions such as (23) create a related difficulty for Greed.

(23) I believe John to be clever

There must be some strong feature of non-finite tense driving the overt movement of *John* to subject position. But the relevant feature is not a Case feature, since Case in ECM constructions is checked in the Spec of the higher Agr_O, in association with *believe*. Further, as briefly discussed above, the raising of the NP cannot in general be driven by any agreement needs of the NP itself. There is, thus, no feature of the NP that needs to be satisfied. Yet, Greed, unlike enlightened self-interest, demands that there be one. The same difficulty for Greed arises in even more extreme form in successive cyclic raising constructions:

(24) John is believed [*t* to be likely [*t* to be arrested *t*]]

What features of *John* itself could possibly demand to be checked in *every* subject position it passes through? It is phenomena of this type that require the computationally complex global property of Greed mentioned earlier. Presumably *John* must move through the intermediate *t* positions in order for it to successfully arrive at its ultimate goal: the nominative Case checking position in the highest clause.[8] Given this, the possibility arises that enlightened self-interest is actually a *stronger* constraint than Greed in one regard. If an instance of movement of α to β can be driven by the needs of β (the feature instantiating the EPP, in the instances under discussion), the computation can be strictly local. *Every* step of movement will immediately satisfy a requirement; no look ahead will be required or permitted.

So far, we have seen that several of the arguments for Greed can be overcome in a principled fashion, and, in addition, that Greed (and the more general theory including it) is problematic in a number of respects. However, there is one remaining argument for Greed. Chomsky (1993) considers the ungrammatical (25).

(25) *There seems to [$_\alpha$ a strange man] [that it is raining outside]

In this instance "α has its Case properties satisfied internal to the PP, so it is not permitted to raise, and we are left with a freestanding *there*." Chomsky takes this freestanding *there* to be a legitimate LF object, satisfying all morphological requirements. The derivation of the example is thus legitimate. Only the interpretation

of the example, and not its structural form, is faulty. Once Greed is weakened to enlightened self-interest, Chomsky's account of (25) is no longer available, assuming that *there* is featurally deficient. While it is still true that the morphological requirements of α and of Tense are satisfied without movement of α to *there*, the ϕ-features of Agr are not checked. Hence, movement is motivated under enlightened self-interest. The derivation should converge, and, according to Chomsky, it should be semantically coherent. As noted, Greed would correctly exclude the derivation. However, we have now seen several reasons for rejecting Greed. Further, even Greed would not explain the impossibility of a version of (25) with *to* replaced by *TO*, where *TO*, analogous to Chomsky's *HIT* considered earlier, is like *to*, but lacking a Case feature. To address this new problem, and simultaneously to reconcile (25) with enlightened self-interest, I suggest, as perhaps the simplest possibility, that the semantic difficulty that Chomsky attributes to (25) with *a strange man* in situ might arise even if *a strange man* were to move. In the absence of any precise theory of what the semantic difficulty is, there is no obvious reason to reject this account.[9]

3.4. Feature movement in existentials

Chomsky has recently suggested, in a series of lectures and, in slightly revised form, in Chomsky (1995a), a somewhat different theory of the LF movement involved in expletive constructions, as part of a revised theory of LF movement more generally. Beginning with the standard Minimalist assumption that all movement is driven by the need for formal features to be checked, Chomsky argues that, all else equal, movement should then never be of an entire syntactic category, but only of its formal features. Further, an unchecked feature of the target drives movement ("Attract F"), much in the spirit of enlightened self-interest, rather than Greed. PF requirements will normally force movement of a category containing the formal features, via a sort of pied-piping, under the reasonable assumption that a bare feature (or set of features) is an ill-formed PF object. For LF movement, on the other hand, pied-piping will normally not be necessary, hence, by economy, will not even be possible. Only the formal features will move, and they will move exactly to the heads that have matching features.[10] In a standard existential sentence like (26), then, the associate *someone* does not move to *there*.

(26) There is someone here

Rather, only the formal features of *someone* move, and only to a corresponding functional head (or heads). The affixal account of *there* does not seem statable in such a theory. Rather, we are led to something much more like the Martin/Groat analysis. If *there* lacks agreement features,[11] then the features of Agr will not be checked in overt syntax. The features of the associate will therefore have to move to Agr (in LF, because of Procrastinate). Note, in addition, that this overcomes the

technical problem noted earlier for the Martin/Groat analysis, since the checking configuration is now allowed even under strict assumptions. In effect, it is a head–head configuration.

An example that is problematic for enlightened self-interest, and for Chomsky's related proposal, is the following:

(27) *The belief [a man to seem [t' is [t here]]]

In (27) the movement of *a man* to satisfy the EPP requirement of the infinitive ought to be permitted, given enlightened self-interest, or its incorporation into Chomsky's feature attraction. I tentatively propose a technical solution: Assume with Chomsky that any visible feature of a head can "attract" a corresponding feature, resulting in the movement of the bundle of formal features (LF movement) or a syntactic constituent (overt movement). But in addition suppose that it is exactly a visible Case feature that makes the feature bundle or constituent available for "A-movement." Once Case is checked off, no further movement is possible. In (27), once *a man* checks nominative in the lower clause, it is no longer available for further A-movement.[12] The same analysis extends to *BELIEVE*, a verb like *believe*, but with no Case feature.[13] Consider (28).

(28) *John BELIEVES [a man to seem [t' is [t here]]]

Again, once *a man* has its Case checked, no further movement of it is possible. This proposal is in the nature of a compromise between Greed and enlightened self-interest. As with Greed, the NP that will move (or whose formal features will) must have a Case feature that has not been checked off. But as with enlightened self-interest, any particular instance of movement of that NP might be solely for the satisfaction of requirements of the target.

One major problem for Greed caries over to this theory, so must be dealt with now. If *be* and unaccusatives check Case (as argued by Belletti (1988) and Lasnik (1992a, 1995a,b)), the agreement features of Agr_S in (29) will remain unchecked.

(29) There is a man here

Note that this is precisely the type of sentence where *be* would necessarily participate in Case checking on the Belletti/Lasnik theory.[14] Hence (the formal feature bundle of) *a man* would not be available for further movement to Agr_S. However, even this problem disappears if we follow Belletti and Lasnik that the specific Case borne by the associate of *there* is one with semantic import. It would then not be checked-off even if it participated in checking. Being not merely a formal feature, it would survive to the LF interface level, so would be visible throughout the syntactic derivation.[15]

3.5. Feature movement and scope phenomena

The feature movement theory has the potential to address several problems concerning existential constructions. For example, in connection with (18)–(19) above, I alluded to Chomsky's argument, based on scope facts, against a substitution analysis of expletive replacement. There are, however, two difficulties in Chomsky's presentation. First, as he points out, in his example (30), it is not the case that there is no scope relation between *not* (or *n't*) and *many students*.

(30) There aren't many linguistics students here

Rather, *many linguistics students* necessarily has narrow scope. While it is true that if that NP were to replace *there*, it would be expected to have (at least as one possibility) wide scope, under the adjunction analysis, according to Chomsky no scope relation is established between *not* and *many linguistics students*. Chomsky indicates that under this circumstance, the scope of *many linguistics students* can be assumed to be narrow. Chomsky analogizes this situation to that in (31).

(31) Pictures of many students aren't here

However, in (31), there truly is no scope relation between negation and *many students*. The sentence is clearly not synonymous with (32).

(32) Pictures of few students are here

But such synonymy would be expected on the implied account of (30), since that example is synonymous with (33).

(33) There are few linguistics students here

In addition to this empirical problem, there is a technical problem. Chomsky evidently bases his argument that no scope relation is established between *not* and *many linguistics students* in the LF of (30) on the assumption that there is no c-command relation between those two expressions. However, on the theory of adjunction proposed by May (1985), developed in Chomsky (1986a), and assumed in all of Chomsky's writings since, there *would* be a relevant c-command relation in (30). *Many linguistics students* would c-command *not*, just as much as it would in (34).

(34) Many linguistics students aren't here

This is so because on May's and Chomsky's theory of adjunction, when α adjoins to β, β becomes a segmented category, and α c-commands anything β did prior to the adjunction. Thus, the scope problem that largely motivated the change from expletive substitution to expletive adjunction is actually not resolved by that change.

The feature movement analysis of existential constructions has the potential to solve this scope problem. If in LF, only the formal features of *many linguistics students*, rather than the entire expression, move to a functional head or heads above negation, it is reasonable to conclude that the quantificational properties remain below negation. Then, if it is this structure that determines scope (i.e. if QR either cannot alter these hierarchical relations or does not exist[16]) the desired results are obtained.

There are further phenomena that are suggestive of the same conclusion. For example, Lasnik and Saito (1991) show that with respect to anaphora, the associate of *there* in an ECM configuration behaves as if it is unmoved. They contrast (35) with (36).

(35) The DA proved [two men to have been at the scene] during each other's trials
(36) *The DA proved [there to have been two men at the scene] during each other's trials

Lasnik and Saito observe that in (35) *two men* behaves as if it c-commands an item that is in the higher clause, an argument related to several presented by Postal (1974).[17] If the ECM subject raises into the higher clause, as conjectured by Lasnik and Saito (1991) and Chomsky and Lasnik (1993), this can be explained.[18] However, one would expect that *there* in (36) also raises into the higher clause. And if *two men* has adjoined to *there*, we would incorrectly expect (36) to have the status of (35), given the theory of adjunction discussed just above.

Negative polarity items display similar contrasts, as seen in (37) vs (38).

(37) The DA proved [noone to be at the scene] during any of the trials
(38) *The DA proved [there to be noone at the scene] during any of the trials

Once again, the ECM subject (*noone*) behaves as if it c-commands an item in the higher clause (*any*), while the associate of *there* does not display this behavior.

Not surprisingly, standard subject raising to subject position patterns with the "subject raising to object position" of ECM constructions. Thus, a raised subject can antecede an anaphor in the higher clause, but the associate of a raised expletive cannot:

(39) Some linguists seem to each other [*t* to have been given good job offers]
(40) *There seem to each other [*t* to have been some linguists given good job offers]

Similarly, a raised negative subject licenses a polarity item in the higher clause while the negative associate of a raised expletive does not:

(41) No good linguistic theories seem to any philosophers [*t* to have been formulated]
(42) *There seem to any philosophers [*t* to have been no good linguistic theories formulated]

Weak crossover shows parallel asymmetries:

(43)(a) Some defendant$_i$ seems to his$_i$ lawyer [*t* to have been at the scene]
(b) *There seems to his$_i$ lawyer [*t* to have been some defendant$_i$ at the scene]

All of the pairs involving overt raising to subject position are straightforward on the feature movement view. When the movement is overt (examples summarized in (44)), the *entire* NP moves.

(44)(a) Some linguists seem to each other [*t* to have been given good job offers]
(b) Many linguistics students aren't [*t* here]
(c) No good linguistic theories seem to any philosophers [*t* to have been formulated]
(d) Some defendant$_i$ seems to his$_i$ lawyer [*t* to have been at the scene]

As a result, the properties (referential, quantificational, etc.) relevant to licensing an anaphor or negative polarity item or determining scope will be in the required structural position. On the other hand, when the movement is covert (examples summarized in (45)), only the formal features raise.

(45)(a) *There seem to each other [*t* to have been some linguists given good job offers]
(b) There aren't many linguistics students here
(c) *There seem to any philosophers [*t* to have been no good linguistic theories formulated]
(d) *There seems to his$_i$ lawyer [*t* to have been some defendant$_i$ at the scene]

The properties relevant to anaphora and scope remain below, failing to provide a higher licensing or scope position.

Further, the ungrammatical ECM examples involving an expletive and an associate (summarized in (46)) are easily accounted for in similar fashion.

(46)(a) *The DA proved [there to have been two men at the scene] during each other's trials
(b) *The DA proved [there to be noone at the scene] during any of the trials

What is now mysterious, though, is the original phenomenon that Lasnik and Saito set out to describe. Consider again the examples that were paired above with those in (46):

(47)(a) The DA proved [two men to have been at the scene] during each other's trials
(b) The DA proved [noone to be at the scene] during any of the trials

What Lasnik and Saito were concerned with, along lines similar to Postal (1974), was the apparent higher behavior of subjects of infinitives. That this behavior is related to ECM is seen in the contrast between the infinitival complements in (47) and the corresponding finite complements in (48).

(48)(a) ?*The DA proved [that two men were at the scene] during each other's trials
(b) ?*The DA proved [that noone was at the scene] during any of the trials

Similarly, for scope purposes, an ECM subject behaves as if it is in the higher clause while a nominative subject does not. Postal noted the contrast in (49): the scope of *few students* can include the higher clause in (49a).

(49)(a) The FBI proved few students to be spies
(b) The FBI proved that few students were spies

All of these licensing and scope phenomena indicate that an ECM subject *may* behave as if it is in the higher clause. A further example discussed by Postal suggests that the ECM subject *must* behave that way:

(50)(a) *Joan believes [him$_i$ to be a genius] even more fervently than Bob$_i$ does
(b) Joan believes [he$_i$ is a genius] even more fervently than Bob$_i$ does

If *him* is raised to the higher clause in (50a), while (uncontroversially) *he* remains in the lower clause in (50b), the contrast can be explained by whatever derives Condition C effects.

We have now arrived at a virtual contradiction.[19] The phenomena in (45) argue that when raising is in LF, only the formal features of an NP raise, leaving behind those properties involved in anaphora, scope, etc. But (47), (49), and (50), which involve the same class of phenomena, argue that referential and scopal properties in ECM constructions do raise, along with the formal features.

At this point, I will consider the discussion of Chomsky (1995a). Chomsky examines some of these same phenomena, concluding that "… the features adjoined to Agr$_O$ … have A-position properties, c-commanding and binding in the standard way." In particular, Chomsky takes the establishment of anaphoric relations to involve solely formal features. (47a) and (50a) are then unproblematic. However, the ability of an ECM subject to license a negative polarity item in the higher clause (47b) or to show higher clause scope behavior (49a) is unexpected, given that "… only formal features of the associate raise, leaving its semantic features behind …" Further, one would expect the associate of *there* to license an anaphor just as much as an ECM subject (or a subject overtly raised to subject position, for that matter). But we have seen that that is incorrect ((46a), (45a)). On this latter point, Chomsky offers a suggestion. Assuming the LF anaphor movement analysis of Condition A effects, Chomsky indicates that the LF structure of

the relevant portion of (45a) would be (51a) or b, where AN is the anaphor, FF is the set of raised formal features, and α is the X^0 -complex formed from INFL and the matrix V.

(51)(a) [$_{\text{INFL}}$ AN [FF (*linguists*) α]]
(b) [$_{\text{INFL}}$ FF (*linguists*) [AN α]]

Chomsky concludes that "On reasonable assumptions, neither of these structures qualifies as a legitimate binding-theoretic configuration, with AN taking FF (*linguists*) as its antecedent." The difficulty with this is that while it might correctly distinguish among the subject-raising to subject-position examples, disallowing anaphor binding when it is *there* that overtly raises, it incorrectly excludes *all* anaphor binding with "subject raising to object position." For example, Chomsky, following Lasnik and Saito, takes a sentence just like (47a) to be acceptable, but (52), the relevant portion of its LF, will be indistinguishable from (51). β is the X^0-complex formed from Agr_O and the matrix V.

(52)(a) [$_{\text{Agr}_\text{O}}$ AN [FF (*two men* β]]
(b) [$_{\text{Agr}_\text{O}}$ FF (*two men* [AN β]]

Since Chomsky's account is faulty, I will now consider two other possible ways out of the dilemma.[20] First, there might be a crucial distinction between Case and other formal features. Suppose Chomsky's proposal concerning feature movement is correct, but only for agreement features. That is, agreement features could be checked via adjunction of those features to an agreement head. In existential and unaccusative constructions, if it is only agreement features that need to be checked by movement (under a "partitive" type approach to Case), referential and quantificational properties would be left behind, with only the formal features raising.

(Structural) Case features, on the other hand, might be checkable only in a Spec–head configuration, as in "classic" Minimalist work. ECM constructions would involve Case-driven covert raising, under the standard Minimalist assumption that Spec of an ECM infinitive is not a Case position. On the suggested hypothesis, the raising would be of the entire NP (exactly as with overt subject raising to subject position), yielding all of the observed parallelisms with raising to subject position.

This approach postulates a new asymmetry between two classes of formal features: agreement features vs Case features. A second approach would rely, instead, on the *already* postulated distinction between overt and covert movement. Recall Chomsky's proposal that for PF reasons, overt movement is always of a category, not just formal features, while covert movement is merely of formal features, since the entire category need not (hence, must not) move. The relevant movement in the *there* constructions considered above is covert, so the account of those constructions is exactly as in the first approach. Only the features move, so for all

other purposes, it is as if no movement took place. For ECM constructions, also, the standard Minimalist assumption is that the movement is covert. This was the source of the paradox. But Koizumi (1993), revising and extending ideas of Johnson (1991), argues that accusative Case is checked overtly in English, just like nominative Case. The accusative NP overtly raises to [Spec, Agr_O] (with V raising to a still higher head position). If this is correct, the seemingly paradoxical asymmetry is immediately reduced to the independent pied-piping asymmetry. In the *there* construction, the only movement is the covert movement of the formal features of the associate to the Agr head. For an ECM subject (or, for that matter, the object of a simple transitive) the movement is overt, hence, of the entire NP.

Both of these approaches correctly entail that, among the NPs considered so far, only the associate of *there* shows lower behavior.[21] All the others show higher behavior. It is also worthy of note that on both of the approaches, the Case of the associate of *there* is apparently licensed independently of *there*, as in the approach of Belletti (1988).

3.6. ACD: a case for overt raising

A further distinctive property of the movement hypothesized to be involved in *there* constructions is noted by Hornstein (1994). Hornstein argues that the apparent infinite regress in antecedent contained deletion (ACD) constructions is resolved not by QR, as on the classic account of May (1985), but by raising to [Spec, Agr_O]. Lasnik (1993) and Takahashi (1996) present similar arguments. As Hornstein observes, this provides a derivation for (53).

(53) John expected [[noone that I did [$_{VP}$e]] to be electable]

In the structure given in (53), the null VP seems to be contained within its antecedent, the larger VP headed by *expected*. But *if* [*noone that I did* [$_{VP}$e]] raises to [Spec, Agr_O] above *expected*, the regress can be avoided. Hornstein notes that (53) contrasts sharply with (54).

(54) *John expected [there to be noone that I did electable]

Hornstein concludes that "there is no expletive replacement.... If expletive replacement obtains, then at LF [(53)] and [(54)] should have analogos structures with *there* and *noone that I did* forming a complex and raising to the matrix Spec Agr_O for Case checking." Hornstein does not indicate how he proposes to deal with the standard arguments, going back to Chomsky (1986b), *for* expletive replacement, namely the agreement facts mentioned above and the A-chain like locality between expletive and associated argument. Again, there is a seeming paradox. There is a strong argument that movement is not involved, and an equally strong argument that it is. But the resolution of the paradox is already at hand. In fact, we have two alternative resolutions. On both of the approaches to

Case and agreement sketched above, only the formal features of the associate of *there* move. Consequently, *noone that I did* in (54) remains in situ, leaving the null VP internal to its antecedent. Further, on both of the approaches, in (53) that NP necessarily raises out of the VP, freeing the null VP from its antecedent.[22]

We have so far seen several reasons for believing that agreement features can be checked in an agreement head position while Case features are checked only in a Spec-head relation.[23] At this point, there is no basis for choosing between the overt and covert movement accounts of accusative Case. The ACD facts, though, can form the basis for an argument in favor of the overt movement account. If Hornstein (1994) and Lasnik (1993) are correct that at least some instances of ACD involve raising to [Spec, Agr_O], and if Chomsky and Lasnik (1993) are correct that ellipsis is a PF deletion process rather than an LF copying one,[24] then, at least in those particular ACD constructions, movement for the checking of accusative Case must be overt. One might then expect that it always is.

As discussed by Lasnik (1993), the basic motivation for the Case analysis of ACD comes from examples involving appositive relative clauses, since the distribution of ACDs in restrictive relatives is much freer. May (1985) claimed that appositives do not support ACD at all, citing (55), which contrasts with (56).

(55) *Dulles suspected Philby, who Angleton did
(56) Dulles suspected everyone Angleton did

This was one of May's major arguments for analyzing ACD in terms of QR. The direct object in (56) is quantificational while that in (55) is not, so in the latter, QR cannot apply and the ellipsis regress cannot be resolved. However, Wyngaerd and Zwart (1991) show that examples structurally parallel to (55) are quite tolerable, a conclusion accepted by Fiengo and May (1992, 1994). Two such examples are trivial modifications of (55) itself:

(57) ?Dulles suspected Philby, who Angleton did not
(58) ?Dulles suspected Philby, who Angleton did as well

What is shown in Lasnik (1993) is that the possibility of this type of ACD construction (unlike the standard restrictive instances) rather strongly correlates with the possibility of structural accusative Case. For example, objects of prepositions typically allow restrictive ACD but not appositive ACD:

(59)(a) Mary stood near everyone Emily did
(b) *Mary stood near Susan, who Emily did as well

Further, for many speakers, the possibility of appositive ACD correlates with the possibility of pseudo-passivization ("reanalysis"), providing some additional evidence, under standard assumptions, that accusative Case is playing a role. Representative

paradigms are presented immediately below. In (60)–(6l), we find acceptable ACD and acceptable pseudo-passive.

(60)(a) ?Dulles spoke to Philby, who Angleton did as well
(b) ?Dulles talked about Philby, who Angleton did as well
(c) ?John took advantage of Bill, who Mary will as well

(61)(a) Philby was spoken to
(b) Philby was talked about
(c) Bill was taken advantage of

And in (62)–(63), we find unacceptable ACD and unacceptable pseudo-passive.[25]

(62)(a) *Mary stood near Susan, who Emily did as well
(b) *John showed Bill Harry, who Mary will as well

(63)(a) *Susan was stood near (by Mary)
(b) *Harry was shown Bill *t*

To consider one case, (63a) suggests that *stand near* cannot reanalyze. Plausibly, a consequence of this inability is that the Case of the object of *near* will not be licensed in [Spec, Agr_O], but rather, internal to the PP (or perhaps in the Spec of some functional projection just above the PP). The elided VP internal to that NP will thus remain internal to its antecedent VP and unable to escape the resolution regress.

Notice that two obvious problems remain for a PF deletion account of the ellipsis phenomena under consideration. By hypothesis, in (57), the direct object *Philby, who Angleton did not* has raised overtly to [Spec, Agr_O]. As mentioned earlier, we must therefore conclude that the main verb *suspected* has also raised overtly to a still higher position. Further, internal to the relative clause, we can then assume that the variable bound by the relative operator also has raised to [Spec, Agr_O] in its clause. The verb in the relative clause, however, must *not* have raised to a higher position, since then VP deletion would not result in the elimination of that verb. The simpler of the two problems is as follows: How can the main VP, out of which V has raised, antecede the deletion of the VP in the relative clause, which still has its V in situ (in order that it is deleted along with the VP)? One straightforward possibility is that the trace of the raised V counts as equivalent to the V itself. This is immediate under the "copy theory" of movement, as advocated recently by Chomsky (1993), in a revival of much earlier approaches.

The harder problem is an immediate consequence of the simpler one. We have now seen evidence that raising to [Spec, Agr_O] can be overt in English. This indicates, given the normal word order of English, that raising of V to a higher head is also overt. However, even though the direct object did raise out of the deleted VP in the constructions just examined, the V did *not* raise out of that VP.[26] Hence,

it is not clear why (64) should not also be possible, with overtly raised object *Philby* and V in situ.

(64) *Dulles Philby suspected

Even worse, Procrastinate should then block (65), where, by hypothesis, raising of V is overt

(65) Dulles suspected Philby

Under Chomsky's program, we must identify strong features forcing the overt movement of V and object NP. (The latter is technically straightforward: A strong NP feature can be attributed to Agr_O. In this direction lies a unification of the EPP, the EPP feature being a property of Agr in general in English.) And then, we must find a reason why the failure of the strong feature driving raising of the V to be overtly checked causes the derivation to crash in (64) but not in the ACD examples.

I will leave open here the precise characterization of the strong feature driving raising of the V, and the identity of the head to which V raises.[27] But suppose that the relevant strong feature is a feature of the **V** itself.[28] And suppose, following Chomsky (1993) but contra Chomsky (1994), that an unchecked strong feature is an ill-formed **PF** object (rather than an ill-formed LF object). Under the assumption that ellipsis phenomena truly do involve deletion, we have the following prediction: Deletion of (a category containing) an item with an unchecked strong feature salvages a derivation that would otherwise crash at PF. The portion of the structure that would have caused the PF crash is literally gone at that level. In the present case, the strong feature of *suspected* in (64) is not checked overtly, so the PF is ill-formed. In (57), repeated as (66), on the other hand, the unraised *suspected* does not survive to the level of PF, as it is deleted.

(66) ?Dulles suspected Philby, who Angleton did not

The PF is then predicted to be well-formed. What of the LF? Recall that *all* checkable features, whether strong or weak, must be checked by the LF interface. There is no difficulty here; the LF will also be well-formed, since in the LF component, the V (or its formal feature bundle) can raise, checking its own checkable features and those of the functional heads it raises to.[29]

3.7. Conclusion

In conclusion, I have argued for a version of the last resort condition that incorporates the basic property of enlightened self-interest: an instance of movement must be for the satisfaction of some feature, where that feature might be of the moved item or of the target. This accords reasonably well with the "Attract F"

view of movement. The condition still embodies one residual element of Greed: that an NP (or its set of formal features) is no longer available for "A-movement" when its Case has been checked off. Finally, in terms of the feature movement theory of (LF) movement, I have suggested, based on binding and ellipsis phenomena, that the movement involved in accusative Case checking in English is overt, as argued by Koizumi (1993) roughly following Johnson (1991). While the general direction seems quite promising, more comprehensive Minimalist theories of Case licensing, and, indeed, of anaphora and of ellipsis, await further investigation.

Acknowledgment

I am indebted to Željko Bošković and Roger Martin for numerous helpful suggestions at every stage of this research and to Željko Bošković for detailed comments on an earlier draft.

4

LEVELS OF REPRESENTATION AND THE ELEMENTS OF ANAPHORA[1]

Here I explore the question of what level of representation is relevant to the determination of anaphoric relations. Of particular concern are the implications of the feature movement theory for the theory of anaphora. The overall generalization is that anaphora seems to crucially depend on S-structure configuration. This, of course, is extremely problematic for minimalism, since that framework denies the existence of S-structure as a significant level of representation. The solution to the problem is that the covert portion of the derivation of LF moves only formal features. These feature movement operations "leave behind" the semantic aspects of nominal expressions crucial to anaphora. Thus, anaphoric connection is determined at LF, but it is as if it is determined at "S-structure." Chomsky (1995a) had argued that LF operations *do* affect anaphoric relations, but his only example involves control. Attempts to replicate that example with other relations invariably fails. Further, I show that even the pattern with control is not general.

One of the most fascinating, and most enduring, topics in the generative investigation of anaphora is the question of levels of representation relevant to the determination of anaphoric connection. In terms of the influential theory of anaphora of Chomsky (1981), with its three binding conditions, the question concerns the level or levels of representation that must satisfy those conditions. Over the years, a variety of technological proposals have been put forward, none of them conceptually very satisfactory. Recent "minimalist" characterizations of the general form of syntactic theory render most of those proposals not just unsatisfactory but unstatable, since the proposals crucially rely on a level of representation, S-structure, which is claimed not to exist. I will consider the quite strong evidence motivating those early proposals and ultimately suggest a way of capturing the apparent S-structure effects without direct appeal to S-structure, thus responding to a potentially powerful counterargument to a major tenet of minimalism.

4.1. Are anaphors licensed at S-structure?

The fact that anaphora has obvious semantic aspects has always suggested that its syntax, the binding conditions, should be determined at LF, the syntax–semantics interface level. However, early on, difficulties with such a theory were recognized. Discussing the following examples, Chomsky (1981: 196–7) argues that Condition C must apply at S-structure:

(1) Which book that $John_i$ read did he_i like?
(2) *He_i liked every book that $John_i$ read.
(3) *I don't remember who thinks that he_i read which book that $John_i$ likes.

Chomsky's point is that following QR, the LF of (2) would be structurally parallel to the S-structure (and LF) of (1), where *John* is outside the c-command domain of *he*. Thus, as in (1), there should be no Condition C effect if LF is the level relevant to that condition. Similarly for (3) following LF *wh*-movement. Contrary to the prediction of the LF theory, in both instances the hypothesized LF movement, unlike the overt movement creating (1), has no effect on binding possibilities. This strongly suggests Chomsky's conclusion: that Condition C is specifically a requirement on S-structure.

The preceding argument is based on what might be termed unexpected ungrammaticality. Chomsky also offers an argument, to the same conclusion, based on unexpected grammaticality. Chomsky proposes that (4) is the target of an LF rule of focus movement.

(4) John said that Bill had seen HIM.

When *HIM* raises, it leaves behind a variable, which, according to Chomsky's account of strong crossover, must be A-free as dictated by Condition C. Since in the postulated LF, the variable would be bound by *John* on the relevant (and possible) reading, we have an additional argument that Condition C is sensitive to properties of S-structure rather than of LF.

Barss (1986) draws the same conclusion for Condition A, based on examples like the following:

(5) $John_i$ wonders which picture of $himself_i$ Mary showed to Susan.
(6) *$John_i$ wonders who showed which picture of $himself_i$ to Susan.

(5) shows that an anaphor within the embedded [Spec, C] can be licensed by an antecedent in the matrix subject position. Given this fact, the ungrammaticality of (6) is surprising if anaphors can be licensed by virtue of their LF positions. On the then standard theory, in LF, the WH-phrase in situ, *which picture of himself*, moves to the embedded [Spec, C] position, where it takes scope. Thus, at LF, the configurational relation between *himself* and its antecedent is virtually identical in

(5) and (6). Hence, the ungrammaticality of (6) shows that anaphors must be licensed at a level prior to LF, for example, S-structure. (6) is ruled out because the reflexive fails to be licensed at that level.

Chomsky (1981) explores another phenomenon with potential implications for the issues under discussion here. Assuming that expletive *there* is co-indexed with its associate, Chomsky observes that a perfectly grammatical example like (7) incorrectly constitutes an apparent violation of Condition C.

(7) There is a man in the room.

The proposal of Chomsky (1981) is that the co-indexing between expletive and associate is of a different sort than that involved in Binding Theory (in particular, co-superscripting as opposed to co-subscripting). Chomsky (1986b) reconsiders the phenomenon, and presents a more attractive account: that in the LF component, the associate A moves to the position of the expletive, as illustrated in (8).

(8) A man is *t* in the room.

Since *t* is here the trace of A-movement, it does not fall under Condition C, so the example is correctly let in. It is significant that this account crucially depends on the rejection of the Chomsky (1981) argument. If S-structure must satisfy Condition C, then later movement will be of no benefit. As I will discuss at length below, independent of the status of the Chomsky (1986b) argument, the same problem arises within the Minimalist Program of Chomsky (1995b), since that program denies S-structure any significant status.

Uriagereka (1988), as part of an extensive discussion of Binding Theory and levels of representation, pursues the Chomsky (1986b) expletive replacement proposal further. Based on the acceptability of (9), Uriagereka proposes that anaphors need not be licensed at S-structure, reasoning that the required c-command relation between *two knights* and *each other* holds at LF (10) but not at S-structure (9).[2]

(9) There arrived two knights on each other's horses.
(10) Two knights arrived *t* on each other's horses.

This interesting phenomenon is actually independent of expletive replacement *per se*, since, as discussed in some detail by Lasnik and Saito (1991), even direct objects of transitive verbs generally seem to c-command certain adjuncts.[3]

(11) I saw two men on each other's birthdays.

In certain versions of Case theory, for example, one proposed by Chomsky (1991), and examined further by Lasnik and Saito (1991) and Lasnik (1993), in the LF component a direct object raises to [Spec, Agr_O], where its Case is licensed. On standard assumptions about phrase structure, (11) might then be a further example

of successful licensing of an anaphor at LF (via A-movement) remedying failure in overt syntax. Below, I will have much more to say about examples like (9) and (11). For the moment, I simply point out the apparent contradiction between this phenomenon and the Chomsky (1981)/Barss (1986) argument that binding conditions must be satisfied at S-structure.

4.2. Are anaphors licensed at LF?

As noted in Lasnik (1993), one way to resolve the contradiction is to reject the existence of the LF operations that Chomsky (1981) was assuming. A somewhat less far reaching alternative, also mentioned in Lasnik (1993), is to limit the operation of such rules as QR and LF *wh*-movement so that only the quantificational head moves, and not the entire expression, as proposed by Hornstein and Weinberg (1990).[4] Thus, the generalization put forward by Chomsky (1981) might be captured as a property specifically of A′-movement, without Chomsky's conclusion that the binding conditions hold of S-structure:

> Such examples indicate that syntactic [overt] movement and movement in the LF-component have quite different effects with respect to the binding theory. This theory applies properly after syntactic movement, but each rule of the LF component converts S-structures to which the binding theory applies correctly to LF-representation to which it applies incorrectly.
>
> (Chomsky 1981: 197)

The next question involves LF A-movement, as postulated for (9) and (11). As noted above, Uriagereka suggests expletive replacement as an account for the licit binding of the reciprocal in (9). However, Lasnik and Saito (1991) and den Dikken (1995) show that expletive replacement does *not* in general create new binding possibilities. For example, it is well known that a raised subject can antecede an anaphor in the higher clause (the classic argument that D-structure need not obey Condition A), but the associate of a raised expletive cannot:

(12) Some linguists seem to each other [*t* to have been given good job offers].
(13) *There seem to each other [*t* to have been some linguists given good job offers].

Other phenomena known to involve A-binding pattern similarly. (14) illustrates this for weak crossover.

(14)(a) Some defendant$_i$ seems to his$_i$ lawyer to have been at the scene.
(b) *There seems to his$_i$ lawyer to have been some defendant$_i$ at the scene.

This set of facts indicates that if LF is the level of representation relevant to the licensing of anaphors and bound pronouns, then literal expletive replacement

cannot be correct. Interestingly, Chomsky (1991) already arrived at the same conclusion based on scope properties of existential constructions. Substitution would result in identical LFs for (15) and (16).

(15) A man is likely to be here.
(16) There is likely to be a man here.

But the interpretive possibilities diverge. In (15), *a man* can evidently have wide or narrow scope with respect to *likely*, while (16) allows only narrow scope for a *man*. The same sort of interpretive divergence arises with respect to scope of negation:

(17) Many linguistics students aren't here.
(18) There aren't many linguistics students here.

Partly for these reasons, Chomsky (1991) modified his (1986b) substitution analysis, proposing instead that the associate adjoins to *there*, the latter being a sort of LF clitic.

The generalization that emerges from all of this is that the associate of *there* always displays "low" behavior, while an overtly moved NP displays "high" behavior. It is perhaps tempting to conclude that there actually is no movement relation between the position of *there* and that of the associate. However, the standard argument for movement is a compelling one: that it provides the basis for an account of the familiar superficially bizarre agreement paradigms displayed by these constructions, with the verb agreeing with something that is not its formal subject;[5]

(19)(a) There is/*are a man here.
(b) There are/*is men here.

Hence, I will pursue the problem of low behavior for the associate in another way. In particular, I will explore the nature of the movement, exactly what moves, and how this interacts with the elements of the theory of anaphora.

4.3. Movement and binding in existentials

4.3.1. C-command asymmetrics

Lasnik and Saito (1991) show that existential constructions in Exceptional Case Marking (ECM) configurations display asymmetries parallel to those seen in subject raising constructions:[6]

(20) The DA proved [two men to have been at the scene] during each other's trials.
(21) *The DA proved [there to have been two men at the scene] during each other's trials.

Recall that in connection with Uriagereka's example above I noted that complements of verbs, whether the verbs are unaccusative (9) or transitive (11), are able to bind into adjuncts. I also noted that under standard assumptions about phrase structure, those direct objects must be assumed to undergo A-movement to some higher position for that binding to take place. However, in a theory of phrase structure of the sort proposed by Larson (1988), there might have been an alternative, since in such a theory, adjuncts are actually base generated lower than complements.[7] It is in this connection that the examples examined by Lasnik and Saito become particularly significant, since for those examples no Larsonian alternative is even possible. There is no way that an adjunct modifying the higher clause can be base generated lower than an argument of the lower clause (e.g. the ECM subject in (20)). We are led to the conclusion of Lasnik and Saito, and of Postal (1974) before them (and based on similar arguments): that the ECM subject undergoes raising.[8] The associate of *there* must then undergo raising of a quite different sort.

It is important to note that the asymmetry seen in (20)–(21) is not limited to reciprocal licensing. All standard c-command phenomena display the same pattern. (22)–(23) illustrate the corresponding weak crossover asymmetry:

(22) The DA proved [no suspect$_i$ to be at the scene of the crime] during his$_i$ trial.
(23) *The DA proved [there to be no suspect$_i$ at the scene of the crime] during his$_i$ trial.

And negative polarity item licensing is shown in (24)–(25).

(24) The DA proved [noone to be at the scene] during any of the trials.
(25) *The DA proved [there to be noone at the scene] during any of the trials.

4.3.2. An "expletive adjunction" analysis

I turn now to possible accounts of the several parallel asymmetries seen above. The first possibility is based entirely on Chomsky (1991). Suppose, as hinted above, that accusative Case is licensed in the [Spec, Agr$_O$] above the licensing verb, via LF A-movement.[9] If we assume, in the spirit of the Minimalist program, that all of the c-command phenomena considered above involve LF, this immediately gives the correct results for direct objects, for ECM subjects, and, of course, for overtly raised subjects (where there is no relevant difference between the S-structure and the LF). Further, the version of expletive replacement espoused in Chomsky (1991) – adjunction to *there*, rather than substitution for it – potentially makes the necessary distinction between NPs with high behavior on the one hand and associates of *there* on the other. The latter will *adjoin* to *there*, hence arguably will not be in the appropriate position to c-command the anaphors, Negative Polarity Items (NPIs), etc. in the examples above. This is a natural extension of the Chomsky (1991) account of scope facts in existential constructions. There are, however, problems with that account. First, as Chomsky points out, in his example cited

as (18) above, repeated as (26), it is not the case that there is no scope relation between negation and *many linguistics students*.

(26) There aren't many linguistics students here.

Rather, *many linguistics students* necessarily has narrow scope. While it is true that if that NP were to replace *there*, it would be expected to have (at least as one possibility) wide scope, under the adjunction analysis, according to Chomsky no scope relation is established between negation and *many linguistics students*. Chomsky indicates that under this circumstance, the scope of *many linguistics* students can be assumed to be narrow. Chomsky analogizes this situation to that in (27).

(27) Pictures of many students aren't here.

However, in (27), there truly is no scope relation between negation and *many students*. The sentence is clearly not synonymous with (28).

(28) Pictures of few students are here.

But such synonymy would be expected on the implied account of (26), since that example is synonymous with (29).

(29) There are few linguistics students here.

In addition to this empirical problem, there is a technical problem. Chomsky evidently bases his argument that no scope relation is established between negation and *many linguistics students* in the LF of (26) on the assumption that there is no c-command relation between those two expressions. However, on the theory of adjunction proposed by May (1985), developed in Chomsky (1986a), and assumed in all of Chomsky's writings since, there would be a relevant c-command relation in (26). *Many linguistics students* would c-command negation, just as much as it would in (30).

(30) Many linguistics students aren't here.

This is so because on May's and Chomsky's theory of adjunction, when α adjoins to β, β becomes a segmented category, and α c-commands anything β did prior to the adjunction. Thus, the scope problem that largely motivated the change from expletive substitution to expletive adjunction is actually not resolved by that change. This indicates that the solution to the asymmetries lies elsewhere.

4.3.3. A feature–movement analysis

Chomsky (1995b) suggests a revised theory of the LF movement involved in expletive constructions, as part of a revised theory of LF movement more generally.

Beginning with the standard Minimalist assumption that all movement is driven by the need for formal features to be checked, Chomsky argues that, all else equal, movement should then never be of an entire syntactic category, but only of its formal features.[10] PF requirements will normally force movement of a category containing the formal features, via a sort of pied-piping, under the reasonable assumption that a bare feature (or set of features) is an ill-formed PF object. For LF movement, on the other hand, pied-piping will normally not be necessary, hence, by economy, will not even be possible. Only the formal features will move, and they will move exactly to the heads that have matching features. In a standard existential sentence, then, the associate does not move to *there*. Rather, only the formal features move, and only to a corresponding functional head (or heads). As observed by Chomsky, the feature movement analysis of existential constructions has the potential to solve the scope problem. If in LF, only the formal features of *many linguistics students*, rather than the entire expression, move to a functional head or heads above negation (presumably the Agr_S head), it is reasonable to conclude that the quantificational properties remain below negation. Then, if it is this structure that determines scope, the desired results are obtained.

The feature movement analysis, being in a sense a more extreme version of the Hornstein and Weinberg proposal mentioned above, would seem to properly handle the anaphora and NPI paradigms in *there* constructions as well. Recall the reciprocal facts of (12)–(13), repeated as (31)–(32).

(31) Some linguists seem to each other [*t* to have been given good job offers].
(32) *There seem to each other [*t* to have been some linguists given good job offers].

When the entire NP has raised, as in (31), whatever properties of the NP are relevant to licensing an anaphor are in the appropriate structural position to do so (both at S-structure and at LF). When, by hypothesis, only the formal features have raised, as in (32), it is reasonable to conjecture that the referential properties relevant to licensing an anaphor remain below.[11] Parallel accounts could be provided for NPI licensing and Weak Crossover (WCO) as in (33a,b) vs (34a,b).

(33)(a) No good linguistic theories seem to any philosophers [*t* to have been formulated].
(b) Some defendant$_i$ seems to his$_i$ lawyer to have been at the scene.

(34)(a) *There seem to any philosophers [*t* to have been no good linguistic theories formulated].
(b) *There seems to his$_i$ lawyer to have been some defendant$_i$ at the scene.

4.3.4. Is there more to anaphora than formal features?

On this kind of account, then, the elements of the theory of anaphora are not merely formal features. Interestingly, Chomsky (1995b) proposes just the

opposite: that the elements of anaphora are precisely formal features. As far as I can tell, there are two bases for Chomsky's position. One to which I will return involves control. The other concerns (some of) the Lasnik and Saito ECM facts mentioned above. Recall that the ECM subject licenses elements in the higher clause. Further, direct object licenses elements that are arguably base generated higher than direct object. In both situations, the formal features of the licensing NP are assumed to undergo LF raising to Agr_O. Chomsky thus reasons that "the features adjoined to Agr_O … have A-position properties, c-commanding and binding in the standard way." (Chomsky 1995: 272) Thus, for all purposes (except scope), feature movement is claimed to have the same consequences as NP movement. However, we have seen overwhelming evidence that this is not so. The (low) associate of *there* cannot bind a (high) reciprocal or license a (high) NPI or a (high) bound pronoun, even though the formal features of the associate raise to the appropriate higher position. For the first of these problems, Chomsky offers a solution. Assuming the LF anaphor movement analysis of Condition A effects, Chomsky indicates that the LF structure of the relevant portion of an example like (32) would be (35(a)) or (b), where AN is the anaphor, FF is the set of raised formal features, and α is the X^0-complex formed from INFL and the matrix V.

(35)(a) $[_{INFL}$AN [FF (*linguists*) α]]
(b) $[_{INFL}$FF (*linguists*) [AN α]]

Chomsky concludes that "On reasonable assumptions, neither of these structures qualifies as a legitimate binding-theoretic configuration, with AN taking FF (*linguists*) as its antecedent" (Chomsky 1995b: 275–6). Chomsky does not mention failure of NPI licensing and parallel failure of bound pronoun licensing (WCO effects). It is not clear that Chomsky's account of anaphora failure would extend to these phenomena, particularly in light of the fact that, other than c-command, the relevant structural relations are presumably different. Further, there is a more fundamental problem with the account. While it might correctly distinguish among the subject raising to subject position examples, disallowing anaphor binding when it is *there* that overtly raises, it incorrectly excludes *all* anaphor binding with "subject raising to object position." For example, as noted above, Chomsky, following Lasnik and Saito, takes a sentence just like (20) to be acceptable, but (36), the relevant portion of its LF, will be indistinguishable from (35). β is the X^0-complex formed from Agr_O and the matrix V.

(36)(a) $[_{Agr_O}$AN [FF (*two men* [β]]
(b) $[_{Agr_O}$FF (*two men* [AN β]]

There is, thus, compelling reason for rejecting the idea that the (sole) elements of anaphora are formal features. Before turning to Chomsky's argument, based on control, for the position that I am rejecting, I would like to sketch a theory (perhaps the only natural theory) that can handle all of the facts considered so far.

Note first that overt raising to subject position is unproblematic on any of the accounts. The raised NP displays high behavior in all respects, exactly as expected, on the uncontroversial assumption that all properties of the NP are raised with it. Next, the associate of *there*, in simple existential constructions, in subject raising constructions, and in ECM constructions always displays low behavior, except for the agreement properties of the higher Agr. This is straightforward under the hypotheses that (i) the formal features, and only the formal features, of the associate raise; and (ii) that anaphors, NPIs, and bound pronouns are licensed by properties (plausibly semantic ones) other than (or in addition to) formal features. The one remaining problem is the disparity between direct objects and ECM subjects, which can license apparently higher items, and *there* in an ECM configuration, which cannot. The latter is unproblematic. The associate of *there* remains low, even if its formal features move high, possibly as a consequence of featural deficiency of *there* itself. The former, on the other hand, is deeply problematic. Under standard Minimalist assumptions, there is movement involved in these constructions, movement for Case checking to Agr_O, but *covert* movement, so only the formal features move. And, as we have seen in detail, features do not suffice to license the items under investigation here. Thus, there is a paradox on those standard assumptions. The crucial standard assumption can be questioned. Koizumi (1993, 1995), revising and extending ideas of Johnson (1991), argues that accusative Case is checked overtly in English, just like nominative Case.[12] The accusative NP overtly raises to [Spec, Agr_O] (with V raising to a still higher head position).[13] If this is correct, the seemingly paradoxical asymmetry is immediately reduced to the independent pied-piping asymmetry. In the *there* construction, the only movement is the covert movement of the formal features of the associate to the Agr head. For an ECM subject or the object of a transitive or unaccusative verb the movement is overt, hence, of the entire NP, with all its semantic/referential properties necessary for licensing anaphors, NPIs, and bound pronouns. Interestingly, this resolves a paradox discussed by Lasnik and Saito (1991), as well. Lasnik and Saito were concerned to explain how an ECM subject, or a direct object, is high enough to c-command all the licensees outlined above. They suggested LF raising to Agr_O as a possibility. However, they also noted the existence of overwhelming evidence that many of these licensing effects crucially involve S-structure configuration. The present approach, based on Koizumi's analysis, immediately resolves this tension (a tension made extreme under the Minimalist claim that there is no significant level of S-structure). The licensing is at LF, but is as if at S-structure, since the only relevant movement is overt. Covert feature-driven movement, involving merely formal features, is incapable of creating new licensing configurations for anaphora, etc.

I must now return to Chomsky's contrary argument that the elements of anaphora are indeed formal features. As noted above, there are actually two aspects to the argument. The first, based on the grammatical examples of Lasnik and Saito, is entirely theory internal, relying on the assumption that raising to Agr_O is covert in English. Since that approach to the ECM phenomena led to

a flat contradiction, I rejected it in favor of Koizumi's *overt* raising account, an account that has considerable independent motivation. Chomsky's empirical argument involves control. He presents an example suggesting that the associate of *there* behaves as if it is high, in its ability to control PRO in an adjunct:

(37) There arrived three men (last night) without [PRO] identifying themselves.

As (38) shows, a typical object cannot control PRO in this construction:

(38) *I met three men (last night) without identifying themselves.

Thus, Chomsky reasons that feature raising, in this instance to Agr_S, does create new control configurations, and, a fortiori, new binding configurations. Momentarily, I will question the generality of the control phenomenon, but even if Chomsky turns out to be correct about control, it will not be necessary to draw a broader conclusion about binding. There are significant and well-known differences between control and binding. For example, as discussed by Lasnik (1992b), even languages unlike English in having "subject oriented" anaphors still have structures of "object control".[14,15] (39)–(40) illustrate this for Polish.

(39) Jan_i opowiadał Marii_j o $\text{swoim}_{i/*j}$ zachowaniu.
John telling Mary about self's behavior
'John was telling Mary about his/*her behavior.'

(40) Jan_i kazał Marii_j [$\text{PRO}_{j/*i}$ napisać artykuł].
John told Mary write article
'John told Mary to write an article.'

Further, control has thematic aspects that are lacking in other instances of anaphora, yet another reason for drawing a distinction.[16]

4.3.5. More on control

Before concluding this discussion, I would like to further explore the intriguing control phenomenon that Chomsky notes. Chomsky implies that the associate of *there* in (37) is behaving just as an overtly raised subject, as in (41), would.

(41) Three men arrived (last night) without PRO identifying themselves.

But already there is some difference. While (41) is perfect, (37) is somewhat degraded for many speakers.[17] This contrast is heightened if the adverbial is fronted:

(42) Without PRO identifying themselves, three men arrived.
(43) ?*Without identifying themselves, there arrived three men.

Further, under raising, the contrast between structures like (37) and (41) greatly intensifies. In the following examples, the adverbial is intended as being in the higher clause, along with the raised subject or *there*:

(44) Someone seems to be available without PRO seeming to be eager to get the job.
(45) *There seems to be someone available without PRO seeming to be eager to get the job.

There is no reason this should be so if raised features of the associate of *there* can control. Finally, while (38) is clearly bad, it is not only complements of unaccusatives that are reasonably acceptable as controllers. The following example is considerably better than (38), even if not quite as good as (37):

(46) The news upset John while reading the paper.

Perhaps thematic properties are involved in control into an adjunct (as they are often assumed to be in control into a complement): the object is a possible controller because the subject is too low on the thematic hierarchy. In (37), the subject is not thematic at all. In (46), it is low relative to the subject.

Until this array of facts is sorted out, an interesting typological claim made by Chomsky must be held in abeyance. Chomsky indicates that languages with expletives of the *there* type (i.e. no agreement features of their own) allow control in the constructions at issue, while languages with expletives of the *il* type (with agreement features of their own) do not.[18] This is because there will be raising of the features of the associate in the former instance but not the latter. Thus, he claims that Italian (with its *there* type expletive) and French contrast:

(47) Sono entrati tre uomini senza identificarsi.
(48) (*)Il est entré trois hommes sans s'annoncer.

But again, there is some question about the data. Every French speaker I have collected data from finds control in (48) fairly acceptable or completely so.[19] Thus, it is not so clear that French and English contrast. Resolution of these important issues awaits further exploration of control in general and in the Minimalist framework in particular.

4.4. Conclusion

The tentative conclusion of this investigation is that for phenomena known to involve c-command, LF movement creates no new licensing possibilities.[20] This result is exactly in accord with that of Chomsky (1981), though now for a somewhat different reason. Then, it was because, by stipulation, the licensing conditions had to be satisfied at S-structure. Such a stipulation is unavailable within the

Minimalist framework. Instead, there is the independently plausible asymmetry between covert movement (strictly of formal features) and overt movement (of an entire category, via pied piping). Under the assumption that these licensing phenomena involve referential and quantificational properties, and not just formal features, the correct result is obtained.[21] For control, a phenomenon that might or might not involve c-command, the pattern is apparently different in certain respects. Whether its licensing is as expected under LF feature movement remains to be determined.

5

PSEUDOGAPPING PUZZLES

In Lasnik (1995d), I proposed an analysis of Pseudogapping in which the right-hand remnant has raised to [Spec, Agr_O] and the VP from which it has escaped subsequently undergoes deletion. Here, I consider some further implications of that analysis, and some problems it raises. One of the problems is the unacceptability of certain English Pseudogapping-like constructions, but where the verb, too, has escaped the VP that will be deleted. I tentatively suggest an account in terms of a version of relativized minimality, after rejecting an analysis that would directly prohibit XP ellipsis when the head X has raised out of XP. This latter possibility is shown to be inconsistent with the behavior of VP-ellipsis in a number of V raising languages. Finally, I explore interactions between Pseudogapping and antecedent contained deletion (ACD), showing how some, but definitely not all, apparent instances of ACD can be analyzed as Pseudogapping. I should point out that the entire analysis in this chapter is stated in terms of the PF crash theory of strong features of Chomsky (1993). In Chapter 6, I reconsider the nature of strong features, showing how some facts that seemed to motivate the PF crash account can be restated in terms of the "online" crash theory of Chomsky (1995a), a theory that Chomsky argues is conceptually superior. Most of the phenomena considered in the present chapter are amenable to a treatment in terms of this revised theory of strong features. There is, however, one exception: The account I offer of the marginality of all Pseudogapping crucially relies on the PF crash theory of strong features. I have not yet been able to find an alternative.

In this chapter, I explore some of the properties of the so-called Pseudogapping construction. This construction is in important respects reminiscent of VP-ellipsis, except that it leaves behind an element of the VP as a remnant. I begin by summarizing the analysis of Lasnik (1995d), which is based on the important proposal of Jayaseelan (1990) that Pseudogapping is simply VP-ellipsis, with the remnant having moved out of the VP. I argue that Jayaseelan's basic proposal is correct,

except for the specific movement rule he invokes, heavy NP shift. I argue, instead that the movement rule involved is "object shift," overt raising to [Spec, Agr_O]. In the course of the presentation, I deal with a number of puzzles that arise, among them certain cases of apparent overgeneration. Finally, I consider the proposal of Bouton (1970) and Lappin (1992) that ACD is derived via Pseudogapping, and argue, along with Fiengo and May (1992), that the proposal is only half right, accounting for some ACD instances but not others. I conclude by discussing alternative hypotheses for the latter type of ACD.

5.1. General properties of Pseudogapping

The ellipsis phenomenon in (1) displays some properties of Gapping (there is a right-side remnant) alongside some properties of VP-ellipsis (there is a finite auxiliary):

(1) John will select me, and Bill will you.

Sag (1976) presents a number of instances, suggesting that they relate to VP deletion, and tentatively concluding that VP deletion must therefore be formulated as a rule deleting a variable (rather than specifically a VP), since a portion of the VP survives the deletion. The following are all from Sag (1976), with (4) and (5) cited from Halliday and Hasan (1973):

(2) John could pull you out of a plane, like he did Ø his brother.
(3) Mary hasn't dated Bill, but she has Ø Harry.
(4) Is she suing the hospital? She is Ø the doctor.
(5) Has he sold his collection yet? He has Ø some of his paintings; I'm not sure about the rest.
(6) Gee, I've never seen you on campus before. Yea! Neither have I Ø you.

Levin (1978, 1979/1986) provides an extensive examination of this type of ellipsis, and employs the name it is now standardly associated with: Pseudogapping. Among her many examples are the following, all from Levin (1978), and all marked? by her:

(7) If you don't believe me, you will Ø the weatherman.
(8) I rolled up a newspaper, and Lynn did Ø a magazine.
(9) Kathy likes astronomy, but she doesn't Ø meteorology.

By and large, the best instances of Pseudogapping involve an NP or PP remnant. Levin (1978) cites the following unacceptable examples with adjectival remnants:

(10) *You probably just feel relieved, but I do Ø jubilant.
(11) *Rona sounded annoyed, and Sue did Ø frustrated.
(12) These leeks taste terrible.
*Your steak will Ø better.

5.2. Toward an analysis

With this much as background, I turn now to a consideration of just what Pseudogapping is. While in many instances it might appear that the process is simply elision of the main verb, there is considerable evidence that more is involved. There are clear instances in which far more than just the main verb is elided:

(13) The DA proved Jones guilty and the Assistant DA will ~~prove~~ Smith ~~guilty~~.
(14) ?John gave Bill a lot of money, and Mary will ~~give~~ Susan ~~a lot of money~~.

Examples (2) and (6) above also display elision of more than the verb.

Rejecting the possibility of an ellipsis rule affecting a discontinuous portion of the structure (as seen in (13) and (14), for example), Jayaseelan (1990) proposes that Pseudogapping constructions result from VP-ellipsis, with the remnant having moved out of the VP by heavy NP shift. I will argue that this proposal is correct in its essentials, though wrong in certain details. In particular, I will begin by providing evidence that Pseudogapping does not entirely correlate with the possibility of heavy NP shift. I have already illustrated Pseudogapping with the first object in a double object construction as remnant. But the first object in a double object construction is resistant to undergoing HNPS:

(15) ?John gave Bill a lot of money, and Mary will ~~give~~ Susan ~~a lot of money~~.
(16) *John gave *t* a lot of money the fund for the preservation of VOS languages.

Conversely, the second object is a poor Pseudogapping remnant, but freely undergoes HNPS:

(17) *John gave Bill a lot of money, and Mary will ~~give Bill~~ a lot of advice.
(18) John gave Bill *t* yesterday more money than he had ever seen.

Jayaseelan's core idea, that Pseudogapping involves VP-ellipsis with prior movement of the remnant out of the VP, is very attractive, but an alternative to HNPS must be found if it is to be preserved. Note that in all the acceptable examples considered so far the remnant is accusative: either the direct object in a simple transitive construction, or the first object in a double object construction, or an exceptionally Case-marked subject of a complement. Given this, it is very tempting to posit raising to [Spec, Agr_O], as first suggested for accusative Case checking by Chomsky (1991), as the sought-after alternative to HNPS.

Under standard assumptions (though ones I will question shortly), raising of accusative NP to [Spec, Agr_O] is covert, taking place in the LF component. Given Jayaseelan's goal, adopted here, of analyzing Pseudogapping as affecting a constituent, this ellipsis process must then be analyzed as copying in the LF component, rather than deletion in the PF component. Consider (13), repeated as (19).

(19) The DA proved Jones guilty and the Assistant DA will ~~prove~~ Smith ~~guilty~~.

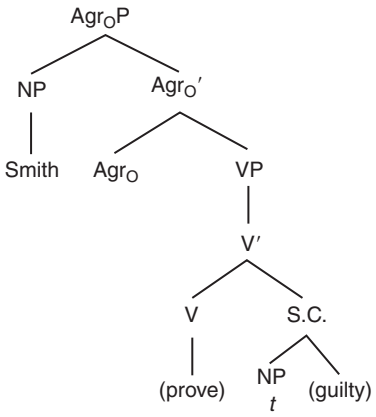

Figure 5.1 The relevant stage where the accusative NP has raised out of the "small clause" but the V has not yet raised.

By hypothesis, *prove guilty* is not a constituent in overt structure. However, in the LF component, following raising of *Smith*, the elided material could form a constituent. If the LF copying process can peer into the LF derivation, a possibility discussed by Hornstein (1994), then potentially there is a relevant stage where the accusative NP has raised out of the "small clause" but the V has not yet raised, as illustrated in Figure 5.1. The ellipsis site could then be VP.

One argument offered by Jayaseelan for his HNPS account has potentially significant implications for any [Spec, Agr$_O$] analysis of the remnant, so I turn to that argument now. Jayaseelan notes, following van Riemsdijk (1978) among others, that the object of a preposition may not undergo HNPS:

(20) *John counted on *t* for support a total stranger.

He then indicates that Pseudogapping is similarly impossible:

(21) *You can't count on a stranger; but you can ~~count on~~ a friend.

As Jayaseelan further notes, typical A-movement can strand the preposition in this construction:

(22) A total stranger was counted on *t* for support.

All else being equal, given the reanalysis responsible for (22), it is not clear why the object of the preposition cannot be a Pseudogapping remnant on a [Spec, Agr$_O$] account. The logic of the situation is clear, the facts perhaps less so. Levin

(1979/1986) indicates that objects of prepositions *can* be Pseudogapping remnants, and makes the provocative suggestion that the best cases "are likely those whose preposition forms a constituent with the verb rather than the following NP." Presumably she has reanalysis in mind. Levin's example (23), offered by her as acceptable, is consistent with that speculation.

(23) You have to sign onto it [the printer] like you do Ø the terminal.

The possibility of pseudopassive with this predicate indicates that reanalysis is available:

(24) The terminal must be signed onto.

The general patterning of data reported by my informants is in accord with Levin's suggestion. Judgments are delicate, since even the best instances of Pseudogapping are somewhat degraded, but they find a consistent correlation between Pseudogapping and pseudopassive. Their judgments, and my own, are that (25a) and (26a) are more acceptable than (27a) and (28a), in rough accord with the possibility of pseudopassive, as seen in the (b) examples.

(25)(a) John spoke to Bill and Mary should Susan.
(b) Bill was spoken to by John.

(26)(a) John talked about linguistics and Mary will philosophy.
(b) Linguistics was talked about by John.

(27)(a) * John swam beside Bill and Mary did Susan.
(b) *Bill was swum beside by John.

(28)(a) * John stood near Bill and Mary should Susan.
(b) *Bill was stood near by John.

Even more extreme instances of reanalysis, as in (29), support Pseudogapping (30) in the manner of (25) and (26) rather than (27) and (28).

(29) Bill was taken advantage of by John.
(30) John took advantage of Bill and Mary will Susan.

None of these structures, either the better ones, like (25), (26), (29), and (30), or the worse ones, like (27) and (28), support HNPS:

(31)(a) *John spoke to yesterday the man he met at the beach.
(b) *John talked about yesterday the man he met at the beach.
(c) *John took advantage of yesterday the man he met at the beach.
(d) *John swam beside yesterday the man he met at the beach.
(e) *John stood near yesterday the man he met at the beach.

These phenomena thus provide some additional evidence against an HNPS account of Pseudogapping and in favor of an A-movement account.

Earlier I indicated that under standard assumptions, the [Spec, Agr_O] analysis advocated here would require an LF copying theory of ellipsis, since the structure necessary for ellipsis is not created in overt syntax. However, on the theory of LF movement advocated by Chomsky (1995a), and further defended by Lasnik (1995b,c), the necessary structure would not even be created in *covert* syntax. On that theory, when movement is triggered by the need for formal features to be checked, all else equal only formal features move. When movement is overt (triggered by a *strong* feature), PF requirements demand that an entire constituent move, via a sort of pied-piping. However, when movement is covert, PF requirements are irrelevant, so economy dictates that movement *not* be of the entire constituent. But then it is very difficult to see how covert raising of (the formal features of) accusative NP to [Spec, Agr_O] could possibly create an ellipsis licensing configuration.

It seems then that if movement creates a configuration licensing ellipsis, the movement must be overt rather than covert. Before I explore how that might be possible in the present instance, I note that if the movement is overt, then the conclusion above, that ellipsis must involve LF copying, no longer follows. If the licensing configuration must be created prior to the LF/PF split, then ellipsis could just as easily be a PF deletion phenomenon. Interestingly, that sort of analysis of ellipsis has been consistently advocated by Chomsky (1994, 1995a), or, much earlier, in a 1971 lecture cited by Wasow (1972), where, according to Wasow, Chomsky "suggests that VP deletion and Sluicing can be formulated as very late rules which delete unstressed strings."

I have noted that the standard view of accusative Case checking in English is that it is facilitated by covert movement, but for raising to [Spec, Agr_O] to be the process making an NP into a Pseudogapping remnant, it must be overt. Further, the verb in the Pseudogapping construction must remain behind in the VP in overt syntax. This raises an important question: Is the special property of Pseudogapping that the accusative NP does raise overtly, or that the verb doesn't? I suggest the latter. Koizumi (1993, 1995), developing ideas of Johnson (1991), proposes that the relevant NP movement is always overt, and that (given the word order of English) the accusative checking V also raises overtly to a still higher position. Koizumi's specific proposal, which he calls the split VP hypothesis, is that V raises to a higher "shell" V position, as shown in Figure 5.2.

The raising of NP and the raising of V must both be driven by strong features. In Lasnik (1995b,c) I offer several arguments for a Koizumi-type approach, and I suggest that the NP raising is driven by an Extended Projection Principle (EPP) feature that resides in Agr_O. Further, following Chomsky, I assume that Agr_O and Agr_S are really the same category, the distinction merely mnemonic. Overt object shift and oven subject shift are then the same phenomenon: satisfaction of the EPP.

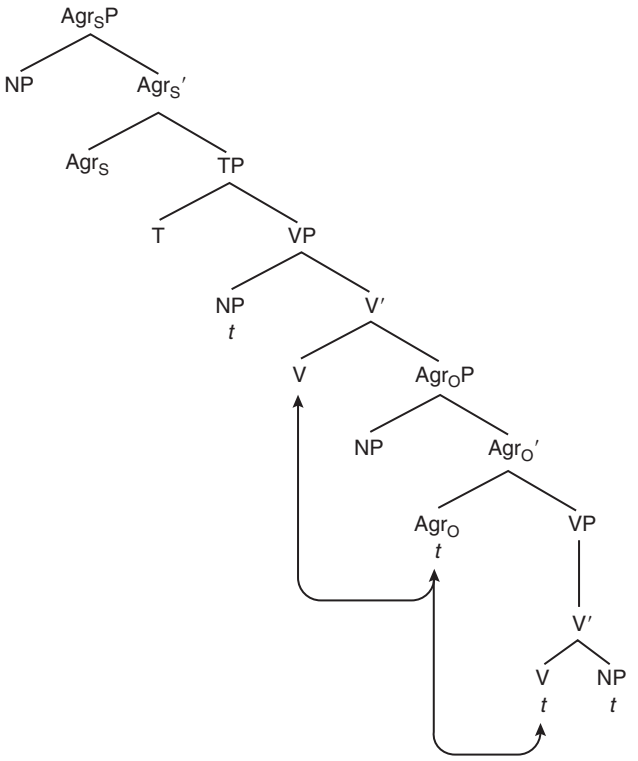

Figure 5.2 Koizumi's split VP hypothesis.

The question that now arises is why the V need not raise in Pseudogapping constructions, given that in nonelliptical sentences it must:

(32) Mary hasn't dated Bill, but she has Harry ~~[$_{VP}$ dated *t*]~~.
(33) *She has Harry dated.

I have not yet discussed the strong feature driving V raising. Suppose that that feature is a feature of the V that raises (rather than of the position it raises to). A promising possibility is that the feature is a θ-feature, given Koizumi's theory that the subject is base generated in the Spec of the higher VP.[1] Now suppose, following Chomsky (1993) but contra Chomsky (1994), that an unchecked strong feature is an ill-formed **PF** object Then we correctly derive the result that deletion of (a category containing) an item with an unchecked strong feature salvages the derivation. The portion of the structure that would have caused a PF crash is literally gone at that level (Figure 5.3).[2]

Consider now sentences with two complements. Given a natural extension of Koizumi's approach, there will be *three* VPs, one for each of the arguments.

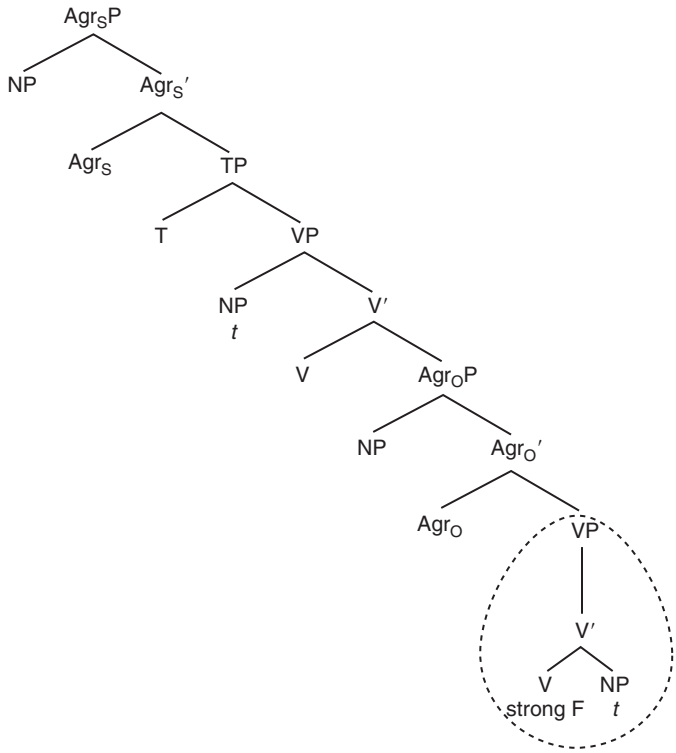

Figure 5.3 Portion of the structure that would cause a PF crash is deleted.

The structure for (34), a double object construction, is shown in Figure 5.4, with the VPs and AgrPs labeled with subscripts for ease of exposition.

(34) John gave Bill a lot of money.

Recall that the first object in a double object construction makes an acceptable Pseudogapping remnant, as seen in (15), repeated as (35).

(35) ?John gave Bill a lot of money, and Mary will ~~give~~ Susan ~~a lot of money~~.

The analysis of this example is fairly straightforward. [Spec, V_2], the "first object," overtly raises to [Spec, Agr_2], and VP_2 undergoes VP deletion in the PF component. So far, it is impossible to tell whether the NP, the "second object," overtly raises to [Spec, Agr_3], and V_3 overtly raises to V_2, via Agr_3. By LF, of course, these raisings must take place, along with the further raising of V to V_1. But since VP_2 will delete, any checked *or* unchecked features it contains will be absent at the level of PF, so they could not cause a PF crash.[3] Thus, the raisings could evidently be covert.

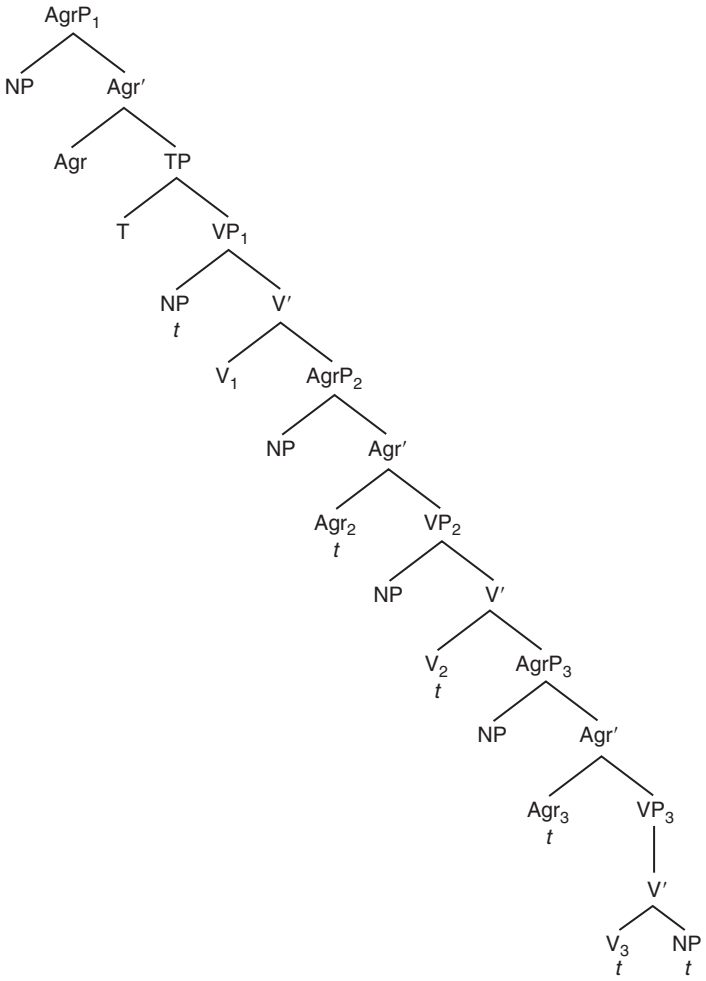

Figure 5.4 Double object structure.

Apart from the still mysterious failure of HNPS to create Pseudogapping remnants, the inability of the second object to be a remnant is accounted for. Consider (17), repeated as (36).

(36) *John gave Bill a lot of money, and Mary will ~~give Bill~~ a lot of advice.

Under the assumption, tacitly adopted above, that the first object begins higher than the second,[4] relativized minimality will guarantee that the first object remains higher. The consequence of this is that there could not be a VP (or any other constituent) to delete which includes the first object but excludes the second.

5.3. On the marginal character of the construction

At this point I take up a property of Pseudogapping constructions that I have put aside: The construction has a certain marginal character. Recall that my PF deletion analysis, coupled with the Chomsky (1993) position that a strong feature not overtly checked causes a **PF** crash, explains why Pseudogapping is possible at all. The unchecked strong feature of the V that fails to raise is remedied by deletion of VP containing that V. Chomsky (1994), though, replaced the PF crash analysis of strong features with an **LF** analysis, proposing that unless a strong feature "is checked before Spell-Out it will cause the derivation to crash at LF." Chomsky instantiates this in the following way: "a checked strong feature will be stripped away by Spell-Out, but is otherwise ineliminable." If the marginal character of Pseudogapping is to be analyzed as a consequence of grammatical properties, I offer the speculation that the proposals of Chomsky (1993, 1994) are *both* correct. Then a strong feature that is not checked in overt syntax will cause the derivation to crash at both PF and LF. A standard EPP violation will fall under this analysis, as will a sentence in which a verb fails to raise overtly, yet survives to the level of PF, as in (33) above, repeated as (37).

(37) *She has Harry dated.

When, on the other hand, a constituent containing the verb is deleted, the PF violation is avoided, but the LF violation persists. The relevant example is the Pseudogapping analog of (37), namely (32), repeated as (38).

(38) Mary hasn't dated Bill, but she has Harry ~~[$_{VP}$ dated *t*]~~.

Now the question is whether this example has the predicted status. What is the predicted status? Not that of a standard EPP violation, obviously, since that would cause both a PF and an LF violation. In fact, the only sort of example currently available for comparison is the type that leads Chomsky (1994) to modify his either (1993) theory. Chomsky is concerned to prevent, without stipulation, lexical insertion in the LF component. For a lexical item bearing phonological features, the mechanism is straightforward: The phonological features will cause the derivation to crash at LF. But this will not be the case for a lexical item lacking phonological features. Chomsky indicates that "empirical consequences seem to arise only in connection with functional heads that have 'strong features'." The one case he considers (though only in the abstract) is that of C with a strong feature that requires overt *wh*-movement. If such a C is introduced covertly, it could not constitute an ill-formed PF object. Hence, Chomsky's modified (1994) theory, but not his earlier (1993) theory, would provide an account of the unacceptability of (39) as a *wh*-question.

(39) (*)You read what.

Note that the much more extreme unacceptability of (40) is not at issue, given Chomsky's requirement that lexical insertion, whether overt or covert, is always at the root

(40) *I wonder you read what.

Example (39) is undoubtedly somewhat degraded. Its exact status is open to question; but the same could be said of Pseudogapping examples. Further research is required, but at this point there is no clear basis for rejecting the possibility suggested here that a strong feature that is unchecked in overt syntax potentially causes an LF crash and a PF crash. Example (39) and baseline Pseudogapping examples instantiate only the LF crash, though for different reasons. In the former case, the strong feature is not introduced until the LF component; in the latter, the strong feature is deleted (along with the VP containing it) on the way to the PF interface.

5.4. Pseudogapping vs standard VP-ellipsis

Since I have analyzed Pseudogapping as VP deletion, one might wonder how classic VP deletion is then to be treated, particularly with a transitive verb. In fact, the analysis is straightforward. Consider (41) with underlying structure (Figure 5.5).

(41) Mary will hire Susan.

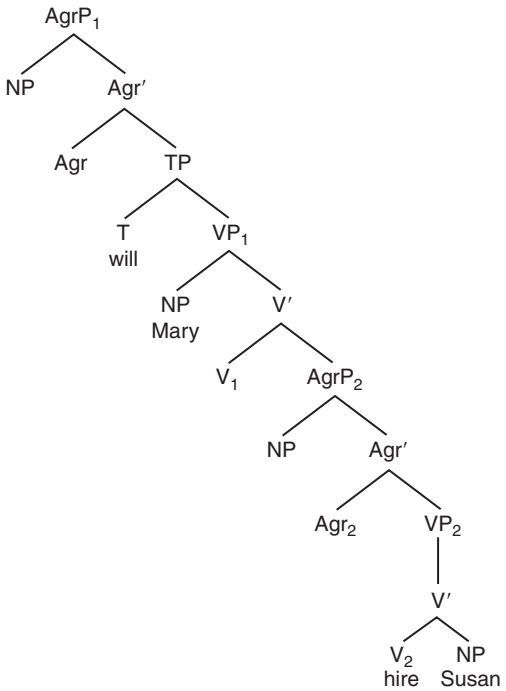

Figure 5.5 Underlying structure of a simple transitive.

As I have shown, if *Susan* raises to [Spec, Agr_2] and *hire* remains in VP_2, then deletion of VP_2 results in Pseudogapping:

(42) … Mary will Susan.

Further, note that if *Susan* does not raise and VP_2 deletes, though "Mary will" would be generated, the resulting structure would have two unchecked strong features, the EPP feature of Agr_2 and the strong (θ-)feature of *hire*. Given that classic VP-ellipsis is completely acceptable (given the appropriate discourse circumstances, etc.), there must be some other derivation for it. And in fact an alternative derivation is readily available. Beginning again with (Figure 5.5), suppose *Susan* raises to [Spec, Agr_2] and *hire* raises to V_1 via Agr_2. With raising of *Mary* to [Spec, Agr_1], all relevant features are checked. Deletion of VP_1 now yields (43) in a violation-free way.

(43) … Mary will.

5.5. Potential problems for the account

5.5.1. Double object constructions

While the account of Pseudogapping sketched so far has accounted for a substantial range of facts, and has done so without the need for a new ellipsis rule, it is not entirely unproblematic. In particular, there seems to be significant overgeneration. I turn now to this problem.

Consider first (44), a Pseudogapping example with two remnants.

(44) *Mary gave Susan a lot of money, and John will ~~give~~ Bill a lot of advice.

As far as I can tell, all examples with this pattern are seriously degraded. The question is why (44) should not be well-formed with structure (Figure 5.6), with VP_3 elided.

However, as Roger Martin (personal communication) first observed, on the general account of Pseudogapping I have given, a rather natural explanation is available. Recall that I explained the general marginality of even the best instances of Pseudogapping by proposing that the strong feature driving overt raising of the V to the higher V position causes both a PF and an LF violation, if the overt raising does not take place. Further, I proposed that the strong feature resides in the lexical V itself, rather than in the shell V that it raises to. Deletion eliminates the (PF) violation, but not the LF one. Now notice that in Figure 5.6, there are *two* shell Vs to which *give* has not overtly raised, hence two strong features that have not been overtly checked, not just one, so plausibly the violation should be more severe.[5]

Considerably more problematic is (45).

(45) *Mary gave Susan a lot of advice, and John will give Bill ~~a lot of advice~~.

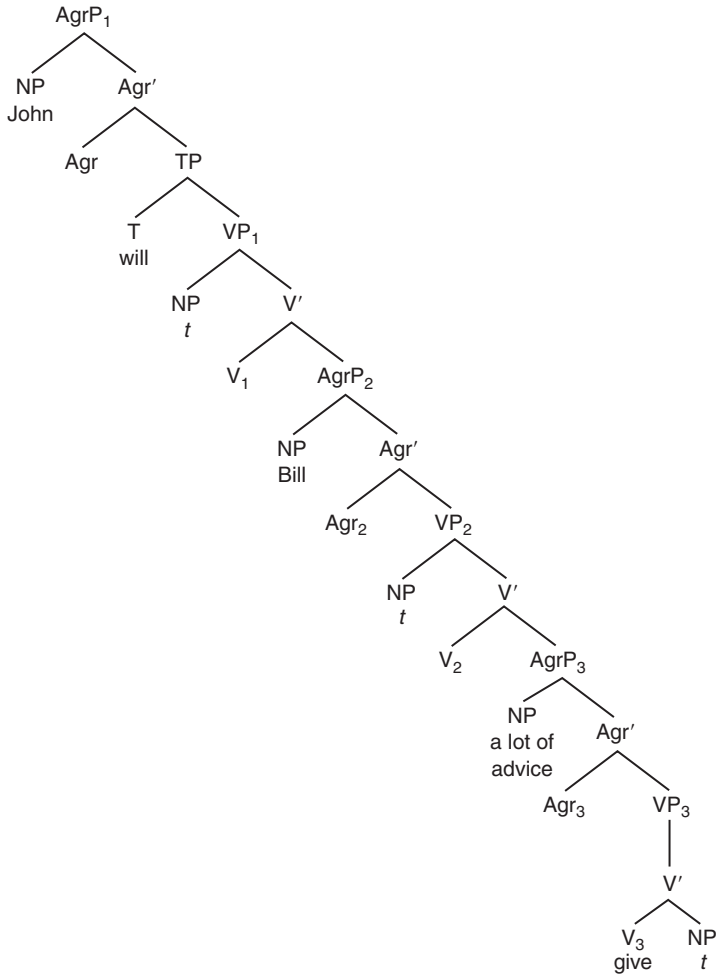

Figure 5.6 Pseudogapping structure for double object construction with two remnants; lowest VP to be deleted.

Example (45) is somewhat similar to (44) except that the violation should have been remedied. The verb *give* has overtly raised to its correct ultimate destination, passing through the intermediate shell V in transit, as illustrated in Figure 5.7.

Note that it is the intermediate VP, VP_2, that has been deleted. If VP_3 had been deleted instead, the absence of *a lot of advice* in the phonetic output would entail that that NP had not overtly raised to [Spec, Agr_3]. But then the EPP feature of Agr_3 would not be checked in overt syntax, causing both an LF and a PF violation (and the latter would not be remedied, since the deletion site would not include Agr_3). However, with VP_2 deleted, *a lot of advice* could have raised to [Spec, Agr_3],

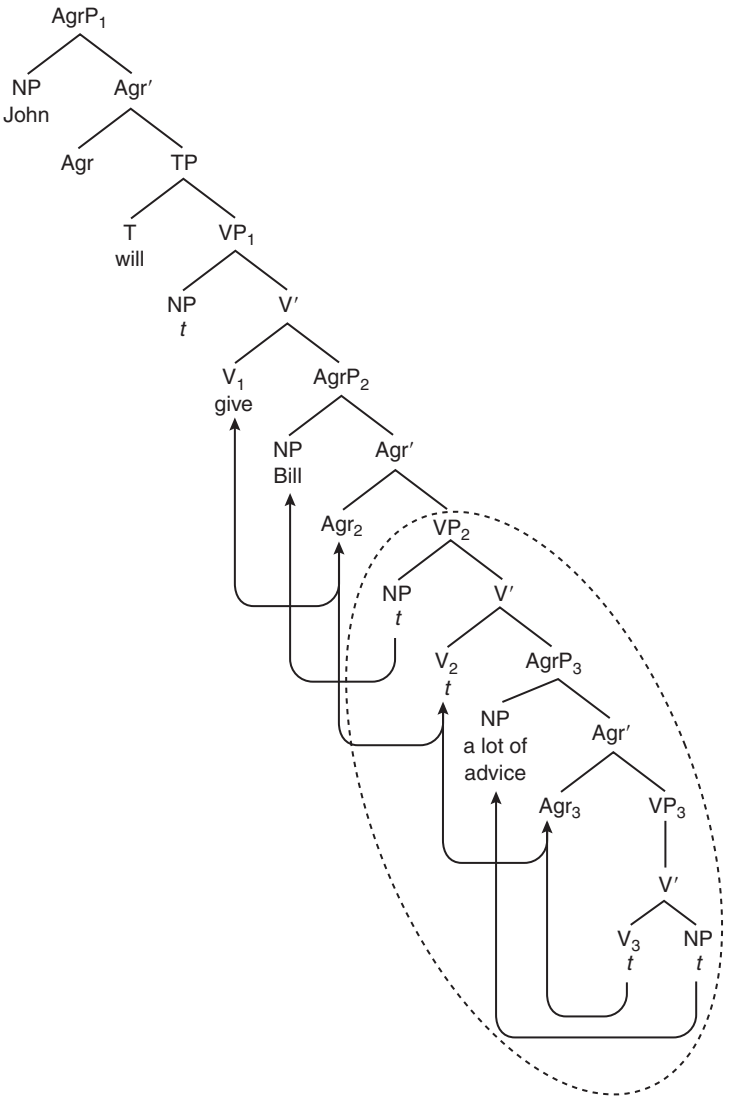

Figure 5.7 Verb *give* has overtly raised to its correct ultimate destination, passing the intermediate shell V in transit; intermediate VP is deleted (taking with it the lowest VP).

evidently avoiding all strong feature violations. Yet, the result is clearly bad. What is responsible?

Descriptively speaking, the situation is somewhat perverse. When the verb remains in VP, the VP can delete, as in standard VP-ellipsis, and its Pseudogapping

alternate. But when the verb has raised out of VP, the VP it has left behind apparently cannot delete. One might hypothesize a constraint to this effect (though almost immediately we will be forced to reject it):

(46) VP ellipsis constraint: VP ellipsis is prohibited if VP has lost its head.

Intriguingly, another ellipsis process seems to obey a similar constraint. Sluicing, a process first investigated by Ross (1969), is standardly (and plausibly) analyzed as *wh*-movement followed by IP deletion:

(47) Speaker A: Mary saw someone.
Speaker B: I wonder who ~~Mary saw~~.

Sluicing is also possible in the matrix:[6]

(48) Speaker A: Mary saw someone.
Speaker B: Who?

Surprisingly, though, if *Infl* has raised to C, Sluicing is blocked:

(49) Speaker A: Mary saw someone.
Speaker B: *Who did ~~Mary~~ ~~see~~?

This is abstractly parallel to what we just saw with Pseudogapping, suggesting a generalization of (46):

(50) XP ellipsis is prohibited if XP has lost its head.

However, there is evidence even against the more limited version of the constraint (46). A number of languages with overt V raising to I nonetheless allow VP-ellipsis, with the effect that everything in the VP except the V is deleted. Doron (1990) shows this for Hebrew:

(51) Q: Salaxt et ha-yeladim le- beit-ha-sefer
you-sent Acc the kids to school
Did you send the kids to school?
A: Salaxti
I sent.
I did.

Martins (1994) shows the same thing for Portuguese and McCloskey (1990) does for Irish:

(52) A Martas deu um livro ao João? Sim, deu.
the Martha gave a book to-the John yes gave
Did Martha give a book to John? Yes, she did.

(53) Q: Ar chuir tú isteach air
INTERR COMP put [PAST] you in on it
Did you apply for it?
A: Chuir
put [PAST]
Yes.

It seems clear, then, that ellipsis of a VP whose head V has raised away is not generally prohibited. The explanation for (45) must lie elsewhere.

In order to pursue this problem further, I would like to briefly examine the licensing condition for VP-ellipsis. Since Zagona (1982, 1988),[7] it has been assumed that there is an Empty Category Principle (ECP)-like constraint on VP-ellipsis: The ellipsis site must be governed by an appropriate head. I will adopt this assumption as well, though how to capture the effects of such a constraint within a minimalist framework is a difficult and important question that I will have to put aside here.[8] Saito and Murasugi (1990) explicitly argue that not just any head, even any lexically realized one, can function as a proper governor in this *sense*. Martin (1992b, 1996a)[9] provides very strong evidence that in the instance of VP-ellipsis, the licensing head is a particular sort of *Infl*, with tense being the crucial feature. Consider then the licensing configuration in grammatical instances of VP-ellipsis, first a simple case as in (54).

(54) Mary left, and John did too.

Here, under any imaginable notion of government, the Tense head governs the VP that is to be deleted (Figure 5.8).

Next consider a baseline instance of Pseudogapping, a process I have analyzed as VP-ellipsis.

(55) Mary hired Susan, and John did Bill.

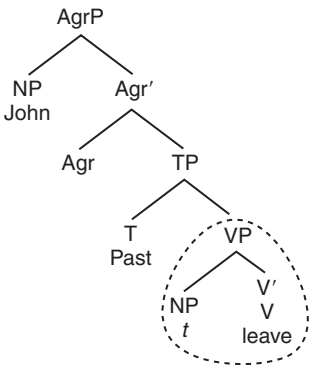

Figure 5.8 Tense head governs the VP that is to be deleted.

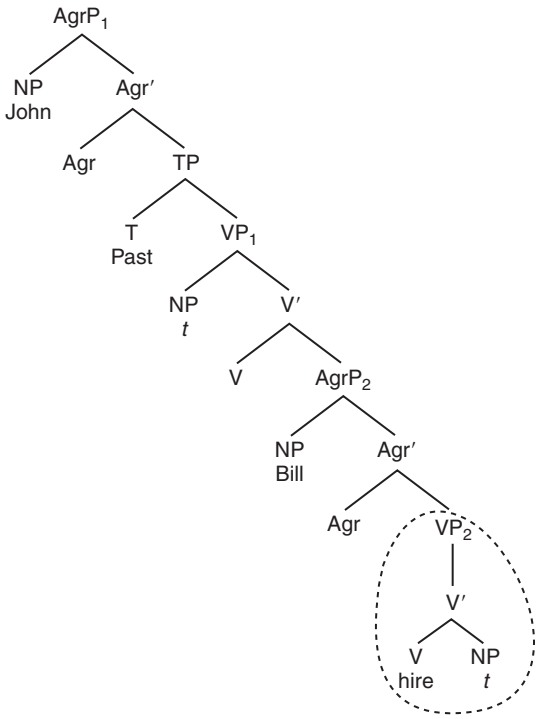

Figure 5.9 Deletion of lower VP when the licensing head, Past, is more remote, with two maximal projections intervening.

This time the licensing head, Past, is the same, but the syntactic connection is considerably more remote, with two maximal projections intervening. Yet, deletion of the lower VP is reasonably acceptable (Figure 5.9). At this point, a comparison of this structure with that of the completely unacceptable (45), slightly modified here, is necessary.

(56) *Mary gave Susan a lot of advice, and John gave Bill ~~a lot of advice~~.

Once again, two maximal projections intervene between Past and the target VP, VP_2. Further, the maximal projections appear to be the same as they were in Figure 5.9, VP and AgrP. However, there is one difference: In the more-or-less acceptable Figure 5.9 the intervening V head is empty, while in the unacceptable (Figure 5.10) the intervening V is the lexical verb *give*, which has raised from the lowest VP. I speculate that this is, in fact, the relevant difference, and I suggest that it is some version of relativized minimality that states this difference.

I tentatively offer two possible ways to instantiate this intuition. First, under the assumption that Agr lacks substantive content, in Figure 5.9 the nearest potentially

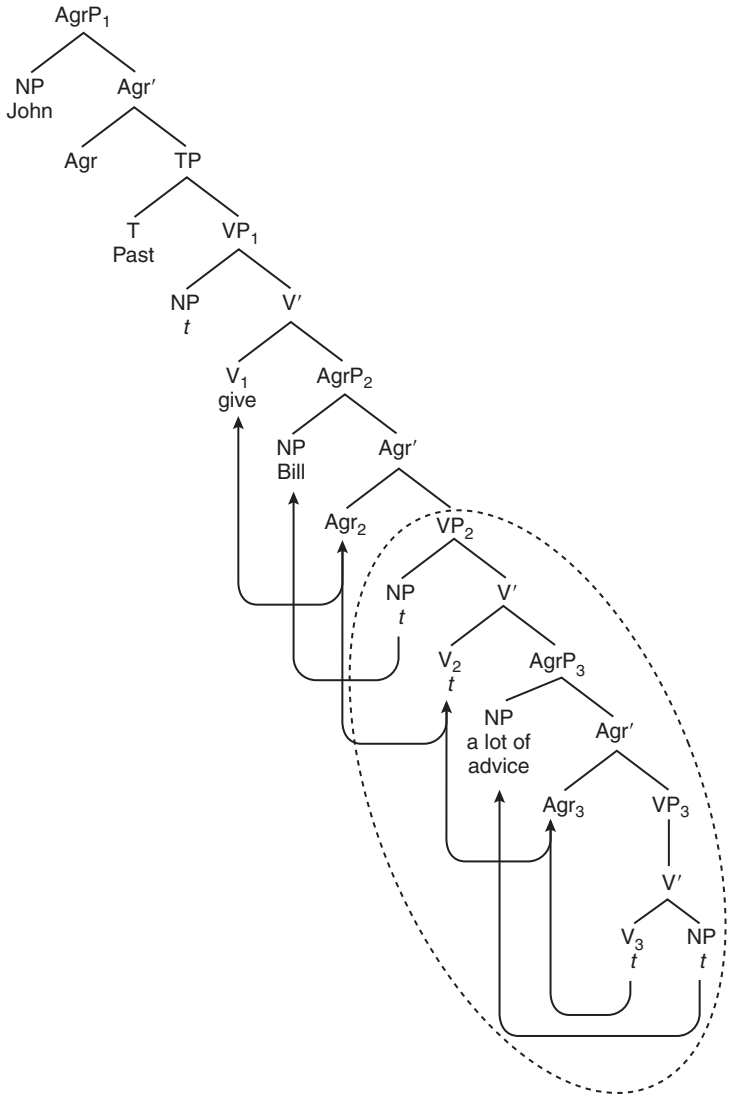

Figure 5.10 Bad Pseudogapping with two maximal projections intervening between Tense and the deleted VP.

governing substantive head is the crucial governor Tense. In Figure 5.10, on the other hand, the nearest potentially governing substantive head is the verb *give*, and that head is not (for whatever reason) an appropriate governor for VP-ellipsis.

The second possibility, more speculative (but perhaps more in keeping with recent trends in syntactic theorizing), would put the relativized minimality requirement

on the head licensing the ellipsis. Suppose the head licensing VP-ellipsis does so by attracting (in the sense of Chomsky 1995a) a feature of the head of the VP. As a consequence of having "lost" this feature, the VP would now be PF defective unless it is deleted.[10] Chomsky argues that attraction seeks the nearest *c*-commanded item with a feature of the appropriate type. In standard simple VP-ellipsis, that feature resides in the immediate complement of the licensing head. And in the ill-formed (56), attraction has "skipped" the V heading the complement of the licensing head and instead attracted a feature of the initial trace of that V, in violation of relativized minimality. Alternatively, a feature of the raised lexical V has been attracted, but that V has not been deleted, resulting in a PF crash. Finally, in the reasonably acceptable Pseudogapping example (Figure 5.9), even though *hire* is geometrically rather remote from the licensing Tense, there is no nearer V with a feature for Tense to attract, so, in the spirit of relativized minimality, it can attract a structurally distant feature.

At this point, I return briefly to consideration of the Hebrew, Portuguese, and Irish examples in (51)–(53). Recall that in those examples, VP-ellipsis was possible even though the head V had raised out of VP. This property of the constructions necessitated the rejection of the hypothesized ellipsis constraint (50) repeated here as (57).

(57) XP ellipsis is prohibited if XP has lost its head.

While a detailed examination of the ellipsis phenomena in the three languages mentioned would take us too far afield, the alternative account I offered as a replacement for (57) seems compatible with the facts. Assuming that, as in English, the licensing head for VP-ellipsis is Tense, a suitable licensing configuration does exist, even though the V head of VP has raised. This is so since, unlike the situation in the ill-formed English (Figure 5.7), the V has raised to the Tense licenser, so that it does not *intervene* between licensing T and target VP.[11]

5.5.2. Heavy NP shift

I must now further explore a set of facts that provided some motivation for the analysis of Pseudogapping I have presented, but which I have not actually explained so far. Recall that my analysis is based on Jayaseelan's fundamental insight, that Pseudogapping is VP-ellipsis with the remnant having moved out of the VP. However, I rejected Jayaseelan's specific implementation, in which the relevant movement rule is heavy NP shift. The basis for the rejection was the following set of data, which indicates that a possible Pseudogapping remnant is not necessarily a possible target of HNPS, and conversely:

(58) ?John gave Bill a lot of money, and Mary will ~~give~~ Susan ~~a lot of money~~.
(59) *John gave *t* a lot of money [the fund for the preservation of *VOS* languages].

(60) *John gave Bill a lot of money, and Mary will ~~give Bill~~ a lot of advice.
(61) John gave Bill *t* yesterday [more money than he had ever seen].

The raising to [Spec, Agr_O] analysis I offered straightforwardly accommodated this paradigm. But there is still an unanswered question: Even if, as I have argued, there is a process other than HNPS creating Pseudogapping remnants, why can not HNPS *also* create them?[12] Given my extension of Koizumi's split-VP hypothesis, one structure that would potentially provide the source for HNPS and Pseudogapping (i.e. deletion of the residual VP) as in (61) is Figure 5.11.

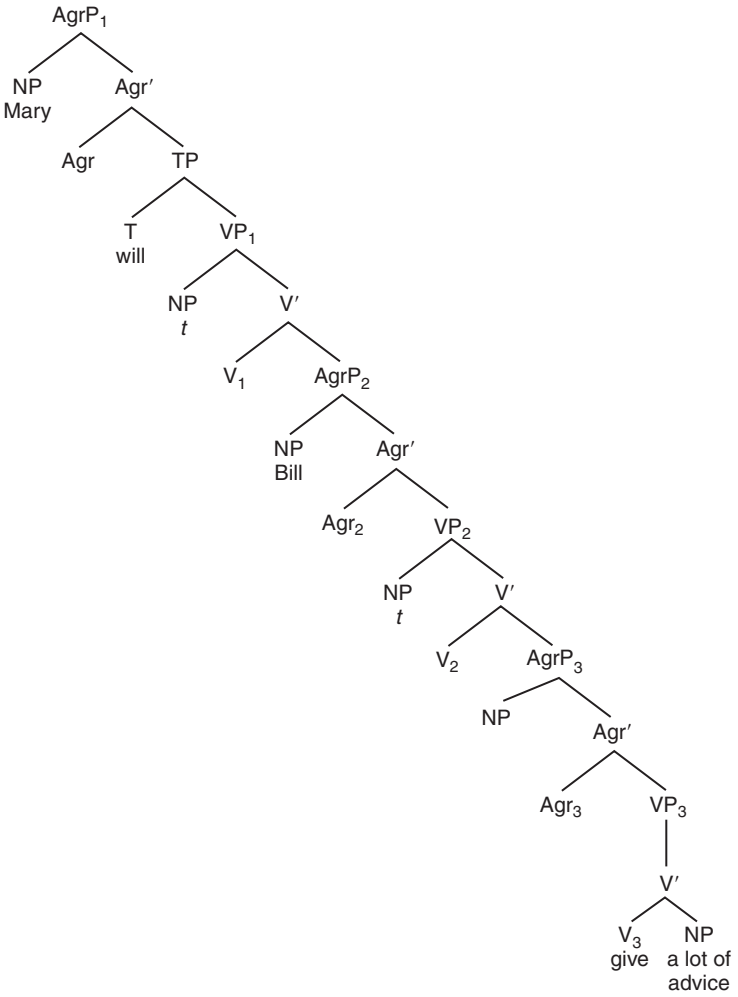

Figure 5.11 A potential source for HNPS and Pseudogapping.

Suppose a *lot of advice* in Figure 5.11 undergoes HNPS to some position higher than *Bill* and the residual VP_1 deletes (taking *Bill* with it). Note that on this derivation, the EPP feature of Agr_3 is not checked overtly, nor are two of the strong θ-features of *give* checked overtly, and we have seen that failure to check strong features overtly leads to considerable cumulative degradation even if the carrier of the feature is deleted. But there is an alternative derivation that must still be considered, one in which all features are checked overtly. Starting again from Figure 5.11, *a lot of advice* can raise to [Spec, Agr_3] and *give* can raise to V_1 via Agr_3 and Agr_2. *A lot of advice* undergoes HNPS to a position outside VP_1, perhaps adjoined to TP, VP_1 itself, or $AgrP_1$; and finally, VP_1 deletes. So far, this appears to be a flawless derivation (illustrated in Figure 5.12 for the unacceptable (60)).

I suspect that the violation in this derivation is independent of the considerations of this chapter, stemming, rather, from the (admittedly ill-understood) strong locality constraints on rightward movement.[13] Assuming that the landing site is VP_1, a closer VP, VP_2 has been skipped. Similarly, if $AgrP_1$ is the landing site,

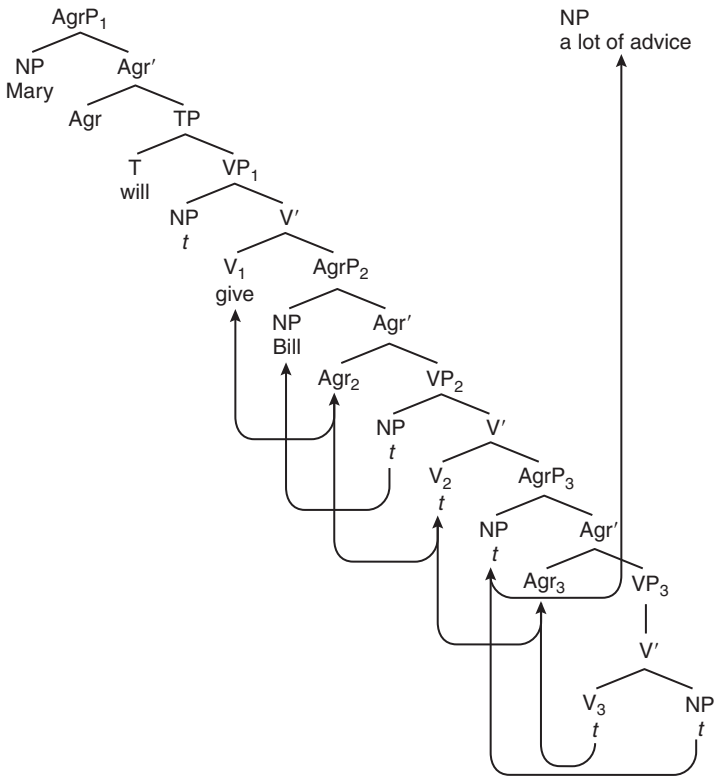

Figure 5.12 *A lot of advice* undergoes HNPS to a position outside VP_1 adjoined to TP, VP, or $AgrP_1$ and finally VP_1 deletes.

$AgrP_3$ and $AgrP_2$ have been skipped. Obviously, if VP and AgrP are both suitable landing sites, numerous closer targets exist. Plausibly, this would render the required longer movement ungrammatical.[14]

One immediate consequence of this line of reasoning is that the shifted heavy NP in (61) is not very high, which entails that the adverb is also not very high. One workable position for the adverb is adjunct to the lowest VP (at least as one option). Given my analysis of Pseudogapping, an example like the following provides support for this conjecture:

(62) John saw Bill yesterday and Mary did ~~see~~ Susan ~~yesterday.~~

On my account, *Susan* has raised out of the lower of two VPs, and the residual VP, evidently including *yesterday*, has deleted. Notice that adverbs that, by their semantic character, would be assumed to be very high in the structure do not undergo "small" deletion (i.e., Pseudogapping), or even VP deletion:

(63) *John saw Bill, fortunately, and Mary did ~~see~~ Susan, ~~fortunately.~~
(64) *John saw Bill, fortunately, and Mary did ~~see Bill~~, ~~fortunately,~~ (too).

Correspondingly, HNPS around such high adverbs seems much less available than around lower ones:

(65) John saw yesterday his old friend from Philadelphia.
(66) ?*John saw fortunately his old friend from Philadelphia.

Thus, the strict locality on HNPS posited to explain the inability of that rule to create Pseudogapping remnants receives some independent support.

5.6. Antecedent contained deletion (ACD)

I turn now to ACD, a much discussed phenomenon often related implicitly to Pseudogapping. Lappin (1992) makes such a proposal,[15] and in this, Lappin follows the earliest investigator of ACD, Bouton (1970). Both these researchers suggest that in a sentence like (67), synonymous with (68), the ellipsis site does not include the position from which the *wh*-movement involved in relativization took place.

(67) John saw everyone you did.
(68) John saw everyone you saw.

In the terms of the present chapter, the *wh*-trace is a right remnant. This potentially resolves the notorious infinite regress that Bouton (followed by Sag (1976), May (1984), and Lappin (1992), among many others) originally saw in true instances of antecedent-contained pro forms. The initial difficulty of such constructions, ACD

among them, is well-known. For example, in the derivation of (68), the antecedent of the missing VP seems to be a larger VP containing that very same missing VP:

(69) John [$_{VP}$ saw everyone [*Op* [you did [$_{VP}$ *e*]]]]
[$_{VP}$ saw everyone [*Op* [you did [$_{VP}$ *e*]]]]

Note that while the issue is most often discussed in terms off LF copying in the recent literature, the problem is symmetric between copying and deletion. Given the tentative conclusions of my discussion above, I will here talk in terms of deletion. Now, observe that if the *wh*-trace is a remnant, rather than part of the ellipsis site, the regress problem disappears:

(70) John [$_{VP}$ saw everyone [*Op* [you did [$_{VP}$ *e*]]]]
saw/see

This was the proposal of Lappin, and of Bouton before him.

Hornstein (1994) offers what appears to be a radically different account even though Hornstein, like Lappin, is concerned to present an alternative to the classic QR account of May (1985). What Hornstein proposes is that raising to [Spec, Agr$_O$] is the process moving the object out of the VP, hence moving the null VP contained inside that NP out of its antecedent. Hornstein takes it for granted that this type of ellipsis involves LF copying, presumably based on the assumption that raising to [Spec, Agr$_O$] is covert. But, as noted above, if that raising is overt, PF deletion becomes a viable possibility.

Given the hypothesis that Pseudogapping involves raising to [Spec, Agr$_O$], Hornstein's proposal can now be seen as quite similar to Lappin's. And on the face of it, both proposals successfully address a problem, originally pointed out by Wyngaerd and Zwart (1991), for Quantifier Raising (QR)-based approaches. May (1985) argued that the process removing the null VP from its antecedent in ACD constructions is QR. His argument was based, in part, on contrasts like the following:

(71) Dulles suspected everyone Angleton did.
(72) *Dulles suspected Philby, who Angleton did.

When the null VP is contained in a quantificational expression, as in (71), the result is acceptable, but when it is in the nonquantificational expression (72), it is not. This is precisely predicted if QR is the (only) available mechanism for resolving the regress inherent in these constructions. In (71), but not (72), the object will undergo QR, thus transporting the null VP out of its antecedent. However, Wyngaerd and Zwart show that examples indistinguishable from (72) in relevant respects are acceptable:

(73) Dulles suspected Philby, who Angleton did not.
(74) Dulles suspected Philby, who Angleton did as well.

They conclude, as does Hornstein later, that something other than QR must be at work, something that can affect nonquantificational expressions. For Hornstein, that something is raising to [Spec, Agr_O], a possibility nicely consistent with the present analysis of Pseudogapping. In addition, a further range of facts is at least roughly in accord. Recall that objects of reanalyzing prepositions constitute somewhat acceptable Pseudogapping remnants. Correspondingly, they constitute somewhat acceptable carriers of ACD sites:

(75) ?Dulles spoke to Philby, who Angleton did not.
(76) ?Dulles spoke to Philby, who Angleton did as well.
(77) ?Dulles talked about Philby, who Angleton did not.
(78) ?Dulles talked about Philby, who Angleton did as well.
(79) ??John took advantage of Bill, who Mary will also.

Recall also that objects of nonreanalyzing prepositions constitute very poor Pseudogapping remnants. Significantly, they are also unacceptable ACD site hosts:

(80) *John stood near Bill, who Mary did not.
(81) *John stood near Bill, who Mary did as well.

Even the double object asymmetry found in Pseudogapping is approximately paralleled in the ACD constructions under constructions under consideration:

(82)(a) ?? John showed Bill, who Mary did as well, the new teacher.
(b) *John showed Bill the new teacher, who Mary did as well.

Thus, there is considerable support for the reduction of ACD to Pseudogapping, and, again, for the reduction of the latter to the combination of raising to [Spec, Agr_O] and VP-ellipsis. This is, in essence, in accord with the proposals of Lappin and Hornstein. However, on closer inspection, it becomes evident that the reduction is not complete. Hornstein, like Wyngaerd and Zwart, assumes that ACD is a unified process, in particular, that the appositive examples they present (in arguing against a QR account) have just the same analysis as the standard restrictive examples. But, as discussed by Lasnik (1993), this is not correct. In fact, none of the constraints on Pseudogapping/(appositive) ACD seen so far hold of the restrictive type standardly in the discussed in the literature:

(83) *John stood near Bill, who Mary did as well.
(84) John stood near everyone Bill did.
(85) *John showed Bill the new teacher, who Mary did as well.
(86) John showed Bill everyone Mary did.

This state of affairs strongly supports the claim of Fiengo and May (1992) that while Pseudogapping[16] is the sole process responsible for appositive ACD, such is not the case for restrictive ACD.[17] What the alternative source can be is a very complicated question, one that I will only be able to touch on here.

As mentioned earlier, the classic analysis of ACD as in May (1985) relies on QR to move the carrier of the null VP out of the antecedent VP, under the assumption that ellipsis involves LF copying (rather than PF deletion).[18] Following QR, the source for LF copying in (87) would be (88).

(87) Dulles suspected everyone Angleton did.
(88) [everyone Angleton did [$_{VP}$ *e*]] [$_{IP}$ Dullest[$_{VP}$ suspected *t*]].

In this theory, QR obviously moves full quantificational expressions (rather than, say, just the quantificational head). As Fiengo and May (1994: 296) note, this entails that, at least under some circumstances, binding conditions must be satisfied at S-structure. In this regard, their argument precisely replicates one of Chomsky (1981: 197). Chomsky observes that (89) exhibits a Condition C effect even though following QR its LF has no A-bound R-expression.

(89) He liked every book that John read.

Thus, Condition C must be satisfied (at least) at S-structure. Lasnik (1993) and Hornstein (1994) point out that under minimalist assumptions about the organization of the grammar, this conclusion is untenable, since there is no level of S-structure in that framework.[19] Fox (1995) proposes a sort of minimalist version of QR wherein the rule applies only if it has to. What would make it necessary is resolution of a scope ambiguity, as in (90), or avoidance of an ellipsis regress, as in (87).

(90) Someone loves everyone.

In (89), on the other hand, nothing makes QR necessary, so it is inapplicable. The LF is therefore indistinguishable from the S-structure in relevant respects, so the Condition C violation is not remediated.

The one other syntactic approach I am familiar with is the extraposition analysis of Baltin (1987). On this analysis, the relative clause containing the missing VP has extraposed (sometimes vacuously) to a position outside the antecedent VP. Thus, when the antecedent is copied, regress can be avoided.[20] Larson and May (1990) point out several difficulties with such an analysis. First, the relative clauses in ACD constructions do not have the outward form of extraposed relatives. As

is well known, in situ relatives allow three possibilities, an overt *wh*-form, an overt complementizer, or neither:

(91) I visited a man who / that / Ø John mentioned recently.

With extraposed relatives, on the other hand, the third possibility is apparently excluded:

(92)(a) I visited a man recently who / that John mentioned.
(b) (c) ?*Ø

Larson and May allude to Stowell's (1981) proper government analysis of null complementizers to explain this contrast.[21] Whatever the precise nature of the constraint, under Baltin's account it would be predicted that the null form is excluded from (ACD) constructions. In direct conflict with this prediction, the null form is freely allowed. Many of the ACD examples cited thus far display the null form. Example (87), repeated as (93), is representative.[22]

(93) Dulles suspected everyone Angleton did.

It should be noted, though, that the null complementizer constraint is not ironclad. The following example, of a type pointed out to me by Mark Baltin, possibly involves extraposition, yet it is quite acceptable:

(94) I threw something out I had no further use for.

Even more similar to (92) is (95), but the latter is considerably better.

(95) ?I visited a man yesterday John had told me about.

The null complementizer phenomenon is, as Larson and May argue, potentially of great relevance to the issue at hand, but it clearly demands further investigation.

In addition to the null complementizer paradigm, Larson and May point out a further difficulty for an extraposition account of ACD. Consider the LF structure of (84) following extraposition.[23]

(96) John [$_{VP}$ [$_{VP}$ stood near everyone] [$_{CP}$ *op* [Bill did [$_{VP}$ e]]]].

LF copying of the VP (which Larson and May call reconstruction), results in a structure that appears to be completely incorrect, lacking a variable to be bound

by *Op*, the relative operator:

(97) John [$_{VP}$ [$_{VP}$ stood near everyone] [$_{CP}$ *Op* [Bill (did) [$_{VP}$ stood near everyone]]]].

Thus, QR is needed regardless, they argue. With *everyone* in (109) raised, the variable it leaves behind is correctly copied as a variable:

(98) everyone [$_{IP}$ John [$_{VP}$[$_{VP}$ stood near *t*] [$_{CP}$ *Op* [Bill (did) [$_{VP}$ stood near *t*]]]]].

Note, though, that this kind of QR would not run afoul of the minimalist binding theory problem, since here, just a simple quantifier is raised. There is no pied-piping. If even in ACD constructions QR can be limited in this way, one of the major difficulties disappears.

There is one other approach to Larson and May's missing variable problem that might also be worth considering. As Larson and May observe, what is needed in the elided VP is a trace. But this leaves open exactly what a trace is. Chomsky (1993) provides discussion bearing on this question. Considering a variety of factors, and in particular reconstruction effects,[24] Chomsky suggests that a trace is initially a full copy of the moved item. Chomsky shows how this provides the basis for an account of the grammaticality of (99):

(99) Mary wondered which pictures of himself Bill saw.

The structure of (99) following *wh*-movement and prior to other operations is as in (100):

(100) Mary wondered [*wh*- which picture of himself] [Bill saw [*wh*- which picture of himself]].

Himself is assigned an appropriate antecedent by virtue of its position in the trace. Now note that the same reconstruction effects show up in relative clauses:

(101) Mary mentioned the pictures of himself that Bill saw.

Thus, the trace in this instance also is presumably a copy of the head:[25]

(102) Mary mentioned the pictures of himself that Bill saw the pictures of himself.

But given this analysis, Larson and May's example (97) is not, after all, incorrect: The trace is precisely a copy of the head, at the relevant point in the derivation.

I close this investigation by observing that there is one new problem with the extraposition analysis. Recall that the second object in a double object construction is a fine restrictive ACD host, as in (86), repeated as (103).

(103) John showed Bill everyone Mary did.

The missing VP is [show Bill *t*]. Thus, the relative clause must extrapose to a position quite high in the structure. But earlier I showed that such long distance

extraposition is *not* possible in the case of HNPS. If the latter effect reflects a general constraint on rightward movement, then Baltin's extraposition account is excluded on those grounds. A more definitive conclusion on this matter awaits a better understanding of movement constraints of this type.

Acknowledgment

I am indebted to Cedric Boeckx, Željko Bošković and an anonymous reviewer for valuable suggestions on this research, and to Cedric Boeckx for editorial assistance.

6

ON FEATURE STRENGTH

Three minimalist approaches to overt movement[1]

Chapters 3 and 4 both argued that certain phenomena that had previously been analyzed in terms of covert movement should be reanalyzed in terms of overt movement. Given a condition like Procrastinate, overt movement must be forced by some requirement, a requirement that Chomsky has stated in terms of "strong features." This chapter compares several possibilities that Chomsky has suggested for the condition demanding that strong features must be checked in overt syntax. Based on the analysis of Pseudogapping of Chapter 5, I show that ellipsis, which I analyze as deletion, can "repair" a structure damaged by failure of normally obligatory overt movement (V raising, in this case) to apply. I then give a parallel analysis of Sluicing, with corresponding implications for Infl raising to C. I show how these analyses fit neatly into the Chomsky (1993) PF crash theory of strong features, although only if the strong features in the constructions at issue reside in the item that is to move, rather than in the "attracting" head (the latter being Chomsky's position on the issue). I then proceed to show how the repair phenomenon can actually be reinterpreted to bring it into line with Chomsky's position on the locus of feature strength. However, I conclude with one phenomenon that seems to argue powerfully against the claim that strong features are invariably in the attracting head: multiple *wh*-fronting in Slavic. Why should the attracting head need more than one *wh*-phrase? Interestingly, Bošković (1999) now directly addresses this question, positing an "Attract-all" feature, thus preserving the possibility that Chomsky's position is correct.

Procrastinate (Chomsky (1993)) favors covert movement; therefore, when movement is overt, it must have been forced to operate "early" by some special requirement, one that Chomsky codes into "strong features." I compare Chomsky's three successive theories of strong features and argue that two ellipsis phenomena, Pseudogapping and Sluicing, provide evidence bearing on the nature of strong features. I argue that movement *or* ellipsis can rescue a derivation with a strong feature, and I conclude that PF crash is relevant either directly, as in Chomsky (1993),

or indirectly, as in the theory presented in Chomsky (1995a) augmented by the multiple-chain theory of pied-piping (especially as interpreted by Ochi (1998)).

6.1. Three theories of feature checking

Given an economy condition like Procrastinate (Chomsky (1993)), which is designed to favor covert movement over overt, we expect no movement to be overt, all else being equal. When movement is overt, rather than covert, then, it must have been forced to operate "early" by some special requirement. Chomsky (1993, 1994, 1995a)[2] codes this requirement into "strong features" and presents three successive, distinct theories of precisely how strong features drive overt movement.

PF crash theory

A strong feature that is not checked in overt syntax causes a derivation to crash at PF (Chomsky (1993)).

LF crash theory

A strong feature that is not checked (and eliminated) in overt syntax causes a derivation to crash at LF (Chomsky (1994)).

Virus theory

A strong feature must be eliminated (almost) immediately upon its introduction into the phrase marker; otherwise, the derivation cancels (Chomsky (1995a)).

In this article I will bring some ellipsis facts to bear on the question of the nature of strong features. I will begin my exploration by briefly summarizing Chomsky's successive justifications for these three proposals, and the technical implementations of them.

The justification for the PF crash theory is as follows:

> … the position of Spell-Out in the derivation is determined by either PF or LF properties, these being the only levels, on minimalist assumptions. Furthermore, parametric differences must be reduced to morphological properties if the Minimalist Program is framed in the terms so far assumed. … we expect that at the LF level there will be no relevant difference between languages with phrases overtly raised or in situ (e.g., *wh*-phrases or verbs). Hence, we are led to seek morphological properties that are reflected at PF.
>
> (Chomsky 1993: 192)

In the text and an accompanying note, Chomsky suggests two possible implementations of this approach:

> … "strong" features are visible at PF and "weak" features invisible at PF. These features are not legitimate objects at PF; they are not proper components of phonetic matrices. Therefore, if a strong feature remains after Spell-Out, the derivation crashes.[3]
>
> (Chomsky 1993: 198)

> 38. Alternatively, weak features are deleted in the PF component so that PF rules can apply to the phonological matrix that remains; strong features are not deleted so that PF rules do not apply, causing the derivation to crash at PF.
>
> (Chomsky 1993: 216)

Whereas the justification for the PF crash theory is conceptual, the justification for the change to the LF crash theory is, as far as I can tell, completely empirical. The relevant discussion is not fully explicit, but what is at issue is evidently the ungrammaticality of sentences like (1).

(1) *John read what?

Assuming that the strong feature forcing overt *wh*-movement in English resides in interrogative C,[3] the potential concern is that that C might be introduced in the LF component, where, checked or not, it could not possibly cause a PF crash. Yet, (1) is bad, so such a derivation must be blocked. I quote Chomsky's discussion.

> … Spell-Out can apply anywhere, the derivation crashing if a "wrong choice" is made … . If the phonological component adds a lexical item at the root, it will introduce semantic features, and the derivation will crash at PF. If the covert component does the same, it will introduce phonological features, and the derivation will therefore crash at LF … . Suppose that root C (complementizer) has a strong feature that requires overt *wh*-movement. We now want to say that unless this feature is checked before Spell-Out it will cause the derivation to crash at LF to avoid the possibility of accessing C after Spell-Out in the covert component.
>
> (Chomsky 1994: 60)

Note that for Chomsky, the problem is specifically limited to the root – that is, to examples like (1), rather than (2).

(2) *Mary wonders John read what.

This is so since Chomsky assumes that lexical material can only be added at the root. Consequently, a C in the complement of *wonder* in (2) must have been added

in the overt syntax, prior to the merger of the complement with *wonder*. But then the PF account would suffice. In (1), on the other hand, C could be added at the root covertly. The new LF account is therefore necessary.[4] Chomsky states it as follows:

> Slightly adjusting the account in Chomsky (1993), we now say that a checked strong feature will be stripped away by Spell-Out, but is otherwise ineliminable.
>
> (Chomsky 1994: 60)

Under this approach, it is not the checking operation itself that eliminates a (strong) feature. Rather, checking renders the strong feature eligible to be eliminated by Spell-Out, the latter now being construed as a sort of operation, instead of just the branch point of a derivation. On this account, whether interrogative C is introduced overtly, as in (3), or covertly, as in (4), the strong feature will persist to the LF interface level since it was not checked prior to Spell-Out.[5]

(3) Spell-Out: C [strong Q] John read what *LF
(4) Spell-Out: John read what
LF: C [strong Q] John read what *LF

Chomsky (1995a) rejects the PF crash theory on conceptual grounds, and the conceptual argument he gives applies equally to the LF crash theory. Thus, he rejects any such account as an "evasion" and proposes what he claims is a more straightforward statement of the phenomenon, here called the virus theory. (Juan Uriagereka (personal communication) suggests this felicitous term. This use of *virus theory* is distinct from Sobin's (1997) use of the same term.)

> … formulation of strength in terms of PF convergence is a restatement of the basic property, not a true explanation. In fact, there seems to be no way to improve upon the bare statement of the properties of strength. Suppose, then, that we put an end to evasion and simply define a strong feature as one that a derivation "cannot tolerate": a derivation $D \rightarrow \Sigma$ is canceled if Σ contains a strong feature …
>
> (Chomsky 1995a: 233)

Chomsky summarizes this approach as follows:

> A strong feature … triggers a rule that eliminates it: [strength] is associated with a pair of operations, one that introduces it into the derivation … a second that (quickly) eliminates it.
>
> (Chomsky 1995a: 233)

Later I will discuss this approach in further detail, showing, in particular, that it does not adequately address the empirical argument Chomsky gave for rejecting the PF crash theory.

Given that the PF crash theory concerns PF, ellipsis potentially provides new evidence bearing on its correctness, if, as Chomsky has consistently maintained over the years, ellipsis involves a PF deletion process.[6] I am aware that a PF analysis of ellipsis is not uncontroversial. In fact, it is quite widely rejected in favor of an LF copying theory, so to the extent that the arguments I will present assume a PF theory of ellipsis, they will have to be regarded as highly tentative. However, I might note that there is actually little in the way of conclusive evidence against (or for) the PF theory. Perhaps the most important argument for an LF approach is the one developed by May (1985 and other works). The essence of the argument is that an LF process, Quantifier Raising (QR), feeds ellipsis resolution in antecedent-contained deletion (ACD) constructions. Therefore, ellipsis resolution must itself be an LF process. Although this is, on the face of it, a very powerful argument, I might note that Hornstein (1994) argues that the crucial process is not actually QR but raising to [Spec, Agr_O], and, as briefly discussed below (and at greater length in Lasnik 1995b,d), there is reason for thinking that that process operates in overt syntax.[7]

Although an LF copying theory is now rather standard, Chomsky and Lasnik (1993) suggest, and Tancredi (1992) develops, a PF theory, largely based on interpretive parallels between elliptical constituents and deaccented one.[8] In this, Chomsky and Lasnik and Tancredi were actually resurrecting an old account of Chomsky's from the late 1960s and early 1970s. For example, Wasow (1972) cites a 1971 lecture where Chomsky "sugges[ed] that VP deletion and Sluicing can be formulated as very late rules which delete unstressed strings." As it happens, these are just the two ellipsis processes I will be considering: first Pseudogapping (a variant of VP-ellipsis, I believe), then Sluicing.

6.2. Pseudogapping: an argument for the "PF" approach

(5) presents a few examples of Pseudogapping from the classic study by Levin (1978).

(5)(a) If you don't believe me, you will Ø the weatherman.
(b) I rolled up a newspaper, and Lynn did Ø a magazine.
(c) Kathy likes astronomy, but she doesn't Ø meteorology.

Although in many instances it might appear that the process is simply elision of the main verb, there is considerable evidence that more is involved. In the examples in (6), the ellipsis site includes the main verb plus (a) the small clause predicate or (b) the second object in a double object construction.

(6)(a) The DA proved Jones guilty and the Assistant DA will ~~prove~~ Smith ~~guilty.~~
(b) ?John gave Bill a lot of money, and Mary will ~~give~~ Susan ~~a lot of money.~~

Rejecting the possibility of an ellipsis rule affecting a discontinuous portion of the structure, Jayaseelan (1990) proposes that Pseudogapping constructions result from VP-ellipsis, the remnant NP having moved out of the VP by heavy NP shift. In Lasnik (l995d) I argue that this proposal is correct in its essentials, though wrong in certain details. In particular, I modify Jayaseelan's analysis by positing raising to [Spec, Agr_O], instead of heavy NP shift, as the process removing the remnant from the ellipsis site.

Before proceeding, I would like to discuss a bit further the general analysis of Pseudogapping as a special case of VP-ellipsis. In the first detailed discussion of Pseudogapping that I have seen, Levin (1978) notes certain apparent differences (to which I will return) between Pseudogapping and VP-ellipsis, but nonetheless concludes that Pseudogapping is in fact VP-ellipsis, suggesting that the differences might follow from properties outside of the syntax.[9] Levin also cites Stump (1977), a work I have not seen, as arguing for a VP-ellipsis account. But later Levin (1979/1986) expresses skepticism. She voices a number of concerns, perhaps most significant among them that "backward" Pseudogapping is very severely degraded, unlike backward VP-ellipsis. The following pair is representative:[10]

(7) *Because John did Ø Clinton, Mary interviewed Gingrich.
(8) Because John didn't Ø, Mary interviewed Gingrich.

Interestingly, though, in notes that Levin added for the published version of Levin (1979), she states

> …I now believe it doesn't make much difference whether [Pseudogapping] is given separate treatment or collapsed with VP Deletion. Some of the restrictions on Pseudogapping would not need to be reflected in the rule, but could be relegated to the discourse component.
> (Levin 1986: 89)

This vague hint could, I believe, be turned into a plausible line of inquiry. The Pseudogapping construction involves strong contrastive focus, and it is conceivable that that property conflicts with the backward version of the construction, though there is not space here to try to give that speculation substance. I might note, though, that a purely syntactic account making Pseudogapping ellipsis of something other than VP does not seem particularly promising, on the face of it. Suppose Pseudogapping turns out to be YP-ellipsis for some $Y \neq V$. And suppose ellipsis can operate backward, as evidenced by the facts of VP-ellipsis. What purely syntactic factor could then prevent YP-ellipsis from operating backward? In light of these considerations, I regard the (admittedly robust) contrast in (7)–(8) as rather weak evidence against a VP-ellipsis account.[11]

Returning to the raising process rescuing the remnant from deletion, I note that under standard assumptions, raising to [Spec, Agr_O] is covert, taking place in the LF component. Given Jayaseelan's (1990) goal, adopted also in Lasnik (l995d) of

analyzing Pseudogapping as affecting a constituent, the ellipsis process must then be analyzed as copying in the LF component, rather than deletion in the PF component. However, on the theory of LF movement advocated by Chomsky (1995a), and further defended in Lasnik (l995b,c), the necessary structure would not even be created in *covert* syntax. On that theory, since movement is triggered by the need for formal features to be checked, all else being equal only formal features move. When movement is overt (triggered by a strong feature), PF requirements demand that an entire constituent move, via a sort of pied-piping. However, when movement is covert, PF requirements are irrelevant, so economy dictates that movement *not* affect the entire constituent. But then it is very difficult to see how covert raising of (the formal features of) accusative NP to [Spec, Agr_O] could possibly create an ellipsis-licensing configuration.

It seems then that if (feature-driven) movement newly creates a configuration licensing ellipsis, the movement must be overt rather than covert. Before I indicate how that is possible in the present instance, I note that if the movement is overt, then the conclusion above, that ellipsis must involve LF copying, no longer follows. If the licensing configuration must be created prior to the LF/PF split regardless, then ellipsis could just as easily be a PF deletion phenomenon.

Now early Minimalist Program literature (e.g. Chomsky (1991, 1993)) (Chomsky and Lasnik (1993)) did have accusative NP raising to [Spec, Agr_O], but covertly rather than overtly. However, Koizumi (1993, 1995), developing proposals by Johnson (1991), argues instead that that raising is always overt, driven as usual by a strong feature. In Lasnik (l995d) I suggest that the strong feature in this instance is an "Extended Projection Principle feature" (EPP) residing in Agr, hence the same feature that drives overt subject raising.[12] I will have little more to say here about this particular strong feature. I will, however, address another strong feature that must be involved in simple transitive sentences without ellipsis. Given that word order in English is V–O rather than O–V, if the complement raises out of VP, the verb must normally raise still higher. Koizumi's specific proposal, which he calls the *split VP hypothesis*, is that V raises, via Agr_O, to a higher "shell" V position, as shown in Figure 6.1 for the sentence in (9)

(9) You will believe Bob.

Under this general hypothesis, consider a simple Pseudogapping example such as (10).

(10) You might not believe me but you will Bob.

If *Bob* overtly raises to [Spec, Agr_O] while *believe* remains in situ, then deletion of the residual VP produces (10). The relevant structure is shown in Figure 6.2.

Deletion of the lower VP yields the Pseudogapping example in (10).

The question that now arises is why the V need not raise in Pseudogapping constructions, given that in nonelliptical sentences it must.

(11) *You will Bob believe.

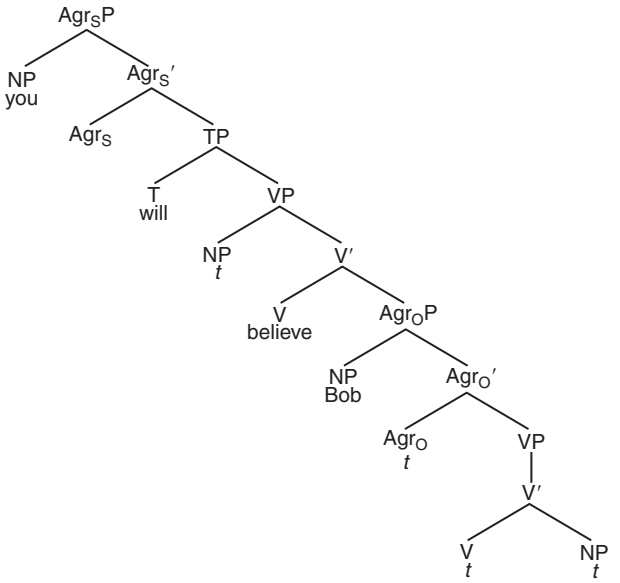

Figure 6.1 Koizumi's split VP hypothesis.

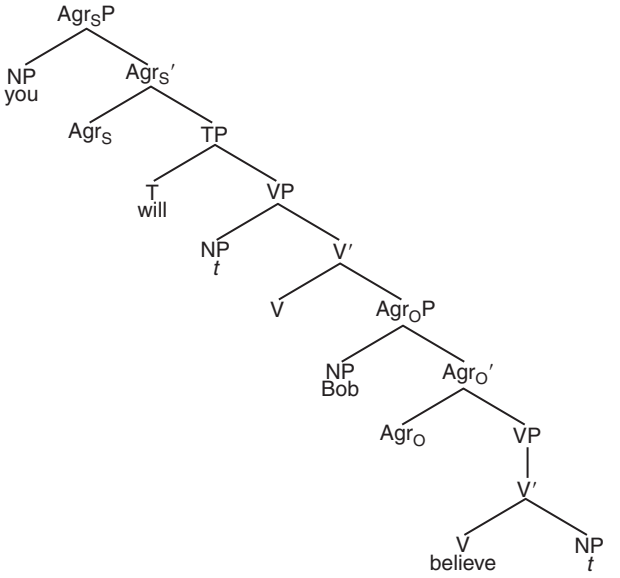

Figure 6.2 Source for simple Pseudogapping.

By hypothesis, a strong feature is involved. Yet, there seem to be two possibilities for a convergent derivation. The V can raise as in Figure 6.1, presumably checking the relevant strong feature. Alternatively, the V can be deleted along with its containing VP as in Figure 6.2. This mysterious state of affairs receives a rather straightforward account under the PF crash theory of strong features, under the new hypothesis that the strong feature forcing the V to raise overtly is a feature of the lexical V itself, rather than of the target position it raises to. The overt raising derivation is essentially unaffected by this change in perspective. The ellipsis structure is much more interesting. Consider Figure 6.2 again, but from the point of view offered by Figure 6.3.

If *believe* fails to raise, and no other relevant process takes place, the strong feature that is not overtly checked causes Figure 6.3 to crash at PF. But if the lower VP containing *believe* is deleted in the PF component, then, patently, the strong feature cannot cause a PF crash, since the (category containing the) feature will be gone at that level.[13] It is not obvious how to capture this result under the LF crash

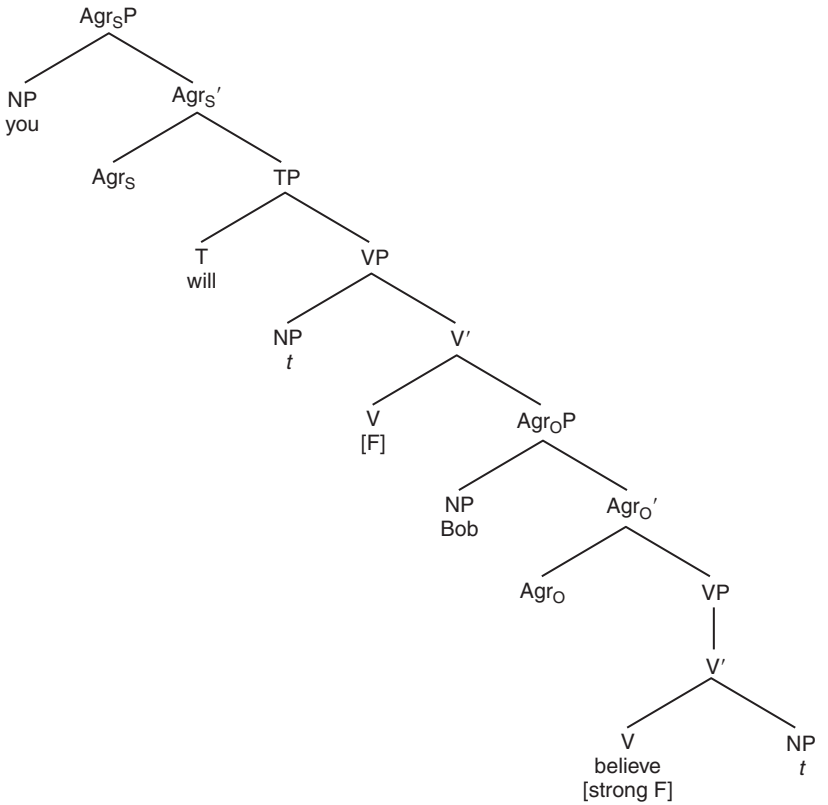

Figure 6.3 The strong feature not overtly checked causes a PF crash.

theory of strong features or the virus theory (though later I will suggest that the latter might be rendered compatible). Thus, we apparently have an argument for the PF crash theory.

6.3. Sluicing

Sluicing, an ellipsis phenomenon first investigated in detail by Ross (1969), displays an abstractly similar paradigm. Saito and Murasugi (1990) and Lobeck (1990) very plausibly analyze Sluicing as *wh*-movement followed by IP-ellipsis. (12) displays a representative example.

(12) Speaker A: Mary will see someone.
Speaker B: I wonder who ~~Mary will see~~.

Ross characterized Sluicing as an embedded-question phenomenon, and the standard examples, like (12), accord with that characterization. However, Sluicing also shows up in matrix interrogative contexts, as in (13).

(13) Speaker A: Mary will see someone.
Speaker B: Who ~~Mary will see?~~

The structure of the matrix Sluicing example (13) is presumably Figure 6.4 (with irrelevant details suppressed).

Note that under the assumption that Sluicing is, as standardly assumed, IP-ellipsis,[14] the source for the Sluicing example must be as indicated above, rather than, for example (14).

(14) Who will Mary see?

For (14) to be the source, the elided material would have to be C′, and not IP. But now an interesting, and interestingly familiar, question arises. Since the source of matrix Sluicing does not have I-raising to C, why is the source ungrammatical without Sluicing?[15]

(15) *Who Mary will see?

Clearly, in matrix questions like (14) there must be a strong feature driving the overt raising of I. But in the matrix Sluicing example (13) that strong feature has evidently not been checked in overt syntax. This is highly reminiscent of the situation with Pseudogapping discussed above, where the lexical verb is normally required to raise to the higher shell V position, but does not have to raise if the VP containing it is elliptical. Similarly, in the matrix Sluicing case I-raising to C, normally obligatory, does not take place if IP is elliptical. A parallel account is available. A priori, the strong feature forcing overt I-raising in matrix interrogatives might be a feature of C or of I. If it is a feature of I, and if we continue to assume that ellipsis is a PF deletion operation (Figure 6.4), then the facts fall nicely

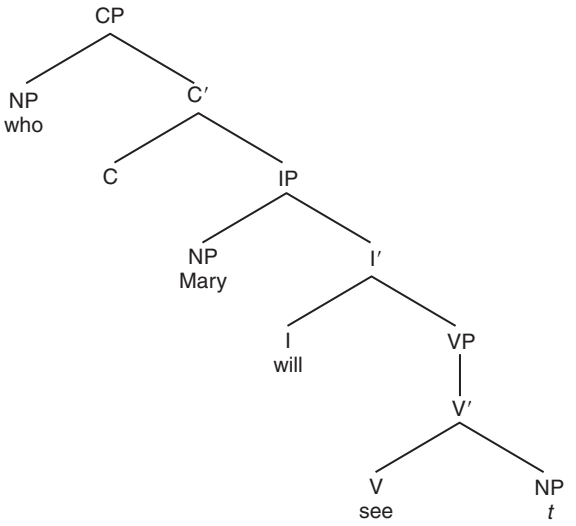

Figure 6.4 Source for matrix Sluicing.

within the purview of the PF crash theory. To see this, consider again but with the strong feature indicated as in Figure 6.5.

If I fails to raise and nothing further happens in the derivation, a PF crash results (15), the strong feature of I remaining unchecked. But if the IP is deleted, then nothing remains of the strong feature, checked or unchecked, at PF, so well-formed Sluicing (13), results.

Thus, for two separate constructions, the generalization is the same: an instance of normally obligatory overt movement does not take place (overtly) if the moved item is contained in an ellipsis site. And for both, the same natural account emerges:

(16)(a) Ellipsis is PF deletion.
(b) An unchecked strong feature causes a PF crash.
(c) In the constructions investigated here, the relevant strong feature resides in the item that (potentially) moves, rather than in the target.

With respect to (16c), I hypothesize that the strong feature is in matrix interrogative I, for matrix interrogatives with or without Sluicing; and in lexical V, for transitive constructions with or without Pseudogapping. The hypothesis regarding lexical V is based on Koizumi's (1993, 1995) approach to clause structure whereby accusative NP always raises overtly. For Koizumi, V invariably raises overtly as well, but I have argued that there are circumstances where this does not happen.

Although I have claimed that the two situations are parallel, that does not reflect the more common view. I-raising to C is quite widely taken to be overt in normal matrix interrogatives in English. On the other hand, apart from

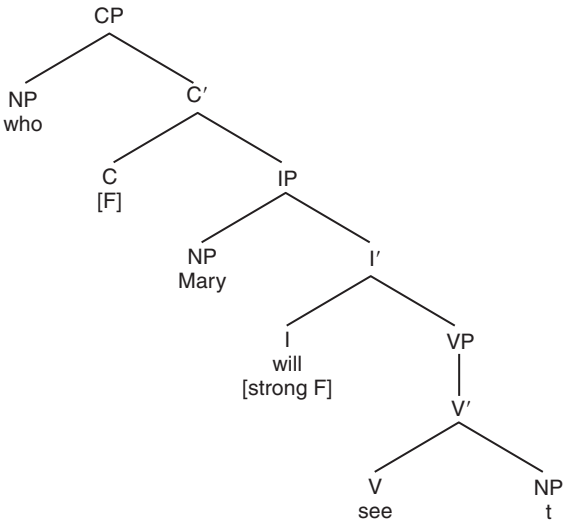

Figure 6.5 Structure for matrix Sluicing with strong feature indicated.

Koizumi's work, NP-raising to [Spec, Agr_O], and hence V-raising to a higher position, is standardly assumed to be covert in English. However, in Lasnik (1995c), based on Lasnik and Saito (1991) (see also Postal (1974) and Wyngaerd (1989)) and Dikken (1995), I have argued at length that such movement is indeed overt. Rather than repeating all of the arguments here, I will merely summarize their thrust and give a representative example. First, I have already argued above that if raising is to create a new ellipsis configuration, the raising must be overt, given the feature movement theory of covert movement. Second, there is an overwhelming generalization that with uncontroversial overt raising, the raised NP displays "high" behavior in all respects, whereas with covert raising (in existential constructions), the "associate" of the expletive displays "low" behavior. This contrast is shown for reciprocal binding in (17).

(17)(a) Some linguists seem to each other [*t* to have been given good job offers].
(b) *There seem to each other [*t* to have been some linguists given good job offers].

The explanation offered in Lasnik (1995c) for this contrast is as follows: When movement is covert – hence, following Chomsky (1995a), affecting only formal features – the referential and quantificational properties needed to create new binding and scope configurations are left behind, so no such new configurations are created.

Crucially, both simple direct objects and exceptional-Case-marked subjects pattern with the overtly raised NP in (17a) rather than with the covertly raised NP in (17b).

(18) The DA questioned two men during each other's trials.
(19) The DA proved [two men to have been at the scene] during each other's trials.

Arguably in (18), and almost certainly in (19), the base position of the antecedent of the reciprocal is too low for binding to obtain. Hence, raising is almost certainly involved. Given (17), there is strong reason to believe that the raising is overt. And then, given the word order of English, the verb must also have raised overtly. Thus, there is a reasonably firm basis for the proposals made in this article.

6.4. The PF theory vs the virus theory

With the ellipsis argument for the PF approach to strong features in hand, at this point it is necessary to consider possible arguments against that approach. One argument is that "look-ahead" is needed. At a given point in the overt portion of a derivation, it is apparently necessary to inspect the PF representation to see whether Procrastinate can be evaded.[16]

A second argument, Chomsky's original one, is that the ungrammatical (1), repeated here, has the possible well-formed derivation in (20).

(1) *John read what?
(20) Spell-Out: John read what
LF: C [strong Q] John read what

As noted above, if C with its strong feature is inserted in the covert component, at the level of PF that strong feature will not exist, hence cannot possibly cause a PF crash.

Chomsky's LF crash theory addresses the second of these arguments but has nothing to say about the first. The virus theory, repeated here, purports to deal with the first argument and indirectly with the second, as well as with the claimed conceptual problem that the PF and LF crash theories are just evasions.

Virus theory A strong feature must be eliminated (almost) immediately upon its introduction into the phrase marker.

Momentarily, we will see how Chomsky makes the virus theory precise. Note first, though, that the virus theory demands something I have already necessarily rejected: that a strong feature is always a property of the target of movement, never of the moved item. This is so since if an item that is to move were to have a strong feature, that feature could not, in general, be eliminated immediately. In some derivations the target that it would check against would be far, even indefinitely far, away. Thus, to the extent that the ellipsis analyses I have presented are

well supported, there is already reason to reject the virus theory. However, Chomsky's instantiation of this theory is of considerable import, so I will turn to it now.

Chomsky makes the following suggestion:

> The intuitive idea is that the strong feature merged at the root must be eliminated before it becomes part of a larger structure by further operations.
>
> (Chomsky 1995a: 234)

After considering how derivations work in general, he indicates that

> the descriptive property of strength is [(21)]. Suppose that the derivation D has formed Σ containing α with a strong feature F. Then
>
> [(21)] D is canceled if α is in a category not headed by α.
>
> (Chomsky 1995a: 234)

Chomsky observes two very interesting properties of this approach: (a) that cyclicity follows,[17] (b) that a strong feature is checked by an overt operation.

> We … virtually derive the conclusion that a strong feature triggers an *overt* operation to eliminate it by checking. This conclusion follows with a single exception: covert merger (at the root) of a lexical item that has a strong feature but no phonological features … .
>
> (Chomsky 1995a: 233)

This exception involves a kind of example we have seen before.

(1) *John read what?

Recall that it was this sort of example that led Chomsky to reject the PF crash theory in favor of the LF crash theory.[18] But, as Chomsky in effect acknowledges, the problem now arises anew in the virus theory. How can derivation (20), repeated here, be blocked?[19]

(20) Spell-Out: John read what
LF: C [strong Q] John read what

To prevent this, covert insertion of strong features must be barred. Chomsky proposes to do this with the economy principle (22).

(22) α enters the numeration only if it has an effect on output.

Chomsky elaborates on this as follows:

> With regard to the PF level, *effect* can be defined in terms of literal identity. … α is selected only if it changes the phonetic form. At the LF level the condition is perhaps slightly weaker, allowing a narrow and readily computable form of logical equivalence to be interpreted as identity.
>
> (Chomsky 1995a: 294)

This immediately raises a question concerning the central argument for the virus theory – that it eliminates the look-ahead inherent in the PF and LF crash theories. There seems to be considerable look-ahead here, all the way from the very beginning of the derivation, the numeration, to the very end, phonetics and semantics.

> Under [(22)], the reference set [for economy comparisons] is still determined by the numeration, but output conditions enter into determination of the numeration itself….
>
> (Chomsky 1995a: 294)

Apart from this conceptual question, there is an empirical question about whether the correct result is in fact obtained. There is reason to think that it is not.

Consider the situation at issue, insertion in the LF component of interrogative C in English, a language in which C has a strong *wh*-feature. (22) purports to prevent this. The first question is whether this C has an effect on output. Clearly, covert insertion of a C will have no phonetic effect. Will it have an effect at the LF output? Either it will or it will not. If it will, then covert insertion is allowed, and we generate (1) with structure (23).

(23) C [$_{IP}$ John read what]

Since this is not the correct result, suppose instead that C will not have a semantic effect. Then we cannot generate (1) with structure (23), so the problem is apparently solved under the assumption that insertion of interrogative C has no effect on semantic output. As Chomsky states the situation:[20]

> … the interface representations (π, λ) are virtually identical whether the operation takes place or not. The PF representations are in fact identical, and the LF ones differ only trivially in form, and not at all in interpretation.
>
> (Chomsky 1995a: 294)

But our goal is actually more general than just ruling out (1) with structure (23). Rather, it is ruling out (1) altogether. Under the assumptions just spelled out, (1) is successfully excluded with C covertly inserted. But what if C is not inserted at all? That is, what if the structure is the same at both LF and "S-Structure?"

(24) [$_{IP}$ John read what]

(24) violates no morphological requirements, and, if C has no effect on output, the assumption that was necessary in order to exclude (1) *with* C inserted, then it should mean exactly *What did John read?* So if C has a semantic effect, inserting it in LF should be permitted. And if it does not have a semantic effect, *not* inserting it should be of no consequence. Thus, even given the new economy condition (22), (1) is allowed, and allowed as a standard interrogative, presumably an incorrect result. In this regard too, then, the PF crash theory of strong features fares no worse than the virus theory. Either one demands an additional stipulation, perhaps just that lexical insertion is prohibited in the covert component[21] (a result Chomsky was trying to deduce, but, as we have just seen, not completely successfully).

6.5. The virus theory and ellipsis

At this point there are no clear arguments in favor of the virus theory of strong features over the PF crash theory. Further, the ellipsis paradigms discussed above seem to provide considerable support for the latter approach. Interestingly, though, there is a potential way to reconcile the ellipsis facts with the virus theory. I will end my investigation by showing how such an account would work. Recall that Chomsky (1995a) proposes that strength is always a property of an "attracting" head, never a property of the item that moves. This is necessary under the virus theory (at least as Chomsky articulates it) since a strong feature in an item to be moved would never be checked quickly enough to keep the derivation from terminating. The above analyses of Pseudogapping and Sluicing are incompatible with that proposal, demanding, as they do, that the moved item sometimes have the strong feature. There is a possible alternative analysis, based on Chomsky's (1995a) theory of pied-piping, particularly as explicated by Ochi (1998).[22]

Ochi, following Chomsky, considers the nature of pied-piping, the usual reflex of movement triggered by a strong feature. Chomsky (1995a) gives the following characterization:

> For the most part – perhaps completely – it is properties of the phonological component that require such pied-piping. Isolated features and other scattered parts of words may not be subject to its rules, in which case the derivation is canceled; or the derivation might proceed to PF with elements that are "unpronounceable," violating FI [Full Interpretation].
>
> (Chomsky 1995a: 262)

Overt movement consists of a complex of operations under this approach.

> Applied to the feature F, the operation Move thus creates at least one and perhaps two "derivative chains" alongside the chain $CH_F = (F, t_F)$

> constructed by the operation itself. One is $CH_{FF} = (FF[F], t_{FF[F]})$, consisting of the set of formal features FF[F] and its trace; the other is $CH_{CAT} = (\alpha, t_{\alpha})$, α a category carried along by generalized pied-piping and including at least the lexical item LI containing F. CH_{FF} is always constructed, CH_{CAT} only when required for convergence .... As noted, CH_{CAT} should be completely dispensable, were it not for the need to accommodate to the sensorimotor apparatus.
>
> (Chomsky 1995a: 265)

Note that this seems to assume the second of the two possibilities Chomsky mentioned in the prior passage, that is, that failure of pied-piping causes a violation specifically at PF.[23] Chomsky goes on to observe that even overt movement might be possible without pied-piping under certain circumstances, if no phonological requirement is violated.

> Just how broadly considerations of PF convergence might extend is unclear, pending better understanding of morphology and the internal structure of phrases. Note that such considerations could permit raising without pied-piping even overtly, depending on morphological structure ....
>
> (Chomsky 1995a: 264)

Consider now how the ellipsis phenomena examined above might be reanalyzed in terms of this theory. Recall my analysis of Pseudogapping in terms of the PF crash theory of strong features. Assuming the split VP hypothesis, in a nonelliptical transitive sentence, for example, the object raises to [Spec, Agr_O] and the lexical V raises to the higher shell V position in order that a strong feature of the lexical V will be checked. If the V does not raise, a PF crash will ensue, but only if the offending item exists at that level. Under a deletion account of ellipsis, ellipsis provides another way to salvage the derivation. When the lower VP is deleted without the V having raised, a PF crash is avoided and the result is acceptable Pseudogapping.

The alternative account preserves the idea of deletion averting a PF crash, but the potential crash now has another cause. The feature driving overt V-raising could be a strong feature of the *higher* V. Once the matching feature of the lower lexical V is "attracted" out of the lower V, the lower V becomes defective. A PF crash will be avoided if either pied-piping or deletion of a category containing the lower V (VP deletion = Pseudogapping in the relevant instances) takes place. This is illustrated in Figure 6.6.[24] Thus, even under the virus theory there is a way to capture the saving effect of ellipsis in the Pseudogapping construction.

Sluicing can be reanalyzed in parallel fashion. Suppose that in accord with the virus theory the strong feature driving overt I-raising in matrix interrogatives resides in C (the usual, and arguably more natural, assumption, at any rate), rather than in I. In a normal matrix interrogative, then, the matching feature of

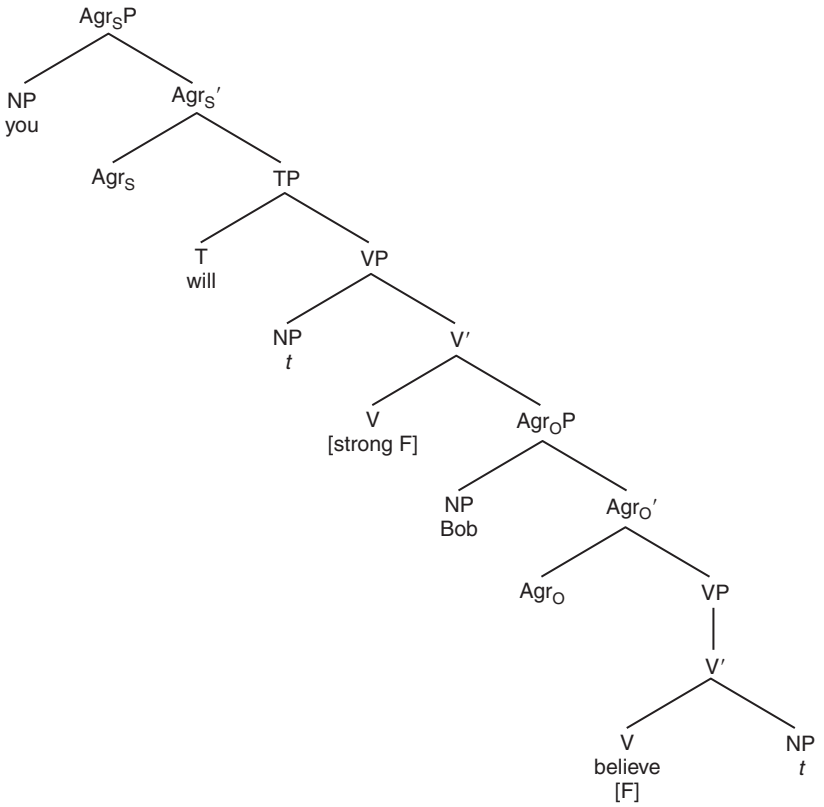

Figure 6.6 A PF crash will be avoided if either pied-piping or deletion of a category containing V takes place.

I raises overtly to check the strong feature of C. This leaves behind a phonologically defective I, which will cause a PF crash unless either pied-piping (i.e. overt raising of I) or deletion of a category containing that I (Sluicing) takes place. This is illustrated in Figure 6.7.

Thus, the essence of the PF account of the ellipsis facts based on the PF crash theory of strong features can be captured under the virus theory as well, a rather surprising result – and perhaps a welcome one if Chomsky's conceptual arguments for the virus theory are accepted. Before concluding, though, I will briefly mention one last, and rather well-known, argument that strong features reside in some moving categories and that the basic premise of the virus theory is therefore incorrect. There is a great deal of literature, going back to Toman (1982) and Rudin (1982, 1988), discussing the phenomenon of multiple *wh*-movement in the Slavic languages. Bošković (1998) presents a treatment of Serbo-Croatian multiple *wh*-movement in terms directly relevant to the present discussion. He argues that

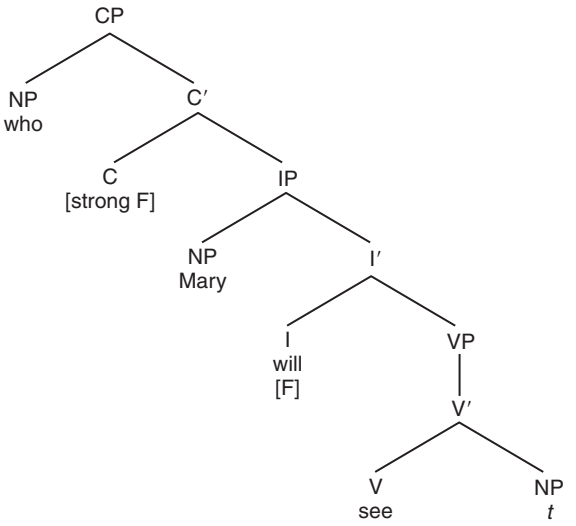

Figure 6.7 A PF crash will be avoided if either pied-piping or deletion of a category containing Infl takes place.

in Serbo-Croatian, *wh*-phrases have a strong focus feature and that that is why they all have to move overtly.

(25)(a) Ko šta gdje kupuje?
who what where buys
'Who buys what where?'
(b) *Ko kupuje šta gdje?
(c) *Ko šta kupuje gdje?
(d) *Ko gdje kupuje šta?

Whether he is right about the precise identity of the feature will not be of concern here. But the fact that all of the *wh*-phrases must move overtly strongly suggests Bošković's basic conclusion, that the strong feature driving the movement resides in the *wh*-phrases themselves. If, instead, it resided in the head to which they move, why wouldn't the movement of just one of them suffice (as in English)?

To summarize, I set out to compare three theories of strong features. Contrary to expectation, existing conceptual arguments based on the computational complexity of look-ahead and on the problem of LF insertion of a head with a strong feature turned out to be inconclusive. An examination of two ellipsis paradigms provided new potential evidence. I suggested a new generalization: that movement *or* ellipsis can rescue a derivation with a strong feature. In light of this, I argued that PF crash is relevant, either directly, as in Chomsky's PF crash theory of strong

features (Chomsky (1993)), or indirectly, as in the virus theory augmented by the multiple-chain theory of pied-piping (especially as interpreted by Ochi (1998)). The existence of constructions, such as Slavic multiple interrogation, where multiple items apparently must move to the same position, tips the balance in favor of the PF crash theory, given present (admittedly limited) understanding.

7

A GAP IN AN ELLIPSIS PARADIGM[1]

Some theoretical implications

The analysis of English verbal morphology of Lasnik (1995e) and Chapter 2 above has been challenged by Roberts (1998) and Potsdam (1997). Here, I respond to those challenges. Lasnik (1995e) was actually primarily concerned with the basic ordering of verbal elements with respect to negation. In particular, the concern was the classic Chomsky (1955, 1957) adjacency requirement on Affix Hopping preventing the association of Infl and V across an intervening *not*, as in **Mary not hired John*/**Mary hired not John*. I showed how a strictly lexicalist account of verbal morphology, like that in Chomsky (1993) has great difficulty in dealing with the distribution of negation. I therefore proposed a partial return to the Chomsky (1955, 1957) account, where in underlying structure, Infl is a syntactically independent item. On the other hand, I argued that auxiliary verbs are "lexicalist." This hybrid theory handles the full range of facts about negation in English. Additionally, I argued, it handles a surprising gap in the paradigm of English VP-ellipsis: *John slept and Mary will too*/**John was here and Mary will too*. Roberts and Potsdam, concentrating just on this ellipsis fact, propose an alternative that relies on a special constraint on ellipsis. I show that that alternative faces serious empirical and technical difficulties.

Lasnik (1995e) puts forward a "hybrid" theory of English verbal morphology in which inflected auxiliary verbs (particularly finite forms) are analyzed in "lexicalist" fashion while main verbs are argued to be derivationally constructed out of syntactically separate stem and affix, very much as in Chomsky (1955) and Chomsky (1957). There were two central concerns motivating this theory. First, and most importantly, the long known paradigmatic differences between auxiliary and main verbs are shown to follow rather naturally from the proposal that auxiliary verbs are drawn from the lexicon fully inflected and raise to Infl to check their inflectional features, while main verbs are drawn from the lexicon as bare stems

(hence lacking inflectional features to check) and therefore do not raise. The contrast between auxiliary and main verbs was shown to be problematic for both of Chomsky's recent analyses of verbal morphology, Chomsky (1991) and Chomsky (1993), as was the distribution of *not* with main verbs. Second, the Lasnik (1995e) analysis provided an account of a surprising gap in a VP-ellipsis paradigm noted by Warner (1986). I will begin by summarizing the ellipsis facts, since two recent papers have presented interesting alternative accounts of those facts.

7.1. A gap in a paradigm

It has long been known that VP-ellipsis can ignore certain inflectional differences between antecedent verb and elided verb. For example, Quirk *et al.* (1972) reported by Sag (1976), observe that a finite form of a verb can antecede the deletion[2] of the bare form that follows a modal, as in the following examples:

(1) John slept, and Mary will too

(2)(a) *John *slept*, and Mary will *slept* too
(b) John *slept*, and Mary will *sleep* too

(3) John sleeps (every afternoon), and Mary should too

(4)(a) *John *sleeps*, and Mary should *sleeps* too
(b) John *sleeps*, and Mary should *sleep* too

In (1) and (3), the past tense form *slept* and the present form *sleeps* serve as antecedent for the deletion of the bare form *sleep*. This suggests that verbal morphology is simply irrelevant for VP-ellipsis, that one inflectional form of a verb counts as "identical" to another. However, as Warner (1986) notes, ellipsis with auxiliaries has a stricter requirement. For example, finite forms of *be* cannot antecede the bare form, as seen in the following examples:

(5) *John was here, and Mary will too

(6)(a) *John was here and Mary will *was* here too
(b) John was here and Mary will *be* here too

Given the central proposal of Lasnik (1995e), this contrast follows under the assumption that a form of a verb V can only be deleted under identity with the very same form. If forms of "main" verbs are created out of lexically introduced bare forms and independent affixes, (1) and (3) will still be allowed, since prior to "Affix Hopping," identity does obtain:

(7) John [Af] sleep, and Mary will ~~sleep~~ too

On the other hand, under the Lasnik (1995e) proposal, *was* (or *is*) is never identical to *be*, so examples like (5) are excluded.

7.2. An alternative treatment of the gap?

Given that finite forms of *be* raise, while finite forms of main verbs do not, could it be that, for some reason, a trace can not serve as (part of) an antecedent for ellipsis? This possibility was considered, and rejected, in Lasnik (1995e). First, it seemed stipulative. Second, there are a number of constructions where a trace can serve as part of an ellipsis antecedent, under plausible assumptions about the transformational derivations of the examples:

(8) Linguistics, I [like *t*], and you should ~~like linguistics~~ too
(9) ?Someone will [be *t* in the office]. Yes there will ~~be someone in the office~~.
(10) That this approach will fail can [be proven *t*]. No it can't ~~be proven that this approach will fail~~.

Roberts (1998) and Potsdam (1997) nonetheless reject the Lasnik (1995e) proposal in favor of a version of the alternative.[3] Both Roberts and Potsdam suggest, limiting the trace constraint to the head of VP, thus evading the potential counterexamples in (8)–(10). Their respective statements are as follows:

(11) "[$_{VP}$[$_{V}$ e] X] cannot antecede VP-ellipsis." (Roberts 1998: 118).
(12) "… a trace of verb movement cannot serve as part of a VPE antecedent." (Potsdam 1997: 362).

As stated, (11) and (12) are far too strong. A number of languages with overt V raising to I nonetheless allow VP-ellipsis, with the effect that everything in the VP except the V is deleted. Doron (1990) shows this for Hebrew:

(13) Q: Salaxt et ha-yeladim le- beit-ha-sefer
you-sent Ace the kids to school
"Did you send the kids to school?"

A: Salaxti
I sent
"I did"

Martins (1994) shows the same thing for Portuguese and McCloskey (1990) does for Irish:

(14) A Martas deu um livro ao João? Sim, deu.
the Martha gave a book to- the John yes gave
"Did Martha give a book to John? Yes, she did."

(15) Q: Ar chuir tú isteach air
INTERR COMP put [PAST] you in on it
"Did you apply for it?"

A: Chuir
put [PAST]
"Yes."

In all of these cases, it appears that the antecedent VP has as its head the trace of the raised V, yet VP-ellipsis is still possible. Simple English examples might also have this property. A modified version of (5) above arguably is relevantly parallel to the Hebrew, Portuguese, and Irish examples just presented:

(16) John was here and Mary was too

I assume that what both Roberts and Potsdam actually intend is something along the lines of (17).

(17) $[_{VP}\ [_{V}\ e]\ X\ ]$ cannot antecede VP-ellipsis of $[_{VP}\ [V\ ]\ X\ ]$, where V is lexical.

In (13) through (16), both the antecedent and the ellipsis site contain the trace of V, so (17) is not violated. On the other hand, in the ill-formed (5), only the antecedent contains a trace, so (17) successfully makes the distinction.
Roberts offers the following justification for a constraint along these lines

> regarding it as an identity condition on formal features of the two Vs: a raised V has different features from a nonraised V, at least in that the raised V has the feature F that triggers movement by needing to be checked, and the copy does not.
>
> (Roberts 1998: 118)

Eventually, I will return to more detailed discussion of (17) and its motivation, but first I discuss in the next section one likely counterexample to (17), and another likely counterexample to a generalization of (17) that would be expected under Roberts's justification quoted just above.

7.3. Problems for the alternative

In this section, I will first show that Pseudogapping is a very strong candidate for a VP headed by verb trace anteceding deletion of a VP headed by a lexical verb. Pseudogapping is an ellipsis phenomenon that displays some properties of gapping (there is a right-side remnant) alongside some properties of VP-ellipsis (there is a finite auxiliary):

(18) John will select me, and Bill will you

Sag (1976) presents a number of instances, suggesting that they relate to VP deletion, and tentatively concluding that VP deletion must therefore be formulated as a rule deleting a variable (rather than specifically a VP), since a portion of the VP survives the deletion. The following is from Sag (1976):

(19) Mary hasn't dated Bill, but she has Ø Harry.

Levin (1978, 1979/1986) provides an extensive examination of this type of ellipsis, using the name it is now standardly associated with: Pseudogapping. Among her many examples are the following, all from Levin (1978), and all marked? by her:

(20) If you don't believe me, you will Ø the weatherman
(21) I rolled up a newspaper, and Lynn did Ø a magazine
(22) Kathy likes astronomy, but she doesn't Ø meteorology

By and large, the best instances of Pseudogapping involve an NP remnant. Levin (1978) cites the following unacceptable examples with adjectival remnants:

(23) *You probably just feel relieved, but I do Ø jubilant
(24) *Rona sounded annoyed, and Sue did Ø frustrated
(25) These leeks taste terrible. *Your steak will Ø better.

While in many instances, it might appear that the process is simply elision of the main verb, there is considerable evidence that more is involved. There are clear instances in which far more than just the main verb is elided:

(26) The DA proved Jones guilty and the Assistant DA will ~~prove~~ Smith ~~guilty~~
(27) ?John gave Bill a lot of money, and Mary will ~~give~~ Susan ~~a lot of money~~

Rejecting the possibility of an ellipsis rule affecting a discontinuous portion of the structure (as seen in (26) and (27) for example), Jayaseelan (1990) proposes that Pseudogapping constructions result from VP-ellipsis, with the remnant having moved out of the VP by heavy NP shift (HNPS). In Lasnik (1995d). I argue that this proposal is correct in its essentials, though wrong in certain details. In particular Pseudogapping does not entirely correlate with the possibility of HNPS. I have already illustrated Pseudogapping with the first object in a double object construction as remnant. But the first object in a double object construction is resistant to undergoing HNPS:

(28) ?John gave Bill a lot of money, and Mary will ~~give~~ Susan ~~a lot of money~~
(29) *John gave *t* a lot of money the fund for the preservation of VOS languages

Conversely, the second object is a poor Pseudogapping remnant, but freely undergoes HNPS:

(30) *John gave Bill a lot of money, and Mary will ~~give Bill~~ a lot of advice
(31) John gave Bill *t* yesterday more money than he had ever seen

In these respects, Pseudogapping correlates more closely with passive (an instance of A-movement) than with HNPS:

(32) Bill was given *t* a lot of money
(33) *A lot of money was given Bill *t*

Partly based on these considerations, in Lasnik (1995d) I propose that Pseudogapping is overt NP raising to [Spec, Agr_O] (A-movement), followed by VP-ellipsis. This is illustrated in (34), which relies on the proposal of Koizumi (1993), following Johnson (1991), that "object shift" is overt (in fact, invariably so) in English.

(34)(a) John hired Bill and Mary will Susan
(b) John [$_{VP}$ hired [$_{AgrP}$ Bill [$_{VP}$ *t t*]]] and Mary will [$_{AgrP}$ Susan ~~[$_{VP}$ hire *t*]]~~

Note that given the usual word order of English, in "John hired Bill," the assumption that the object raises entails that the verb raises to a still higher position. For Koizumi, that higher position is the head V of a VP shell, under his "split VP" hypothesis. Under this analysis, which has considerable support, Pseudogapping is thus a direct counterexample to the ellipsis constraint given in (17) above, and repeated here as (35).

(35) [$_{VP}$ [$_{V}$ e]X] cannot antecede VP-ellipsis of [$_{VP}$ [V] X], where V is lexical.

There is a second potential problem. Given Roberts's rationale for (35), it should presumably generalize to all heads, not be limited to V and trace of V:

(36) [$_{YP}$ [$_{Y}$ e] X] cannot antecede YP-ellipsis of [$_{YP}$ [Y] X], where Y is lexical.

Sluicing (Ross (1969)), now standardly analyzed as IP-ellipsis (Lobeck (1990) and Saito and Murasugi (1990)), provides a potential counterexample to (36). Sluicing involves a fronted *wh*-phrase as the sole remnant of a *wh*-question. Standard examples Sluicing in an embedded question:

(37) Speaker A: Mary will see someone
Speaker B: Tell me who ~~Mary will see~~

Now notice that the antecedent for Sluicing can display Infl raising to Comp:

(38) Speaker A: Never will [$_{IP}$ Harry *t* go to a linguistics lecture again]
Speaker B: Tell me why [$_{IP}$ ~~Harry will never go to a linguistics lecture again~~]

(39) Speaker A: Never will [$_{IP}$ Susan *t* understand some linguists]
Speaker B: Tell me which linguists [$_{IP}$ ~~Susan will never understand~~]

On standard assumptions, than, we have another set of counterexamples to the generalization of (35)–(36).

7.4. Why isn't Roberts's line of reasoning valid?

Given that the trace of a raised X^0 has had a feature (or set of features) checked and deleted, why *can* it antecede the deletion of an XP with its head in situ (as in Pseudogapping and Sluicing)? One would not a priori expect this. Further, Oku (2001) shows on empirical grounds that the features of the antecedent for ellipsis must be a superset of the features of the elided item. Before attempting to answer this question, I will outline an independent question that is ultimately related to this one.

I have claimed that NP raises but V does not raise in the Pseudogapping construction. But then, why must V raise in the corresponding nonelliptical version (and, more generally, in all nonelliptical sentences where there is a complement to the verb)?

(40)(a) … and Mary will Susan
(b) … and Mary will [$_{\text{AgrP}}$ Susan ~~[$_{\text{VP}}$ hire *t*]~~]

(41) *Mary will Susan hire

A parallel question arises with respect to certain instances of Sluicing. Although the standard examples of Sluicing are in embedded questions, Sluicing is also possible in matrix questions:

(42) Speaker A: Mary will see someone
Speaker B: Who ~~Mary will see~~

(43) Speaker A: Never will [$_{\text{IP}}$, Harry *t* go to a linguistics lecture again]
Speaker B: Why [$_{\text{IP}}$ ~~Harry will never go to a linguistics lecture again~~]

(44) Speaker A: Never will [$_{\text{IP}}$ Susan *t* understand some linguists]
Speaker B: Which linguists [$_{\text{IP}}$ ~~Susan will never understand~~]

Given that Infl does not raise to Comp in the Sluicing construction, why must Infl raise in the corresponding matrix nonelliptical version (and, more generally, in virtually all matrix *wh*-questions)?

(45) *Who Mary will see?
(46) *Why Harry will never go to a linguistics lecture again?
(47) *Which linguists Susan will never understand?

As for why overt movement is necessary, I will assume here the theory of Chomsky (1995a). Under this theory, overt movement is driven by a "strong feature" of a head, which attracts a matching feature within the complement of that head.[4] Further, according to this theory, the "attraction" takes place immediately upon the introduction of the strong feature into the phrase marker. All movement, whether covert or overt, is fundamentally feature movement. When movement is

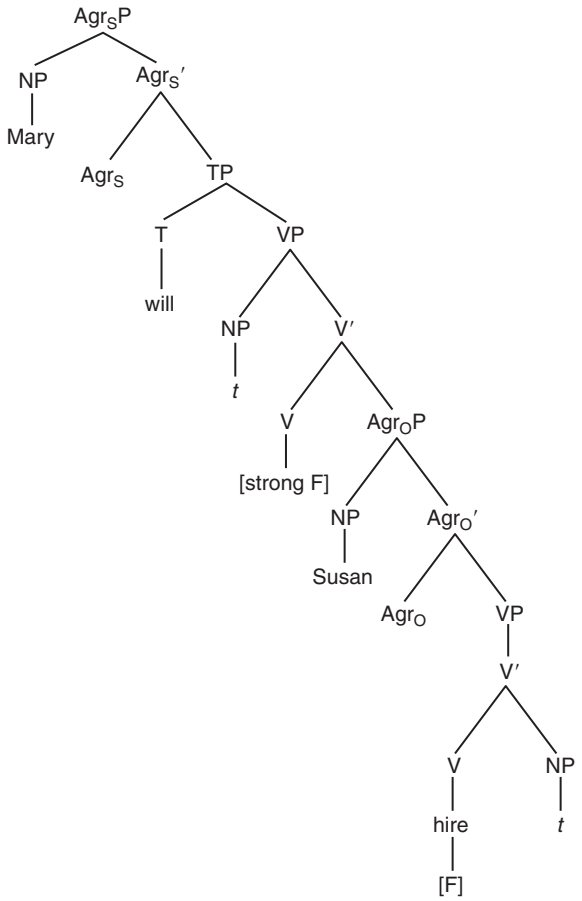

Figure 7.1 Split VP structure with the strong feature of the shell V attracting a matching feature of the lower V.

overt, though, the attracted feature typically drags along an entire lexical item (head-movement) or phrase (XP-movement), via a sort of pied-piping, as motivated in the following passage:

> For the most part – perhaps completely – it is properties of the phonological component that require pied-piping. Isolated features and other scattered parts of words may not be subject to its rules, in which case the derivation is canceled; or the derivation might proceed to PF with elements that are 'unpronounceable,' violating FI.
>
> (Chomsky 1995a: 262)

Given the "split VP" hypothesis of Koizumi (1993, 1995), which I assume here, the structure of (48) is as in Figure 7.1.[5]

(48) Mary will hire Susan

The "shell" V in Figure 7.1 attracts a feature of the lexical V *hire*. In the derivation, of (48), the V pied-pipes along with the attracted feature, creating a head chain along with the basic feature chain.

> Applied to the feature F, the operation Move thus creates at least one and perhaps two "derivative chains" alongside the chain $CH_F = (F, t_F)$ constructed by the operation itself. One is $CH_{FF} = (FF[F], t_{FF[F]})$, consisting of the set of formal features FF[F] and its trace; the other is $CH_{CAT} = (\alpha, t_\alpha)$, α a category carried a long by generalized pied-piping and including at least the lexical item containing F. CH_{FF}, is always constructed, CH_{CAT} only when required for convergence … As noted, CH_{CAT} should be completely dispensable, were it not for the need to accommodate to the sensorimotor apparatus.
>
> (Chomsky 1995a: 265)

Thus, if pied-piping were not to take place when attraction is overt, the result would be phonologically defective. Chomsky speculates about the possibility of avoiding this phonological defect without pied-piping:

> Just how broadly considerations of PF convergence might extend is unclear, pending better understanding of morphology and the internal structure of phrases. Note that such considerations could permit raising without pied-piping even overtly, depending on morphological structure … .
>
> (Chomsky 1995a: 264)

Ochi (1997) explores this issue further, proposing that the phonological defect resides specifically in the category left behind by the feature movement. Raising of that category (i.e. pied-piping) remedies the defect by placing the category in close enough proximity to the raised features. From this perspective, there might indeed be an alternative to pied-piping: deletion of (a category containing) the item that "lost" features to raising. If the potentially offending item is deleted, it cannot cause a PF violation, since it will no longer exist at that level. More specifically, in Figure 7.1, if only the attracted features raise, but the V does not raise, a PF crash will ensue, but only if the offending item exists at that level. Deletion provides another way to salvage the derivation. When the lower VP is deleted without the V having raised, a PF crash is avoided and the result is acceptable Pseudogapping:

(49) … Mary will Susan

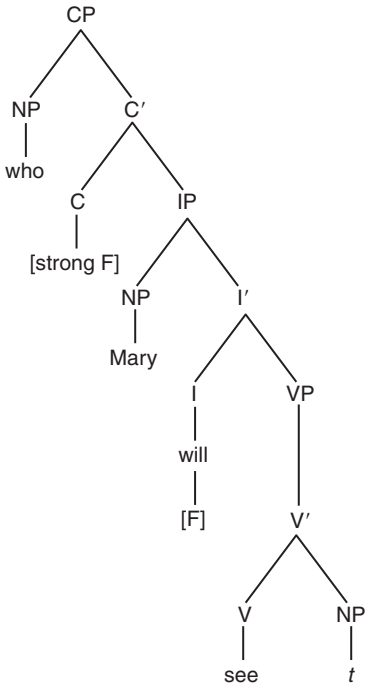

Figure 7.2 Strong feature of C attracting a matching feature of Infl.

The same line of analysis is available for matrix Sluicing. Recall the problem: In a matrix *wh*-question, Infl raising to C is normally obligatory, but it does not apply when the IP is elided. Consider the structure of a matrix *wh*-question.

When C attracts Infl in Figure 7.2, *will* becomes phonologically defective. A PF crash will ensue unless either *will* raises, producing (50), or IP is deleted (and defective *will* along with it) producing (51), an instance of matrix Sluicing.

(50) Who will Mary *t* see
(51) Who ~~Mary will see~~

Note that now, the major prima facie counterexamples to Roberts's proposal ((17), as generalized to (36)), are completely compatible with it, or more accurately, are compatible with Roberts's justification for the constraint. True, we still have the trace of a raised item X serving as part of the antecedent XP for ellipsis of a corresponding XP′ with unraised X′. But under the approach sketched here, the trace and unraised item are identical in syntactic features. Following Roberts's assumption, the trace has lost the feature(s) that triggered raising and thus participated in checking. My examples of Pseudogapping and matrix Sluicing therefore seemed to run afoul of the plausible constraint prohibiting an item with fewer

features from anteceding ellipsis of an item with more. But I have suggested that the apparently unraised V in Pseudogapping and Infl in matrix Sluicing actually have participated in feature raising, just not in pied-piping. In terms of features, there is identity; the same checking has taken place in X and X′.[6]

7.5. Roberts's account reconsidered

With Pseudogapping and matrix Sluicing eliminated as counterexamples to a theory in the spirit of Roberts's, can Roberts's account of the gap in the VP-ellipsis paradigm (repeated as (52)–(53)) now be accepted?

(52) John slept, and Mary will ~~sleep~~ too
(53) *John was here, and Mary will ~~be here~~ too

In (53), *be* does not raise at all, with or without pied-piping, whereas *was* obviously does raise, resulting in features being checked and deleted. So some version of Roberts's constraint would seem to correctly exclude ellipsis here. However, on deeper inspection, there is still a difficulty. Recall Roberts's justification for his constraint, in the spirit of recoverability of deletion:

> A trace left behind by raising has lost the feature that triggered raising while an item that never raises has not. Thus, the feature of antecedent *was* and target *be* differ in features.
>
> (Roberts 1998)

But in this instance, what are those features? It is hard to see how they could be anything other than inflectional features. Yet, checking and deleting the inflectional features of *was*, and hence of its trace, makes it **more** like *be*, not **less** like *be*, since the latter lacks inflectional features altogether. It is thus very difficult to see how any version of recoverability based on formal features will correctly distinguish between (52) and (53).

7.6. Another kind of justification for (36)

Potsdam (1997), apparently independently of Roberts, proposes the same constraint on ellipsis and for the same reason: to account for the ellipsis paradigm gap. Like Roberts, he limits the constraint to X^0 traces, acknowledging that XPs do not conform. Potsdam, unlike Roberts, does not attempt to deduce the constraint, but he does suggest that the X^0 vs XP asymmetry in this case follows from a more general distinction. In particular, Potsdam claims that "… corresponding X^0 traces must have the same binder in both the antecedent and target clauses … in contrast to the behavior of XP traces …" (Potsdam 1997: 365).[7] To illustrate this property of XP traces, he gives (54), among other examples.

(54) Chicken, she'll eat, but ostrich, she won't

In contrast, he claims that in Hebrew and Irish, both V-raising languages that have VP-ellipsis, "the raised verbs in ellipsis antecedent and target clauses must be the same" (Potsdam 1997: 365). Potsdam suggests that (55) is a universal condition, indicating that he knows of no language where this constraint does not hold.[8]

(55) Corresponding X^0 traces must have the same binder in both the antecedent and target clauses.

The universality is questionable, however. Serbo-Croatian seems to be a language lacking the constraint.[9] In the following example, "carefully" is understood in both conjuncts, strongly suggesting that what is at issue is ellipsis, and not merely a null object in the second conjunct.

(56) Ivan piše rad pažldivo, a njegov asistent čita
Ivan writes paper carefully and his assistant reads
"Ivan is writing a paper carefully, and his assistant is reading it carefully."

Further, if, following standard assumptions, we take the possibility of sloppy identity as diagnostic of ellipsis,[10] the following example provides additional evidence for VP-ellipsis with nonidentical raised verbs in Serbo-Croatian:

(57) Marko gradi sebi kucu, a Marija kupuje
Marko builds himself house and Marija buys
"Marko is building himself a house, and Maria is buying herself a house."

Indeed, there is serious doubt that (55) holds even in Hebrew, one of Potsdam's only two example languages.[11] Consider the major argument of Doron (1990). Taking it for granted that only ellipsis (and not null objects) gives rise to sloppy identity readings, she presents the following pair of discourses (her example 12).[12]

(58) Q: dina soreget et ha- svederim Se- hi loveSet
Dina knits ACC the sweaters that she wears
"Does Dina knit the sweaters that she wears?"

A1: lo, aval ima Sela soreget
no, but mother hers knits
"No, but her mother does."

A2: lo, ima Sela kona (la)
no, mother hers buys (to-her)
"No, her mother buys them (for her)."

According to Doron,

> The first answer, [A1], is ambiguous just as its English translation indicates: Dina's mother knits Dina's sweaters under one reading, but she knits her own sweaters under the other reading. This construction therefore shows a characteristic of VP-ellipsis. The second answer, [A2], is not ambiguous: Dina's mother buys Dina's sweaters, not her own. This non-ambiguity is to be expected, since this answer, having a distinct verb from the question, does not involve VP-ellipsis but a missing object … .
>
> (Doron 1990: 9–10)

A fuller explication of the missing (sloppy) reading of A2 reveals an interfering factor:

(59) Q: Does Dina knit the sweaters that she wears?
A: No her mother buys the sweaters that she$_i$ wears

The putative answer (59A) is strikingly unresponsive to the question. This factor alone could render sloppy identity highly inaccessible in (58A2). Once this discourse oddity is controlled for, the picture becomes very different.[13] Consider first an example very similar to Doron's:

(60) dina soret et ha-svederim Se- hi loveSet,
Dina knits the sweaters that she wears
be-?od ima Sela kona
while mother hers buys

This does have a sloppy reading whereby Dina's mother buys the sweaters that she herself wears. However, as Danny Fox points out, this is not surprising regardless of whether the correct syntactic analysis of the example involves ellipsis or null object, since with an *overt* object (*otam* "them"), such a reading is still possible, presumably stemming from an "e-type" reading of the pronoun, in the sense of Evans (1980). As Fox also notes, the same is true in English:

(61) Dina knits the sweaters that she wears while her mother buys them

This, of course, makes it even more evident that the difficulty with the sloppy reading in Doron's (58A2) has nothing to do with the syntax of the sentence. To probe possible syntactic effects more accurately, Idan Landau suggests a test example where an e-type sloppy reading is unavailable:

(62) dina ohevet ko sveder Se- hi loveSet
Dina loves every sweater that she wears
aval ima Sela sonet otam
but mother hers hates them
"Dina loves every sweater that she wears but her mother hates them."

(62), with an object pronoun, has only a strict reading. On the other hand, when there is no object pronoun, a sloppy reading is as acceptable as a strict reading:

(63) dina ohevet ko sveder Se- hi loveSet
Dina loves every sweater that she wears
aval ima Sela sonet
but mother hers hates
"Dina$_j$ loves every sweater that she wears but her mother$_i$ hates every sweater that she$_{i/j}$ wears."

This example, like the Serbo-Croatian examples cited earlier, provides very strong evidence against Potsdam's constraint (55) as a universal.[14] Thus, a reduction of the Roberts–Potsdam constraint (17), repeated here as (64), is highly suspect.

(64) [$_{VP}$ [$_V$ e]X] cannot antecede VP-ellipsis of [$_{VP}$ [V]X], where V is lexical.

7.7. Summary

In this brief study of a gap in the English VP-ellipsis paradigm, I have considered two versions of a strictly lexicalist alternative to the account suggested in Lasnik (1995e). Both proposals, that of Roberts (1998) and that of Potsdam (1997), rely on a constraint that I restated as (65).

(65) [$_{VP}$ [$_V$ e]X] cannot antecede VP-ellipsis of [$_{VP}$ [V]X], where V is lexical.

Roberts argued, in effect, that (65) follows from the principle of recoverability of deletion as follows. A raised V must have checked a feature (which has thereby deleted), and, by hypothesis, checking/deletion is uniform throughout the V chain. As a consequence, a V trace will have only a proper subset of the features of an unraised V, and thus be unable to antecede its deletion. I showed that Pseudogapping is a prima facie counterexample to (65), and that a certain type of matrix Sluicing is a prima facie counterexample to a virtually necessary extension of (65). I then proceeded to show how both the Pseudogapping facts and the Sluicing facts can be rendered compatible with Roberts's justification for (65), if not actually with (65) itself, under the view of pied-piping suggested by Chomsky (1995a) and developed in more detail by Ochi (1997) and Lasnik (1999b). Ironically, though, it turned out that on this line of reasoning the original gap that Roberts had set out to explain was no longer handled.

Potsdam (1997) offered another kind of justification for (65), arguing that it was an instance of a more general (and universal) V trace identity requirement for ellipsis. However, I showed that the requirement did not hold in Serbo-Croatian, or even in Hebrew, one of the two languages mentioned by Potsdam.

Descriptively speaking, what we are left with is a more limited version of (65), to the effect that a verb raised to check inflectional features (Tense and possibly

agreement) leaves behind a trace that cannot antecede deletion of a (non-raised) verb lacking those inflectional features. Lasnik (1995e) provides one way of capturing this generalization, in terms of a "hybrid" theory of verbal morphology, largely motivated by problems with the strictly lexicalist theory of Chomsky (1993). Ultimately, there will surely be a better way, but that remains for further research to reveal.

8

ON A SCOPE RECONSTRUCTION PARADOX

Chapters 3–6 all involved, to some extent, the syntax of A-movement. This chapter considers one aspect of the semantics of A-movement – the scope of an A-moved quantifier. In the early 1980s, Chomsky observed that while a universal quantifier in subject position can take scope under clausal negation, if such a quantifier undergoes subject raising to subject position, it cannot take scope under negation in the clause where it originated. Based on the conclusions of Chapters 3 and 4, I show that Chomsky's observation extends to "subject raising to object position" as well, thus supporting Chomsky's claim that there is no A-movement scope reconstruction. However, the phenomenon of "quantifier lowering" is widely assumed to be exactly A-movement scope reconstruction. To attempt to resolve this paradox, I suggest here that quantifier lowering might not exist, and I display several configurations where it would be expected to obtain but does not.

Chomsky has observed that a universal quantifier that has undergone subject raising loses the ability to be in the scope of lower clause negation, as seen in the following contrast:

(1) (it seems that) everyone isn't there yet
(2) everyone seems [t not to be there yet]

Partly on this basis, he argues that there is no A-movement reconstruction at all. He therefore suggests, following May (1977, 1985), that familiar instances of "Quantifier Lowering" (QL) as in (3) are not instances of reconstruction, but involve a literal lowering operation.

(3) Some politician is likely to address John's constituency

Hornstein (1995) argues that QL should be treated as reconstruction, and he resolves the resulting paradox by suggesting that the phenomenon in (2) has nothing to do with A-movement *per se*. In this note, I will accept Hornstein's

initial position but attempt to resolve the paradox by questioning the existence of QL.

Zubizarreta (1982) attributes to Noam Chomsky the observation that while a universal quantifier in subject position can be understood inside the scope of clausal negation, as in (4), this possibility disappears when the subject undergoes raising to subject position, as in (5).

(4) (it seems that) everyone isn't there yet
(5) everyone seems [t not to be there yet]

Zubizarreta takes this as a potential argument against QL, and, in fact, takes Chomsky to be making such an argument. QL, discussed especially by May (1977; 1985), accounts for the apparent ambiguity of such examples as (6).

(6) Some politician is likely to address John's constituency

May (1977) observes that "[(6)] may be taken as asserting either (i) that there is a politician, that is, Rockefeller, who is likely to address John's constituency, or (ii) that it is likely that there is some politician (or other) who will address John's constituency." Since May's ground-breaking work, that ambiguity has generally been taken, following May, to stem from whether the surface subject "lowers" or not.

On May's account, this lowering is a variant of Quantifier Raising (QR). Like QR, it moves a quantifier and attaches it to S. QL differs from QR only if that it attaches the quantifier to a lower clause. Another view of QL has also been considered in the literature. Under this alternative view, recently discussed by Hornstein (1995), QL is a reconstruction effect. The quantifier is interpreted as if it were in a position it occupied at any earlier point in the derivation. In cases like (6), that position is embedded subject.

Chomsky (1995a), in effect, distinguishes these two theories. He argues, partly on the basis of (4)–(5), that there are no reconstruction effects at all with A-movement, but he allows for the possibility of QL. Chomsky suggests that "lowered" readings

> … could result from adjunction of the matrix quantifier to the lower IP (c-commanding the trace of raising and yielding a well-formed structure if the trace of quantifier lowering is deleted, along the lines of May's original proposal). But reconstruction in the A-chain does not take place …
>
> (Chomsky 1995a: 327)

Under this approach, QL stands in sharp contrast to other reconstruction-type phenomena, since the others are all taken to result from copy-movement and

complementary deletion. As Chomsky says, "The basic assumption here is that there is no process of reconstruction …" (Chomsky 1995a: 326). For QL, on the other hand, there does seem to be a literal operation of reconstruction. The question then arises why it should not be more generally available.

With this in mind, I would like to consider the possibility that there is no actual process of QL and thus that apparent lowered readings of examples like (6) would have to arise from reconstruction in the minimalist sense: interpretation of a trace. On this understanding, as on Zubizarreta's, the fact in (4) is what Zubizarreta (and, much later, Hornstein) took it to be, namely a potential argument against QL. Both Zubizarreta and Hornstein maintain that the QL phenomenon is real. Zubizarreta notes the apparent counterargument, but puts it aside. Hornstein argues that the lack of narrow scope for *everyone* in (4) is independent of raising, hence irrelevant to the issue of A-movement reconstruction. Hornstein (personal communication) suggests that the crucial property is contraction – the wide scope for negation is possible only when negation has contracted. However, for my informants (and myself) examples with uncontracted negation (but without raising) seem to allow wide scope negation rather readily. Consider the following examples, in a situation where a teacher is being reprimanded for giving all the students As:

(7) School policy requires that everyone not get an A
(8) It is important for everyone not to get an A

Both examples seem reasonably appropriate to the situation, indicating that the reading in question, with the universal understood under negation, is available.[1]

If we therefore adopt Chomsky's point of view that the absence of a reading for (4) with the universal under negation does, indeed, reflect absence of "reconstruction" with A-movement, the question now arises whether that point of view can be reconciled with Hornstein's natural analysis of standard QL phenomena, as in (5), as A-movement reconstruction. On the face of it, that seems an impossible task: How can one reconcile a contradiction? If there is no A-movement reconstruction, and if QL is A-movement reconstruction, then we are led to the completely bizarre conclusion that QL does not exist.

Amazingly, there is some reason to believe that that completely bizarre conclusion is correct.[2] Precise semantic characterizations of "lowered" readings in QL configurations are surprisingly difficult to find in the literature. Standard descriptions generally rely on paraphrase by a similar sentence with pleonastic *it* as subject and a finite complement, as in the following, where (10) is the paraphrase for the "lowered" reading of (9):

(9) Someone is likely to solve the problem
(10) It is likely that someone will solve the problem

But for an interesting and rather wide range of raising examples, paraphrase of this sort fails. Consider the following example:

(11) No large Mersenne number was proven to be prime

(11) cannot accurately be paraphrased by (12).

(12) It was proven that no large Mersenne number is prime

Similar paraphrase failure occurs in the following pair.

(13) Noone is certain to solve the problem
(14) It is certain that noone will solve the problem

(13) describes a situation where the problem under discussion is of at least middling difficulty, and the potential problem solvers are not omniscient. (14), on the other hand, is a sentence about either an impossible problem or a hopelessly inept group of solvers.

This paraphrase failure is not limited to negative contexts. Consider (15) in a situation where there are five fair coins, flipped in a fair way.

(15) Every coin is 3% likely to land heads

Note that the situation strongly biases the sentence towards the lowered reading, but that reading still is not possible. (15) cannot be accurately paraphrased as (16).

(16) It is 3% likely that every coin will land heads

The only reading for (15) is one describing the much less plausible situation in which each coin is weighted in such a way that it is 33 times more likely to land heads than tails.

So far, we have seen that for two scope phenomena, raising to subject position fails to display reconstruction. Interestingly, there is a "raising-to-object" construction that displays similar lack of reconstruction. On a very plausible analysis of the *make … out* construction (a construction discussed by Kayne (1985)), *John* in (17) has raised into the higher clause (since it appears to the left of the particle *out*, which is associated with the matrix verb).

(17) They made John out to be a fool

Evidence that *John* is the underlying subject of the lower clause comes from the fact that (18) seems to be a perfect paraphrase of (17).

(18) They made out that John is a fool

Further, the NP can be a pleonastic:

(19) They made there out to be a solution
(20) They made out that there is a solution

Now notice that the scope contrast between Chomsky's (5) and (4) is mirrored in this particle construction:

(21) The mathematician made every even number out not to be the sum of two primes
(22) The mathematician made out that every even number isn't the sum of two primes

In contrast with (22), the only reading available for (21) is the implausible one where the mathematician was engaged in the absurdly futile activity of trying to convince someone that no even number is the sum of two primes (and not the far more plausible one where she is merely trying to convince someone that Goldbach's conjecture is false). Thus, even with strong pragmatic bias towards wide scope for the negation, it still is not available.

Significantly, though the judgments are subtle, the failure of "quantifier lowering" seen in classic raising examples like (11) is also mirrored in the raising particle construction. Compare (23) with (24):

(23) The DA made no defense witnesses out to be credible
(24) The DA made out that no defense witnesses were credible

On pragmatic grounds, the only remotely plausible interpretation of (23) would be one synonymous with (24). But that interpretation is very difficult to obtain. Instead, the sentence has a bizarre interpretation where the DA perhaps had the intention of trying to show that (some of) the defense witnesses were credible, but never acted on that intention.

Thus, there is substantial evidence that Chomsky is correct about absence of (one type of) A-movement reconstruction. Further, there is even some reason to believe that the stronger interpretation of Chomsky's claim, whereby even "quantifier lowering" does not exist, is correct. But for the latter, there is, of course, all the standard contradictory evidence that QL does exist, including the famous ambiguity of (6), repeated here as (25).

(25) Some politician is likely to address John's constituency

What is the difference between examples like (25), where lowering seems to obtain, and examples like (13) and (15), repeated here, where it does not?

(26) Noone is certain to solve the problem
(27) Every coin is 3% likely to land heads

As far as I can tell, all of the standard examples displaying a QL ambiguity have indefinite subjects, unlike my examples of lowering failure. However, on standard accounts, either of the reconstruction variety or of the lowering variety, it is totally unclear why that should matter.

There is one familiar argument that "lowered" readings do, indeed, involve some sort of syntactic operation. While compelling on the face of it, the argument actually seems to me inconclusive. The argument, due to May (1985), is that a "lowered" reading for the quantifier is incompatible with the binding of a pronoun in the upper clause. May gives the following example:

(28) No agent$_i$ was believed by his$_i$ superior to be a spy for the other side

(28) cannot be paraphrased as (29).

(29) It was believed by his$_i$ superior that no agent$_i$ was a spy for the other side

There are, however, two interfering factors here. First, I do not believe that "lowering" is possible for (28) even without pronoun binding, as in (30).

(30) No agent was believed by Dulles to be a spy for the other side

In (30), it is not easy to see this, perhaps because the two readings are close. Changing the matrix predicate to *known* or *proved* sharpens the difference and makes it clear that the lowered reading does not exist:

(31) No agent was known by Dulles to be a spy for the other side
(32) No agent was proved by Dulles to be a spy for the other side

The second interfering factor stems from the very characterization of the lowering phenomenon. Taking paraphrasability by an expletive construction as the defining property, it becomes trivial that lowering is not available in (28), since the paraphrase is ungrammatical in its own right, constituting a Weak Crossover (WCO) violation. And, as now expected, this second interfering factor persists even when the first one is controlled for (i.e. by substituting an indefinite subject):

(33) Some agent$_i$ was believed by his$_i$ superior to be a spy for the other side

Once again, paraphrase fails (on the indefinite reading), but once again, the expletive example is ungrammatical, a WCO violation, whether as a paraphrase or not:

(34) It was believed by his$_i$ superior that some agent$_i$ was a spy for the other side

Thus, it is difficult to draw any firm conclusion from the pronoun binding argument.

What should we make of all of this? As far as I have been able to tell, the standard lowering examples differ from the new "anti-lowering" ones I have presented in that the truth conditional difference between the hypothesized lowered

reading and the non-lowered reading is much less clear in the standard examples. The clearer the truth conditional difference, the less accessible the lowered reading seems to be. This suggests that there is no lowering at all, the apparent lowered reading of raised indefinites having some other source. Conceivably, the basic difference between the two readings of sentences with raised indefinites lies in the speaker's point of view with respect to the raised subject. On one reading, the speaker has a particular individual in mind, but, for some discourse reason or other, does not identify that individual. On the other reading (the "lowered" one), the speaker does not have any particular individual in mind. The ambiguity might then fall under theme–rheme properties, the "wide scope" quantifier being a theme or topic.[3] Or perhaps some as yet unidentified semantic property of indefinites is relevant.[4] If either of these possibilities turns out to be correct, the strong interpretation of Chomsky's claim that there is no A-movement reconstruction might yet be viable.[5]

9

SOME RECONSTRUCTION RIDDLES

Two reconstruction "riddles" are examined here. The third riddle considered is the apparent conflict, introduced in Chapter 8, between standard quantifier lowering and lack of "reconstructed" interpretation of a universal quantifier that has raised out of a clause containing clausal negation. Here I present some further data, and further consider the question (though still inconclusively) of the source of the apparent lowered interpretations in quantifier lowering situations. The first two riddles concern binding theoretic reconstruction effects (or their absence). Freidin (1986) and Lebeaux (1988) pointed out an apparent argument vs adjunct Condition C reconstruction contrast, as in (i) vs (ii).

(i) *Which report that $John_i$ was incompetent did he_i submit?
(ii) Which report that $John_i$ revised did he_i submit?

The noun complement in the fronted NP in (i) seems to produce "forced" reconstruction, while the relative clause in the fronted NP in (ii) does not. I examine the similar accounts of this contrast presented by Lebeaux and by Chomsky (1993), but then I question the empirical basis for the analyses, showing that examples structurally parallel to (i) in all relevant ways do not show Condition C effects. I then go on to explore the (mostly positive) consequences of this for the theory. The second riddle, which sets the stage for the third, is Chomsky's (1995a) observation that A-movement reconstruction cannot rehabilitate certain Condition B violations. The proposed answer to both the second and third riddles is that, as Chomsky suggested, there is no A-movement reconstruction.

The investigation of reconstruction phenomena has played a significant role in theory construction in several modules of syntax, including those determining phrase structure, the nature of transformational derivations, and, of course, anaphoric connection. In this chapter, I will discuss two of the most interesting proposals about reconstruction that have appeared in recent years. The first is

the widely accepted Lebeaux (1988) account of certain complement/adjunct asymmetries, first noted by Freidin (1986), in terms of generalized transformations. The second is the almost universally rejected claim of Chomsky (1995a) that there is no reconstruction with A-movement. In the first instance, I will suggest that though the argument is plausible, it is actually incorrect. And in the second, I will suggest the reverse: that while the argument is implausible, its conclusion might actually be correct.

9.1. Condition C complement/adjunct reconstruction asymmetries (the "Freidin–Lebeaux effect")

Until the mid-1980s, it was rather standardly assumed that overt movement can salvage what would otherwise have been a Condition C violation. (1), derived from something like (2), is a representative type of example.

(1) Which report that John_i revised did he_i submit?
(2) he_i submitted which report that John_i revised

Note that (2) would be in violation of Condition C, just like (3).

(3) *He_i submitted the report that John_i revised

Freidin (1986) observes an interesting apparent exception. In (4), movement does not seem to rehabilitate the violation.

(4) *Which report that John_i was incompetent did he_i submit?

Freidin notes that the difference stems from the status of the clause internal to the fronted NP. In the good (1), that clause is a relative clause, while in the bad (4), it is a complement. Freidin suggests that the process of reconstruction can somehow make this distinction.

Lebeaux (1988) presents similar contrasting examples:

(5)(a) *He_i believes the claim that John_i is nice.
(b) *He_i likes the story that John_i wrote.
(c) *Whose claim that John_i is nice did he_i believe?
(d) Which story that John_i wrote did he_i like?

Lebeaux provides a detailed account, in terms of a new (in part, very old) theory of phrase structure. Following Chomsky (1981), Lebeaux proposes that D-structure is a representation of thematic relations. Given this, complements, being necessarily θ-marked, have to be represented at this level, while relative clauses, being adjuncts of some sort, need not be, Lebeaux reasons. Relative clauses can be inserted anywhere in the course of the derivation, via a generalized transformation. Lebeaux goes on to propose that Condition C must be satisfied

not (just) at S-structure or LF, but everywhere in the course of the derivation. (5c) is then ruled out, since at D-structure (and, in fact, at every point of the derivation prior *wh*-movement), Condition C is violated, as *he*$_i$ c-commands *John*$_i$. (5d), on the other hand, has a possible derivation where the relative clause is not inserted until after *wh*-movement. On such a derivation, there is no point at which *he*$_i$ c-commands *John*$_i$, so there is no Condition C violation.

Chomsky (1993) gives further similar examples, presented here in (6), questions certain properties of Lebeaux's account, and then proposes an extension of it.

(6)(a) *Which claim [that John$_i$ was asleep] was he$_i$ willing to discuss
(b) Which claim [that John$_i$ made] was he$_i$ willing to discuss

Chomsky states that

> The approach is appealing, if problematic. For one thing, there is the question of the propriety of resorting to generalized transformations. For another, the same reasoning forces reconstruction in the case of A-movement. Thus, [(7)] is analogous to [(6a)]; the complement is present before raising and should therefore force a Condition C violation.
>
> (Chomsky 1995b: 204)

(7) the claim that John was asleep seems to him [$_{IP}$ *t* to be correct]

The second of these objections, I will return to. As for the first, Chomsky's concern is not with generalized transformations *per se*. Rather, the worry is that they are used in too narrow a set of circumstances. Chomsky, instead, proposes "a full-blown theory of generalized transformations." All aspects of structure are so derived. Then, of course, the asymmetry Lebeaux was concerned with is no longer explained. Chomsky thus replaces Lebeaux's account with the following:

(8)(a) The "Extension Condition": structure must be built strictly cyclically.
(b) Adjuncts are exempt from the Extension Condition; relative clauses are adjuncts.
(c) "Reconstruction" is essentially a reflex of the formation of operator–variable constructions.
(d) An operator chain (a sequence of copies) undergoes complementary deletion.
(e) Condition C is an LF requirement.

The contrast between (6a) and (6b) is handled in the following way. First, the grammatical (6b), with the relative clause, has one of the following two derivations, consistent with (8):

(9)(a) [Which claim [that John made]] was he willing to discuss ~~which~~ claim PF
(b) Which claim [that John made]] was he willing to discuss ~~which~~ claim LF
(c) For which x that John made, he was willing to discuss x claim Interpretation (?)

OR?

(10)(a) [Which claim [that John made]] was he willing to discuss ~~which claim~~ PF
(b) [Which claim [that John made]] was he willing to discuss ~~which claim~~ LF
(c) For which x, x a claim that John made, he was willing to discuss x Interpretation (?)

Notice that in both of these, no Condition C violation arises. Now consider (6a), with a noun complement. The adjunct exemption is not relevant, so the derivation must be cyclic, as in the following:

(11)(a) Which claim [that John was asleep] was he willing to discuss [~~which claim that John was asleep~~] PF
(b) [Which ~~claim [that John was asleep]~~] was he willing to discuss [~~which~~ claim that John was asleep] LF
(c) For which x, he was willing to discuss x claim that John was asleep Interpretation (?)

In LF (12), *he* c-commands *John* yielding a Condition C violation, evidently the correct result. Note, though, that two other potential derivations, seemingly made available by the theory, must be excluded:

(12)(a) Which claim [that John was asleep] was he willing to discuss [~~which claim that John was asleep~~] PF
(b) [Which claim [that John was asleep]] was he willing to discuss [~~which~~ claim ~~that John was asleep~~] LF
(c) For which x that John was asleep, he was willing to discuss x claim Interpretation (?)

(13)(a) Which claim [that John was asleep] was he willing to discuss [~~which claim that John was asleep~~] PF
(b) [Which claim [that John was asleep]] was he willing to discuss [~~which claim that John was asleep~~] LF
(c) For which x, x a claim that John was asleep, he was willing to discuss x Interpretation (?)

I continue to hedge on the precise status of the (c) examples as interpretations, but surely if (9c) and (10c) are appropriate interpretations for their corresponding LFs, then (12c) and (13c) are for theirs. Given that, LFs (12b) and (13b) cannot be excluded on semantic grounds. Yet, evidently incorrectly, those representations are not in violation of Condition C. To rule out derivations like (12) and (13), Chomsky proposes (14).

(14) "… preference principle for reconstruction: Do it when you can (i.e., try to minimize the restriction in the operator position)." (Chomsky 1995a: 208).

(14) leaves us with (11), whose LF violates Condition C, the desired result.

Thus far, we have considered the adjunct–argument Condition C reconstruction asymmetry, and two plausible analyses: that of Lebeaux in terms of a partial theory of generalized transformations, and that of Chomsky in terms of the "preference principle." Before examining those analyses further, I would like to consider the facts in more detail. While the contrast seems clear enough in the standard examples cited above, there are examples structurally parallel to the noun complement cases, which, surprisingly, are reasonably, even fully, acceptable. A few samples are as follows:

(15) Which piece of evidence that John was guilty did he successfully refute?
(16) How many arguments that John's theory was correct did he publish?
(17) Which proof that Mary's theory is superior to John's did she present?

In all of these, contrary to now standard assumptions, coreference between the pronoun and the name is acceptable. In questioning those standard assumptions, Kuno (1997) gives comparable examples:

(18)(a) Whose allegation that John_i was less than truthful did he_i refute vehemently?
(b) Whose claim that the Senator_i had violated the campaign finance regulations did he_i dismiss as politically motivated?

All of the examples considered so far, both the standard bad ones and the surprising good ones, have involved interrogation. By the logic of the standard analyses, topicalization would also be expected to show the "Freidin–Lebeaux effect." And this has, in fact, been claimed in the literature. Chomsky and Lasnik (1993) give the following pair:

(19)(a) *The claim that John_i was asleep, he_i won't discuss
(b) The claim that John_i made, he_i won't discuss

And Chomsky (1993) gives these:

(20)(a) *The claim that John_i is asleep, he_i was willing to discuss
(b) The claim that John_i made, he was willing to discuss

However, as with the interrogative examples, the generality of the forced reconstruction effect is open to question. The following examples, the first two of them due to Postal (1997), who, like Kuno, questions the standard assumptions, all seem much better than would be expected under the Lebeaux or Chomsky accounts:

(21) The claim that the director_i was corrupt, he_i was unwilling to discuss

(22) That the director$_i$ was corrupt, everyone knew that he$_i$ would always be able to deny with a straight face
(23) The widespread belief that John$_i$ is incompetent, he$_i$ deeply resents.
(24) This argument that John's$_i$ theory is correct, he$_i$ is now ready to publish.
(25) Mary's attempt to hire John's$_i$ student, he$_i$ heartily endorsed.
(26) John's request to attend Mary's$_i$ lecture, she$_i$ immediately granted.

Postal and Kuno thus seem justified in questioning the standard generalization. But what of the standard examples? Why are the complement instances degraded? Consider first the topicalization instances, repeated here:

(27) *The claim that John$_i$ is asleep, he$_i$ was willing to discuss
(28) *The claim that John$_i$ was asleep, he$_i$ won't discuss

(27) is indeed substantially degraded, in fact virtually bizarre. But, there seems to be a serious tense mismatch between the main and complement clauses, one that persists even in the absence of intended coreference. And with coreference, the situation denoted is pragmatically strange. The Chomsky and Lasnik example (28) controls for both of these properties, and I now believe that we were incorrect in calling it bad. And the several native informants I have consulted concur that it is far better than advertised (except where the topicalized NP is taken as referring to a claim that John made, an effect that can presumably be characterized by the operation of Condition C completely internal to that NP).

Consider now the standard interrogative examples, first Lebeaux's (5c), repeated as (29).

(29) *Whose claim that John$_i$ is nice did he$_i$ believe?

I suspect that the undeniable peculiarity of (29) stems, at least in large part, from factors independent of forced reconstruction. First, it is at least somewhat unusual for someone (John in this case) to rely on others' claims in order to determine his or her own personality characteristics (niceness in this instance). Further, it is not easy to imagine a situation where a set of claims that John is nice can be sufficiently individuated that some can be believed and others not. To illustrate this point, I present the following one scene play, with three characters:

(30) Susan: John is nice.
Mary: John is nice.
!!John: I believe Susan but I don't believe Mary.

John's line of dialogue is very strange in this context. But if (30) is not the kind of situation that would make (29) felicitous, what would be? I suspect that Chomsky's interrogative example (6a), repeated here as (31), has this property as well, though to a lesser extent (and, in fact, the example seems less bad than (29)).

(31) *Which claim [that John$_i$ was asleep] was he$_i$ willing to discuss

A very similar example from Munn (1994) is somewhat worse than (31):

(32) *Which claim that John$_i$ was asleep did he$_i$ later deny

As with (31), the individuation that is presupposed is somewhat unlikely. Additionally, the *later* raises a question: Later than what? And with the example given in isolation (as Munn gave it), the only plausible response to that question is, later than he made the claim. But then there is arguably a Condition C effect completely internal to the interrogative NP, with an "understood" *John* as the subject of *claim*.

There might also be an interfering pragmatic factor in Freidin's example (4), repeated as (33).

(33) *Which report that John$_i$ was incompetent did he$_i$ submit?

It is not customary for an individual (say, John) to be in a position where he would submit reports (even more peculiarly, one selected out of several) on his own incompetence.

If this discussion is on the right track, we would expect that when these interfering pragmatic factors are controlled, the resulting sentences are improved. In fact, we have already seen that this is so. All of the following examples, repeated from above, are syntactically indistinguishable from the standard bad cases, but are far better:

(34) Which piece of evidence that John$_i$ was guilty did he$_i$ successfully refute?
(35) How many arguments that John's$_i$ theory was correct did he$_i$, publish?
(36) Which proof that Mary's$_i$ theory is superior to John's did she$_i$ present?
(37) Whose allegation that John$_i$ was less than truthful did he$_i$ refute vehemently?
(38) Whose claim that the Senator$_i$ had violated the campaign finance regulations did he$_i$ dismiss as politically motivated?
(39) The claim that the director$_i$ was corrupt, he$_i$ was unwilling to discuss.
(40) That the director$_i$ was corrupt, everyone knew that he$_i$ would always be able to deny with a straight face.
(41) The widespread belief that John$_i$ is incompetent, he$_i$ deeply resents.
(42) This argument that John's$_i$ theory is correct, he$_i$ is now ready to publish.
(43) Mary's attempt to hire John's$_i$ student, he$_i$ heartily endorsed.
(44) John's request to attend Mary's$_i$ lecture, she$_i$ immediately granted.

All of this suggests that the complement/relative clause reconstruction asymmetry is illusory. Suppose this is so. How problematic is that for syntactic theory? Postal (1997) implies that if the complement cases do not show Condition C reconstruction, that would constitute an argument against trace theory (i.e. of the "copy" version Chomsky assumes). But it is not really that. Rather, it is, at most, an argument against a potential argument for trace theory; no conclusion can be drawn from the denial of the factual basis for the potential argument. Consider now Chomsky's specific account of the asymmetry. What problem would arise there? None, as far as I can tell. What is doing the crucial work is the "preference principle" for reconstruction, repeated here:

(45) "…preference principle for reconstruction: Do it when you can (i.e., try to minimize the restriction in the operator position)."

But there is nothing a priori desirable about this principle as opposed to, say, one that would minimize the restriction in the variable, or one that would freely allow minimization in either position. Further, if there is no complement/relative clause reconstruction asymmetry, the stipulated exemption to the extension condition for relative clauses is no longer needed in Chomsky's approach (as in (8a,b) above). And on Lebeaux's account briefly summarized above, the reconstruction asymmetry could be eliminated by denying that Condition C must be satisfied everywhere in the course of the derivation, instead limiting its application to a specific level of representation (necessarily LF if central claims of Chomsky (1995b) are correct).

9.2. On lack of reconstruction with A-movement

I turn now to certain questions of A-movement reconstruction (or lack thereof). Recall one of Chomsky's concerns about Lebeaux's account of the (alleged) complement/relative clause reconstruction asymmetry: that it incorrectly predicts forced reconstruction for A-movement as well as for A′-movement, with noun–complement constructions, as in (46).

(46) The claim that John_i was asleep seems to him_i [$_{IP}$ *t* to be correct]

Before considering Chomsky's account, I should note that Lebeaux (1988, 1991) actually does have something to say about examples like (46). Lebeaux proposes that lexical material is inserted only in the head position of an A-chain. On this proposal, an empty category is inserted into the complement subject θ-position. The lexical subject *The claim that John was asleep* is not inserted until after the empty category is raised to its Case position in the higher clause. *John* is then never in the domain of *him*, so there is no Condition C effect. Chomsky's account given above in (8c) and repeated here is possibly, but not obviously, superior.

(47) "Reconstruction" is essentially a reflex of the formation of operator–variable constructions.

Why this should hold is unclear. Chomsky (1993) offers the following discussion, which does not entirely clarify matters:

> The reconstruction process … applies only to operator–variable constructions. What about A-chains, which we may assume to be of the form CH = (α, *t*) at LF (α the phrase raised from its original position *t*, intermediate traces deleted or ignored)? Here *t* is a full copy of its antecedent, deleted in the PF component.
>
> (Chomsky 1995b: 210)

Chomsky assumes that A-movement leaves a full copy as a trace, and further, that the copy persists throughout the computation. Absence of "reconstruction" is thus little better than a stipulation.

So far, Chomsky's claim of lack of reconstruction with A-movement is based on the apparent "Freidin–Lebeaux effect" with A′-movement but not with A-movement. But I have suggested that even with A′-bar movement there is no forced Condition C reconstruction. If that is correct, there is nothing special to say about A-movement in this regard. At this point, I turn to two other interesting arguments that Chomsky (1995a) offers for lack of reconstruction with A-movement. The first, like the one discussed already, is based on Binding Theory, but this time Condition B. And this time, the issue is not reconstruction refraining from *creating* a violation, but rather reconstruction not being able to *rehabilitate* a violation. Chomsky gives the following example, observing that it has the status of a Condition B violation:

(48) *John$_i$ expected [him$_i$ to seem to me [$_\alpha$ *t* to be intelligent]]

Chomsky reasons that "… under reconstruction the violation [of Condition B] should be eliminated, with *him* interpreted in the position of *t*…" (Chomsky 1995a: 326) There is a question here about the correct characterization of the domain relevant to pronominal obviation (the "Governing Category" (GC) of the Government-Binding (GB) framework). But under plausible assumption, *John* would, indeed, be outside of the relevant binding domain of *him* were the latter "reconstructed." In fact, given the predicate internal subject hypothesis, which Chomsky assumes, the reconstructed position would presumably be even lower than Chomsky indicates in (48).

The next argument involves scope interaction between clausal negation and subject universal quantifier. Chomsky (1995a) gives the following paradigm:

(49)(a) (it seems that) everyone isn't there yet
(b) Everyone seems [*t* not to be there yet]

In (49a), *everyone* is easily understood as within the scope of negation, while in (49b) it cannot be so understood. While the relevant interpretive mechanism at work in (49a) is mysterious, Chomsky reasonably argues on the basis of the contrast that in (49b), "… there is no reconstruction to the trace position …" (Chomsky 1995a: 327) As Chomsky indicates, this conclusion raises questions about standard "quantifier lowering effects" as in (50) from May (1977).

(50) Some politician is likely to address John's constituency

I turn now to a brief discussion of those effects.

9.3. Quantifier Lowering

To reconcile his conclusion based on (49) with standard Quantifier Lowering (QL) as in (50), Chomsky suggests that the "lowered" reading in the latter instance

> … could result from adjunction of the matrix quantifier to the lower IP (c-commanding the trace of raising and yielding a well-formed structure if the trace of quantifier lowering is deleted, along the lines of May's original proposal). But reconstruction in the A-chain does not take place, so it appears.
>
> (Chomsky 1995a: 327)

QL has been widely discussed in the literature, but precise characterization of the ambiguity remains surprisingly elusive. Paraphrase is pretty much all one finds, and the following, from May (1977), is as good as any:

> [(50)] may be taken as asserting either (i) that there is a politician, e.g., Rockefeller, who is likely to address John's constituency, or (ii) that it is likely that there is some politician (or other) who will address John's constituency.
>
> (May 1977)

Interestingly, just as Chomsky presented evidence based on a (surprising) missing scope possibility that "reconstruction" does not take place with A-movement, there are also surprising missing scope possibilities with QL. Some of the examples, like Chomsky's, involve negation:

(51) No large Mersenne number was proven to be prime
(52) Noone is certain to solve the problem

Neither of these has a "lowered" reading, at least on the paraphrase characterization above. (51) is not accurately paraphrasable as (53), nor is (52) as (54).

(53) It was proven that no large Mersenne number is prime
(54) It is certain that noone will solve the problem

This paraphrase failure is not limited to negative contexts. Consider (55) in a situation where there are five fair coins, flipped in a fair way.

(55) Every coin is 3% likely to land heads

This situation strongly biases the sentence towards the lowered reading, but that reading still is not possible. (55) cannot be accurately paraphrased as (56).

(56) It is 3% likely that every coin will land heads

These facts suggest that there is no QL. That is, as Chomsky claims, there is no reconstruction to the position of trace of raising. But additionally, there is no lowering of a quantifier (by a QR type rule) to adjoin to a lower IP. This might be because there are no rules of the QR type at all, or because the Move operation is strictly one of raising. Given one of those prohibitions, I can only speculate about what is going on with the standard examples like (50), repeated as (57).

(57) Some politician is likely to address John's constituency

As far as I know, examples showing an apparent lowering effect always have an indefinite as subject, unlike the new examples I have presented. I assume that this property is crucial, perhaps in the following fashion: On the first reading of (57) discussed by May, the speaker has a particular individual in mind (a politician, in this instance), but, for some discourse reason or other, does not identify that individual. On the second reading (the "lowered" one), the speaker does not have any particular individual in mind. The apparent ambiguity might then fall under theme–rheme properties, the "wide scope" quantifier being a theme or topic. Notice in this regard that even in a completely transparent context, we can find something strangely reminiscent of the two readings May presents for raising sentences:

(58) Some politician addressed John's constituency
(59)(a) … namely Rockefeller
(b) … I can tell by all the balloons and flags on the green

May (1985) presents an important argument that actual syntactic lowering must be involved in the second reading of examples like (57): namely that such

a "lowered" reading for the quantifier is incompatible with the binding of a pronoun in the upper clause. May gives the following example:

(60) No agent$_i$ was believed by his$_i$ superior to be a spy for the other side

It does seem correct that there is no lowered reading for the subject in (60). However, the same is apparently true even with no bound pronoun, as in examples (51)–(52) above. A more relevant test would involve an indefinite subject, as in (61).

(61) Some professor$_i$ is believed by his$_i$ students to be a tyrant

It is certainly correct that (61) cannot be paraphrased as (62).

(62) *It is believed by his$_i$ students that some professor is a tyrant

But it is not clear what we can conclude from the fact that a sentence cannot be paraphrased by an ungrammatical sentence (in this case, one that violates the Weak Crossover constraint). That leaves the discourse test summarized above: Does the speaker have an individual in mind or not? And it does seem that (61) can be felicitously uttered under either circumstance, as illustrated in (63).

(63)(a) Howard Lasnik is believed by his students to be a tyrant
(b) Some professor (or other), I have no idea exactly who, is believed by his students to be a tyrant

The context for (63b) might be the discovery of graffiti scrawled on the lavatory wall saying "Our professor is a tyrant," or it might even be mere general background knowledge about the typical sociology of a university.

There is another class of apparent A-movement reconstruction effects, that, as far as I know, Chomsky has not addressed at all. This class includes Condition A reconstruction of the sort discussed in detail by Belletti and Rizzi (1988), among many others. Standard examples involve classic raising predicates and also "psych" predicates (which are argued, in part on this basis, to be raising predicates of a particular sort). Examples of the familiar sort are as follows:

(64) Each other's supporters frightened the candidates
(65) Each other's supporters seem to the candidates to be unscrupulous

The contrast between such examples, on the one hand, and ones with simple transitive or control predicates on the other, constitutes one widely accepted argument for A-movement reconstruction. Examples contrasting with (64) and (65) are given in (66) and (67).

(66) *Each other's supporters attacked the candidates
(67) *Each other's supporters asked the candidates to be more honest

I have marked (64)–(67) with the standard judgments, but I must confess that I am no longer confident that (64)–(65) are as good as they are always claimed to be or that (66)–(67) are as bad. If they do not really contrast, then, needless to say, there is no clear argument for reconstruction. If, on the other hand, the contrasts in (64)–(67) are genuine, the satisfaction of Condition A in (64)–(65) might be handled "on-line," as in Belletti and Rizzi (1988) or in Lebeaux's theory. Formally, this makes sense if satisfaction of Condition A involves a formal feature, a not implausible assumption.

Arguably, determination of scope is not satisfaction of a formal feature, but rather, is a matter of interpretation at the interface. The absence of scope reconstruction in Chomsky's example (49b), and perhaps in my (51), (52), and (55), then follows, as noted above, if there is no actual lowering of the relevant sort, and (Chomsky's assumption) traces of A-movement are not visible at LF. But, recall that that assumption follows only from the stipulation that reconstruction is a property solely of operator–variable constructions. A more principled possibility might be that A-movement, unlike Ā-movement, does not leave a trace, where a trace is, following Chomsky, a copy of the item that moves, and LF reconstruction effects result from failure to delete (a portion of) a lower copy. This distinction is conceptually plausible: Ā-movement typically creates an operator– variable relation, so at least an "initial" trace is necessary. For A-movement, on the other hand, the trace is seemingly a theoretical excrescence. There are not two separate interpretive roles for a moved NP and its trace to fulfill.

Chomsky (1995a), in effect, argues against this proposal, by arguing that the trace of A-movement must be present at the LF level:

> In the phonological component, traces delete. We have found no reason to extend that convention to the $N \rightarrow \lambda$ computation, and indeed cannot; were we to do so, θ-positions would be invisible at LF …
>
> (Chomsky 1995a: 301)

I suggest that the undesirable consequence that concerns Chomsky can be avoided. Suppose that instead of being determined specifically at the LF level, θ-roles are "checked" in the course of a derivation. The moved argument is then itself a record of the crucial part of the history of its derivation. This view of θ-roles as features is argued for on independent grounds by Bošković and Takahashi (1998) and Lasnik (1995d). To the extent that such an analysis can be maintained, we can circumvent Chomsky's conclusion above that the trace of A-movement must persist to the interface level (hence, a fortiori, must exist in the first place). There is then no clear empirical objection to the suggestion that

A-movement does not leave a trace, and some reason to think that the suggestion is correct. But we are left with perhaps the biggest reconstruction riddle of all: Given a "bare phrase structure" of the approach articulated and motivated by Chomsky (1994) and Chomsky (1995a) how is it even possible for movement not to leave a trace?

10

CHAINS OF ARGUMENTS

The missing A-movement reconstruction effects of Chapters 8 and 9 are examined here in further detail. Many more phenomena are considered, including several involving "raising to object." After critically considering Chomsky's (1995a) explanation, I present a simple alternative: There is no A-movement reconstruction because A-movement does not leave a "trace" (i.e. a copy). Finally, I explore consequences of the proposal for certain island constraints, and for θ-theory.

In this chapter, I will be concerned with some of the properties of configurations resulting from A-movement and of the chains – A-chains – created by such movement. I will concentrate particularly on several arguments about A-chains presented by Chomsky (1995a), and the consequences, sometimes surprising, of those arguments.

10.1. Trace deletion and reconstruction in A-chains

As my point of departure, I take Chomsky's (1995a: 326) claim about a contrast between Ā-movement and A-movement: "That reconstruction should be barred in A-chains is … plausible on conceptual grounds." I begin by examining those conceptual grounds. Chomsky's immediate concern at this point of his exposition is trace deletion, particularly intermediate trace deletion. He suggests that certain analyses proposed in Chomsky (1991) and Chomsky and Lasnik (1993) based on intermediate trace deletion are incorrect and that there is, in fact, no process of trace deletion. Rather, the effects of trace deletion follow from reconstruction "understood in minimalist terms."[1]

First, a very brief reminder of those earlier analyses and the role of trace deletion in them. Chomsky, and Chomsky and Lasnik, were concerned with the fact that a range of "island" violations do not have the severely degraded status of Empty Category Principle (ECP) violations, even though the derivations of the examples seem to produce intermediate traces that are not properly governed.

One such example is (1), first discussed in roughly these terms by Lasnik and Saito (1984).

(1) ??who do you wonder [$_{CP}$ whether [$_{IP}$ John said [$_{CP}$ *t′ e* [$_{IP}$ *t* solved the problem]]]]

The intermediate trace *t′* is marked * since it is too distant from its nearest antecedent (by any standard measure) to be properly governed, but the example is merely marginal. The accounts of this fact at issue incorporated a process of deletion, constrained (like all transformational operations) by economy. Deletion is possible only to turn an illegitimate LF object into a legitimate one, where the legitimate LF objects are uniform chains (whose members are all in A-positions, all in Ā-positions, or all in X^0-positions) and operator–variable pairs.

Deletion in the chain (*who*, *t′*, *t*) is permissible since the chain is neither uniform (*who* and *t′* are in Ā-positions; *t* is in an A-position) nor an operator–variable *pair*. More generally, in the case of successive-cyclic Ā-movement of an argument, an intermediate trace (starred or otherwise) can (in fact must) be deleted in LF, voiding an ECP violation when the trace to be deleted is starred. On the other hand, long movement as in (2) will violate the ECP, since the movement chain in this instance is uniformly an Ā-chain and economy therefore prevents the deletion of *t′*.

(2) *how do you wonder [$_{CP}$ whether [$_{IP}$ John said [$_{CP}$ *t′ e* [$_{IP}$ Mary solved the problem *t*]]]]

Similarly, ultralong A-movement will also be properly excluded, even when the first step is "short," as in (3).

(3) *John seems [that [it is likely [*t′* to be arrested *t*]]]

With this much background, I return to the main point. Chomsky (1995a) is concerned with long A-movement, but via an intermediate Ā-position.

(4) *John seems [that [t_2 [it was told t_1 [that …]]]]

Notice that the chain of *John* in (4) is nonuniform, so the deletion process outlined above should be applicable – incorrectly, it appears. Chomsky (1995a: 326) concludes,

> We do not want to permit the intermediate (offending) trace t_2 to delete, unlike what happens in [long *wh*-movement of an argument]. The distinction suggests a different approach to intermediate trace deletion: perhaps it is a reflex of the process of reconstruction, understood in minimalist terms. … The basic assumption here is that there is no process of reconstruction; rather, the phenomenon is a consequence of the formation

> of operator–variable constructions driven by F[ull] I[nterpretation], a process that may (or sometimes must) leave part of the trace – a copy of the moved element – intact at LF, deleting only its operator part.
>
> (Chomsky 1995a: 326)

In fact, it does seem that the only successful uses of economy-constrained deletion in chains involve long *wh*-movement of arguments, where a nonuniform chain is turned into an operator–variable pair. I know of no comparable instances where an ECP violation is voided by deletion of an offending intermediate trace turning a nonuniform chain into a uniform chain. Since this new approach predicts this fact, there is some justification for the approach. On the other hand, it raises questions of its own. First, it is not clear that t_2 in (4) is an offending trace in the relevant sense (i.e. in the sense of the earlier theory). Is movement from that intermediate position to the surface position of *John* too long? Maybe, or maybe not. And even if it is, that could presumably be remedied by further adjunction steps. There is, however, a way to retain the essence of the new analysis. Suppose we were to continue to assume that there is no trace deletion. But suppose we retained from the earlier approach the idea that only operator–variable pairs and uniform chains are legitimate LF objects. Then (4) would be correctly excluded, but not because of an offending trace *per se*. Rather, the whole *chain* would be the offender.

The account of (3) remains unchanged. Though the chain is legitimate, it does contain an offending trace, one that now cannot be eliminated under any circumstances, since (3) does not involve an operator chain. Ironically, though, the major phenomenon originally motivating the uniform chain approach now loses its account. Recall that the offending intermediate trace in the case of argument movement (1) was deletable by virtue of being part of a nonuniform chain, whereas the corresponding offending trace in the case of adjunct movement (2), as part of a uniform chain, was not deletable. But in the new approach, deletability has nothing to do with uniformity. Rather, the intermediate trace in (1) deletes as a direct consequence of operator–variable formation. Similarly, the intermediate trace in (2) should be able to delete. As a consequence, the difference in status of the two examples is no longer explained. I put this question aside and return to aspects of the analysis more directly related to properties of A-chains.

10.2. Chomsky's arguments against A-reconstruction

10.2.1. A "conceptual" argument

As Chomsky observes, his (1995a) approach to intermediate trace deletion makes it a subcase of "reconstruction." Further, "[t]he reconstruction process would then be restricted to the special case of Ā-movement that involves operators. That reconstruction should be barred in A-chains is thus plausible on conceptual

grounds" (Chomsky 1995a: 326). In passing, I note that the argument actually seems more empirical than conceptual, based, as it is, on an acceptability contrast between long A-movement and long (argument) *wh*-movement. Momentarily, I will discuss some of Chomsky's several arguments that there is, in fact, no reconstruction with A-movement. Before that, though, I want to briefly examine the somewhat curious account that Chomsky has offered for the claimed fact. Chomsky assumes that movement invariably leaves behind a "trace" in the form of a copy of the moved item. Further, as just discussed, he argues that there is no process of trace deletion *per se*. Rather traces are (sometimes) eliminated as part of the process of operator–variable creation. Thus, traces in other types of constructions are never eliminated. That they are not eliminated in A-constructions provided part of Chomsky's account of the extreme ungrammaticality of "improper" movement, as in (4). Reconstruction phenomena are simply situations where a (portion of a) moved item behaves, for some purposes, as if it were in some position it occupied earlier in the derivation. Yet, almost paradoxically, Chomsky concludes that the impossibility of *eliminating* an A-trace makes it plausible that reconstruction should be *barred* in A-chains. If anything, one might expect, on the contrary, that on this theory reconstruction is always *necessary* in A-chains. Thus, although the intuition that reconstruction should be specifically a property of operator–variable constructions is fairly clear, Chomsky's mechanism for instantiating the intuition is far from straightforward.

10.2.2. An argument from binding

Alongside the above argument, which Chomsky regards as conceptual, he gives some empirical arguments. The first is a binding-theoretic argument based on (5).

(5) *John expected [him to seem to me [$_{\alpha}$ *t* to be intelligent]]

Chomsky observes that "[u]nder the relevant interpretation, [(5)] can only be understood as a Condition B violation, though under reconstruction the violation should be obviated, with *him* interpreted in the position of *t* …" (Chomsky 1995a: 326). Notice that this argument tacitly assumes a formulation of *governing category* (GC), the domain in which a pronoun must be free, such that the "reconstructed" (6) would not also violate Condition B.

(6) John expected [to seem to me [$_{\alpha}$ him to be intelligent]]

The formulation in Chomsky (1981: 209–11), is not obviously consistent with the argument.

(7) β is a governing category for α if and only if β is the minimal category containing α, a governor of α, and a SUBJECT accessible to α.
(8) SUBJECT = Agr in a finite clause; NP of S in an infinitival; NP of NP in an NP.
(9) γ is accessible to α iff α is in the c-command domain of γ and …

α in (6) is not the GC for *him*, since there is no SUBJECT accessible to *him* in that domain. Further, by hypothesis, there is no SUBJECT at all in the intermediate clause, after reconstruction. Thus, the GC for *him* would actually be the matrix, and the required Condition B effect is obtained after all. On the other hand, under the formulation in Chomsky (1986b), or the related one in Chomsky and Lasnik (1993: 552), Chomsky's argument does go through. Consider the latter formulation:

(10) The GC for α is the minimal complete functional complex (CFC) that contains α and in which α's binding condition could, in principle, be satisfied.

The requirement on *him*, that it be A-free in a local domain, could, in principle, be satisfied in α in (6). And since *him* is, in fact, A-free in that domain (which is a CFC), Condition B is satisfied, incorrectly so, as Chomsky implies. Chomsky's argument thus seems valid. It should be pointed out, though, that the type of formulation of GC that makes it valid might have difficulty with another construction.

(11) *John$_i$ believes him$_i$ to be intelligent

If *him*, the "exceptional-Case-marking" (ECM) subject, is in the lower clause, then, by the above line of reasoning, (11) is incorrectly not a Condition B violation. Below, I will present an analysis of ECM in which the ECM subject raises into the matrix clause. The analysis has as a by-product the elimination of the problem just noted.

Chomsky's argument against A-movement reconstruction based on (5) is abstractly reminiscent of a problem that Belletti and Rizzi (1988) deal with in their theory of anaphora. To account for certain instances of apparently "backward" binding, Belletti and Rizzi propose that Condition A can be satisfied anywhere in the course of a derivation.[2] Given this theory, something needs to be said about examples like the following:

(12) *himself seems to him [*t* to be clever]

Prior to movement, Condition A is presumably satisfied, since as (13) (from Chomsky (1995: 304)) shows, the (NP in the) *to* phrase c-commands into the complement infinitival.

(13) *they seem to him$_i$ [*t* to like John$_i$]

For Belletti and Rizzi, (12) satisfies Condition A, but it violates Condition B, which, according to Belletti and Rizzi, must be satisfied specifically at S-Structure, unlike Condition A. Observe that Chomsky's (5) could also be ruled out in the same way, if, as is plausible (though not logically necessary), its S-Structure

configuration violates Condition B. However, one of Chomsky's major minimalist tenets is that there is no S-Structure as a significant level of representation: all binding conditions are assumed to apply at the LF level. In particular, Condition B must be satisfied at LF. So far, then, Chomsky's Condition B argument against A-movement reconstruction stands.

10.2.3. A "scope" argument

Chomsky's next major argument is based on a curious scope interaction in English between clausal negation and a universal quantifier in subject position. Chomsky presents the following paradigm:[3]

(14)(a) (it seems that) everyone isn't there yet
(b) I expected [everyone not to be there yet]
(c) everyone seems [*t* not to be there yet]

He then argues as follows: "Negation can have wide scope over the quantifier in [(14a)], and it seems in [(14b)] but not in [(14c)]," concluding that "reconstruction in the A-chain does not take place, so it appears" (Chomsky 1995a: 327).

Hornstein (1995: 239) challenges the empirical basis for this argument.[4] He acknowledges that there is a contrast of the kind seen in (14). He gives the following examples, parallel to (14a,c):[5]

(15)(a) everyone didn't leave
(b) everyone seems not to have left

However, he claims that there is an empirical flaw in the argument. He gives the following examples, claiming that they do not allow negation to take scope over *everyone* either:

(16)(a) John would prefer for everyone not to leave
(b) John wanted very much for everyone not to leave

Thus, we would not expect such a reading in (15b) regardless. Hornstein (personal communication) suggests that the crucial property is contraction: the wide scope for negation is possible only when negation has contracted, as in (14a) and (15a). My informants do find wide scope for negation in (16a,b) somewhat less accessible than in (14a) and (15a). However, they do not find it as inaccessible as in (14c) and (15b). Their judgment on (14b) is similar: wide scope for negation is possible, at least to a significant extent. Some other examples with uncontracted negation seem to allow wide scope negation rather readily. Consider the following examples, in a situation where a teacher is being reprimanded for giving all the students As:

(17) school policy requires that everyone not get an A
(18) it is important for everyone not to get an A

Both examples seem reasonably appropriate to the situation, indicating that the reading in question is available. Thus, there is reason to think that Chomsky is correct in suggesting that there is something special about raising constructions in their inability to allow the reading. It is not implausible to take this, as Chomsky does, as evidence that there is no A-movement reconstruction.

10.3. Binding, ECM and overt raising

It will be of interest to examine ECM constructions in this connection. I suspect that Chomsky intended his (14b) to instantiate ECM. It well might; however, the situation is somewhat equivocal, since, as shown by Bresnan (1972), *expect* has multiple subcategorization frames. "True" ECM constructions will be particularly interesting since there is substantial evidence that they exhibit raising – in fact, overt raising. I will briefly summarize the arguments, which fall into two basic classes. First, there are the several paradigms discussed by Lasnik and Saito (1991), some of them following Postal (1974), indicating that an ECM subject can bind into a matrix–clause adverbial. When combined with Lasnik and Saito's arguments that covert raising does not create new binding configurations (and the modern minimalist versions of those arguments in Lasnik (1995b,c, 1997)), those paradigms strongly argue for overt raising in ECM constructions. Second, there is the combination of Chomsky and Lasnik's (1993) argument that (at least some) ellipsis is a PF deletion phenomenon and the argument from Lasnik 1995d that an ECM subject can "escape" an ellipsis site via raising.

10.3.1. Binding into adverbials

Representative examples showing "high" binding behavior for ECM subjects are as follows, displaying, respectively, Condition A satisfaction, weak crossover mitigation, and negative polarity item licensing:

(19) the DA proved [two men to have been at the scene of the crime] during each other's trials
(20) the DA proved [no suspect to have been at the scene of the crime] during his_i trial
(21) the DA proved [no one to have been at the scene of the crime] during any of the trials

For most speakers I have interviewed, (19)–(21) contrast significantly with corresponding examples with finite complements.

(22) ?*the DA proved [that two men were at the scene of the crime] during each other's trials
(23) ?*the DA proved [that no $suspect_i$ was at the scene of the crime] during his_i trial
(24) ?*the DA proved [that no one was guilty] during any of the trials

This contrast strongly suggests that the ECM subject is in the higher clause at some relevant level of representation, the conclusion reached by Lasnik and Saito, and by Postal much earlier. Working within the Government-Binding (GB) framework of assumptions, Lasnik and Saito argue that the relevant level of representation is S-Structure, hence, that raising is overt. In Lasnik (1995b,c, 1997), I present arguments that that conclusion is still valid under minimalist assumptions. The movement is presumably feature driven, the relevant feature being Case, or, more likely, an Extended Projection Principle (EPP) feature in Agr_O.[6] Given Chomsky's (1995a) powerful conceptual argument that covert (formal) feature-driven movement should affect (formal) features alone, the question arises whether raising of mere features creates new binding-type relations. There is very good reason to believe that it does not.

Beginning with Chomsky (1986b), a series of arguments have been made that in existential constructions such as (25), a movement relation connects the expletive and the "associate."

(25) there is a man here

The standard argument for movement is a compelling one: that movement provides the basis for an account of the familiar superficially bizarre agreement paradigms displayed by these constructions, where the verb agrees with something that is not its formal subject.

(26)(a) there is/*are a man here
(b) there are/*is men here

The several approaches under which the associate raises have all had a major empirical defect: they have failed to capture the fact that the associate is always interpreted in situ with respect to scope, as in (27) from Chomsky (1991).

(27) there aren't many linguistics students here

Chomsky (1995a) argues that this scope fact can be captured if the movement affects only the formal features (including the agreement features) of the associate, a consequence in line with a general economy condition mandating that a movement operation move as little as possible. For covert movement, nothing would require that more than the formal features move.

Given this analysis, it is straightforward to show that covert movement (at least of the feature-driven variety) does not create new binding-type relations. (19)–(21), repeated here as the (a) cases of (28)–(30), sharply contrast with corresponding examples with *there*, the (b) cases of (28)–(30).[7]

(28)(a) the DA proved [two men to have been at the scene of the crime] during each other's trials
(b) *the DA proved [there to have been two men at the scene of the crime] during each other's trials

(29)(a) the DA proved [no suspect$_i$ to have been at the scene of the crime] during his$_i$ trial
(b) *the DA proved [there to have been no suspect$_i$ at the scene of the crime] during his$_i$ trial

(30)(a) the DA proved [no one to have been at the scene of the crime] during any of the trials
(b) *the DA proved [there to have been no one at the scene of the crime] during any of the trials

We are led to the conclusion that the high behavior of the ECM subject in the (a) cases is the result of overt raising. Koizumi (1993, 1995), developing ideas of Johnson (1991), outlines a theory that makes this possible, whereby the ECM subject raises to [Spec, Agr$_O$] of the matrix clause and the matrix verb raises to a still higher "shell" V position.

10.3.2. ECM in Pseudogapping

The second argument for overt raising of an ECM subject has to do with the Pseudogapping ellipsis construction. This construction is exemplified in (31).

(31) Mary hired John, and Susan will ~~hire~~ Bill

Following Jayaseelan (1990), in Lasnik (1995d) I argue that the construction involves VP-ellipsis, the remnant having escaped from the ellipsis site via a movement operation. Departing from Jayaseelan, I argue that the movement operation is not heavy NP shift (HNPS) but raising to [Spec, Agr$_O$]. In part, the argument is based on a divergence between Pseudogapping and HNPS possibilities in double object constructions (for discussion, see Lasnik (1995d, 1999c)). In particular, the first object is a good Pseudogapping remnant yet resists HNPS.

(32) ?John gave Bill a lot of money, and Mary will ~~give~~ Susan ~~a lot of money~~
(33) *John gave *t* a lot of money [the fund for the preservation of VOS languages]

Conversely, the *second* object is a poor Pseudogapping remnant but freely undergoes HNPS.

(34) *John gave Bill a lot of money, and Mary will ~~give Bill~~ a lot of advice
(35) John gave Bill *t* yesterday [more money than he had ever seen]

Note that there *is* a correlation with typical A-movement. The first object, but not the second, can undergo passive.

(36) Bill was given *t* a lot of money
(37) ?*a lot of money was given Bill *t*

Now notice that an ECM subject makes a good Pseudogapping remnant.

(38) the DA proved Jones (to be) guilty and the Assistant DA will ~~prove~~ Smith ~~(to be) guilty~~

Thus, we have additional reason for thinking that an ECM subject (like an object, as a matter of fact) undergoes overt raising.[8]

10.3.3. Scope interaction in ECM: is raising optional?

I now return to the main question of this part of the discussion: how do universal ECM subjects interact with clausal negation? The prediction is now that negation cannot take wide scope, just as it cannot in raising-to-subject constructions (Chomsky's observation), since both constructions involve overt A-movement into the higher clause, and Chomsky's account of his observation should carry over: there is no A-movement reconstruction. The prediction is clearly correct for certain ECM constructions. For example, in the rather unusual ECM particle construction (based on *make … out*) discussed by Kayne (1985), a universal ECM subject is clearly outside the scope of negation, as seen in (39).

(39) the mathematician made every even number out not to be the sum of two primes

Note that the only reading is the implausible one where the mathematician was engaged in the futile activity of trying to convince someone that no even number is the sum of two primes (and not the far more plausible one of merely trying to convince someone that Goldbach's conjecture is false). Thus, even with strong pragmatic bias toward wide scope for the negation, it still is not available, consistent with the raising analysis combined with Chomsky's claim.

It is perhaps unsurprising that (39) should pattern with Chomsky's (14c), repeated here.

(40) everyone seems [*t* not to be there yet]

(40) clearly displays overt raising; indeed, it is a paradigmatic instance of the phenomenon. But the interest of the construction exemplified in (39) is that it, too, clearly displays overt raising. *Every even number* is the thematic subject of the lower clause, yet it appears to the left of the particle *out*, which is part of the higher predicate.

Since I have argued that even more conventional ECM constructions involve overt raising, they too would be expected to require wide scope for the thematic subject over lower clausal negation. Curiously, most of my informants find narrow scope possible, even if somewhat disfavored, unlike the situation with raising

to subject or with the *make*-NP-*out* construction. Some representative examples are as follows:[9]

(41) I believe everyone not to have arrived yet
(42) I proved every Mersenne number not to be prime

Those same informants (along with every other native English speaker, I believe) disallow narrow scope for the universal when it undergoes passive/raising to subject position.

(43) everyone is believed not to have arrived yet
(44) every Mersenne number was proved not to be prime

In (44), there is strong bias toward narrow scope, but it is still not available. Only the wildly false wide scope reading exists.

The task immediately at hand is to reconcile the substantial evidence that ECM subjects undergo overt raising with the scope fact in (41)–(42). Narrow scope for the universal subject should not be possible if it has raised, just as narrow scope is not possible in Chomsky's (14c) or in (39) and (43)–(44).[10] Descriptively, the situation so far is that when it is completely clear from the word order that raising has taken place, narrow scope for a universal ECM subject is impossible. But when the word order is equivocal, narrow scope is possible. This suggests that in the latter circumstance, overt raising has not necessarily taken place. The question now is whether this is a flat contradiction or merely a problem. The former will be true if *in the very same sentence* there is evidence both for and against raising. Some of the tests *for* raising are difficult to carry out with a universal quantifier as ECM subject. But Pseudogapping does seem to be possible in that circumstance.

(45) Mary proved every Mersenne number not to be prime, and John will every Fibonacci number

The fact that *every Fibonacci number* is a Pseudogapping remnant indicates that it has overtly raised. If, simultaneously, it could take scope under the (elided) negation as it can in (46), we would have a contradiction.

(46) John proved every Fibonacci number not to be prime

However, it seems that unlike the situation in (46), narrow scope is not possible for *every Fibonacci number* in (45).[11] Thus, there is no direct contradiction. Rather, we are led to the conclusion that raising must be *optional*. Note that that conclusion is consistent with all the binding and ellipsis evidence for raising summarized above. All of these phenomena simply indicate that raising is *possible*, available when necessary but not necessarily obligatory.

There is actually one phenomenon discussed by Postal (1974) and Lasnik and Saito (1991), but not yet discussed here, that argues that raising *is* obligatory. Postal's statement of the argument is based on "a fundamental pronominalization constraint" due to Langacker (1969) that states that a pronoun cannot both precede and command its antecedent. There are a number of more recent formulations of this constraint, including my (1976) noncoreference rule and Chomsky's (1981) Condition C. Any of these formulations can correctly distinguish (47) from (48), but only if the embedded subject in (47) has (necessarily) raised into the higher clause.

(47) *John believes him_i to be a genius even more fervently than Bob_i does
(48) Joan believes he_i is a genius even more fervently than Bob_i does

This phenomenon is, perhaps, not as problematic as it might first appear. After all, it is not uncommon for "object shift" to be obligatory with pronouns even when it is optional with lexical NPs (for discussion, see Johnson (1991) and Diesing (1996)). In fact, even in English, there is a bit of independent evidence for this state of affairs. In the *make out* construction discussed above, many speakers allow the ECM subject to the right of *out* as well as to the left.

(49) Mary made John out to be a fool
(50) Mary made out John to be a fool

If, as assumed above, an example like (49) exhibits overt raising, it is reasonable to conjecture that there has been no overt raising in (50). Significantly, the analog of (49) with a pronoun as ECM subject is still good, but the analog of (50) is bad, even for speakers who find (50) itself fully acceptable.

(51) Mary made him out to be a fool
(52) *Mary made out him to be a fool

This is abstractly quite parallel to the situation seen above, which also argued that raising is optional, yet obligatory with pronouns. In fact, the parallel is still deeper. Recall that a universal to the left of *out* cannot have narrow scope under clausal negation in the infinitival, as seen in (39), repeated here.

(53) the mathematician made every even number out not to be the sum of two primes

Those speakers I have interviewed who accept the alternative word order for the construction, as in (54), do allow narrow scope for the universal in that instance.

(54) the mathematician made out every even number not to be the sum of two primes

I believe that other "height" tests summarized above also conform to the pattern just outlined, though more data collection is clearly in order. For example, a negative ECM subject to the left of *out* seems much more comfortable with a negative polarity item in the matrix clause than does one to the right of *out*.

(55)(a) the lawyer made no witnesses out to be idiots during any of the trials
(b) ?*the lawyer made out no witnesses to be idiots during any of the trials

And, though judgments are subtle, I find similar effects with the two other height tests considered earlier, anaphor binding and weak crossover.

(56)(a) the DA made the defendants out to be guilty during each other's trials
(b) ?*the DA made out the defendants to be guilty during each other's trials

(57)(a) the DA made no suspect$_i$ out to have been at the scene of the crime during his$_i$ trial
(b) ?*the DA made out no suspect$_i$ to have been at the scene of the crime during his$_i$ trial

Thus, it is still reasonable to think that (overt) raising of an ECM subject is optional, obligatory only with pronouns. When raising of a universal does take place, as evidenced by position to the left of *out* or high binding or survival as a Pseudogapping remnant, that universal cannot "reconstruct" so as to take scope under lower clausal negation. That is, "raising to object" parallels "raising to subject" and provides further evidence for Chomsky's claim that "reconstruction in [an] A-chain does not take place…"

10.3.4. What is "optional" about raising?

The next question is how the optionality of "raising to object" is to be instantiated. I continue to assume that when the ECM subject raises, it raises to [Spec, Agr$_O$].[12] This is essentially as in Chomsky (1991), except that I take the movement to be overt rather than covert, for all the reasons outlined above. As briefly mentioned earlier, given the word order of English, the fact that the raising is overt further entails that the verb normally raises to a still higher position, as in Koizumi's (1993, 1995) "split-VP" hypothesis, which I adopt in its essentials. The relevant portion of an ECM structure with raising, (58), is as shown in Figure 10.1.

(58) she will prove Bob to be guilty

In Lasnik (1995c), I argued that the driving force for the overt movement of the NP is a strong EPP feature in Agr$_O$. Following Chomsky (1991), I took Agr$_O$ to be the same item as Agr$_S$, the labels being merely mnemonic. Overt "object shift" is then analyzable as the same phenomenon as the standard EPP. I assume that Case checking is just a side-effect of EPP satisfaction;[13] there is no requirement

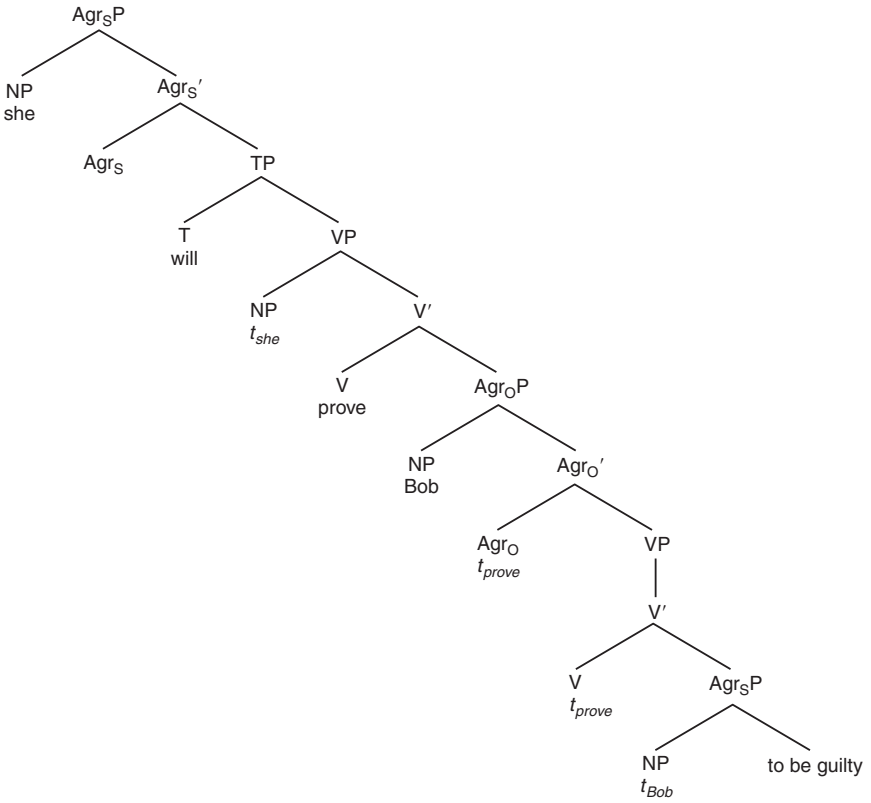

Figure 10.1 The relevant portion of an ECM structure with ECM subject and matrix V both raising.

that Case *per se* be checked overtly. One way to make the raising optional might be to abandon the idea that Agr_O is the same item as Agr_S, assuming, instead, that only the latter obligatorily has an EPP feature. Agr_O would only optionally have the feature. Some of the discussion in Chomsky (1995a) hints at another possibility. Chomsky reasons that "[i]f Agr has no strong feature, then PF considerations, at least, give no reason for it to be present at all, and LF considerations do not seem relevant" (Chomsky 1995a: 350). He thus suggests that "Agr exists only when it has strong features" (Chomsky 1995a: 351). Along these lines, suppose that the optionality of raising is the optionality of Agr_O.[14] If Agr_O is present, overt raising will be forced by its strong EPP feature. If Agr_O is absent, there will be no overt raising; the nominal's Case will be checked by covert raising of its formal features to the V.[15] Under that circumstance, the nominal will not participate in high binding, nor will it survive as Pseudogapping remnant. On the other hand, it will be able to take low scope, as in the instances of ambiguous interaction between universal and negation discussed earlier.

10.4. Quantifier Lowering

I return now to another scope "reconstruction" phenomenon briefly alluded to above: so-called Quantifier Lowering (QL). (59) is a classic example, from the groundbreaking discussion in May (1977).

(59) some politician is likely to address John's constituency

The intuition about (59) is that it is ambiguous. To quote May "[(59)] may be taken as asserting either (i) that there is a politician, for example, Rockefeller, who is likely to address John's constituency, or (ii) that it is likely that there is some politician (or other) who will address John's constituency" (May 1977: 189). Since May (1977), that ambiguity has been taken to stem from whether the surface subject "lowers" or not. Chomsky distinguishes this phenomenon from the one found in low scope under negation for a universal quantifier subject. Recall that Chomsky observes that a raised NP cannot reconstruct to take narrow scope in that way, as in (14c), repeated here.

(60) everyone seems [*t* not to be there yet]

On the other hand, Chomsky accepts the standard claim that (59) is ambiguous and that the ambiguity is a structural property. As mentioned in note 4, to reconcile his conclusion based on (60) with standard QL as in (59), Chomsky suggests that the "lowered" reading in the latter instance

> could result from adjunction of the matrix quantifier to the lower IP (c-commanding the trace of raising and yielding a well-formed structure if the trace of quantifier lowering is deleted, along the lines of May's original proposal). But reconstruction in the A-chain does not take place, so it appears.
>
> (Chomsky 1995a: 327)

However, as also mentioned in note 4, Zubizarreta (1982) takes the two phenomena to be related, indicating that the lack of low scope in an example like (60) is a potential argument against QL. Hornstein (1995) also takes the two phenomena to be related and regards (60) as potentially arguing against QL (which he treats as A-movement reconstruction). Although this is clearly not Chomsky's position, it does, in fact, seem to be the null hypothesis.

Under the null hypothesis, some other way of resolving the apparent contradiction must be found. In this connection, it must first be noted that it is not entirely clear precisely what the phenomenon of QL is. It is often taken as paraphrasability by a sentence with an expletive subject, as perhaps intended in the quotation from May (1977) given above. Thus, the QL version of (61) is taken to be synonymous with (62).

(61) some politician is likely to address John's constituency
(62) it is likely that some politician will address John's constituency

Given this characterization, QL turns out to be far more limited than it is generally taken to be. Consider first the following two subject-raising examples with negative subjects:

(63) no large Mersenne number was proven to be prime
(64) no one is certain to solve the problem

Neither of these has a "lowered" reading, at least on the paraphrase characterization. (63) is not accurately paraphrasable as (65), nor is (64) as (66).

(65) it was proven that no large Mersenne number is prime
(66) it is certain that no one will solve the problem

This paraphrase failure is not limited to negative contexts. Consider (67) in a situation where there are five fair coins, flipped in a fair way.

(67) every coin is 3% likely to land heads

This situation strongly biases the sentence toward the lowered reading, but that reading still is not possible. (67) cannot be accurately paraphrased as (68).

(68) it is 3% likely that every coin will land heads

Thus, at least on this characterization, there is reason to believe that Zubizarreta and Hornstein are correct in taking absence of a low reading in an example like (69) to potentially argue for failure of QL, and that Chomsky is correct that that absence indicates impossibility of reconstruction with A-movement.

(69) everyone seems [*t* not to be there yet]

What of the successful instances of QL, such as (70)?

(70) some politician is likely to address John's constituency

As far as I can tell, these obtain exclusively with indefinite subjects. If QL is an instance of A-movement reconstruction, it is not obvious how this restriction is to be captured. For that matter, the same is true under Chomsky's suggested characterization of the phenomenon, following May. I therefore offer the tentative speculation that there is no QL at all,[16] and that the reason, following Chomsky in essence (though not in specific detail), is that there is no A-movement reconstruction. The apparent paraphrasability with indefinites might then be a consequence of the meaning of indefinites, rather than the result of a syntactic operation, though I readily concede that much further investigation is needed.

May (1985) presents an important, and often cited, argument that actual syntactic lowering must be involved in the second reading of examples like (70): namely, that such a "lowered" reading for the quantifier is incompatible with the binding of a pronoun in the upper clause. He gives the following example:

(71) no agent$_i$ was believed by his$_i$ superior to be a spy for the other side

It does seem correct that there is no lowered reading for the subject in (71). However, the same is apparently true even with no bound pronoun, as in examples (63)–(64). A more relevant test would involve an indefinite subject, as in (72).

(72) some professor$_i$ is believed by his$_i$ students to be a tyrant

It is certainly true that (72) cannot be paraphrased as (73).

(73) *it is believed by his$_i$ students that some professor$_i$, is a tyrant

But it is not clear what we can conclude from the fact that a sentence cannot be paraphrased by an ungrammatical sentence (in this case, one that violates the Weak Crossover Constraint). I will thus continue to tentatively assume that it is not a syntactic operation that is responsible for the "lowered" reading of raised indefinites.

10.5. A-movement does not leave a trace

The most fundamental question remains: if, indeed, there is no A-movement reconstruction, why should that be? Recall that for Chomsky, there is simply the stipulation that reconstruction is a property solely of operator–variable constructions. Further, the mechanism for instantiating the property – no deletion of traces in A-chains – does not seem to capture it at all. As an alternative, I would like to suggest that A-movement, unlike Ā-movement, does not leave a trace, where a trace is, following Chomsky, a copy of the item that moves, and LF reconstruction effects result from failure to delete (a portion of) a lower copy. This distinction is conceptually plausible: Ā-movement typically creates an operator–variable relation, so at least an "initial" trace is necessary. For A-movement, on the other hand, the trace is seemingly a theoretical excrescence. There are not two separate interpretive roles for a moved NP and its trace to fulfill. It might therefore be conceptually desirable for A-movement not to leave a trace. However, Chomsky (1995a) offers an indirect argument against that possibility, to which I now turn.

Chomsky's argument specifically concerns trace deletion, but it straightforwardly extends to not leaving a trace in the first place. The argument is that at least the initial trace of A-movement is needed for θ-theoretic reasons.

> In the phonological component, traces delete. We have found no reason to extend that convention to the $N \rightarrow \lambda$ computation, and indeed cannot; were we to do so, θ-positions would be invisible at LF …
>
> (Chomsky 1995a: 301)

In effect, Chomsky seems to take θ-role assignment as an A-movement reconstruction effect (so his stipulation barring reconstruction becomes more specific). An alternative approach is imaginable. Suppose that instead of being determined specifically at the LF level, θ-roles are "checked" in the course of a derivation.[17] The moved argument is then itself a record of the crucial part of the history of its derivation. This view of θ-roles as features is argued for on independent grounds in Bošković and Takahashi (1998) and Lasnik (1995d). To the extent that such an analysis can be maintained, we can circumvent Chomsky's conclusion that the trace of A-movement must persist to the interface level (hence, a fortiori, must exist in the first place). The absence of scope reconstruction[18] would then follow from the fact that, plausibly, determination of scope is not satisfaction of a formal feature, but a matter of interpretation at the interface.[19]

Chomsky himself, in a different (but not unrelated) context, presents an argument that would appear to have the consequence for θ-role assignment that I just arrived at (and paradoxically so, since it ostensibly argues just the opposite). Section 4.6 of Chomsky (1995a) is devoted to arguing that "there should be no interaction between θ-theory and the theory of movement" (Chomsky 1995a: 312). In particular, according to Chomsky, movement can never create a θ-configuration. In a theory with D-Structure, this is virtually automatic. But within a minimalist approach where LF is assumed to be the sole interface with semantics, the consequence that "θ-relatedness is a 'base property'" would be considerably more surprising and interesting. Chomsky suggests that this follows from the "configurational" view of θ-roles: "A θ-role is assigned in a certain structural configuration …" (Chomsky 1995a: 313). The reasoning is as follows:

> If α raises to a θ-position Th, forming the chain CH = (α, *t*), the argument that must bear a θ-role is CH, not α. But CH is not in any configuration, and α is not an argument that can receive a θ-role.
>
> (Chomsky 1995a: 313)

If this argument is valid, it appears to have an obviously unintended consequence. A-movement of an argument should *never* be permitted. This is so since such movement will always create a chain; at LF, by the reasoning in the quoted passage, the moved argument would not be able to receive a θ-role. That is, whether the movement is *to* a θ-position or *from* a θ-position, the result is a chain, and chains are not in "configurations." If the constraint Chomsky suggests is valid, the only obvious way it can be satisfied is for θ-roles to be assigned prior to movement. But then, as discussed above, whether A-movement leaves a trace or not is irrelevant to θ-assignment. Ironically, this conclusion undermines the argument against movement into a θ-position. If an A-trace is not only not helpful in the assignment of a θ-role, but would actually make such assignment impossible, then, obviously, the argument that such a trace must exist for θ-theoretic reasons fails. But if A-traces do not exist, then an argument will invariably be a single-membered chain no matter how many times it moves. Thus, even if it were to

move into a θ-position, it would still be in a "configuration" in the relevant sense, so the θ-role should be assignable. And, indeed, there are a number of recent arguments that movement into a θ-position does occur (see, for example, Bošković (1994a); Bošković and Takahashi (1998); Lasnik (1995d); Hornstein (1999)).

10.6. Concluding remarks

This discussion is, admittedly, very far from conclusive. To the extent that it is on the right track, it partly resurrects older ideas, but within a more modern framework of assumptions. For example, the idea that θ-assignment is a "base property," whereas scope belongs to late derived structure, is familiar from the Extended Standard Theory, and even from the Standard Theory (Chomsky (1965)). In the theory sketched here, as in Chomsky (1995a), this is captured without a level of D-Structure, but rather with what might be regarded as a "D-Structure component," to use a phrase suggested to me by Juan Uriagereka and Roger Martin. If, as claimed by Chomsky (1995a), and further supported above, there is no A-movement scope reconstruction, this receives a principled account if A-movement leaves no trace (i.e. copy). I will conclude with a problem: how can movement without a trace possibly be reconciled with a "bare phrase structure" theory of structure building? A-movement not leaving a trace means that a "term" in the sense of Chomsky (1994) is eliminated. Whether this is a serious problem only further research can reveal.

Acknowledgment

I am indebted to Željko Bošković and to all the participants in my 1998 syntax seminar at the University of Connecticut for very helpful comments, questions, and suggestions; to Arthur Stepanov for invaluable assistance in the preparation of the manuscript; and to Norbert Hornstein and Juan Uriagereka for spirited encouragement.

NOTES

2 PATTERNS OF VERB RAISING WITH AUXILIARY "BE"

1 I return to the question of why verb raising takes precedence over Affix Hopping. As for the apparent "last resort" nature of *do*-support, see Watanabe (1993) and Baker (1991) for proposals. See also Bobaljik (1994) for discussion in terms of a Chomsky (1957) style analysis of English verbal inflection, similar to one I will argue for below.

2 Chomsky (1991) takes Affix Hopping to be a standard instantiation of syntactic Move α. Hence, no simple adjacency requirement of the sort in earlier analyses could be relevant. Below, we will see that the earlier analysis is superior in just that regard.

3 Agr_S is the subject agreement projection, and Agr_O is the object agreement projection.

4 As Roger Martin observes, the mechanism by which re-raising remedies the violation is not entirely clear. Chomsky states that "Subsequent LF raising … to the position of *t* is required to create a proper chain" (Chomsky 1991: 426). Note though, that if the raising is adjunction, and if the chain is originally improper because the *t* is marked [−γ] in the ECP notation of Lasnik and Saito (1984, 1992), then the violation is not obviously eliminated. One possible approach to this problem is to allow γ-marking to freely apply anywhere in the derivation, rather than insisting that it apply immediately as the trace is created. Once all raising is completed, every trace is arguably in a configuration of antecedent government.

5 The illicit A′-movement in question is long adjunct movement, for example, of the sort discussed by Lasnik and Saito (1984):

(i) *Why_1 do you believe the claim that John said [Bill left t_1]
The derivation to be excluded involves a short move to the lowest [Spec, C] followed by a long move out of the complex NP to the matrix [Spec, C]. Similarly, the relevant illicit A-movement would be such "super-raising" as (ii):

(ii) *John seems that it is likely to be arrested *t*
with a short move to the lowest [Spec, I] followed by a long move to the matrix. For head movement, Željko Bošković suggests (iii), where *have* has moved through Agr_O on its way to matrix Infl:

(iii) *You have not believed Peter to *t t* gone there

6 Interestingly, this checking view more precisely captures the insight of the earliest modern version of Case theory, that of Vergnaud (1977), than does the Case assignment approach of Chomsky (1980, 1981).

7 These examples are taken from Wilder and Čavar (1994). See also Bošković (1994b) for discussion of a Serbo-Croatian construction that allows fully inflected finite auxiliary verbs to remain in situ.

8 See Epstein (1998) for discussion, and a possible analysis.

9 See also Halle and Marantz (1993), Bobaljik (1994).

10 I continue to assume that movement is driven solely by features of the appropriate sort. In particular, I assume that the property of being an affix is not a feature relevant to syntactic head-raising. Thus, as Roger Martin notes, under either Chomsky's (1993) GREED constraint, or the weaker ENLIGHTENED SELF-INTEREST of Lasnik (1995a), movement of bare V will be blocked by general economy considerations.

11 See Oku (1996) for some discussion.

12 See Lasnik (1995e) for more detailed discussion.

13 I use the term "deletion" merely for ease of exposition. As far as I can tell, the arguments I present here are neutral between a PF-deletion approach to ellipsis and an LF-copying one.

14 This is essentially the formal analog of the insightful semantic proposal of Warner (1986), though he argued that such a treatment is appropriate for *be* but not for *have* (based on subtle acceptability differences that I am putting aside).

15 (i) indicates that even in a positive imperative *be* behaves strictly like a bare verb, and does not undergo raising:

(i)(a) Should I be quiet? Please do.
(b) *Please be.

If imperative *be* could raise, (ib) could arise from raising, and deletion of the residual VP, just as in (ii):

(ii) John is

16 The characterization of *do* support is a surprisingly recalcitrant problem within recent transformational theories. On the face of it, the process would seem to be in full accord with economy approaches: the rule applies if and only if it has to. However, if Chomsky (1994) is correct that derivations are only compared if they involve all the same lexical choices (the same "numeration") derivations with and without *do* will not be compared. As a result, we seem to have no way to capture the intuition that has guided work on English verbal morphology for forty years. I hope to explore this issue in detail in future work. Here, I offer the speculation that *do* is not, in fact, a distinct lexical choice. Rather, forms of *do* are simply the spell-out of certain morphemes (Tense, Imp, Hab) when they fail to merge with a verb. Merger, being a low-level phonetic process, would be automatic whenever applicable. As far as I can tell, such an approach, when grafted onto the theory of verbal morphology outlined above, is consistent with the bulk of the relevant facts. The remaining problems involve the impossibility of *do* with *have* and *be* in finite clauses, even when there is negation or inversion:

(i) *John doesn't be here
(ii) *Does John be here

Recall that the bare form of *be* does exist: it shows up in imperatives, in infinitives, and following modals. How to prevent it from occurring in finite negatives and interrogatives is unclear. See Lasnik (1995e, fn. 20) for a speculation.

17 Given the discussion above, it would be helpful to examine how VP-ellipsis patterns in the habitual construction. Unfortunately, I do not have the relevant data at present.

3 LAST RESORT AND ATTRACT F

1 The relevance of (4) is unclear, even under all of Chomsky's assumptions, since the failure of the higher subject position to be filled in overt syntax will inevitably cause the derivation to crash.

2 As observed by Martin (1992a), the line of analysis pursued here resolves a puzzle for last resort more generally. On the lexicalist view demanded by the now standard LF theory of Case, an NP is inserted into a structure with its Case feature. Case is not "assigned," but is "checked." The puzzle is that even if an NP is in an inappropriate Case position, one that does not match the Case of the NP, movement to an appropriate Case position is still blocked. If it were only last resort that prevented movement from a Case position to another, it would not be clear how this instance is prevented.

3 See Bošković (1994a, 1995c) for several arguments that movement into a θ-position must not be excluded in general.

4 This does not entirely resolve the scope issue, as I will discuss below. A related matter is that, as noted by Lasnik and Saito (1991), substitution creates binding theoretic difficulties as well. I will return to this issue also.

5 Note that the familiar agreement paradigm with existentials and similar constructions, which provided the primary motivation for movement in the first place, potentially raises the same technical problem as Case licensing. Below I will sketch an account that avoids this problem.

6 Since the agreement in sentences like (20) is also manifested on the verb, we have yet more evidence that agreement features of an NP are not checked off, even when the NP participates in a checking relation.

7 Again, this is as expected if the feature has semantic content.

8 For present purposes, I will continue to follow Chomsky on this point, but it is not obviously correct. Chomsky (1991) argues that in overt syntax, VP internal subject moves directly to [Spec, I]. This move seems to violate shortest move by skipping [Spec, Agr_O]. However, Chomsky claims that [Spec, Agr_O] need not exist until LF, so it is not actually skipped. This leaves open the possibility that [Spec, I] similarly need not exist overtly. Granted, if it did not, the EPP would be violated. But under Greed, that should be of no concern to the NP that is moving.

9 There is actually reason to believe that the complement of *to* in raising type constructions is simply unavailable for A-movement. If this were not so, standard raising in such instances as (i) would be prevented by Chomsky's recent version of relativized minimality (essentially, Oka's (1993) Shallowness), assuming (as is standard) that the complement of *to* c-commands the clausal complement.

(i) John seems to Mary [*t* to be clever]

I defer consideration of this issue.

10 Note that this provides hope of making Procrastinate a true economy condition: waiting until LF to move entails moving less material.

11 Chomsky actually claims that *there* not only lacks agreement features but also Case features. I reject here the latter claimed deficiency, as it would seem to preclude an account of (i).

(i) *There someone laughed

If *There* does not check nominative Case, that Case should still be available for (the Case feature of) *someone*. Further, as John Frampton observes, (ii), from Lasnik (1992a), should also then be well-formed.

(ii) *There is likely [there to be a man here]

12 As all alternative, Ormazabal (1995) argues that complements of ECM verbs are CPs (while control complements are IPs), the reverse of the widely accepted proposal of Chomsky (1981). The impossibility of nominalizations of such constructions then follows, on Ormazabal's account, from general constraints on the incorporation of the zero complementizer.

13 See Bošković (1994c, 1995c) for extensive discussion of such verbs.

14 One additional problem in Belletti's original proposal, and carried over in the successive revisions in Lasnik (1992a) and (1995a), persists: the apparent optionality of "partitive" Case.

In (i), and even more strikingly in (ii), it evidently must be true that *be* need not discharge its Case.

(i) A man is here
(ii) John is here

I know of no fully satisfactory solution to this problem, so for present purposes I will give a merely technical answer: *be* and unaccusative verbs come in doublet pairs with and without the Case feature. So-called middle constructions might involve similar doublets, thereby permitting the object of a normally transitive verb to move to a Case checking subject position. A possibility that I am exploring in work in progress is that Case (whether partitive or even nominative or accusative) need not be checked, the apparent Case checking needs of verbs and Tense actually reflecting a more general EPP feature. See below for brief discussion of a generalized EPP.

15 Something like Ormazabal's proposal of footnote 12 might be necessary for (i).

(i) The belief [a man to be [*t* here]]…

Suppose a *man* bears partitive Case, and *be* checks that Case. By hypothesis, *a man* is still available for further movement, so it should be able to move to [Spec, I] in (i) thus satisfying the EPP. All formal requirements are then met. Only one other possibility occurs to me. If, as argued in Lasnik (1995a), the configuration for partitive Case checking is not created until LF (*here* raising to the "light" verb *be*, and the complex raising to Agr_O; *a man* in [Spec, Agr_O]) then *a man* would not have its Case checked overtly. To have its Case checked, *a man* would have to lower covertly. If (contra Lasnik and Saito (1984, 1992)) there is a general prohibition of such lowering, then *a man* could not get its Case checked at all.

16 Among recent work on this, see Lasnik and Saito (1992) for the former possibility and Kitahara (1992) for the latter.

17 I put aside consideration of the precise nature of the constraint on anaphor–antecedent relations, and of other similar relations to be discussed immediately below.

18 Further, as Lasnik and Saito also discuss, given that objects of simple transitives also behave as if they c-command adverbial adjuncts, even objects must raise to a higher position (on standard, but not completely uncontroversial, assumptions about phrase structure and c-command).

19 Lasnik and Saito's conclusion was similar, and, as will be seen, for similar reasons.

20 Bošković (1994c), operating under a slightly different set of assumptions, suggests an additional way out. He proposes, following Chomsky (1991) and Lasnik (1995a), that *there* is an affix. Further, following Lasnik, he takes that affixal nature to be a formal inadequacy driving movement. However, departing from Lasnik, he proposes that rather than the associate raising and adjoining to *there, there* lowers and adjoins to the associate (a sort of Affix Hopping). This immediately gives the result that for LF purposes, the behavior of the associate is in accord with its S-structure position.

21 Chomsky (1995a) denies that the associate of *there* has lower behavior in general. He indicates that "Covert raising to Agr_S places the features of the associate in a structural position with the essential formal properties of [Spec, Agr_S]. We therefore expect the associate to have the binding and control properties of the overt subject …" We have seen overwhelming evidence that this expectation is not fulfilled. Chomsky offers little evidence that it is. The major apparently confirming phenomenon he presents involves control:

(i) There arrived three men (last night) without [PRO] identifying themselves

As (ii) shows, a typical object cannot control PRO in this construction:

(ii) *I met three men (last night) without identifying themselves

As extensively discussed above, we cannot conclude that the associate of *there* normally has higher behavior. One possibility is that control involves a different configuration from binding. In support of this, it might be noted that the configurations licensing control remain rather mysterious. Alternatively, one might question the generality of the control phenomenon. Chomsky implies that the associate of *there* is behaving just as an overtly raised subject, as in (iii), would.

(iii) Three men arrived (last night) without PRO identifying themselves

But already there is some difference. While (iii) is perfect, (i) is somewhat degraded. This contrast is heightened if the adverbial is fronted:

(iv) Without PRO identifying themselves, three men arrived
(v) ?*Without identifying themselves, there arrived three men

Further under raising, the contrast between structures like (i) and (iii) intensifies. In the following examples, the adverbial is intended as being in the higher clause, along with the raised subject or *there*:

(vi) Someone seems to be available without PRO seeming to be eager to get the job
(vii) *There seems to be someone available without PRO seeming to be eager to get the job

Until this array of facts is sorted out, the following interesting typological claim made by Chomsky must be held in abeyance. Chomsky indicates that languages with expletives of the *there* type (i.e. with no agreement features of their own) allow control in the constructions at issue, while languages with expletives of the *il* type do not. (He seems to indicate that the null subject parameter is also somehow relevant, but according to his analysis, that factor is orthogonal.) Thus, he indicates that Italian and

French contrast:

(viii) Sono entrati tre uomini senza identificarsi
(ix) (*)Il est entré trois hommes sans s'annoncer

But again, there is some question about the data. Several French speakers find (ix) reasonably acceptable. Thus, it is not so clear that French and English contrast. Resolution of these important issues awaits further exploration of control in the Minimalist framework.

22 The theory of Bošković (1994c) also correctly handles the ACD paradigms, as Bošković observes.

23 In this chapter, only structural Case has been directly considered. See Lasnik (1995a) for a proposed extension of this conclusion to inherent Case as well.

24 Wasow (1972) attributes the identical proposal to a lecture of Chomsky's in 1971: Chomsky (class lectures,1971) "… suggests that there is an independently necessary rule which reduces stress on certain repeated string…. He suggests that VP deletion and Sluicing can be formulated as very late rules which delete unstressed strings." It is sometimes argued that ellipsis must involve LF copying rather than PF deletion based on "missing ambiguities." For example, while (i) is two ways ambiguous, (ii) is not four ways ambiguous, but still only two.

(i) John wants to catch a fish
(ii) John wants to catch a fish and Mary does too

The ambiguity of a *fish* must be resolved the same way in the missing VP as in the overt one. Given plausible assumptions about LF processes, this follows. However, as already noted in Lasnik (1972), the missing ambiguities phenomenon arises even in the non-elliptical analogue of (ii):

(ii) John wants to catch a fish and Mary wants to catch a fish too

Thus (i) seems not to bear on the choice between a copying and a deletion account. See Tancredi (1992) for recent discussion.

25 It is significant that the more common sort of ACD, that involving restrictive relativization, does not have this property, as noted above. Even when reanalysis is impossible, as in (63), restrictive ACD is virtually perfect:

(i) Mary stood near everyone Emily did
(ii) John showed Bill everyone Mary did

This fact casts serious doubt on any theory of ACD that provides the same treatment for restrictive and appositive instances, such theories as those of Wyngaerd and Zwart (1991), Lappin (1992), and Homstein (1994). Fiengo and May (1992) is a rare instance of an analysis that both provides for the possibility of appositive ACDs and still distinguishes them from restrictive ones. I leave open here the analysis of examples like (i) and (ii). See Lasnik (1993) for discussion.

26 Note that this analysis is flatly inconsistent with "Holmberg's Generalization," hence arguably untenable. However, the factual basis for the generalization has been persuasively challenged by Bobaljik and Carnie (1996) and Guilfoyle (1993) for Irish; and by Déprez

(1989), Bobaljik (1994), and Bobaljik and Jonas (1996) for German and Dutch. Given Chomsky's (1993) deduction of Holmberg's Generalization via shortest move, the question arises how any language could fail to conform to it. See Bošković (1997b), Watanabe (1993), Bobaljik (1994), and Takahashi (1994) for a range of answers to this question.

27 See Koizumi (1993) and Ura (1993) for relevant discussion. Koizumi's "split VP hypothesis" proposes that even a simple transitive construction involves two VPs separated by Agr_O. The verb raises to the head of the higher shell-like VP. Given that the subject is generated as the Spec of the higher VP, a reasonable conjecture is that the feature driving the raising is a θ-feature. See Bošković and Takahashi (1998) for evidence that θ-features are strong in English.

28 For Chomsky, on the other hand, feature strength is apparently always a property of the head to which raising takes place. See Lasnik (1999b) for a reanalysis in Chomsky's terms.

29 Lasnik (1999c) presents an analysis of Pseudogapping in these terms.

4 LEVELS OF REPRESENTATION AND THE ELEMENTS OF ANAPHORA

1 This research was supported in part by NSF grant SBR-951088. I would like to acknowledge the very helpful suggestions of Željko Bošković and those of a reviewer.

2 Actually, according to the specific analysis of expletive–argument pairs in Chomsky (1986b), (9) would straightforwardly satisfy the binding requirement of the anaphor at S-structure, since Chomsky claimed that throughout the derivation expletives are coindexed with their associated arguments. But see Lasnik (1992b) for arguments against this position.

3 I do not give a minimal pair here, such as "I saw two knights on each other's horses" because of the possibility of taking the material after the verb as an NP, "two knights on each other's horses." No such possibility exists for the example in the text.

4 It is not immediately obvious how to extend this to the focus movement case. Also, while it is easy to imagine how scope facts can be handled on the Hornstein and Weinberg theory (or, for that matter, on a theory with no LF A′-movement at all), one of the major arguments for QR, May's (1985) account of antecedent contained deletion, demands that the entire expression move. See Lasnik (1993), Hornstein (1994), and Lasnik (1995d) for discussion.

5 This argument depends on the "minimalist" assumption that the configurations relevant to syntactic relations are severely limited, essentially to Spec–head and head–head (via head adjunction). Thus, agreement at a distance would not be possible, a limitation in descriptive power, all else equal.

6 In examples (20)–(25), the adjunct is to be taken as modifying the matrix predicate, as indicated by the brackets. With the adjunct modifying the embedded predicate, even the *there* examples are acceptable, as expected in parallel to Uriagereka's example.

7 See also Kayne (1994, ch. 7).

8 And, again, short of the Larsonian structure, we are led to the conclusion that direct objects also raise. Postal did not arrive at this latter conclusion, primarily because he took the structural requirement on binding phenomena to involve command (weaker than c-command in specifying clausal nodes only).

9 Similarly for the Case of the associate of *there*, given Uriagereka's example above. This is expected on the Belletti (1988) theory, developed further by Lasnik (1992a, 1995a), in which *be* and unaccusatives are Case licensers. The movement to (a position near) *there* is then not for Case reasons, but for agreement reasons, as in Martin (1992a), and Groat (1995).

10 See Martin (1992a) and Groat (1995) for related proposals.

11 I do not mean to deny the relevance of formal agreement features to the licensing of anaphors, but merely to claim that more is involved.

12 See Lasnik (1995c,d) for further arguments. And see Wyngaerd (1989) for what was, as far as I know, the earliest proposal that overt raising to [Spec, Agr_O] creates binding possibilities.

13 A reviewer notes a potential problem for a Koizumi type analysis of English: the Germanic VO languages (English, Scandinavian) contrast with the Germanic OV languages (Dutch, German) with respect to verb–object adjacency, the lack of adjacency in the latter occasionally explained by overt raising to [Spec, Agr_O]. Zwart (1993), revised as Zwart (1997), is a major example of such an analysis. Interestingly, though, Zwart bases his major arguments for raising to Agr_O in Dutch on Wyngaerd (1989). But Wyngaerd explicitly, and persuasively, argues that English does not differ from Dutch in this regard. The account of the adjacency asymmetry must then lie elsewhere.

14 Subject orientation remains a mysterious phenomenon. An intriguing possibility is that it relates to how Case is licensed (as suggested by Lasnik (1993)), or to *when* Case is licensed, with covert movement of (formal features of) objects to their Case position resulting in inability of those objects to license anaphors. At this point, however, this is just wild speculation.

15 Norvin Richards points out that there is another way to look at this property of control on which it is less obviously relevant to the point at issue: "Subject orientation" seems to be a property of particular anaphors, rather than of languages *per se*. Given that perspective, PRO could be regarded as a lexical item without that specific property.

16 See Lasnik (1992b) for discussion.

17 Further, as Bob Fiengo pointed out in the discussion following the oral presentation of this chapter, with a slight change in the adverbial, (37) itself degrades substantially:

(i) ?*There arrived three men without saying hello

18 Chomsky indicates that the null subject parameter is also somehow relevant, but according to his analysis, that factor is orthogonal. A priori, a null subject could be of the *there* type or of the *il* type, and that is all that should be relevant on Chomsky's account.

19 Chomsky in fact notes that "… the French examples … have a more equivocal status than in the idealization here" (Chomsky 1995a: 384).

20 I postpone for another occasion considerations of "reconstruction": essentially, the question of whether there is licensing *prior* to S-structure.

21 Another possibility consistent with the facts examined here is that what is at issue is not formal licensing but rather interpretation. Earlier approaches to anaphora, such as RI of Chomsky (1973) and the disjoint reference rule of Lasnik (1976), utilized interpretive rules rather than syntactic filters or licensing conditions. Chomsky and Lasnik (1993) suggest that such an approach should be resurrected.

5 PSEUDOGAPPING PUZZLES

1 See Bošković and Takahashi (1998) for a powerful argument that θ-features are strong formal features in English.

2 The V is, of course, still present in the LF component, and in that component is free to raise, checking its feature.

3 Later, we will see some indirect evidence that the overt raising of V to V_2 does, indeed, take place, as does that of the second complement to [Spec, Agr].

4 Larson (1988) argues against this, but the binding theoretic phenomena (Barss and Lasnik 1986) that provided the major impetus for his analysis are easily accommodated on the present point of view.

5 A range of "degree of grammaticality" phenomena can be similarly analyzed. For example, Bošković (1997c) considers the contrast between the very bad (i) and the even worse (ii).

(i) *John$_i$ is likely t_i sleeps often.
(ii) **Is likely John sleeps often.

Bošković observes that while (ii) violates both the Inverse Case Filter and the EPP, (i) violates only the first of these conditions. The Inverse Case Filter is the requirement that a Case licenser actually license a Case, thus discharging its Case feature. If the EPP is also the requirement that a feature be checked, perhaps the *D*-fearure suggested by Chomsky (1995a), the situation at hand is strikingly parallel to the one in the text: *Infl* has one unchecked feature in (i), and two in (ii). Epstein (1990) provides another instance in a classic study of degrees of grammaticality. Epstein argues against the "Visibility" reduction of the Case Filter to the Theta Criterion on the grounds that speakers systematically distinguish between examples that violate both the Case Filter and the Theta Criterion (iii), on the one hand, and those that violate only the Case Filter (iv), on the other hand.

(iii) **I hope John to be likely that John left.
(iv) *I hope John to think that Bill left.

If, as proposed in the text, arguments have theta features that need to be checked, in (iii) *John* has two unchecked features (Case and theta), while in (iv) *John* has only an unchecked Case feature. So once again, there is reason to think that two unchecked features cause more extreme deviance than one. I am grateful to an anonymous reviewer for pressing me to make this connection more explicit, and to Cedric Boeckx for extremely helpful discussion.

6 Such matrix Sluicing instances raise an interesting question. If the source of (i) is (ii), then what is deleted must be C′ rather than IP, since I has raised to C.

(i) Who?
(ii) Who did Mary see?

This is problematic under the fairly standard assumption that rules of grammar do not target intermediate projections. If, on the other hand, I has not raised, the mystery is why the hypothesized source is unacceptable:

(iii) *Who Mary saw.

Further, the phenomenon is independent of *do*-support:

(iv) Mary will see someone.
Who?
(v) *Who Mary will see?

This state of affairs is strikingly reminiscent of what we saw earlier with the hypothesized source of simple Pseudogapping:

(iv) John saw Bill and Mary did/will Susan (*see).

The same solution might also be possible. It is standardly assumed that there is a strong feature forcing the movement of I to C in matrix *wh*-questions. Suppose, contrary to the standard assumption, that the strong feature resides in I rather than in C. (i.e. there is something special about matrix I in an interrogative.) Then, raising will be necessary unless ellipsis eliminates the offending feature. Still unexplained, though, is the fact that while Pseudogapping is marginal sluicing is perfect.

7 See Lobeck (1990) as well.
8 See Park (1995) for an interesting suggestion.
9 See also Lasnik (1993).
10 See Lasnik (1999b) for the outlines of a theory of features and ellipsis with roughly this character.
11 With (50) rejected as a constraint on ellipsis, the Sluicing effect in (49), repeated as (i), remains unexplained.

(i) Speaker A: Mary saw someone.
Speaker B: *Who did ~~Mary see~~?

The proposal of Saito and Murasugi (adopted by Martin as well) about the specific way ellipsis is licensed might be relevant here. They suggest that the licensing head must agree with its specifier. In the Sluicing example in (i), the licensing head is C. Now the content of C is the raised *Infl* (T and Agr), which obviously agrees with the subject, but does not obviously agree with the specifier of CP. I leave for future investigation the task of making this speculation more precise.

12 There is, of course, another unanswered question as well: Why can't the first object in a double object construction undergo HNPS? I will not have anything to say about this here.
13 See Baltin (1982) for important early discussion of such constraints.
14 It is interesting to note that depending on the precise nature of the locality constraint on HNPS, Jayaseelan's analysis might be an option for simple instances of Pseudogapping where the remnant would not have to move far to escape from the elided VP.
15 Though, curiously enough, he claims that Pseudogapping is otherwise impossible.
16 Fiengo and May do not actually discuss Pseudogapping in any detail, but nothing they say is inconsistent with the analysis presented here, as far as I can tell.
17 Pseudogapping is, of course, an available source for the restrictive instances that parallel the appositives. But there must be an additional source as well.
18 Given my argument that at least some instances of VP-ellipsis (those involved in Pseudogapping) are PF deletion, such an LF approach would demand that VP-ellipsis can be deletion *or* copying, a possibly problematic consequence.

19 Hornstein's solution to the problem – that ACD (almost) always involves raising to [Spec, Agr_O] – was rejected above. The additional mechanism Hornstein proposes, base generating most PPs outside of the VP-ellipsis site, is shown by Kennedy (1997) to be unworkable.

20 I am translating slightly, as Baltin's view of the ACD problem is somewhat different from the one adopted here, essentially May's.

21 Larson and May attribute to Stowell the claim that the trace of a deleted complementizer must be properly governed. In fact, Stowell does not claim that deletion leaves a trace at all. Rather, for him, it is a base-generated null complementizer that must be properly governed.

22 Curiously, in ACD constructions, it is the overt operator that is degraded, as illustrated in (i):

(i) Dulles suspected eveyone ??who / that Angleton did. / Ø

23 Larson and May actually used an example involving a direct object:

(i) John saw everyone that you did.

I use a slightly more complicated example to avoid the possibility, discussed extensively above, of raising to [Spec, Agr_O].

24 In the conventional usage of the term, rather than that of Larson and May.

25 I leave open the question of whether the null operator is also a copy of the head. Note, in passing, that the reconstruction effects at issue show up with an *overt* relative operator as well:

(i) Mary mentioned the pictures of himself which Bill saw.

6 ON FEATURE STRENGTH: THREE MINIMALIST APPROACHES TO OVERT MOVEMENT

1 Portions of this material were presented at the 1997 Open Linguistics Forum at the University of Ottawa. I am grateful to the audience there, to the participants in my 1997 seminar at the University of Connecticut, and to Željko Bošković, Bob Freidin, Masao Ochi, and an anonymous *Linguistic Inquiry* reviewer for very helpful suggestions.

2 Page references to Chomsky (1993) will be to pages in Chomsky (1995b), where the paper was reprinted. Page references to Chomsky (1994) will be to pages in Campos and Kempchinsky (1995), one of two books where the paper was published (the other being Webelhuth (1995)). I use the (1993) and (1994) citations for ease of exposition and to keep clear the historical development of the ideas I am exploring.

3 Notice that the relevant strong feature could not reside in the *wh*-phrase, since in multiple interrogation all but one of the *wh*-elements remain in situ.

(i) Who gave what to who?

4 Željko Bošković (personal communication) reminds me that Chomsky's (1993) theory of the organization of the grammar, as actually stated, avoids this problem, since

Chomsky simply stipulates that "[a]fter Spell-Out, the computational process … has no further access to the lexicon…" (Chomsky 1993: 189). In Chomsky (1994) he was explicitly concerned to eliminate that stipulation.

5 Note that it is irrelevant whether the strong feature in C is checked by covert movement, as it will still not be "stripped away."

6 Eventually, I will suggest a reinterpretation of the virus theory under which it, too, will have a significant PF aspect.

7 Actually, I do not believe that Hornstein's approach can be the entire solution. See Lasnik (1993, 1999c), and especially Kennedy (1997) for discussion. But even its partial success suggests that an alternative to the QR analysis might ultimately be possible. Note, by the way, that if covert movement processes affect only formal features, as is sometimes proposed on the basis of Chomsky (1995a), then an LF process such as QR could not possibly newly create a configuration licensing ellipsis, an argument made in Lasnik (1999c).

Two other recent arguments are worthy of comment. First, Kennedy and Merchant (1997), following Haik (1987), note that certain ellipsis constructions show sensitivity to island effects, and they seem to tentatively conclude from that fact that ellipsis involves LF copying. This is somewhat ironic, since Ross (1969) presents a classic argument for deletion, and against copying, based on obedience to island constraints.

Next, Chung *et al.* (1995) argue that Sluicing must be an LF process since it is sensitive to purely semantic properties of the antecedent. However, Romero (1997) shows that the phenomenon discovered by Chung, Ladusaw, and McCloskey depends on focus properties and that it actually shows up in the nonelided counterpart of the example they consider, contrary to their claim.

8 Such parallels had already been noted in Lasnik (1972), where numerous examples were presented, including the following:

(i)(a) John wants to catch a fish.
(b) John wants to catch a fish and so does Bill.
(c) John wants to catch a fish and Bill wants to catch a fish also.

It is a standard observation that (ia) is ambiguous, *a fish* being specific or nonspecific on some accounts. Yet (ib) is just two-ways ambiguous, not four. The interpretive similarity between (ib) and (ic) shows that this missing ambiguity phenomenon is not limited to ellipsis. Chomsky and Lasnik (1993) note that some condition (which they call PARR; in Lasnik (1972) it was called the Parallel Principle) is needed to guarantee the parallelism observed in an example like (ic). They go on to suggest that under a PF account of ellipsis, nothing further need be said about an example like (ib). This constitutes an (admittedly minor) argument for the PF approach.

Bošković (2000a) presents another type of argument that at least one ellipsis process is a PF phenomenon. Bošković provides compelling evidence that one crucial aspect of the distribution of Serbo-Croatian clitics is determined by a phonological constraint. He then shows that constructions that violate the requirement can be rescued by VP-ellipsis. This is abstractly very similar to an argument I will present below.

9 Sag (1976), too, had briefly considered Pseudogapping (not yet known by that title) and had also tentatively suggested that it is VP-ellipsis, while, as Levin later did, acknowledging certain differences.

10 I am somewhat misrepresenting Levin's point here, since her claim is that the ungrammatically of backward Pseudogapping is simply a consequence of the ungrammaticality of Pseudogapping in subordinate clauses. But here, I (and my informants) strongly disagree with the factual claim. (7) seems far worse than its forward counterpart.

(i) ?Mary interviewed Gingrich because John did Ø Clinton.

11 As an anonymous *Linguistic Inquiry* reviewer observes, though, it will not be crucial to the following discussion that YP actually be VP. All that will be crucial is that the remnant move out of some phrase that undergoes subsequent deletion.

12 In that paper I further suggest that the raising is not, in principle, limited to accusative NPs. Rather, all else being equal, complements more generally undergo such raising, including PP complements (an instance I discuss there) and clausal complements.

13 An anonymous reviewer for *Linguistic Inquiry* provides several examples suggesting that sometimes the Pseudogapping remnant must raise very high and that, therefore, sometimes considerably more than just a V must raise. Consider the following instances of Pseudogapping:

(i) While she didn't want to read *War and Peace*, she did *Bleak House*.
(ii) I tried to steal the Rembrandt, but I didn't the Picasso.

The point is that in the nonelided versions (or, for that matter, in the antecedent clauses), seemingly what must be raised is *want to read* or *try to steal*, rather than just a simple verb, since those are the sequences that are missing in the elided versions. Actually, I am not certain just how accessible those readings are, as opposed to ones where just the lower verb is missing. Assuming that they are possible, I will briefly speculate on what the derivations might be. I note first that long A-movement out of control clauses is hardly unprecedented; see, for example, Nemoto (1993) for extensive discussion of such movement in Japanese. So the ellipsis itself is not necessarily problematic. As for the raising in the nonelided clauses, two possibilities come to mind. First, as a result of "restructuring," verb sequences in control structures might be able to behave as if they constituted simple verbs and raise accordingly. Alternatively, suppose that such raising of complex verbs is *not* possible. Then one might assume that deletion is the only option available, once the remnant has raised high. Nonelided versions would then be the result of normal short raising of complement and verb, internal to the control clause. Needless to say, there is much more to be said about this type of optionality. I hope to explore it further in future work.

14 A virtual necessity if an ellipsis site must be a maximal projection.

15 There is another mystery surrounding matrix sluicing. Suppose that I raises to C and that the ellipsis site is IP.
The result should be as in (i), but this is ungrammatical.

(i) Speaker A: Mary will see someone.
Speaker B: *Who will ~~May see?~~

The proposal of Saito and Murasugi (1990) and Lobeck (1990) about the specific way ellipsis is licensed might be relevant here. They suggest that for a head to license ellipsis of its complement, that head must agree with its specifier. In the sluicing example

in (i), the licensing head is C. Now the content of C is the raised I (T and Agr), which obviously agrees with the subject, but does not obviously agree with the specifier of CP. I leave for future investigation the task of making this speculation more precise. Note, by the way, that (ii) is much improved.

(ii) Speaker A: Someone will see Mary.
Speaker B: Who will ~~see Mary?~~

But that is to be expected, since here there is a possible derivation via VP-ellipsis, one that does not involve sluicing at all.

16 The LF crash theory shares this problem.

17 At least for overt movement, though Chomsky does not add this qualification.

18 The situation is actually more complicated, since there are languages, such as French, that have *wh*-movement of the English sort, but only optionally in matrix questions. See Bošković (2000b) for extensive discussion.

19 As Máire Noonan (personal communication) has pointed out, even *overt* insertion of C in the matrix without overt *wh*-movement seems to be incorrectly allowed by Chomsky's formulation.

> [A] strong feature merged at the root must be eliminated before it becomes part of a larger structure by further operations.
>
> (Chomsky 1995a: 234)

Chomsky elaborates this as follows:

> Suppose that the derivation D has formed Σ containing α with a strong feature F. Then … D is canceled if α is in a category not headed by α.
>
> (Chomsky 1995a: 234)

When, as in the example now under consideration, the interrogative will not be embedded, hence will never be part of a larger structure, nothing demands that the strong feature be checked overtly.

20 Here I am somewhat reinterpreting what Chomsky actually said, since prior context indicates that he was referring to the operation of "insertion of strong features." But I do not see how to fit such an operation (insertion of strong features independently of the item of which they are features) into the theory. Possibly I am missing something crucial.

21 Or at least lexical insertion of an item with a strong feature.

22 Ochi's concern is the locality of movement, in particular, the fact that only Relativized Minimality effects follow in any natural way from Attract F. Other island effects seem to make sense only from the point of view of the moving item, rather than the target. Ochi proposes that the feature chain, created by Attract F, is responsible for the Relativized Minimality effects whereas the pied-piping chain, created by Move α in order to remedy the defect in α created by the movement of the formal features out of α, is responsible for other island effects.

23 As Željko Bošković (personal communication) observes, "globality" is thus still present in the pied-piping process, just as it was (implicitly or explicitly) in the PF and LF crash theories.

24 The entire tree is shown in Figure 6.6 just for expository purposes. In the actual derivation, the strong feature of the higher V would attract the corresponding feature of *believe* immediately upon the introduction of the former into the phrase marker, in accord with the virus theory.

7 A GAP IN AN ELLIPSIS PARADIGM: SOME THEORETICAL IMPLICATIONS

1 Portions of this material have been presented in colloquia at Yale University and at the University of Pennsylvania. I am indebted to those audiences for their questions and suggestions.

2 At this point, I use the term "deletion" merely for ease of exposition. Certain of the arguments I present below do, indeed, seem to favor a PF deletion approach to ellipsis over an LF copying one.

3 Roberts and Potsdam both advocate a fully lexicalist approach to morphology. Potsdam, unlike Roberts, does at least mention the apparent advantage of the hybrid account in dealing with the distribution of negation in English. In fact, he claims in his conclusion that "A lexicalist approach to verbal morphology was shown to be better suited to capturing the patterns." However, this is a difficult claim to evaluate, since there is no discussion whatsoever of the facts of negation, and not even a hint of how any lexicalist approach would capture those facts.

4 See Lasnik (1999b) for detailed discussion and comparison of three theories of strong features.

5 The entire tree is shown in (Figure 7.1) just for expository purposes. In the actual derivation, the strong feature of the higher V would attract the corresponding feature of *hire* immediately upon the introduction of the former into the phrase marker, in accord with the demand of Chomsky (1995a) that a strong feature be checked immediately upon its introduction into the structure.

6 Note that the identity constraint still must be limited to heads, given examples like (8)–(10), repeated here as (i)–(iii).

(i) Linguistics, I [like *t*], and you should ~~like linguistics~~ too
(ii) ?Someone will [be *t* in the office]. Yes there will ~~be someone in the office~~.
(iii) That this approach will fail can [be proven *t*]. No it can't ~~be proven that this approach will fail~~.

It is still unclear why the constraint should make this distinction.

7 It is not implausible that there might be some connection between this constraint and the one responsible for the ellipsis paradigm gap. However, it is not entirely clear how the latter follows from the former.

8 Potsdam observes that Otani and Whitman (1991) assert that no such constraint holds for Chinese, Japanese, and Korean. However, he notes that "Hoji (1995) offers a convincing analysis of the Japanese data which does not appeal to ellipsis" (Hoji 1991: 365). In particular, Hoji argues that the relevant constructions in Japanese involve null objects rather than ellipsis. Hoji shows that, contrary to standard assumptions, a sloppy-type reading is available with null objects. This is somewhat ironic (and a potential flaw in Potsdam's reasoning) since Potsdam's claim about Hebrew is based on Doron (1990). But Doron insists that sloppy identity is conclusive evidence for ellipsis. Below, I will examine Doron's argument in more detail.

9 Thanks to Sandra Stjepanović for the data, and for helpful discussion.

10 But see note 8.

11 I have not had the opportunity yet to explore the phenomenon in Irish, the other language Potsdam alludes to. He did not actually provide data from either language, but I had access to Doron (1990), though not to Potsdam's cited source for Irish, McCloskey (in preparation).

12 The symbol "S" is apparently used to stand for "š" in Doron's examples, so I will follow that convention here.
13 I am indebted to Danny Fox and Idan Landau for data and discussion.
14 Whether it holds at all remains to be seen.

8 ON A SCOPE RECONSTRUCTION PARADOX

1 In fact, in the examples originally cited by Zubizarreta contraction is not involved:

(i) Everyone will not come
(ii) Everyone is likely not to come

2 Postal (1974) already claimed that quantifiers that have undergone subject raising have only high scope.
3 Or, alternatively, the specific/non-specific ambiguity of indefinites. These two distinctions overlap to a significant extent.
4 I would like to think that this property would relate to the "pseudoscope" of indefinites enlighteningly discussed by Kratzer (1997), but at the moment I see no obvious way to connect the phenomena.
5 See Lasnik (2000b) for further examination of A-movement reconstruction.

10 CHAINS OF ARGUMENTS

1 Chomsky (1993: 34–5) discusses another apparent difference between Ā-movement and A-movement with respect to reconstruction. He contrasts (i) with (ii), indicating that (i) shows a Condition C effect between *John* and *he* whereas (ii) does not.

(i) which claim that John was asleep was he willing to discuss
(ii) the claim that John was asleep seems to him to be correct

Lebeaux (1988, 1991) also grapples with a contrast like this. Although the contrast could be accommodated by Chomsky's approach, or by the one I will develop below, I am not certain how significant that fact is, since I am not certain how clear the facts are. (i) does seem peculiar to me, but other examples that are syntactically parallel, like (iii), seem fine.

(iii) which piece of evidence that John was asleep was he willing to discuss

In Lasnik (1998), I discuss this issue further.
2 Below, I will consider implications of the factual basis of Belletti and Rizzi's proposal for Chomsky's claim that there is no A-movement reconstruction.
3 To my knowledge, the contrast between examples like (14a) and ones like (14c) was first discussed by Zubizarreta (1982), who attributes the observation to Chomsky. Zubizarreta presents the phenomenon as a possible argument against Quantifier Lowering (QL) in the sense of May (1977). Interestingly, Chomsky (1995a) suggests that these facts might be compatible with QL. Chomsky hints that "'lowering effects' of the kind first discussed by Robert May … could result from adjunction of the matrix quantifier to the lower IP (c-commanding the trace of raising and yielding a well-formed structure if the trace of quantifier lowering is deleted, along the

lines of May's original proposal)" (Chomsky 1995a: 327). I will return to further consideration of QL.

4 Hornstein actually attributes the argument to Chomsky and Lasnik (1993; cited by Hornstein as 1991). But Chomsky and Lasnik do not discuss the phenomenon at all. Perhaps what Hornstein has in mind is Chomsky's (1995a) argument that I summarized above. On the other hand, he reports it as an attempted argument against QL, but, as noted above, Chomsky's (1995a) argument explicitly allows QL.

5 Hornstein attributes these two examples to Chomsky and Lasnik.

6 See Lasnik (1995d) for discussion.

7 Den Dikken (1995) points out a parallel contrast with raising to subject position.

(i) some applicants$_i$ seem to each other$_i$ to be eligible for the job
(ii) *there seem to each other$_i$ to be some applicants$_i$ eligible for the job

Under the general assumptions adopted here, this is particularly revealing, since subject raising to subject of *there* shows the same agreement effects seen in (26).

(iii) there seems/*seem to be a man here
(iv) there seem/*seems to be men here

I take this to indicate covert raising of the formal features of the associate to matrix Agr_S. Symmetrically, I take it that the formal features of the associate raise to Agr_O in the (b) cases of (28)–(30).

8 Bošković (1997a) provides another argument, showing that under very plausible assumptions, sentences like (i) must involve Agr_OP coordination and overt raising, with the matrix subject and verb undergoing across-the-board extraction from matrix Agr_OP.

(i) John believes Peter to be crazy and Mary to be smart

9 I am not sure how the coordination structures mentioned in note 8 fare in this regard. Given Bošković's argument, the prediction is that in an example like (i), narrow scope under negation should not be possible for the ECM subjects.

(i) Mary proved every Mersenne number not to be prime and every Fibonacci number not to be even

Unfortunately, judgments fail at this point.

10 It should be kept in mind that so far there is no clear *explanation* for Chomsky's fact, but there is a strong descriptive generalization. I return to the question of how the generalization is to be explained.

11 The ECM subject in the first conjunct in (45), *every Mersenne number*, also cannot take narrow scope under the negation in its clause. I assume this is a parallelism effect of the sort investigated in Lasnik (1972) and, more recently and more interestingly, in Fox (1995).

12 I also assume that that is the position an object in a simple transitive construction raises to, given Lasnik and Saito's (1991) binding arguments showing that objects (like ECM subjects) can bind into adverbials, and given that an object can be a Pseudogapping remnant.

13 That is, Agr has no Case feature of its own. When T or V along with its Case feature merges with Agr, then a nominal expression that has raised to [Spec, Agr] will be in a Case-checking configuration, even though the raising was driven by the EPP.

14 See Kim (1997) for arguments that the position of a Pseudogapping remnant has a focus feature. Given the optionality of (the target of) raising that I have suggested, a fruitful line of inquiry would center on focus effects with raised ECM subjects in general.

15 There is now the question of why Agr_S is obligatory. This is exactly the question of why the standard EPP holds. At this time, I have no more to contribute to this than anyone else.

16 Interestingly, Postal (1974) claims that a quantifier that has undergone subject raising to subject position invariably takes high scope, that is, that there is no QL. He suggests that the same is true for subject raising to object position (i.e. ECM constructions), but that seems much less clear. First, there are the universal-negative interactions discussed earlier. Second, quantificational subjects do seem to allow scope beneath ECM verbs, as in (i), which, in contrast to (ii), has a pragmatically sensible reading.

(i) the defense attorney proved none of the defendants to be guilty
(ii) none of the defendants were proved to be guilty by the defense attorney

17 A-movement Condition A reconstruction of the sort discussed by, among others, Barss (1986), Belletti and Rizzi (1988), and Lebeaux (1988, 1991) might be treated in a similar on-line fashion, as proposed by Belletti and Rizzi. This makes sense if satisfaction of Condition A involves a formal feature, a not implausible assumption. Such a possibility is hinted at by Chomsky (1995a: 381) (though not for A-movement). I must offer a slight disclaimer, though, as I am no longer entirely confident that psych and raising constructions, as in (i)–(ii), sharply contrast with ones with simple transitive or control predicates, as in (iii)–(iv). The standard judgments are in parentheses.

(i) (✓) each other's supporters frightened the candidates
(ii) (✓) each other's supporters seem to the candidates to be unscrupulous
(iii) (*) each other's supporters attacked the candidates
(iv) (*) each other's supporters asked the candidates to be more honest

18 Here and throughout, I abstract away from scope interactions between subject and object in sentences such as (i).

(i) someone loves everyone

The classic treatment, developed by May (1977), captures the ambiguity via an LF Ā-movement operation, QR: both quantifiers move and either can move to a position above the other. However, Kitahara (1992) and Hornstein (1995) argue that we are instead dealing with an A-movement scope reconstruction effect, with the subject possibly reconstructing to its base position inside the VP and lower than the LF position of the object. Note that on the phrase structure I adopt, of the type proposed by Koizumi, subject is higher than object throughout the derivation, so Kitahara's and Hornstein's account is unavailable.

19 Another tempting alternative is provided by the strictly derivational approaches of Epstein (2000) and Uriagereka (2000). Of necessity, θ-roles would be determined online in such an approach, so no trace would be necessary at some late derived representation. At the moment, however, I see no way under these approaches to draw the desired distinction between θ-role assignment and scope determination.

REFERENCES

Baker, C. L. (1991) "The syntax of English *not*: the limits of core grammar," *Linguistic Inquiry* 22: 387–429.

Baltin, M. R. (1982) "A landing site theory of movement rules," *Linguistic Inquiry* 13: 1–38.

—— (1987) "Do antecedent contained deletions exist?," *Linguistic Inquiry* 18: 579–95.

Barss, A. (1986) "Chains and anaphoric dependence: on reconstruction and its implications," unpublished PhD dissertation, MIT.

Barss, A. and Lasnik, H. (1986) "A note on anaphora and double objects," *Linguistic Inquiry* 17: 347–54.

Belletti, A. (1988) "The Case of unaccusatives," *Linguistic Inquiry* 19: 1–34.

Belletti, A. and Rizzi, L. (1988) "Psych-verbs and theta theory," *Natural Language and Linguistic Theory* 6: 291–352.

Bobaljik, J. (1994) "What does adjacency do? The morphology–syntax connection," in *MIT Working Papers in Linguistics* 22, 1–32. Cambridge, MA: MITWPL, Department of Linguistics and Philosophy, MIT.

Bobaljik, J. and Carnie, A. (1996) "A minimalist approach to some problems of Irish word order," in R. D. Borsley and I. Roberts (eds) *The Syntax of the Celtic Languages: A Comparative Perspective*, Cambridge: Cambridge University Press.

Bobaljik, J. and Jonas, D. (1996) "Subject positions and the roles of TP," *Linguistic Inquiry* 27: 195–236.

Bošković, Ž. (1994a) "D-structure, theta-criterion, and movement into theta-positions," *Linguistic Analysis* 24: 247–86.

—— (1994b) "Participle movement in Serbo-Croatian and related issues," unpublished manuscript, University of Connecticut.

—— (1994c) "*Wager*-class verbs and existential constructions," unpublished manuscript, University of Connecticut.

—— (1995a) "Case properties of clauses and the Greed principle," *Studia Linguistica* 49: 32–53.

—— (1995b) "Participle movement and second position cliticization in Serbo-Croatian," *Lingua* 96: 245–66.

—— (1995c) "Principles of economy in nonfinite complementation," unpublished PhD dissertation, University of Connecticut.

—— (1997a) "Coordination, object shift and V-movement," *Linguistic Inquiry* 28: 357–65.

—— (1997b) "On certain violations of the superiority, Agr_O, and economy of derivation," *Journal of Linguistics* 33: 227–54.

—— (1997c) *The syntax of nonfinite complementation: an economy approach*, Cambridge, MA: MIT Press.

Bošković, Ž. (1998) "Multiple *wh*-fronting and economy of derivation," in *Proceedings of the West Coast Conference on Formal Linguistics* 16, Stanford, CA: Stanford University.

—— (1999) "On multiple feature checking: multiple *Wh*-fronting and multiple head movement," in S. D. Epstein and N. Hornstein (eds) *Working minimalism*, Cambridge, MA: MIT Press.

—— (2000a) "Second position cliticization: syntax and/or phonology," in F. Beukema and M. den Dikken (eds) *Clitics in Europe*, Amsterdam: John Benjamins.

—— (2000b) "Sometimes in SpecCP, sometimes in-situ," in R. Martin, D. Michaels, and J. Uriagereka (eds) *Step by step: essays on minimalism in honor of Howard Lasnik*, Cambridge, MA: MIT Press.

Bošković, Ž. and Takahashi, D. (1998) "Scrambling and last resort," *Linguistic Inquiry* 29: 347–66.

Bouton, L. F. (1970) "Antecedent-contained pro-forms," in *Proceedings of the Sixth Regional Meeting of the Chicago Linguistic Society*, Chicago, IL: Chicago Linguistic Society, University of Chicago.

Branigan, P. (1992) "Subjects and complementizers," unpublished PhD dissertation, MIT.

Bresnan, J. W. (1972) "Theory of complementation in English syntax," unpublished PhD dissertation, MIT.

Campos, H. and Kempchinsky, P. (eds) (1995) *Evolution and revolution in linguistic theory: essays in honor of Carlos Otero*, Washington, DC: Georgetown University Press.

Chomsky, N. (1955) "The logical structure of linguistic theory," unpublished manuscript, Harvard University, published in part by Plenum, New York, 1975, University of Chicago press, Chicago, 1985.

—— (1957) *Syntactic structures*, The Hague: Mouton.

—— (1962) "A transformational approach to syntax," in A. A. Hill (ed.) *Proceedings of the 1958 Conference on Problems of Linguistic Analysis in English*, Austin, TX.

—— (1965) *Aspects of the theory of syntax*, Cambridge, MA: MIT Press.

—— (1973) "Conditions on transformations," in S. Anderson and P. Kiparsky (eds) *A festschrift for Morris Halle*, New York: Holt, Rinehart, and Winston.

—— (1980) "On binding," *Linguistic Inquiry* 11: 1–46.

—— (1981) *Lectures on government and binding*, Dordrecht: Foris.

—— (1986a) *Barriers*, Cambridge, MA: MIT Press.

—— (1986b) *Knowledge of language*, New York: Praeger.

—— (1987) *Generative grammar: its basis, development and prospects*, Kyoto: Kyoto University of Foreign Studies.

—— (1991) "Some notes on economy of derivation and representation," in R. Freidin (ed.) *Principles and parameters in comparative grammar*, Cambridge, MA: MIT Press, reprinted in Chomsky (1995b).

—— (1993) "A minimalist program for linguistic theory," in K. Hale and S. J. Keyser (eds) *The view from building 20: essays in linguistics in honor of Sylvain Bromberger*, Cambridge, MA: MIT Press, reprinted in Chomsky (1995b).

—— (1994) "Bare phrase structure," *MIT Occasional Papers in Linguistics* 5, Cambridge, MA: MIT, also in Webelhuth (1995) and Campos and Kempchinsky (1995).

—— (1995a) "Categories and transformations," in Chomsky (1995b).

—— (1995b) *The minimalist program*, Cambridge, MA: MIT Press.

—— (2001) "Beyond explanatory adequacy," *MIT Occasional Papers in Linguistics* 20.

Chomsky, N. and Lasnik, H. (1977) "Filters and control," *Linguistic Inquiry*, 8: 425–504.

Chomsky, N. and Lasnik, H. (1993) "The theory of principles and parameters," in J. Jacobs, A. v. Stechow, W. Sternefeld, and T. Vennemann (eds) *Syntax: an international handbook of contemporary research*, Vol. 1, Berlin: Walter de Gruyter, reprinted in Chomsky (1995b).

Chung, S., Ladusaw, W., and McCloskey, J. (1995) "Sluicing and logical form," *Natural Language Semantics* 3: 1–44.

Déprez, V. (1989) "On the typology of syntactic positions and the nature of chains," unpublished PhD dissertation, MIT.

Diesing, M. (1996) "Semantic variables and object shift," in H. Thráinsson, S. D. Epstein, and S. Peter (eds) *Studies in comparative Germanic syntax*, Vol. II, Dordrecht: Kluwer.

Dikken, M. den (1995) "Binding, expletives, and levels," *Linguistic Inquiry* 26: 347–54.

Doron, E. (1990) "V-movement and VP-ellipsis," unpublished manuscript, Hebrew University of Jerusalem.

Emonds, J. (1978) "The verbal complex V′–V′ in French," *Linguistic Inquiry* 9: 151–75.

Epstein, S. (1990) "Differentiation and reduction in syntactic theory: a case study," *Natural Language and Linguistic Theory* 8: 313–23.

—— (1998) "Overt scope marking and covert verb-second," *Linguistic Inquiry* 29: 181–227.

—— (2000) "Un-principled syntax and the derivation of syntactic relations," in S. D. Epstein and N. Hornstein (eds) *Working minimalism*, Cambridge, MA: MIT Press.

Evans, G. (1980) "Pronouns," *Linguistic Inquiry* 11: 337–62.

Fiengo, R. and May, R. (1992) "Ellipsis and apposition," unpublished manuscript, CUNY and UC Irvine.

—— (1994) *Indices and identity*, Cambridge, MA: MIT Press.

Fox, D. (1995) "Economy and scope," *Natural Language Semantics* 3: 283–341.

Freidin, R. (1986) "Fundamental issues in the theory of binding," in B. Lust (ed.) *Studies in the acquisition of anaphora*, Vol. 1, Dordrecht: Reidel.

Green, L. (1993) "Topics in African American English: the verb system analysis," unpublished PhD dissertation, University of Massachusetts.

Groat, E. (1993) "English expletives: a minimalist approach," in *Harvard working papers in linguistics* 3, 81–8. Cambridge, MA: Department of Linguistics, Harvard University.

—— (1995) "English expletives: a minimalist approach," *Linguistic Inquiry* 26: 354–65.

Guilfoyle, E. (1993) "Verbal nouns and non-finite clauses in Modern Irish and Old English." *Paper presented at the Canadian Linguistic Association Annual Meeting*, Carleton University, Ottawa, Ont.

Haïk, I. (1987) "Bound VPs that need to be," *Linguistics and Philosophy* 10: 503–30.

Halle, M. and Marantz, A. (1993) "Distributed morphology and the pieces of inflection," in K. Hale and S. J. Keyser (eds) *The view from building 20: essays in linguistics in honor of Sylvain Bromberger*, Cambridge, MA: MIT Press.

Halliday, M. A. K. and Hasan, R. (1973) *Cohesion in spoken and written English*, London: Longmans.

Hoji, H. (1995) "Null object and sloppy identity in Japanese," unpublished manuscript, University of Southern California.

Hornstein, N. (1994) "An argument for minimalism: the case of antecedent contained deletion," *Linguistic Inquiry* 25: 455–80.

—— (1995) *Logical form: from GB to minimalism*, Cambridge, MA: Blackwell.

—— (1999) "Movement and control," *Linguistic Inquiry* 30: 69–96.

Hornstein, N. and Weinberg, A. (1990) "The necessity of LF," *The Linguistic Review* 7: 129–67.

Huang, C.-T. J. (1981/82) "Move *wh* in a language without *wh*-movement," *The Linguistic Review* 1: 369–416.

Jayaseelan, K. A. (1990) "Incomplete VP deletion and gapping," *Linguistic Analysis* 20: 64–81.

Johnson, K. (1991) "Object positions," *Natural Language and Linguistic Theory* 9: 577–636.

Kayne, R. (1985) "Principles of particle constructions," in J. Guéron, H.-G. Obenauer, and J.-Y. Pollock (eds) *Grammatical representation*, Dordrecht: Foris.

—— (1994) *The antisymmetry of syntax*, Cambridge, MA: MIT Press.

Kennedy, C. (1997) "Antecedent-contained deletion and the syntax of quantification," *Linguistic Inquiry* 28: 662–88.

Kennedy, C. and Merchant, J. (1997) "Attributive comparatives and bound ellipsis," unpublished manuscript, Linguistics Research Center, University of California at Santa Cruz.

Kim, J.-S. (1997) "Syntactic focus movement and ellipsis: a minimalist approach," unpublished PhD dissertation, University of Connecticut.

Kitahara, H. (1992) "Checking theory and scope interpretation without Quantifier Raising," in *Harvard Working Papers in Linguistics* 1, 51–71. Cambridge, MA: Department of Linguistics, Harvard University.

Koizumi, M. (1993) "Object agreement phrases and the split VP hypothesis," in *Papers on Case and Agreement I: MIT Working Papers in Linguistics* 18, J. D. Bobaljik and C. Phillips (eds) 99–148. Cambridge, MA: MITWPL, Department of Linguistics and Philosophy, MIT.

—— (1995) "Phrase structure in minimalist syntax," unpublished PhD dissertation, MIT.

Kratzer, A. (1997) "Scope or pseudoscope? Are there wide-scope indefinites?," in S. Rothstein (ed.) *Events and grammar*, Dordrecht: Kluwer.

Kuno, S. (1997) "Binding theory in the minimalist program," unpublished manuscript, Harvard University.

Labov, W. (1969) "Contraction, deletion, and inherent variability of the English copula," *Language* 45: 715–76.

—— (1972) *Language in the Inner City*, Philadelphia: University of Pennsylvania Press.

Langacker, R. W. (1969) "On pronominalization and the chain of command," in D. A. Reibel and S. A. Schane (eds) *Modern Studies in English*, Englewood Cliffs, NJ: Prentice-Hall.

Lappin, S. (1992) "The syntactic basis of ellipsis resolution," in S. Berman and Hestvik (eds) *Proceedings of the Stuttgart Ellipsis Workshop, Arbeitspapiere des Sonderforschungsbereichs* 340, Stuttgart.

Larson, R. K. (1988) "On the double object construction," *Linguistic Inquiry* 19: 335–91.

Larson, R. and May, R. (1990) "Antecedent containment or vacuous movement: reply to Baltin," *Linguistic Inquiry* 21: 103–22.

Lasnik, H. (1972) "Analyses of negation in English," unpublished PhD dissertation, MIT.

—— (1976) "Remarks on coreference," *Linguistic Analysis* 2: 1–22, reprinted in Lasnik (1989).

—— (1981) "Restricting the theory of transformations: a case study," in N. Hornstein and D. Lightfoot (eds) *Explanations in linguistics*, London: Longmans, reprinted in Lasnik (1990).

—— (1985) "Illicit NP movement: locality conditions on chains?," *Linguistic Inquiry* 16: 481–90, reprinted in Lasnik (1989).

—— (1989) *Essays on anaphora*, Dordrecht: Kluwer.

—— (1990) *Essays on restrictiveness and learnability*, Dordrecht: Kluwer.

—— (1992a) "Case and expletives: notes toward a parametric account," *Linguistic Inquiry* 23: 381–405.

Lasnik, H. (1992b) "Two notes on control and binding," in R. K. Larson, S. Iatridou, U. Lahiri, and J. Higginbotham (eds) *Control and grammar*, Dordrecht: Kluwer.

—— (1993) "Lectures on minimalist syntax," *University of Connecticut Occasional Papers in Linguistics* 1, University of Connecticut, Storrs, reprinted in Lasnik (1999a).

—— (1995a) "Case and expletives revisited: on Greed and other human failings," *Linguistic Inquiry* 26: 615–33, reprinted in Lasnik (1999a).

—— (1995b) "Last resort," in S. Haraguchi and M. Funaki (eds) *Minimalism and linguistic theory: Proceedings of the First Numazu Conference on Formal Linguistics*, Tokyo: Hituzi Syobo, reprinted in Lasnik (1999a).

—— (1995c) "Last resort and attract F," in L. Gabriele, D. Hardison, and R. Westmoreland (eds) *Proceedings of the Sixth Annual Meeting of the Formal Linguistics Society of Mid-America*, Bloomington, IN: Indiana University, reprinted in this volume as Chapter 3.

—— (1995d) "A note on pseudogapping," in *Papers on minimalist syntax, MIT working papers in linguistics* 27, R. Pensalfini and H. Ura (eds), 143–163. Cambridge, MA: MITWPL, Department of Linguistics and Philosophy, MIT, reprinted in Lasnik (1999a).

—— (1995e) "Verbal morphology: *Syntactic Structures* meets the minimalist program," in H. Campos and P. Kempchinsky (eds) *Evolution and revolution in linguistic theory: essays in honor of Carlos Otero*, Washington, DC: Georgetown University Press, reprinted in Lasnik (1999a).

—— (1997) "Levels of representation and the elements of anaphora," in H. Bennis, J. Rooryck, and P. Pica (eds) *Atomism and binding*, Dordrecht: Foris, reprinted in this volume as Chapter 4.

—— (1998) "Some reconstruction riddles," in *University of Pennsylvania working papers in linguistics* 5.1, A. Dimitriadis, H. Lee, C. Moisset, and A. Williams (eds), 83–98. Philadelphia, PA: Penn Linguistics Club, University of Pennsylvania, reprinted in this volume as Chapter 9.

—— (1999a) *Minimalist analysis*, Malden, MA: Blackwell.

—— (1999b) "On feature strength: three minimalist approaches to overt movement," *Linguistic Inquiry* 30: 197–217, reprinted in this volume as Chapter 6.

—— (1999c) "Pseudogapping puzzles," in E. Benmamoun and S. Lappin (eds) *Fragments: studies in ellipsis*, Oxford: Oxford University Press, reprinted in this volume as Chapter 5.

—— (2000a) *Syntactic Structures revisited: contemporary lectures on classic transformational theory*. Cambridge, MA: MIT Press.

—— (2000b) "Chains of arguments," in S. D. Epstein and N. Hornstein (eds) *Working minimalism*, Cambridge, MA: MIT Press, reprinted in this volume as Chapter 10.

—— (2001) "Subjects, objects, and the EPP," in W. Davies and S. Dubinsky (eds) *Objects and other subjects: grammatical functions, functional categories, and configurationality*, Dordrecht: Kluwer Academic Publishers.

—— (2002) "Clause-mate conditions revisited," *Glot International* 6: 94–6.

—— (In press) "Grammar, levels, and biology," in J. McGilvray (ed.) *The Cambridge Companion to Chomsky*, Cambridge: Cambridge University Press.

Lasnik, H. and Freidin, R. (1981) "Core grammar, Case theory, and markedness," in A. Belletti, L. Brandi, and L. Rizzi (eds) *Theory of Markedness in Generative Grammar – Proceedings of the 1979 GLOW Conference* 407–421. Scuola Normale Superiore, Pisa, reprinted in Lasnik (1990).

Lasnik, H. and Saito, M. (1984) "On the nature of proper government," *Linguistic Inquiry* 15: 235–89, reprinted in Lasnik (1990).

Lasnik, H. and Saito, M. (1991) "On the subject of infinitives," in L. M. Dobrin, L. Nichols, and R. M. Rodriguez (eds) *Papers from the 27th Regional Meeting of the Chicago Linguistic Society Part One: The General Session*, Chicago, IL: Chicago Linguistic Society, University of Chicago, reprinted in Lasnik (1999a).

—— (1992) *Move α: conditions on its application and output*, Cambridge, MA: MIT Press.

Lebeaux, D. (1988) "Language acquisition and the form of the grammar," unpublished PhD dissertation, University of Massachusetts.

—— (1991) "Relative clauses, licensing and the nature of the derivation," in S. Rothstein (ed.) *Syntax and semantics 25: perspectives on phrase structure*, New York: Academic Press.

Lee, R. K. (1993) "Constraints on A-movement, negative polarity items licensing, and the checking theory." *Paper Presented at the 1993 Seoul International Conference on Generative Grammar*.

—— (1994) "Economy of representation," unpublished PhD dissertation, University of Connecticut.

Levin, N. (1978) "Some identity-of-sense deletions puzzle me. Do they you," in *Proceedings of the Fourteenth Annual Meeting of the Chicago Linguistic Society*, Chicago, IL: Chicago Linguistic Society, Chicago Univerisity.

—— (1979/1986) "Main verb ellipsis in spoken English," unpublished PhD dissertation, Ohio State University. (Published 1986 by Garland, New York.)

Lobeck, A. (1990) "Functional heads as proper governors," in *Proceedings of North Eastern Linguistic Society* 20, Amherst, MA: GLSA, University of Massachusetts.

Martin, R. (1992a) "Case theory, A-chains, and expletive replacement," unpublished manuscript, University of Connecticut.

—— (1992b) "On the distribution and Case features of PRO," unpublished manuscript, University of Connecticut.

—— (1996a) "A minimalist theory of PRO and control," unpublished PhD dissertation, University of Connecticut.

—— (1996b) "On LF *wh*-movement and *wh*-islands," in *University of Connecticut working papers in linguistics* 5, E. Laurençot, R. Lee, and M.-K. Park (eds), 25–44. Storrs: University of Connecticut.

Martins, A. M. (1994) "Enclisis, VP-deletion and the nature of Sigma," *Probus* 6: 173–205.

May, R. (1977) "The grammar of quantification," unpublished PhD dissertation, MIT.

—— (1985) *Logical form: its structure and derivation*, Cambridge, MA: MIT Press.

McCloskey, J. (1990) *Clause structure, ellipsis and proper government in Irish*, Santa Cruz, CA: Syntax research center, Cowell College, University of California at Santa Cruz.

—— (in preparation) "Irish syntax: clauses and clause types."

Munn, A. (1994) "A minimalist account of reconstruction asymmetries," in *Proceedings of the North Eastern Linguistic Society* 24. Amherst, MA: GLSA, University of Massachusetts.

Nemoto, N. (1993) "Chains and case positions: a study from scrambling in Japanese," unpublished PhD dissertation, University of Connecticut.

Ochi, M. (1997) "Move or Attract?: attract F and the pied-piping chain." *Paper presented at the Open Linguistics Forum*, Ottawa.

—— (1998) "Move or Attract?," in *Proceedings of the West Coast Conference on Formal Linguistics* 16, Stanford, CA: CSLI, Stanford University.

Oka, T. (1993) "Minimalism in syntactic description," unpublished PhD dissertation, MIT.
Oku, S. (1996) "VP ellipsis and 'VP adverb' placement," unpublished manuscript, University of Connecticut.
Oku, S. (2001) "A minimalist theory of LF copy," in G. M. Alexandrova and O. Arnandova (eds) *The minimalist parameter: selected papers deriving from the Open Linguistics Forum "Challenges of Minimalism,"* Ottawa, 21–23 March 1997, Amsterdam: John Benjamins.
Ormazabal, J. (1995) "The syntax of complementation: on the connection between syntactic structure and selection," unpublished PhD dissertation, University of Connecticut.
Otani, K. and Whitman, J. (1991) "V-raising and VP-ellipsis," *Linguistic Inquiry* 22: 345–58.
Park, M.-K. (1995) "A morpho-syntactic study of Korean verbal inflection," unpublished PhD dissertation, University of Connecticut.
Pollock, J.-Y. (1989) "Verb movement, universal grammar, and the structure of IP," *Linguistic Inquiry* 20: 365–424.
Postal, P. (1972). "The best theory," in S. Peters (ed.) *Goals of linguistic theory*, Englewood Cliffs, NJ: Prentice-Hall.
—— (1974) *On raising: one rule of English grammar and its theoretical implications*, Cambridge, MA: MIT Press.
Postal, P. (1997) "Strong crossover violations and binding principles." Paper presented at the Eastern States Conference on Linguistics, Yale University.
Potsdam, E. (1997) "English verbal morphology and VP ellipsis," in *Proceedings of the North Eastern Linguistic Society* 27, Amherst, MA: GLSA, University of Masachusetts.
Quirk, R., Greenbaum, S., Leech, G., and Svartvik, J. (1972) *A grammar of contemporary English*, London: Seminar Press.
Riemsdijk, H. van (1978) *A case study in syntactic markedness*, Dordrecht: Foris.
Rizzi, L. (1990) *Relativized minimality*, Cambridge, MA: MIT Press.
Roberts, I. (1993) *Verbs and diachronic syntax*, Dordrecht: Kluwer.
—— (1994) "Two types of head movement in Romance," in D. Lightfoot and N. Hornstein (eds) *Verb movement*, Cambridge: Cambridge University Press.
—— (1998) "*Have/Be* raising, Move F, and Procrastinate," *Linguistic Inquiry* 29: 113–25.
Romero, M. (1997) "Recoverability conditions for sluicing," in F. Corblin *et al.* (ed.) *Empirical issues in formal syntax and semantics*, Bern: Peter Lang.
Rosenbaum, P. S. (1967) *The grammar of English predicate complement constructions*. Cambridge, MA: MIT Press.
Ross, J. R. (1969) "Guess who?," in R. I. Binnick, A. Davison, G. M. Green, and J. L. Morgan (eds) *Papers from the Fifth Regional Meeting of the Chicago Linguistic Society*, Chicago, IL: Chicago Linguistic Society, University of Chicago.
Rudin, C. (1982) "Who what to whom said: an argument from Bulgarian against successive cyclicity," in *Papers from the Seventeenth Regional Meeting of the Chicago Linguistic Society*, Chicago, IL: Chicago Linguistic Society, University of Chicago.
—— (1988) "On multiple questions and multiple *wh*-fronting," *Natural Language and Linguistic Theory* 6: 445–501.
Sag, I. (1976) "Deletion and logical form," unpublished PhD dissertation, MIT.
Saito, M. and Murasugi, K. (1990) "N′-deletion in Japanese," in *University of Connecticut Working Papers in Linguistics* 3, J. Ormazabal and C. Tenny (eds), 87–107. Storrs: University of Connecticut.

Saito, M. and Murasugi, K. (1993) "Subject predication within IP and DP," unpublished manuscript, University of Connecticut.

Sobin, N. (1997) "Agreement, default rules, and grammatical viruses," *Linguistic Inquiry* 28: 318–43.

Stowell, T. (1981) "Origins of phrase structure," unpublished PhD dissertation, MIT.

Stump, G. (1977) "Pseudogapping," unpublished manuscript, Ohio State University.

Takahashi, D. (1994) "Minimality of movement," unpublished PhD dissertation, University of Connecticut.

—— (1996) "On antecedent contained deletion," in *University of Connecticut working papers in linguistics* 5, E. Laurençot, R. Lee, and M.-K. Park (eds), 65–80. Storrs: University of Connecticut.

Tancredi, C. (1992) "Deletion, deaccenting, and presupposition," unpublished PhD dissertation, MIT.

Toman, J. (1982) "Aspects of multiple *wh*-movement in Polish and Czech," in R. May and J. Koster (eds) *Levels of syntactic representation*, Dordrecht: Foris.

Ura, H. (1993) "On feature-checking for *wh*-traces," in *Papers on Case and Agreement I – MIT Working Papers in Linguistics* 18, J. D. Bobaljik and C. Phillips (eds), 215–42. Cambridge, MA: MITWPL, Department of Linguistics and Philosophy, MIT.

Uriagereka, J. (1988) "On government," unpublished PhD dissertation, University of Connecticut.

—— (2000) "Multiple Spell-Out," in S. D. Epstein and N. Hornstein (eds) *Working minimalism*, Cambridge, MA: MIT Press.

Vergnaud, J.-R. (1977) Personal letter to Howard Lasnik and Noam Chomsky, Paris.

Warner, A. (1986) "Ellipsis conditions and the status of the English copula," in *York Papers in Linguistics* 12, 153–72. Heslington, England: University of York.

Wasow, T. (1972) "Anaphoric relations in English," unpublished PhD dissertation, MIT.

—— (1979) *Anaphora in generative grammar*, Ghent: E. Story-Scientia.

Watanabe, A. (1993) "AGR-based Case theory and its interaction with the A′-system," unpublished PhD dissertation, MIT.

Webelhuth, G. (ed.) (1995) *Government and binding theory and the Minimalist program*, Oxford: Blackwell.

Wexler, K. (1994) "Optional infinitives, head movement, and the economy of derivations," in D. Lightfoot and N. Hornstein (eds) *Verb movement*, Cambridge: Cambridge University Press.

Wilder, C. and Čavar, D. (1994) "Word order variation, verb movement, and economy principles," *Studia Linguistica* 48: 46–86.

Wyngaerd, G. V. (1989) "Object shift as an A-movement rule," in *MIT Working Papers in Linguistics* 11, 256–71. Cambridge, MA: MITWPL, Department of Linguistics and Philosophy, MIT.

Wyngaerd, G. V. and Zwart, J. W. (1991) "Reconstruction and vehicle change," in F. Drijkoningen and A. v. Kemenade (eds) *Linguistics in the Netherlands 1991*, Amsterdam: John Benjamins.

Zagona, K. (1982) "Government and proper government of verbal projections," unpublished PhD dissertation, University of Washington.

—— (1988) "Proper government of antecedentless VP in English and Spanish," *Natural Language and Linguistic Theory* 6: 95–128.

Zubizarretta, M.-L. (1982) "On the relationship of the lexicon to syntax," unpublished PhD dissertation, MIT.

Zwart, C. J.-W. (1993) "Dutch syntax: a minimalist approach," unpublished PhD dissertation, University of Groningen.

—— (1997) *Morphosyntax of verb movement. A minimalist approach to the syntax of Dutch*, Dordrecht: Kluwer.

INDEX

A/A′-position 10, 15
A-movement reconstruction 124
A-position 10, 15, 35
adjacency 13, 15–16, 158 n2; PF- 16
adjunct 16, 127
adjunct-argument reconstruction 129, 132
adjunction 32, 46, 48; expletive 47
adverb 15; and Heavy NP Shift 76
affix 2, 7, 10, 13, 30, 103, 159 n10, 162 n20; lowering 11
affixation 27
Affix Hopping 7–8, 10, 13, 16, 18, 104, 158 n2
African-American English 20; *see also* *be*, Habitual
Agr 8–9, 28–31, 37–8, 71
agreement 2, 26–7, 46, 160 n5, 160 n6, 164 n5
anaphor 4, 34, 42, 44–5, 47, 52; LF movement of 35, 50–1
antecedent (for VP ellipsis) 17, 104–6
antecedent-contained deletion (ACD) 37–40, 76, 78, 87, 163 n22, 163 n25; appositive relative 38, 78–9, 163 n25; restrictive relative 38, 78–9, 81, 163 n25
associate (of *there*) 26–8, 30–1, 34–5, 37, 47, 50–3, 162 n21
attract, Attract F 30–1, 73, 98, 171 n22
auxiliary 2, 12, 18

Baker, M.: *do*-support 158 n1
Baltin, M.: ACD, extraposition 79–80, 82, 167 n13
bare form (verb) 16, 18, 104, 159 n16
bare phrase structure 138, 157
bare stem 103
Barss, A.: double object construction 166 n4; reconstruction 43, 175 n17
be (Standard American English) 7, 11, 13–14, 17–19, 113, 159 n14, 159 n16
be/unaccusative 27, 31, 161 n14
Belletti, A.: A-movement reconstruction 136–7, 143, 173 n2, 175 n17; partitive Case 27, 31, 37, 161 n14
Binding 35–6, 42–3, 50, 52, 81, 94, 133, 142, 146, 162 n21, 164 n2; backward 143; of pronoun 123; reciprocal 94
Bobaljik, J.: adjacency 13, 16, 159 n9; *do*-support 158 n1; Holmberg's Generalization 163 n26, 164 n26
Boeckx, C.; degree of grammaticality 166 n5
bound morpheme 6
Bouton, L. F.: ACD 56, 76
Bošković, Ž.: ACD 163 n22; auxiliary 158 n7: clitics 169 n8; coordination 174 n8, 174 n9; degree of grammaticality 166 n5; head 158 n5; *HIT* 24–5; Holmberg's Generalization 164 n26; multiple *wh*-movement, strong feature 100–1; pied-piping 171 n23; strong feature 168 n4, 171 n18; *there* 162 n20; theta theory 137, 157, 160 n3
Bresnan, J.: ECM 145

c-command 32, 35, 44, 47, 50, 73
Carnie, A.: Holmberg's Generalization 163 n26
Case 2, 27; accusative 4, 37–8, 51; checking; checking view of 158 n6; Exceptionally Case Marking 4, 29; Filter 3; licensing head 23; marked-position 23; nominative 23; partitive 27–8, 36; of trace of *wh*-movement 3
category 30, 36, 48

Ćavar, D.: verb raising in Swedish 158 n7
chain 141; A-chain 37; head 111; operator 127; uniformity 10, 140, 141
checking 10; accusative Case 60; configuration 31
checking domain 27
Chinese 3, 172 n8
Chomsky, N.: affixation 27; anaphor 49–51, 165 n21; binding 35–6, 42–5, 79, 143, 150–1, 164 n2; Case 2–3, 10, 158 n6; control 52; copy theory 39, 81, 132, 138; cyclicity 95; driving force for movement 22, 28; economy 10; ellipsis 60, 145; existential/expletive construction 26, 30, 37, 44–8, 53, 162 n21; explanatory adequacy 3; Greed 23–4, 29; *HIT* 25; inclusiveness condition 5; last resort 23; levels of representation 4–5; lexical insertion 98; lexicalism 10; pied-piping 84, 111; Procrastinate 83; reconstruction 118–22, 126–30, 133–4, 137, 141–2, 148, 154–5, 157, 173 n1, 174 n4, 175 n17; S-structure 45, 79; scope 32, 144, 153; strong feature 61, 64, 84–6, 95–6, 99, 101, 109, 164 n28; verbal morphology 6–8, 10, 18, 104, 158 n2, 158 n4; virus theory 96; *wh*-movement 3
Chung, S.: sluicing 169 n7
clause-mate (condition) 4–5
complement of *to* 160 n9
Complementizer 28–9, 64, 69, 80; interrogative 85–6, 97
Condition A 35, 43, 50, 136–7, 143, 145; *see also* Binding
Condition B 142, 143; *see also* Binding
Condition C 43, 79, 126–32, 150, 173 n1; *see also* Binding
contraction 144
contrastive focus 88
control 52–3, 162 n21; object 52
convergence 12
coordination 27, 174 n8, 174 n9
copy 127, 137, 141, 155, 157
covert syntax/component 60, 98
crash 23; LF 11, 14, 64–5, 85, 91, 95; PF 11, 14, 40, 61–7, 83–4, 91, 93, 95, 111–12; *see also* Feature (strong)
crossover: strong 43; weak 34, 45, 47, 49, 123, 136, 145, 151, 155

D-structure component 157
deep structure 4
degree of grammaticality 166 n5
deletion 5, 16, 18, 40; PF 38–9, 57, 60, 64, 89, 92–3, 111, 159 n13, 163 n24; VP 56, 60, 65, 163 n24
deletion under identity 18; trace 9, 140–1
Deprez, V.: Holmberg's Generalization 163 n26
derivation 9
Diesing, M.: object shift 150
Dikken, M. den: binding 174 n7; overt object raising 94
do(-support) 7, 19, 158 n1, 159 n16
Doron, E.: VP ellipsis 69, 105, 114
double object construction 57, 62, 66, 78, 81, 87
driving force for movement 22, 27–8, 40
Dutch 164 n26, 165 n13

e-type pronoun 115
economy 10, 25, 30, 60, 89, 96, 140–1, 159 n10, 159 n16
ellipsis IP- 92, 108; and LF/PF 87, 89, 93, 159 n13, 163 n24; VP- 16, 18, 55–7, 69–70, 72, 78, 88, 104, 147
embedded question 92
Emonds, J.: verb raising 8
Empty Category Principle (ECP) 3, 9, 10, 14, 70, 139, 158 n4; violation 140–1; *see also* Chain Uniformity
enlightened self-interest 28–31, 159 n10; *see also* Greed
Epstein, S.: affix lowering 159 n8; degree of grammaticality 166 n5; theta role 176 n19
Evans, G.: e-type pronoun 115
Exceptionally Case Marking (ECM) 4, 29, 35, 37, 50, 57, 145–6, 148
existential construction 27, 31, 48, 51, 146
explanatory adequacy 3, 5, 7
expletive 27, 30, 34, 37, 146, 162 n21; replacement 37, 45, 47
Extended Projection Principle (EPP) 23–4, 31, 40, 60, 89, 151–2, 175 n15
Extension Condition 127
extraposition 79, 82

feature: agreement 27, 30–1, 36, 38, 160 n11; Case 23–4, 31, 36, 38, 160 n11; feature strength 13, 98, 164 n28; focus 101; formal 30–1, 35, 37, 49, 50–1, 60, 89, 94, 137; inflectional feature 13; Q-feature 28–9; strong

feature 11, 29, 40, 60–1, 64, 66, 68, 83, 85, 89, 91–3, 96, 107, 167 n6
Fiengo, R.: ACD 38, 56, 79, 163 n25; control 165 n17
finite complement 35
Finnish 3
formal requirement 28
Fox, D.: parallelism 174 n11; QR 79; VP-ellipsis 115, 173 n13
Frampton, J.: existential construction 161 n11
Freidin, R.: Case theory 3; complement/adjunct asymmetries 126, 129, 131
French 8, 11–12, 14–15, 53, 163 n21, 165 n19, 171 n18
Full Interpretation 9

gamma-marking 9–10, 12; *see also* Empty Category Principle
Gapping 56
generalized transformation 126–7
German 164 n26, 165 n13
globality 25, 171 n23
Governing Category 142
government 9, 27; antecedent 158 n4; and VP ellipsis 70
Greed 23–31, 159 n10
Green, L.: African-American English verbal morphology 6, 20
Groat, E.: expletive 28
Guilfoyle, E.: Holmberg's Generalization 163 n26

habitual (Hab) 6, 7, 20
habitual *be* (African-American English) 20
Halle, M. 2; verbal morphology 13, 159 n9
Halliday, M. A. K.: Pseudogapping 56
Hassan, R.: Pseudogapping 56
have 7, 11, 13–14, 17–18, 159 n14
head 16; adjunction 164 n5; functional 10, 30; head–head configuration 31, 164 n5; of elided VP 69, 73; licensing (in VP ellipsis) 71–3
Head Movement Constraint (HMC) 9
Heavy NP Shift (HNPS) 56–7, 59, 73, 76, 82, 88, 107
Hebrew 69, 73, 105, 114, 172 n8
HIT 25, 30
Hoji, H.: VP-ellipsis 172 n8
Holmberg's Generalization 163 n26
Hornstein, N.: ACD 37–8, 78, 87, 163 n25, 168 n19, 169 n7; LF-operations 45, 164 n4; quantifier scope 118–20, 144, 153–4, 174 n4, 174 n5, 175 n18; S-structure 79; theta theory 157
Huang, C.-T. J. 3

identity 18, 104, 172 n6
il 53, 162 n21, 165 n18
imperative (Imp) 18–19, 159 n15
inclusiveness 5
incorporation 161 n12
indefinite 122–3, 135, 154–5
infinitive 159 n16
Infl 7–8, 16; affixal 13; featural Infl 13–14; and Sluicing 60; and VP ellipsis 70
interrogation 1; matrix 92; multiple 28, 168 n3
inversion 159 n16
Irish 69, 73, 105, 114, 163 n26, 172 n11
island 139, 169 n7
island violation repair 4, 139
it 25
Italian 53

Japanese 3, 170 n13, 172 n8
Jayaseelan, K. A.: Pseudogapping, Heavy NP Shift 55, 57, 73, 88, 147
Johnson, K.: overt object raising 37, 51, 60, 89, 108, 147, 150
Jonas, D.: Holmberg's Generalization 163 n26

Kayne, R.: *make* … *out* construction 121, 148; position of adjunct 164 n7
Kennedy, C.: ACD, ellipsis 168 n19, 169 n7
Kim, J.-S.: Pseudogapping 175 n14
Kitahara, H.: scope 175 n18, 161 n16
Koizumi, M.: overt object raising, "split VP" hypothesis 37, 51, 60, 89, 108, 147, 151; raising of V 164 n27
Korean 172 n8
Kuno, S.: reconstruction 129

Labov, W.: African-American English 6
Ladusaw, W.: Sluicing 169 n7
Landau, I.: VP-ellipsis 115, 173 n13
Langacker, R.: pronominalization 150
Lappin, S.: ACD 56, 76, 163 n25
Larson, R.: ACD 79–81; adjunct 47; double object construction 166 n4
Lasnik, H. 2; ACD 37–8, 78, 169 n7; binding 44–5, 52, 143, 150, 160 n4; control 52; disjoint reference 165 n21;

Lasnik, H. (*Continued*)
double object construction 166 n4; ellipsis 87, 104, 145, 163 n24; Empty Category Principle 3, 9; Exceptional Case Marking (ECM) 4–5, 34, 47, 50–1, 145–6, 150; existential constructions 46; feature strength 164 n28; intermediate trace 140; island 139; LF-operations 45; NP-raising to [Spec, Agr_O] 94, 108, 174 n6, 174 n12; parallelism 174 n11; partitive Case 27, 31; Pseudogapping 55, 88, 107–8; reconstruction 129–30, 174 n4; S-structure 79; subject orientation 165 n14; theta theory 137, 157; verbal morphology 3, 7, 13, 103; zero morpheme 2
last resort 28, 158 n1, 160 n2
Latin 3
Lebeaux, D.: complement/adjunct asymmetries 126–7, 129, 132; reconstruction 137, 173 n1, 175 n17
levels of representation 4, 11, 42
Levin, N.: Pseudogapping 56, 58, 87, 107, 170 n10
lexical insertion in the LF component 64
lexicalism 10, 12, 14, 103, 172 n3
LF: clitic 27, 46; and expletive construction 26; as a level of representation 43, 45; LF-component 11, 57; LF-copying 5, 38, 57, 60, 77, 79–80, 87, 89, 159 n13, 163 n24, 169 n7; LF crash theory 97
Lobeck, A.: ellipsis 92, 108, 167 n7, 170 n15
locality 37, 75; *see also* Chain, Relativized Minimality
lowering (verbal morphology) 8

McCloskey, J.: sluicing 169 n7; VP ellipsis 69, 105, 172 n11
make … out construction 121–2, 148,150–1
Marantz, A.: verbal morphology 13, 159 n10
Martin, R.: D-structure component 157; driving force for movement 28; expletive constructions 28; last resort 160 n2; Pseudogapping 66; verbal morphology 158 n3, 159 n10
Martins, A. M.: VP ellipsis 69, 105
maximal projection: and VP-ellipsis 71
May, R.: ACD 37–8, 56, 76, 79–81, 87, 163 n25, 164 n4; adjunction 32, 48; quantifier scope 118–19, 123, 134–5, 153–5, 173 n3, 175 n18
Merchant, J.: ellipsis 169 n7
merger: covert 95; PF-merger 13, 15, 159 n16
middle construction 161 n14
modal 1, 104, 159 n16
movement; A-movement 23, 27, 31, 45, 47, 58, 107–8, 160 n9; A-bar movement 45, 158 n5; copy-movement 119; copy theory of 39; covert 9, 15, 34, 36, 38, 51, 57, 60, 83–4, 88, 146; feature 32, 34, 36, 48, 51, 109, 111, 146; feature-driven 89; focus 43, 164 n4; full copy 133; head 158 n5; improper 142; LF- 3, 26, 30–1, 43, 45, 48, 50; long adjunct 158 n5; Move alpha 158 n2; obligatory 93; overt 34, 36–8, 60, 64, 84, 89, 93–4, 108; rightward 75, 82; successive cyclic 25; target 31, 91, 93; into theta position 157, 160 n3; V-movement 10; *wh*- 3, 43, 45, 64, 92
Munn, A.: *HIT* 26; reconstruction 131
Murasugi, K.: ellipsis 70, 92, 108, 167 n11, 170 n15; shortest move 25

NEG 9, 15
negation 1, 14–15, 46, 48, 118–20, 133, 144, 159 n16
Negative Polarity Item (NPI) 34, 47–9, 145, 151
Nemoto, N.: long A-movement 170 n13
never 15
nominalization 161 n12
noncoreference 150
Noonan, M.: strong feature 171 n19
not 32, 104
numeration 24, 97, 159 n16

object 51; direct 51, 94; of preposition or PP complement 58–9, 78, 170 n12; shift 56, 108; of simple transitive 37
objective Case 25
obligatory transformation 7, 22
oblique 24
Ochi, M.: feature movement 84, 98, 111, 171 n22
Oka, T.: Shallowness 160 n9
Oku, S.: VP-ellipsis 109, 159 n11
operator: overt and null 80, 168 n22, 168 n25; relative 39, 80; variable 10, 37; *see also* Chain Uniformity

optionality 7, 149, 151; of Agr_O 152
Ormazabal, J.: ECM 161 n12, 161 n15
Otani, K.: VP-ellipsis 172 n8
overt syntax 11, 24, 64

paraphrasability 120–1, 123, 134–6, 153–5
Park, M.-K.: ellipsis 167 n8
passive 107
PF crash theory 84–6, 93, 97–101
PF reasons for category movement 36
PF requirement 30, 60, 89
phi-feature 27–8, 30
phonological defective element 111
pied-piping 30, 37, 49, 60, 89, 99, 110–11, 171 n23
Pollock, J.-Y.: split Infl hypothesis 8, 15
Polish 52
Portuguese 69, 73, 105
Postal, P.: ECM, overt object raising 4, 35, 94, 145–6, 150, 164 n8; quantifier scope 129, 173 n1, 175 n16
Potsdam, E.: VP-ellipsis 105–6, 113–14, 116, 172 n3
preference principle 128–9, 132
Prepositional Phrase 29
Procrastinate 11–12, 15, 24–5, 30, 40, 83, 160 n10
Pseudogapping 55–7, 59, 61, 64–6, 70, 73, 76, 78, 83, 87–9, 106–7, 111, 147–9, 152, 174 n12, 175 n14, 170 n10; marginal character of 64
pseudo-passive 38–9, 59
psych predicate 136

quantificational expression 79
quantifier 118
Quantifier Lowering (QL) 118–24, 134–6, 153–4, 173n3
Quantifier Raising (QR) 37–8, 43, 45, 79, 81, 87, 118–19, 135, 175 n18, 164 n4
Quirk, R.: Pseudogapping 16; VP-ellipsis 104

R-expression 79
raising: of accusative NP 57; of complement of *to* 23; feature 52–3; head- 159 n10; I-to-C 92–3, 108; of indefinite NP 124; LF- 12, 35, 50; overt 56, 60; re- 9, 158 n4; subject 118; successive cyclic 29; super- 158 n5; to object, to [Spec, Agr_O] 5, 37–8, 44, 47, 50–1, 56–7, 60, 78, 87–8, 108, 151; *see also* Exceptional Case Marking, Pseudogapping; verb 9–10, 15, 37, 40, 60–1, 73, 89, 108–9, 166 n3
reanalysis 38, 59, 78
reciprocal 45, 49
reconstruction: A-movement 118–22, 125–6, 136, 139, 153–5, 173 n1, 175 n17; A-bar movement 81, 127–9, 133, 142, 173 n1
redundancy 24
referential and quantificational property 34, 36, 94
regress (in ACD) 38, 76–7, 79
relative 39, 79–81
relativized minimality 10, 15, 63, 71–3
remnant (Pseudogapping) 55–7, 60, 62–3, 73, 76, 78, 88, 106–7
restructuring 170 n13
Richards, N.: control 165 n15
Riemsdijk, H. van: Heavy NP Shift of the object of a preposition 58
Rizzi, L.: A-movement reconstruction 136–7, 143, 173 n2, 175 n17; relativized minimality 15
Roberts, I.: verb raising 15; VP-ellipsis 105–6, 108, 112–13
Romero, M.: sluicing 169 n7
Rosenbaum, P. S.: ECM 4
Ross, J. R.: ellipsis, sluicing 69, 92, 108, 169 n7
Rudin, C.: multiple *wh*-movement 100

Sag, I.: ACD 76; Pseudogapping 56, 106, 169 n9; VP-ellipsis 16, 18, 104, 169 n9
Saito, M. 3, 4; binding 44–5, 160 n2; ellipsis 70, 92, 108, 167 n11, 170 n15; Empty Category Principle 9, 34; Exceptional Case Marking 46–7, 50–1, 145–6, 150; existential constructions 46; overt raising to [Spec, Agr_O] 94, 174 n12; shortest move 25
Scandinavian 165 n13
scope 27, 32, 34–5, 46, 48, 79, 94, 118–19, 121, 133, 137, 144, 148, 176 n19
Serbo-Croatian 100–1, 114, 158 n7, 169 n8
shortest move 25, 160 n8
Sobin, N.: virus theory 86
sloppy identity 114–16, 172 n8
Sluicing 60, 83, 92–3, 108, 112, 163 n24, 166 n6, 169 n7, 170 n15

small clause 27, 58, 87
spec-head configuration 27, 38, 164 n5
specifier 16
Spell-Out 85–6, 159 n16
split Infl hypothesis 8
split VP hypothesis 89, 108
Standard American English (SAE) 6, 9–18
Stjepanović, S.: VP ellipsis 172 n9
Stowell, T.: null complementizer 80
stranded affix filter 7, 19; *see also* Affix Hopping
strict reading 116
Stump, G.: Pseudogapping 88
Subjacency 10
subject: Exceptional Case Marking 37, 47, 51, 57; of infinitives 35; orientation (anaphora) 165 n14; VP-internal 160 n8
substitution 26, 32, 46–7, 160 n4
Superiority 28
surface structure (S-structure) 4, 7, 42, 44, 79
Swedish 12, 15

Tancredi, C.: VP ellipsis 87, 163 n24
Takahashi, D.: ACD 37; shortest move 164 n26; theta theory 137, 157
Tense (ellipsis) 70, 72
tense/agreement (inflectional) morpheme 1, 6
Tensed Sentence Condition 5
term 157
thematic properties of control 52–3
there 24, 26, 29–30, 46, 53, 160 n11, 162 n20, 162 n21, 165 n18
theta-role 8, 20, 25; assignment 176 n19; configurational view of 156; as feature 61, 137, 156
theta-theory 155
Toman, J.: multiple *wh*-movement 100
trace: of affix lowering 12; and antecedent for VP-ellipsis 105; copy theory 81; intermediate 139–40; and verb raising 9, 39, 158 n4; and VP ellipsis 81, 106; of *wh*-movement 3, 76

Ura, H.: raising of V 164 n27
Uriagereka, J.: D-structure component 157; reciprocal binding in existential construction 44–5; theta role 176 n19; virus theory 86

verb: auxiliary 8, 15, 18, 103–4; bare 13, 159 n15; finite form of 16; main 12, 15, 18–19, 103–4; transitive 37
Vergnaud, J.-R.: Case theory 3, 158 n6
visibility approach to Case 3
virus theory 86, 92, 95, 97; *see also* Strong Feature

Warner, A.: VP-ellipsis 16–18, 104, 159 n14
was 16
Wasow, T.: VP deletion and Sluicing 60, 87, 163n24
Watanabe, A.: *do*-support 158 n1; shortest move 164 n26
weak feature 11
Weinberg, A.: LF-operations 45, 164 n4
Wexler, K.: verb raising in Swedish 12
wh-in situ 28
Whitman, J.: VP-ellipsis 172 n8
Wilder, C.: verb raising in Swedish 12, 158 n7
Wyngaerd, G. V.: ACD 38, 78, 163 n25; overt object raising 94, 165 n13

X-bar theory 27

Zagona, K.: VP ellipsis 70
Zubizarreta, M.-L.: quantifier scope 119–20, 153–4, 173 n1, 173 n3
Zwart, C. J.-W.: ACD 38, 78, 163 n25; overt object raising 165 n13