INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

CYCLIC DOMAINS IN SYNTACTIC THEORY

Arthur Vladimirovich Stepanov, Ph.D.

University of Connecticut, 2001

The dissertation explores a principally non-constraint-based approach to locality phenomena in *wh*-movement, in particular, impossibility of overt extraction of a *wh*-phrase out of certain structural domains ('islands'), and the successive cyclic character of *wh*-movement across a finite clause boundary. The proposed theory of locality utilizes the idea in the earlier works of Cinque, Bresnan and Cattell, that movement is only possible within a well-defined type of configurations -'cyclic domains', and implements this idea in the Minimalist framework (Chomsky 1995, 2000). We argue that under the minimalist view, the notion cyclic domain need not be stipulated; rather, it reduces to a cover term for a phrase marker containing a source and target of movement, in which movement proceeds without intermediate steps. A minimalist theory of cyclic domains is conceptually and empirically superior to the previous transformational theories of locality which are inherently constraint-based.

In the first part, we argue, contra traditional theories (cf. 'Condition on Extraction Domains'), that extractability out of subjects and adjuncts is regulated by different mechanisms of grammar. Overt *wh*-extraction out of a subject is allowed, but leads to a violation at the syntax-phonology interface if the subject has previously moved in syntax, forming a non-trivial chain. Overt *wh*-extraction out of adjuncts is

Arthur Vladimirovich Stepanov—University of Connecticut, 2001

precluded in syntax because by the time extraction is supposed to occur, the source and target of extraction are not (yet) within the same phrase marker. This proposal accounts for the cross-linguistic variation in extractability out of subjects and for the apparently universal ban on overt extraction out of adjuncts, and has other welcome empirical consequences.

In the second part, we address the issue of the local character of *wh*-movement in finite clauses. Particular attention is devoted to investigating the syntax of '*wh*-scope marking' questions, in the framework of Indirect Dependency, developed by Dayal and adopted and modified here. The key proposal is that the syntactic structure of *wh*-scope marking questions and questions involving 'long-distance' *wh*-movement is fundamentally similar. Exploring this similarity, we suggest that successive cyclicity in long-distance *wh*-movement is a residual effect of the underlying *wh*-scope marking structure, and is therefore epiphenomenal.

CYCLIC DOMAINS IN SYNTACTIC THEORY

Arthur Vladimirovich Stepanov

M.A., University of Connecticut, 1999

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

at the

University of Connecticut

2001

UMI Number: 3023466

Copyright 2001 by
Stepanov, Arthur Vladimirovich

All rights reserved.

UMI®

UMI Microform 3023466

Copyright 2001 by Bell & Howell Information and Learning Company.
All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

Bell & Howell Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

Copyright by

Arthur Vladimirovich Stepanov

2001

APPROVAL PAGE

Doctor of Philosophy Dissertation

CYCLIC DOMAINS IN SYNTACTIC THEORY

Presented by

Arthur Vladimirovich Stepanov, M.A.

Major Advisor _____
Howard Lasnik

Associate Advisor _____
Željko Bošković

Associate Advisor _____
William Snyder

University of Connecticut

2001

ii

Table of Contents

Chapter 3   An 'Eclectic' Approach Part II: Late Adjunction

iv

Chapter 6   Long-Distance Extraction as Residual Wh-Scope Marking

INTRODUCTION

1. Background

1.1 The Framework

Transformational syntactic theory views recurrent structural patterns of language as a realization of a particular state of human mind, the state corresponding to one's 'knowledge of language' ('I-language', see Chomsky (1986b)). It seeks a formally robust description of these patterns and their explanation from the point of view of learnability and child language acquisition. Generative-transformational grammar then is a theory of language as part of the biological endowment of human beings.

We adopt the basic assumptions of modern transformational grammar in the form of the principles and parameters framework, implemented in the Minimalist program (cf. Chomsky (1995c), (2000)). Specifically, we assume that, as part of its genetic make-up, the human mind/brain features a computational algorithm designed specifically for generating expressions of natural language - linguistic expressions (LE). Principles governing the operation of this specialized algorithm constitute Universal Grammar (UG); it is those principles that are common to all individual languages like English, Russian or Chinese. We take a LE to be a pair {PF, LF}, PF and LF its *phonetic form* and *logical form*, respectively (unpredictable with respect to each other, cf. Saussurean arbitrariness). A LE is an output of a syntactic *derivation*, a series of computational steps during which lexical units are taken from the lexicon and assembled together into a single structure ('phrase marker'), which can be further modified in the course of the derivation. The LE is then submitted for interpretation at the interfaces between the computational

1

system of language and external cognitive systems. In particular, PF is interpreted at the 'Articulatory-Perceptual' interface, where it is assigned phonetic shape and prosodic contour, among other. LF is interpreted at the interface with our systems of thought, where it is assigned meaning via the rules of compositional semantics. These general assumptions, rooted in the history of transformational grammar, form a conceptual basis for investigation of expressions in particular languages as artifacts of the (universal) linguistic capacity of humans.

There are two fundamental properties of natural language that any syntactic theory must capture. The first one is hierarchical structure, namely, the language's ability to build sentences out of smaller phrases - constituents, which, in turn, may be composed of yet smaller units. Thus the computational algorithm must contain an elementary device responsible for structural composition. We assume that to form LEs, the computational algorithm utilizes a structure building operation *Merge*. Merge takes two (potentially more) lexical units and combines them together into a single phrase marker. Merge applies recursively: it can use the resulting phrase marker for further assembling. Both lexical units and pieces of structure assembled from them are *syntactic objects*. Merge then is an operation applying to syntactic objects.

The second fundamental property of natural language is the property of 'constituent displacement' or movement. A simple instance of movement is illustrated below:

(1)   a. This book, John likes *t*

         b. John likes this book

Thematic relations (*this book* is an object of *likes*) indicate that the NP *this book* 'belongs' in a position following *likes* (cf. 1b), yet in (1a) it is displaced ('topicalized') in front of the sentence. We assume that the computational algorithm also features an operation responsible for movement. Properties of and constraints on this operation have been a subject of considerable debate in the literature, and, indeed, a motivation for a number of theoretical revisions in the last several decades. In this work we investigate these properties further. For now, we simply assume the existence of this operation.

In this framework, variation across individual languages reduces to (unavoidable) lexical differences, as well as differences in morphological properties of functional heads like Tense, Complementizer etc. These differences are finite and limited.

1.2 Locality of Wh-Movement: The Standard Approach

Wh-movement is a transformational rule that applies in *wh*-questions. It moves the *wh*-phrase into Spec-CP, leaving a trace (copy) in the position where it is originated:

(2) Who did they elect *t* as President?

Since the earliest stages of generative syntax it has been observed that syntactic *wh*-movement is constrained: it cannot apply in every construction containing a *wh*-phrase. Researchers traditionally refer to this state of affairs as locality. Locality phenomena constituted the empirical basis for development and major changes of syntactic theory. Ross (1967) has shown, in particular, that a *wh*-phrase may not be moved out of certain

well defined structural domains which Ross termed 'islands'. Those include, among other, complex NPs and sentential subjects. The relevant examples follow:

(3)   ?*What$_i$ did [$_{IP}$ he see [$_{NP}$ a man [$_{CP}$ who [$_{IP}$ was wearing t$_i$]]]]?

(4)   ?*Who$_i$ would [$_{IP}$ [$_{CP}$ for John to like t$_i$] please you]?

In order to account for locality effects of this sort, Ross proposed a series of construction-specific constraints, in particular, the Complex NP constraint and the Sentential Subject constraint, each of which prohibits extraction out of the respective structural configuration.

In the context of a larger enterprise to radically restrict the expressive power of transformations (due mainly to considerations of learnability), Chomsky (1973), (1977) eliminates the need to postulate construction-specific constraints of this kind. Instead, he proposes a general condition restricting the domain of application of transformational rules - *Subjacency*:

(5) *Subjacency*

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

b. Y is subjacent to X if there is at most one cyclic category C≠Y such that C contains Y and C does not contain X.

Subjacency prohibits a single instance of *wh*-movement across more than one 'cyclic node', with a proviso that cyclic nodes are NP and IP in English (cf. Chomsky (1977)).

For instance, *what* in (3) crosses (at least) one NP and one IP node, in violation of Subjacency. In (4), movement of *who* crosses two IPs.[1] Thus instead of a number of construction specific constraints, various 'island' phenomena now receive a uniform account.[2]

Restricting the domain of application of rules in this way imposes a particular view on cases where a *wh*-phrase can apparently be displaced indefinitely far away from its base-generated position:

(6)   a. What did Bill buy *t*?

       b. What do you think Bill bought *t*?

       c. What do you think John said (that) Bill bought *t*?

       d. What do you think Mary remembers that John said (that) Bill bought *t*?   etc.

Given cases like these, *wh*-movement was sometimes argued to be 'unbounded' (cf. Ross (1967), Bresnan (1976), among others). Under a Subjacency-based theory of locality (aka Bounding Theory in the frameworks of Extended Standard Theory and Government and Binding), however, what looks like 'unbounded' *wh*-movement in (6) is in fact a series of iterative applications of *wh*-movement, each of which moves the *wh*-phrase into a local

---

[1] An immediate problem with this version of Subjacency is that it seems unable to distinguish between the ungrammatical (4) and (i), where, apparently, also two IPs are crossed:
(i)   Who would it please you for John to like *t*?

[2] As research progressed, various alternative notions have been proposed. Kayne (1981), (1983), among others, restates the relevant array of phenomena in terms of the (reformulated) Empty Category Principle (ECP). Koster (1978) proposes a theory based on his Principle of Locality and Bounding Condition, to explain both the boundedness of *wh*-movement and the island phenomena. Overall, there seems to be an agreement in the literature that some kind of constraint is necessary to state the locality phenomena.

COMP (Spec-CP) by crossing at most one cyclic node (IP). This movement came to be called successive cyclic.

The modified notion of Subjacency in *Barriers* (Chomsky (1986a)) provides an even more unified approach to 'island' phenomena. In addition to subsuming the standard island cases above, the *Barriers* approach also captures other types of island effects such as those in Huang (1982). Thus, consider the analysis of certain 'island' phenomena in the *Barriers* system of Chomsky (1986a). Subjacency Condition and related definitions are formulated below:

(7) In a well formed chain with a link (α_i, α_{i+1}), α_{i+1} must be 1-subjacent to α_i

(8) β is *n-subjacent* to α iff there are fewer than $n + 1$ barriers for β that exclude α

(9) γ is a *barrier* for β iff (a) or (b):

   a. γ immediately dominates δ, δ a Blocking Category for β

   b. γ is a Blocking Category for β, $\gamma \neq$ IP.

(10) γ is a Blocking Category for β iff γ is not L-marked and γ dominates β

(11) α L-marks β iff α is a lexical category and θ-governs β.

(12) α θ-governs β iff α is a zero-level category that θ-marks β, and α and β are sisters.

Given (7)-(12), consider extraction out of complex NP in (3) repeated here:

(3) ?*What$_i$ did [$_{IP}$ he see [$_{NP}$ a man [$_{CP}$ who [$_{IP}$ was wearing t$_i$]]]]?

Movement of *what* in (3) crosses two barriers. The embedded CP is not L-marked, hence is a Blocking Category (by (10)-(12)) and a barrier (by (9b)). Furthermore, the NP *a man...* is also a barrier since it immediately dominates the CP. By crossing two barriers, movement forms a chain (*what*, t_{what}) in which t_{what} is not 1-subjacent to *what*. Subjacency is thus violated. (4), too, is subsumed under this account: movement of *who* crosses the non-L marked CP boundary, and the (matrix) IP node, both of which are barriers, violating Subjacency.

In addition, the Barriers approach implements even more rigorously the idea of boundedness of movement transformations: in this system, movement proceeds by successive adjunction to maximal projections along the way (which, by hypothesis, voids potential barriers).

Some version of the Barriers approach remains dominant in most current theories of locality. For instance, Takahashi (1994) reformulates the insight behind Subjacency in terms of the 'Shortest Move' condition (cf. Chomsky (1995c), Ch.3) which states that movement must be as short as possible. Thus, roughly, *Barriers*-style successive adjunction to maximal projections along the path of movement satisfies the 'Shortest Move' condition, whereas movement across a 'barrier' amounts to its violation (see Chapter 2 for more discussion).

The Barriers-Subjacency approach (in the wide sense) raises the following three questions, the answers to which set up the stage for our present study:

(13) Why should extractability out of a domain be (ultimately) stated in terms of theta-marking and not some other factor(s)?

(14) What is the status of Subjacency in the grammar? Is it primitive or derivable?

(15) Subjacency (or its descendants) provides a uniform account of island effects. But is the unification idea warranted in the first place?

Question (13) is adduced by the spirit of *Barriers* itself. At the heart of the *Barriers* analysis was the idea that a structural complement (sister of a lexical head) is not a barrier for extraction, whereas a non-complement (subject, adjunct) is. This idea is encoded in the notion of L-marking, which, in turn, is defined, roughly, as theta-marking by a lexical head.[3] But the implied relation between extractability and (lexical) theta-marking is not obvious. A priori one expects nothing in common between the two. In fact, in the system of Chomsky (1981) conditions governing movement (e.g. Bounding Theory) and those responsible for theta-role assignment (Theta-Theory) are autonomous with respect to each other and indeed constitute different modules of the core grammar. Thus until it can be shown in detail within the Barriers approach how the two are related we seem to be missing a generalization, and the relation in question cannot be taken for granted.

Question (14) is a 'minimalist' question, in the sense of the Minimalist program (Chomsky (1995c), (2000), (2001)). The Minimalist program considers the design of language to be optimal. This means, in essence, that in the ideal case, properties of syntax must follow from conditions imposed on the computational system by 1) interfaces with external performance systems - taken to be at least the articulatory system(s) and the

---

[3] Another possibility is to state this distinction in terms of the Projection Principle, requiring 'saturating' the positions within a maximal projection. But this is, in essence, a theta-theoretic requirement. Given that the Projection Principle has no independent status in the current framework, this possibility is not a real alternative.

system(s) of thought - or, alternatively, by 2) the design of the computational system itself (=principles of UG). A priori, Subjacency appears not to be an interface requirement, since neither the articulatory nor conceptual performance systems seem to have anything to do with restricting the domain of application of movement rules. It may be argued that Subjacency is part of the design of the computational system itself, in the sense that restricting the domain of movement helps reduce the 'memory load' on syntactic computation, thus reducing the computational complexity of the system (cf. Chomsky (2000)). This is a plausible direction, but at the present stage of syntactic theorizing it is not clear what the relevant notion of 'memory' is, and whether a notion like that is relevant at all in syntactic computation. In addition, if it can be shown empirically that effects usually attributed to Subjacency are due to some other factors, this would remove the motivation for the concept in the first place. In this work, we intend to show just that.

Question (15) has a wider conceptual and empirical aspect, the former largely depending on an answer to Question (14). In pre-minimalist, constraint-based frameworks, a unified analysis of locality effects (under the umbrella of Subjacency or another condition) is clearly desirable. By a general criterion of "Ockam's razor", a theory that needs a single constraint (cf. Chomsky 1986a) is more desirable than a theory that needs five or more constraints (Ross 1967), in order to explain the same range of phenomena. The Minimalist program, however, offers more than one possibility for looking at locality. As noted above, minimalism maintains that empirical properties are ideally attributed to the design of the grammar, or interface requirements. Note that these two concepts by definition are unrelated and have very different properties. Thus several possibilities become equally plausible: a) all locality phenomena are (uniformly) an effect

of the design; b) all locality phenomena are (uniformly) effects of interface requirements; c) some locality phenomena follow as an empirical effect of the design; others as effects of interface requirements. These three scenarios entail two major possibilities for looking at locality in minimalism - a *unified* approach, similarly to the pre-minimalist theories, and an *'eclectic'* approach, in which various locality effects arise as a consequence of various parts of the grammar. Neither is imposed on us on conceptual grounds. The decisive criterion for choosing the right theory is the empirical one. Indeed, our view in this work will be that an eclectic approach is the correct one.

In this study we set out to develop a theory of locality that answers the above questions. Our general answers will be:

(16) 1. There is no fundamental relation between extractability and (lexical) theta-marking

2. Subjacency has no independent status in the grammar. Its empirical effects is derivable in other ways.

3. A theory of locality based on *any* unifying concept is neither sufficient nor necessary. Various grammatical mechanisms are responsible for various locality effects ('eclectic' approach)

2. Cyclic Domains

Ross' influential work set a particular methodological standard for investigating locality phenomena, which has been followed in subsequent transformational analyses. Namely, it has been a commonly (perhaps implicitly) assumed since Ross that transformations may

generally apply in *any* environment, and only in certain configurations their application is blocked. Effort has then been put to properly define what kind of configurations those are. Ross' study itself is a classical example of this method: his locality constraints explicitly refer to specific constructions (complex NPs, sentential subjects, coordinate structures).

As researchers moved away from relying on highly construction-specific constraints in the direction of more general locality conditions, they nevertheless preserved Ross' perspective in formulating these conditions. Subjacency is the case at hand. Although Subjacency was never a principle targeting specific constructions *per se*, the general tendency was to construe Subjacency either as a property of derivations that produce these constructions (more specifically, a condition on movement rules), or, alternatively, as a property of representations at the relevant levels ('S-Structure', LF). Proponents of the first alternative include Chomsky (1973), and (1977), among others (cf. (5) above); proponents of the second alternative include Freidin (1978), Freidin and Lasnik (1981), Browning (1991), among others. The analysis in Chomsky (1986a) is compatible with both derivational and representational construals of Subjacency (see Browning (1991) for detailed discussion).

Cinque (1978) (see also Cattell (1976)) and, more recently, Postal (1999) recognize that this method of looking at locality misses an important generalization, namely, that both cross-linguistically and within a given language there are many more types of constructions in which extraction is blocked ('islands') than those in which it is allowed ('non-islands'). More specifically, these authors observe that extraction is only allowed out of those finite clauses that are structural complements of the embedding predicate. This is, of course, the familiar complement/non-complement distinction that was at the

heart of Huang's CED and Chomsky's *Barriers*. But whereas the latter implement this distinction in the way that is consistent with Ross' method, Cinque proposes to reverse this method altogether, to capture the above generalization. He suggests that (extraction) transformations may only apply in certain well-defined configurations which he terms *cyclic domains*, and are blocked otherwise (see also Bresnan (1976)). A very similar idea is pursued in Kayne (1981), (1983). The two views on locality are contrasted below:

(17)   *Standard View on Locality (Ross, Subjacency-based theories)*

Extraction is allowed everywhere except in certain configurations ('islands')

(18)   *Cyclic Domain View on Locality (Cinque, Bresnan, Cattell)*

Extraction is *only* allowed within certain configurations -'cyclic domains'.

(19) A *cyclic domain* is any sequence of clauses $c_1,...c_n$ (where ',$...c_n$' may be null) such that c_{i+1} is embedded in c_i as an argument (clause) of the predicate of c_i, in logical form, for each i $(1<i<n)$.[4]   Cinque (1978)

(19) effectively encodes a complement/non-complement distinction: it allows extraction only out of domains that are complements (='arguments') of the predicate.

Under the cyclic domain view (18), the emphasis of investigation of locality shifts from the conditions blocking extraction (cf. (17)), to investigation of conditions that allow extraction. In particular, the task of investigation now becomes characterizing the

---

[4] Cattell (1976)'s notion of 'syntactic configuration' is, virtually equivalent to (19):
(i)   A *syntactic configuration* is a maximal sequence of sentoids [≈ clauses - *A.S.*], $S_1...S_n$, such that each S_i (i≠1) is embedded in the predicate of S_{i-1} and is a function of the verb.

notion of 'cyclic domain' in an explanatory manner, which means, among other things, answering the question of why a particular notion of cyclic domain is to be chosen over any other. We want to argue that this shift of emphasis is desirable for at least three reasons.

First, if we re-formulate the task of investigation in this manner, it becomes possible to explore it in a way that does not force us to postulate any new abstract principles of UG. Indeed, the standard view (17) treating 'islands' as certain 'exceptions' to otherwise unconstrained extraction rules, must concentrate on conditions that rule out certain sentences ('island violations'). This inevitably leads to postulating additional constraints on the rules of grammar that would otherwise generate these sentences. On the other hand, under the cyclic domain view (18) we proceed from the premise that transformational rules may apply within certain domains only; in other words, we define the conditions under which the rules in principle generate good sentences. Under this approach, there is no need to rule out sentences that would otherwise be generated by the grammar; rather, we define the grammar in such a way that these sentences will simply not be generated at all. Consequently, no additional constraints are needed.

One might object to the latter claim by arguing that a theory based on the cyclic domain view in the sense of (18)-(19) is no better than the one based on the standard view. After all, both seem to need a stipulation: the latter 'subjacency' or some other notion (see fn. 2), the former the notion 'cyclic domain'. In fact, Bresnan (1976) shows that the notion of cyclic domain is definable via Subjacency, which would suggest that

the two types of approaches are notational variants.[5] However, recent progress in syntactic theory gives one the reason to believe that this is not so.

In the Minimalist program (Chomsky (1995c)), syntactic properties are dictated either by 1) interface conditions, or 2) the general architecture of grammar (see above). We argue that in the framework of the Minimalist program, it becomes possible to show that transformational rules apply in domains of the kind described in (19), on natural grounds. In other words, it can be shown that movement transformations apply in environments described by (19) because of the way 1) and 2) work, without the need to stipulate the notion cyclic domain. Rather, in the minimalist system cyclic domain is a cover term for a phrase marker containing the source and target of movement, in which movement proceeds without intermediate steps. There are no constraints on movement within cyclic domains in narrow syntax. The only potentially allowed sort of constraints are 'bare output conditions' at the interfaces, which may function as 'filters' on syntactic derivations.

On the other hand, it is not clear how the standard, constraint-based approach, e.g. Subjacency-based theories, can be translated into minimalism on natural assumptions. This state of affairs allows us to tease apart the two approaches, and conclude that the cyclic domain view of locality is conceptually superior to the standard view.

Another, no less important reason to pursue the cyclic domain approach is the empirical one. We want to maintain that it allows one to incorporate into the theory of

---

[5] "A and B are in the same *cyclic domain* if either B is subjacent to A, or there is a sequence of COMP's $c_1,...,c_n$ such that B is subjacent to c_n, c_i is subjacent to c_{i-1} for each i ($1 < i \leq n$), and $c_1 = A$." (p. 364)
    Another version of Bresnan's definition makes use of the notion '(immediately) accessible', which does virtually the same work as Cinque's 'embedded in'.

grammar many empirical phenomena and generalizations concerning various extraction domains that were either unnoticed before or could not be incorporated into previous theories of locality. For one thing, Cinque's original observation that extraction is possible only out of a small subset of finite clauses finds a straightforward explanation under the cyclic domain view. In addition, as will become clear, our version of the cyclic domain view accounts for the cross-linguistic variation in extractability out of certain domains, such as subjects, while at the same time capturing the universality of others.[6]

3. Overview of the Dissertation

The dissertation explores a minimalist theory of cyclic domains for *wh*-movement. Under this theory, the local character of *wh*-movement, and, more generally, A'-movement, follows directly from 1) the radically impoverished architecture of grammar in the minimalist sense and 2) the properties of the interfaces. Thus our theory provides conceptual and empirical substance to the programmatic insights in Chomsky (1995c), (2000). We argue, in particular, that no unified account of 'islands' (based on Subjacency

---

[6] Postal (1999) also acknowledges an existing methodological bias towards the (standard) view, according to which "all constituents are *non*-islands by default". Pointing out that this view is not forced a priori, he takes up an alternative view that all constituents are islands by default, and states the task as to search for common features of 'open' constituents. This, he argues, leads to "a highly general lawful characterization of open constituents...", which he ultimately states in relational terms. Postal then proceeds to argue that the standard view is virtually imposed on the transformational framework: "Since transformations will, if not prevented, apply across island boundaries, blocking the output requires principles analogous to Ross's island constraints".

We believe the task projected by Postal is practically equivalent to that entailed by the cyclic domain view. We thus agree with Postal that this view is more promising than the standard one, for the reasons spelled out above in the text. However, we disagree with him in that the standard view is a virtue of the transformational framework. Our task in this study is precisely to show that it is possible to delimit the domain of application of transformations on natural grounds, in a way that they never apply across island boundaries. Thus Postal's argument against the standard view is not an argument against transformations.

or any other concept, under the standard view) is either necessary or sufficient. Rather, in the ideal case, extraction possibilities in each case are dictated by interface conditions, or the general architecture of grammar. In other words, these two factors are those that determine the domain of application of extraction rules. To the extent this enterprise is successful, we support the conclusion that locality of A'-movement is a direct consequence of the design of human language.

Let us now delimit the scope of our study. The range of phenomena usually employed in most studies of locality includes A'-extraction from:

(20)　　I. 'Islands'

　　　　　1. 'Relativized Minimality' islands

　　　　　2. 'Weak'/'selective' islands: 'inner' islands (negation), factive islands

　　　　　3. Coordinate structures

　　　　　4. Adjuncts

　　　　　5. Subjects (in certain languages)

　　　　　6. Complex NP

　　　　II. Extraction from finite clauses; Successive cyclicity

We assume that an articulated theory of the phenomena in I.1., I.2., I.3. can be provided in terms of the architecture of grammar or interface conditions. Consider these phenomena in turn.

I.1. A common example of a 'relativized minimality' island is a *wh*-island, as in (21a):

(21)   a. ?*What₁ do you wonder [cp who bought t₁]

b. C do you wonder who bought what

In (21a) *what* moves across another *wh*-phrase, *who*, causing ungrammaticality. The standard pre-minimalist analysis of (21a) is that it is a Subjacency violation, whether Subjacency is defined in terms of bounding nodes (Chomsky (1973)), or barriers (Chomsky (1986a)). In recent versions of the minimalist framework (Chomsky (1995c)) the ungrammaticality of (21a) is understood as a result of 'intervention', whereby one *wh*-element 'intervenes' on the way of the other. The intervention effect is accounted for in the following manner. First, it is assumed that movement is triggered by the need to satisfy a feature of the target of movement. Thus the interrogative feature residing in C in (21b) is satisfied by establishing a relation ('Attract' in Chomsky (1995c)) with a matching feature in the c-command domain of C. Second, it is assumed that the computational system of grammar includes an operation such as Attract Closest. In (21b), there are initially two candidates for satisfying the interrogative feature of C, the *wh*-feature of *who* and that of *what*, and *who* initially c-commands *what*. By definition, *who* is closer to C than *what*. Hence, Attract Closest cannot pick *what* to satisfy the feature of C, and the ungrammaticality arises (see Chomsky (1995c) for reasons why *who* cannot raise in (21b)). That is, (21a) is ruled out by considerations pertaining to the architecture of grammar, without the need to impose external constraints. In this regard, we consider the Attract Closest account of 'relativized minimality' islands compatible with the cyclic domain view.

I.2. The term selective or 'weak' islands usually refers to islands which block extraction of certain *wh*-elements, but not others (see Cinque (1990), Postal (1999), cf. also Takahashi (1994)). Most often the relevant distinction is between arguments, which are extractable, and adjuncts, which are not. The standard cases of selective/weak islands include negative islands, factive islands and extraposition islands, exemplified respectively below (cf. Cinque (1990)):

(22) a. I wonder how Peter fixed the car   *Negative island*

      b. *I wonder how Peter didn't fix the car

(23) a. (?) I wonder which car Mary regrets that Peter fixed   *Factive island*

      b. *I wonder how Mary regrets that Peter fixed the car

(24) a. (?) I wonder which car it matters that Peter fixed   *Extraposition island*

      b. *I wonder how tall it matters that Peter is

It is not clear what the relevant criterion defining the class of selective/weak islands is, nor whether selective islands form a homogeneous class that should be accounted for in a uniform way. For our present purposes, we adopt the view that this class of locality phenomena falls mostly under the principles pertaining to the interface with semantics/pragmatics (cf. 2) above), as proposed in Honcoop (1998), Kuno and Takami (1997), Rullman (1995), and Szabolcsi and Zwarts (1992-1993), but contra Manzini (1998) and Rizzi (1990). We will not be concerned with the precise nature of these principles. See the above authors for a range of relevant proposals.

I.3. The impossibility of extraction of and out of a conjunct in a coordinate structure is exemplified below:

(25)   a. *Who$_i$ did John see Mary and t$_i$?

         b. *Who did Peter finish the book and fire?

There is a number of exceptions to Ross's CSC (see, in particular, Postal (1998)). Ross's claim that the CSC is a constraint on movement, explicitly or implicitly maintained in most transformational studies on locality, has been repeatedly challenged in various studies. See, in particular, Goodall (1987), Munn (1993) and Postal (1998).[7] Goodall argues that the CSC follows from considerations usually attributed to the LF interface, in particular, the ban on vacuous quantification and binding theory, as well as the theta-criterion. In a similar vein, Munn argues that the CSC is a semantic constraint enforcing conjunction of phrases of the same semantic type. We assume, with these authors, the LF/semantic nature of a constraint regulating extraction out of conjuncts, although, as in the case I.2., we remain neutral with regard to a correct formulation of this constraint.

This leaves us with the phenomena listed in I.4., I.5., I.6. and II. The main focus of the present study lies in an investigation of these phenomena in the minimalist framework which most naturally implements the cyclic domain view (18).

---

[7] The CSC has long resisted a satisfactory explanation in most versions of the principles and parameters framework. For instance, it is clear that it cannot be reduced to a Subjacency violation in cases like (25), whether Subjacency is defined in terms of bounding nodes or barriers: indeed, movement of *who* in (25) crosses at most one bounding node (in the system of Chomsky (1973)), the coordinate NP node itself; no barriers are crossed (in the system of Chomsky (1986a)). Munn (1993) further shows that the CSC cannot be reduced to Empty Category Principle (cf. Chomsky (1981)), or any constraint on movement.

The dissertation consists of two major parts. The first part addresses the issue of extractability out of subjects and adjuncts. Chapter 1 discusses a traditional 'unified' approach banning extraction out of both these domains, the standard representative of which is the Condition on Extraction Domains (Huang (1982)). A priori, under the unified approach the impossibility of extraction out of adjuncts implies the impossibility of extraction out of subjects in the same language. This expectation is not met, as we demonstrate with an array of facts systematized form a variety of unrelated languages, in which extraction out of adjuncts leads to ungrammaticality, while extraction out of subjects is fully grammatical. The 'unified' approach, especially its minimalist descendant, also faces other conceptual and empirical problems.

In Chapters 2 and 3 we explore an alternative possibility in the form of an 'eclectic' approach, consisting of non-overlapping analyses of extractability out of subjects and adjuncts. Chapter 2 deals with extraction out of subjects. We explore a version of the chain uniformity theory of extractability out of subjects, originally proposed in Takahashi (1994) and adopted and modified here, which formalizes the earlier idea that moved domains become islands for extraction (cf. Freezing Principle of Wexler and Culicover (1981)). The chain uniformity theory relates the impossibility of extraction out of subject to its movement from the base-generated position (Spec-vP) prior to extraction. Under the chain uniformity theory, the fact in the languages discussed in Chapter 1 extraction out of subject is possible is expected given that subjects in those languages do not move prior to extraction. Independent evidence in support of this theory is drawn from the domain of extractability out of 'specific'/'presuppositional' DPs.

Chapter 3 provides a theory explaining the apparently universal ban on overt extraction out of adjuncts. Those include, aside from various type of adjunct phrases, certain types of complex NPs (cf. Stowell (1981)). We develop a structure building algorithm based on the minimalist bare phrase structure, and the economy principle 'Least tampering' (Chomsky (2000)) which disfavors change of basic relations inside the existing structure. The algorithm forces adjuncts to be inserted postcyclically, giving rise to the situation when the extractee and the target of extraction are not in the same phrase marker by the time when extraction is supposed to take place. We discuss a number of consequences of the late adjunction algorithm, involving experiencer constructions, approximative inversion in Russian etc. Taken together, the chain uniformity analysis and postcyclic insertion of adjuncts define domains of extraction - cyclic domains - strictly from the point of view of interfaces (the former) and the architecture of grammar (the latter).

We thus reject the idea, expressed in *Barriers*, that all islands arise as a consequence of some unifying principle. By rejecting it, we automatically question the empirical basis of the locality theory of *Barriers*. In this respect, a note should be made concerning the issue of different degree of (un)grammaticality in extraction of 'argument' *wh*-phrases vs. 'adjunct' *wh*-phrases, out of islands, as shown below:

(26)   a. ??What$_i$ did you meet a man who was wearing t$_i$?

    b. *How$_i$ did you meet a man who was fixing a car t$_i$?

In *Barriers,* the more degraded status of (26b) was considered to be a consequence of the same system that rules out (26a) (more concretely, an ECP violation; see also Lasnik and Saito (1984)). But since we question the empirical basis of locality in *Barriers* altogether, (26a) and (26b) no longer need to be treated on a par. It is possible that the more degraded status of (26b) is a consequence of a different principle.[8] With this in mind, we leave this issue beyond the scope of the present study.

The second part of the dissertation deals with *wh*-extraction from finite clauses, in 'long-distance' *wh*-questions. We investigate locality of long-distance *wh*-movement, encoded in its successive cyclic character, whereby a *wh*-phrase moving long-distance to the matrix interrogative Spec-CP must pass through each intermediate Spec-CP on its way. In this connection, particular attention is devoted to a study of questions involving '*wh*-scope marking' or 'partial *wh*-movement', and their connection with 'long-distance' *wh*-questions.

Chapter 4 argues for a structure of finite clausal complements as 'NP-shells' (the term coined in Müller (1995)), that is, CPs inside a projection of a nominal pro-form. Under the NP-shell analysis, the impossibility of extraction out of finite complements in a number of languages, including Russian, Finnish, some dialects of German, etc, is explained as an instance of 'Complex NP island' and is accounted for along the lines of Chapter 3.

---

[8] Perhaps the continuing lack of success in capturing this argument/adjunct asymmetry in the current framework is suggestive that it should not be accounted for within the same system of assumptions.

Under the view adopted in this work, it must be that the 'ECP effect' associated with the argument/adjunct asymmetry is due to a principle over and above the analysis of locality phenomena proposed here. This principle can possibly be attributed to the LF interface, given the well known facts from *wh*-in situ languages like Chinese (Huang (1982)), in which LF *wh*-dependencies across certain 'islands' can be formed with arguments, but not with adjuncts.

Chapter 5 investigates the *wh*-scope marking strategy. Data from Russian and Polish are introduced into discussion. Russian and Polish provide evidence for a particular theory of *wh*-scope marking, a version of the Indirect Dependency Approach (Dayal (1994), (1996a), (1996b)) which we adopt and modify here. On the basis of these languages, we argue for a particular structure of *wh*-scope marking questions, which mirrors the structure of finite complements (NP-shells). We demonstrate that the proposed structure yields a straightforward compositional semantics of *wh*-scope marking, without need to make additional assumptions concerning the syntax-semantics interface. We also address a number of puzzles for the Indirect Dependency Approach noted in the literature, and show that most of them dissolve under the proposed syntactic structure.

Chapter 6 explores the intimate connection between *wh*-scope marking and 'long-distance' *wh*-movement in languages like English. We argue that the syntactic structure of *wh*-scope marking underlies 'long-distance' *wh*-questions, in languages such as English. We argue that successive cyclicity is a residual effect of the *wh*-scope marking structure in long-distance *wh*-questions. We conclude that locality in the form of successive cyclicity is epiphenomenal, and nothing special needs to be postulated in addition to the existing mechanisms and devices of grammar in order to delimit the domain of application of *wh*-movement out of finite clauses.

CHAPTER 1

Extraction From Non-Complements: Unified Approaches

1. Traditional Unified Approaches

1.1 Outline

It is well known that in English and many other languages overt *wh*-extraction out of the subject of the sentence (either sentential or nominal) leads to a degradation, as opposed to *wh*-extraction out of object:

(1) a. ?*Who$_i$ does [a picture of t$_i$] hang on the wall?

     b. ?*What$_i$ is [to park there t$_i$] illegal?

(2) Who$_i$ did you see [a picture of t$_i$]?

Since Ross (1967) and Chomsky (1973) the questions in (1) are standardly assumed to violate the '(Sentential) Subject Condition'. Another well-observed fact, sometimes referred to as an 'Adjunct Condition effect', is that it is difficult to extract out of structural adjuncts (cf. Cattell (1976)). Those include temporal, result, comparative, concessive, non-restrictive and restrictive relative, appositive, conditional *if*-clauses, among other (see Postal (1999) for a fuller typology of clauses that fall under the adjunct category). The following is an example involving a temporal clause-adjunct:

(3) ?*Who$_i$ did Mary cry after Peter hit t$_i$?

24

Superseding various analyses of (1) and (3) in the EST/early GB literature, Huang (1982)

makes an influential proposal that the unacceptability of both types of constructions

should be accounted for in a unified manner, namely, as violations of a more general

Condition on Extraction Domain (CED):

(4) *Condition on Extraction Domain* (Huang (1982), p. 505):

A phrase A may be extracted out of a domain B only if B is properly governed.

The CED recognizes a crucial distinction between structural complements such as direct

objects and non-complements - subjects and adjuncts - with respect to extraction, noted

originally by Cattell (1976). This distinction is expected under the CED given that

objects, in GB terms, are 'properly governed' domains, while subjects and adjuncts are

not.[1]

Chomsky (1986) incorporates the CED phenomena into the *Barriers* system,

reducing CED to the principle of Subjacency that restricts the number of barriers crossed

upon movement. Just like CED, the *Barriers* system is designed around the

complement/non-complement distinction. This distinction is ultimately encoded in the

notion of L-marking, that is, theta-marking by a lexical head. Complements are L-

marked, while subjects and adjuncts are not, and by their virtue of not being L-marked,

---

[1] Note also that under the CED, structural conditions on extraction domains and those on a trace of an extracted element are fundamentally homogeneous. See also Cinque (1990) for elaboration of this possibility.

non-complements become barriers for extraction (see Introduction). Similarly, Chomsky and Lasnik (1993) [p. 79]"... take a barrier to be an XP that is not a complement".

Implementing the complement/non-complement distinction in the Barriers approach in this manner leaves an important conceptual question unanswered: why should extractability out of a domain depend on its being L-marked? As we pointed out in the Introduction, this connection is unexpected, given the classical GB architecture of grammar dominant at the time (Chomsky (1981)), in which Theta theory and Bounding theory (responsible for locality of movement) are different modules of core grammar, driven by separate sets of principles. In the Barriers theory, on the other hand, the connection is stipulated, and woven into the relevant definitions, but not given an adequate explanation.

In the dominating bias for a 'unified' approach, the 'connectedness' theory in Kayne (1983) deserves a separate mention (see also Kayne (1981)). Kayne's theory accounts for 'subject condition' effects in languages like English by capitalizing on the fact that those and related effects arise when extraction is taking place from a left-branch. Kayne then sets up his system in such a way as to preclude extraction from a left branch, by appealing to the notion of 'canonical government configuration' defined in linear terms.[2] Interestingly, as observed by Longobardi (1985), this theory, as formulated, does not explain an 'adjunct condition' effect in extraction out of adjuncts, most of which are right branching. These effects are thus presumed to be relegated to some different module of grammar. Longobardi (1985) proposes a technical modification to Kayne's theory to

---

[2] More generally, for Kayne, extraction is legitimate if the 'path' between the extracted element in its landing site, and its trace, consists exclusively of maximal projections satisfying certain structural conditions, defined, among other, via 'canonical government configuration'.

extend it to adjunct cases. But this extension, crucially, incorporates 'adjunct condition' cases by referring to a notion *other than* canonical government configuration which regulates extraction out of subjects (more concretely, the notion of government). Thus, inadvertently or not, the 'connectedness' approach in both original and revised formulation actually gives up the 'unified' spirit.[3]

In the minimalist framework, notions like 'government' or 'barrier' are no longer pertinent. On minimalist grounds, several authors, in particular, Nunes and Uriagereka (2000) proposed analyses of the contrast between (1)-(3) vs. (4) which preserve the 'unified' approach to subjects and adjuncts. In those analyses, the distinction between objects, on the one hand, and subjects and adjuncts, on the other, is encoded in terms of their derivational behavior in the course of structure building. According to these authors, a subject or an adjunct, but not an object, is a phrase marker that has become 'inaccessible' to the computational system at the derivational stage preceding the point at which extraction out of this phrase marker is supposed to take place. Note that this 'derivational' approach to CED phenomena dispenses with the stipulated connection between extractability of a domain and its being theta-marked, encoded in Barriers, thus potentially promising a more elegant and appealing theory.

However, virtually all approaches that treat subjects and adjuncts uniformly, as non-complements, including traditional (CED/Barriers) and minimalist, face a serious empirical challenge. The obvious consequence of treating subjects and adjuncts in a

---

[3] In this sense, we can regard the connectedness approach as a precursor of our 'eclectic' approach to extractability, developed in Chapters 2 and 3.

uniform manner is that extraction out of adjuncts is banned whenever extraction out of subjects is, and vice versa.[4]

The expectation is not fulfilled. In a number of languages, listed in Section 1.2, extraction out of subjects is allowed, whereas extraction out of adjuncts is banned. The fact that the languages in question are unrelated suggests that the divergence in the extractability pattern is not an accident but a systematic phenomenon that must be accounted for in any theory of A'-locality.

The existence of such languages presents a challenge for the theories of extractability which treat subjects and adjuncts in a unified manner. At a more fundamental level, it raises a concern as to whether the complement/non-complement distinction is indeed the relevant distinction for extractability. At any rate, the distinction is not imposed on us on conceptual grounds. And *a priori*, it may or may not be relevant. In order to address this concern, we will explore both alternatives within the Minimalist framework. Under the first alternative (the distinction is relevant; only complements are extractable domains) the extractability out of subjects in languages from Section 1.2 implies that subjects in those languages are 'complements'. We call this a 'neo-unified' approach. The other alternative is to abandon the distinction as irrelevant, and seek different and non-overlapping theories of extractability out of subjects and adjuncts. We call this an 'eclectic' approach. This chapter presents and explores the 'neo-unified' approach. Chapters 2 and 3 introduce the 'eclectic approach'. Merits of both approaches will be discussed.

---

[4] Kayne's *Connectedness* approach, as amended by Longobardi (1985), may not necessarily have this consequence. This is expected, because, as we pointed out, it is not committed to the 'unified' view.

In Section 1.3 we consider how pre-minimalist unified accounts fare with respect to the full paradigm of extractability, including examples from languages listed in Section 1.2. Section 2 discusses the 'neo-unified' approach, including its possible extension to languages in Section 1.2 in a way that salvages Cattell's complement/non-complement distinction.

1.2 Empirical challenge

Given two types of non-complement domains (subjects and adjuncts), there are four combinatorial possibilities with respect to which of those are 'permeable' with regard to extraction:

(5) 1)   subjects are non-permeable; adjuncts are non-permeable;

    2)   subjects are permeable; adjuncts are non-permeable;

    3)   subjects are permeable; adjuncts are permeable;

    4)   subjects are non-permeable; adjuncts are permeable;

English and many other languages realize possibility 1). Since both types of domains behave uniformly, it is reasonable to suppose that they form a homogeneous class that must be dealt with in a uniform manner, indeed the direction that led to postulation of CED and the *Barriers*/Subjacency analysis.

However, this possibility does not exhaustively cover cross-linguistic variation. Indeed, possibility 2) - subjects permeable, adjuncts not - is realized as well. The literature contains mentions of various individual languages in which it is possible to

extract out of subjects, but not adjuncts. However, the relevant data have not, as far as we know, been systematized into a larger cross-linguistic pattern. The following is a list of examples from a number of languages that realize possibility 2): Japanese, Basque, Navajo, Turkish, Palauan, Hungarian and Russian.[5] In all the languages cited extraction out of object is grammatical. Examples involving relativization, comparative deletion in Japanese, and topicalization in Palauan, are assumed to involve operator movement, all reminiscent of the regular *wh*-movement. [Su] marks extraction out of subject, [Ad] out of adjunct.

- JAPANESE (Kikuchi (1987), Takahashi (1994), see also Ross (1967), Kuno (1973))

(6)   [OP [Mary-ga   *t*   yonda no]-ga   akirakana yorimo John-wa   [Su]

         Mary-Nom   read   that-Nom is-obvious than   John-Top

         takusan-no hon-o   yonda

         many-Gen book-Acc read

         '(*) John read more books than [that Mary read _] is obvious'

---

[5] Other languages that realize possibility 2) include Hindi and certain Austronesian languages, e.g. Malagasy. Cinque (1978) also cites Armenian as a language not observing (sentential) Subject Condition effects. We have no data at hand to support or refute this claim.

     French has sometimes been claimed to violate the Subject Condition (cf. e.g., Sportiche (1981)). But Tellier (1990) points out that the all alleged violations are restricted to relativizations involving *dont* 'whose', which, she argues, are orthogonal for checking Subject Condition effects. If one controls for this by using examples not involving *dont* (such as *de qui, de lacquelle* 'of whom'), Subject Condition effects obtain, as expected. On *dont* relativizations see also Godard (1985).

     Kitahara (1994) cites an example of extraction out of a non-sentential subject in Icelandic, marked '?' and judged grammatical. This piece of data is puzzling, given that extraction out of sentential subjects in Icelandic is always degraded (cf. Zaenen (1985)).

(7)   *[OP Bill-ga [Mary-ga t   yonda kara]   odoroita   yorimo]   [Ad]

          Bill-Nom Mary-Nom read   because was-surprised than

        John-wa takusan-no hon-o   yonda

        John-Top many-Gen book-Acc read

        '*John read more books than Bill was surprised because Mary read'

- BASQUE (De Rijk (1972), Lara Reglero p. c.)

(8) a. Amorratuak dirala   bixtan dagon   zakur oiekin   ez det   [Su]

        rabid-det.pl are-that sight-at is-that (rel) dog   those-with neg Aux

        ibili nai

        walk want

        'I don't want to walk with those dogs that it is obvious are rabid'

        (*lit.* I don't want to walk with those dogs that [that _ are rabid] is obvious)

    b. (?) Mireni   gustatzen   zaizkiola argi dagoen zakur hoiekin   [Su]

        Mary-dat like-imperf. Aux-rel clear is-rel   dog   those-with

        ez   dut   ibili nahi

        neg Aux walk want

        'I don't want to walk with those dogs that it is obvious that Mary likes'

        (*lit.* I don't want to walk with those dogs that [that Mary likes _] is obvious)

(9) *Mireni   gustatzen   zaizkiolako Jon   harrituta zegoen zakur hoiekin   [Ad]

        Mary-dat like-imperf.   Aux-because John shocked was-rel dog those-with

ez   dut   ibili nahi

neg Aux walk want

'*I don't want to walk with those dogs that John was shocked because Mary liked'

- NAVAJO (Platero (1974), Barss et al. (1991)):

(10) ?Łééchąą'í iisxį -(n)í   gíí   shi-ł   bééhózin-ígíí   nahał'i   [Su]

   dog   perf:3:kill (something)-NOM me-with is:known-REL imp:3:bark

   'The dog that I know to have killed (something) is barking'

   (*lit.* the dog that (the fact) that _ killed something is known to me, is barking)

(11) *Ashkii yah'ííyáa-go   hadeeshghaazh-ęę   sitsilí   át'é   [Ad]

   boy   entered-COMP I:shouted:out-REL my:younger:bro   is

   'The boy that I shouted when t came in is my younger brother"'

- TURKISH (Hankamer and Knecht (1976), Kural (1993), Leyla Zidani-Eroğlu p.c.)

(12) a.   [Opᵢ [Ahmet-in tᵢ   git-me-si]-nin   ben-i   üz-dü-ğ-ü]   ev   [Su]

        Ahmet-GEN go-inf-agr-GEN I-ACC   sadden-pst-comp-agr house

   lit. 'The house [which [that Ahmet went to _] saddened me]'

   b. [Opᵢ [pro [[tᵢ anne-si]-nin   herkes-le konuş-tu-ğ-u]-nu m   [Su]

           mother-agr-GEN everyone-with talk-pst-comp-agr-ACC

      duy-du-ğ-um] adam

      hear-pst-comp-agr   man

lit. 'The man [whose I heard [that [_ mother] talked to everyone]]'

(13) a. *pro [[Ahmet   t$_i$ yediği]   için] sana   kızdim   pastayı$_i$ [Ad]

1sg Ahmet-NOM   eat-past-3sg for you-DAT anger-past-1sg cake-ACC

'I got angry with you because Ahmet ate the cake'

b. *[[Ali sat-týk]-tan   sonra] pro rahatla-dýð-ým ev   [Ad]

Ali sell-nom-abl   after   relax-nom-1sg   house

'The house that I relaxed after Ali sold _'

- PALAUAN (Georgopoulos (1991)):

(14)   Mary [a kltukl [el kmo ng-oltoir   er a John __ $_i$]]   [Su]

Mary R-clear Comp R-3s-Im-love   John

'Mary, [that __ loves John] is clear '

(15) a. *[a di mla se'elik el se'al] a ku-rael] er a party le u'ul ng-mla er ngii   [Ad]

my old boyfirend   IR-1s-left party because 3s-was there

'My old boyfriend, I left the party because _ was there'

b. *ng - ngera a 'omurael er a party le u'ul   rebek el 'ad a meruul   [Ad]

what   IR-2-left   party because every   man R-do

'*What did you leave the party because everyone was doing?'

- HUNGARIAN (Kiss (1987), Eva Bar-Shalom p. c.)

(16) Melyik színésznőnek₁ gondolja János, hogy t₁ a fényképe meglett?   [Su]

which actress's   thinks   Janos that the picture-her turned up

'Which actress does John think that a picture of _ turned up?'

(17) *Milyen színésznőt szalatt el János mielőtt Péter látta?   [Ad]

which actress   ran-away Janos before   Peter saw

'Which actress did John ran away before Peter saw _?'

- RUSSIAN

(18) a. S   kem   by ty xotel   čtoby   govorit' bylo by odno udovol'stvie? [Su]

With whom sbj. you   wanted that-sbj. to speak were sbj one pleasure

lit. 'With whom would you want that [to speak _] were sheer pleasure'

b. Cto by ty xotel čtoby   kupit'   ne sostavljalo by nikakogo truda?   [Su]

what sbj you wanted that-sbj. to-buy not constitute sbj. no   labor

lit. 'What would you want that [to buy _] would not be any trouble'?

(19) *S   kem   Ivan rasserdilsja potomu čto Petr vstretilsja?   [Ad]

With whom John got-angry because   Peter met

'*With whom did John get angry because Peter met?'

The languages represented above are more or less standardly regarded in the literature as thematic *pro*-drop languages (see the relevant references above for discussion).[6] Note that the grammatical instances of extraction out of subjects in the languages reported above cannot be analyzed as involving no extraction at all, and instead, phonologically silent *pro* in the object position. If this were so, then, one would expect that a similar option would salvage the cases of extraction out of adjuncts as well. We take the examples above to represent genuine instances of extraction.

In fact, extraction out of subject is possible in certain cases even in English, provided the subject is postverbal, as shown in the following example:

(20)   Which president was there a picture of *t* on the wall?

The importance of examples like (20) will become clear in the following section.

Possibilities 3) and 4) in (5) imply extractability out of adjuncts. Significantly, there seem to be no languages in which extraction out of adjuncts (not obscured by intervening factors like the availability of resumptive pronouns) is possible. Until convincing evidence to the contrary is found, we regard this state of affairs as an indication that those two possibilities are not realized in natural language.

The following is the resulting cross-linguistic picture, with respect to extractability out of non-complements:

---

[6] Russian is characterized as only a non-thematic *pro*-drop language in Franks (1995) (in other words, it has phonologically null expletives). But see Benedicto (1993) for the view that Russian is a full (thematic and non-thematic) pro-drop language.

(21)

	Theoretical Possibility	Realized?
1	subjects are non-extractable; adjuncts are non-extractable	YES
2	subjects are extractable; adjuncts are non-extractable	YES
3.	subjects are extractable; adjuncts are extractable	NO
4.	subjects are non-extractable; adjuncts are extractable	NO

1.3 Problems With the Traditional Unified Approaches

Note that the need for something like the CED or *Barriers* in the first place follows from the general tendency to <u>constrain</u> the grammar. This tendency is inherent in the standard approach to locality, designed to prevent generation of sentences that would otherwise be generated - in our case, sentences involving extraction out of subjects in languages like English. The standard view then needs to explain why the constraint does not apply in languages in which extraction out of subjects is possible.

Under the analyses based on CED or *Barriers*, the existence of languages displaying the diverging pattern with respect to extractability out of subjects and adjuncts is a priori unexpected. This is so because these analyses fundamentally rely on some concept unifying 'non-complements' - subjects and adjuncts, as opposed to 'complements' - objects. In the *Barriers* system, this concept is L-marking (that is, theta-marking by a lexical head): because subjects are not complements, they are not L-marked, hence, predicted to be barriers for extraction cross-linguistically. The notion of L-marking cannot be (trivially) parameterized so as to allow subjects in the above languages to be L-marked and thus 'void' their barrierhood. The invariantly non-L-marked character of

subjects follows on the standard assumptions concerning VP structure, according to which the subject is located more remotely from the lexical head V than object.

The CED in Huang's original formulation (cf. (4)) allows a technical loophole to accommodate the above languages, as shown in Lasnik and Saito (1992). One can make subjects in those languages 'exempt' from the application of the CED if one stipulates that they are like objects, not like adjuncts, in that they are 'properly governed', in contrast to subjects in languages like English. Lasnik and Saito (1992) indeed suggest that subjects in Japanese remain in Spec-VP, where they are generated (in accord with the internal subject hypothesis), and are properly governed by Infl, whereas subjects in English and other languages raise from Spec-VP to Spec-IP, a position that is not properly governed by Infl. We agree with Lasnik and Saito (1992) in the insight that the extractability out of subject domains in Japanese is ultimately related to their non-movement from a VP internal position. We return to this insight in Chapter 2.

1.4 Toward a novel analysis

Even though the examples in section 1.2 constitute an empirical argument against unifying subjects and adjuncts under one concept, they do not necessarily argue against the complement/non-complement distinction as relevant for extractability. In order to account for the above examples while maintaining the distinction (i.e. only complements are extractable domains), we are led to view subjects in the above languages as 'complements' in some sense, for the purposes of extraction. Let us examine this idea, given the current assumptions concerning the architecture of grammar, and see if it can

correctly capture the 'CED effects' and their lack in the languages listed in the previous section.

2. A 'Neo-Unified' Approach

2.1 Outline

Recent minimalist analyses of the 'CED effects' in English include Nunes and Uriagereka (2000) (henceforth N&U) and Toyoshima (1997) (see also Epstein (1999), Uriagereka (1999)). These analyses share the underlying idea which we informally state in (22) and (23):

(22) A Neo-Unified Approach: If a phrase marker X was assembled in parallel with a phrase marker Y, and then X and Y were Merged, whereupon Y projects, no extraction is ever possible from X, since X becomes 'morphologically opaque'.

(23) X is *assembled in parallel* with Y iff there exists a derivational point at which X and Y co-exist in the derivational space, and are unconnected.[7]

The approach formulated in (22) and (23) in fact, resonates greatly with the 'connectedness' approach of Kayne (1983) (although is not quite the same, because of the 'adjunct condition' cases, see above). Instead of articulating this approach using Kayne's definitions (most of which are not suited for expositional purposes), let us make the idea

---

[7] (23) effectively excludes heads from consideration, under a natural assumption that heads must be Merged immediately upon their selection from the Numeration; they cannot co-exist with the other phrase marker, unconnected.

behind it more clear by invoking a color metaphor.[8] Suppose that each phrase marker is painted a unique color in the course of the derivation. Lexical items - heads - are drawn from the lexicon (into a Numeration) colorless. When two items are Merged together, they form a phrase marker/syntactic object that receives a specific color - say, red, to distinguish it from other phrase markers possibly being built in the same 'derivational space'. This is the insight behind the notion 'assembled in parallel' in (23). Further projections of the same head receive the same color (say, if V' is red, so is VP) to emphasize the fact that each projection belongs to the same phrase marker as the head. Since heads are taken from the lexicon colorless, Merging a head with the existing phrase marker creates an object that assumes the color of that phrase marker (e.g., Merging V with a red NP results in a red V').

Consider now (24), keeping to the standard minimalist assumptions concerning clause structure (cf. Chomsky (1995c), Ch.4):

(24)   *Who does [a picture of t_{who}] hang on the wall?

In the course of the derivation of (24), two phrase markers are built, the subject [$_{DP}$ *a picture of who*] and [$_{V'}$ *hang on the wall*]. Since those are built independently (in parallel) they each acquire a different color. Suppose the DP is painted red, and the V' green. The red DP is Merged with the green V', projecting the VP. Since the VP is a projection of (green) V', the VP is green and is now the only phrase marker in the derivational 'space'. The VP is then Merged with I (or T), becoming I', still green, as it is still the same phrase

---

[8] A different, more technically involved, metaphor can be drawn along the lines of Kayne (1983).

marker (and I/T is colorless). The red DP moves to Spec-I, and IP is projected, a projection of I', hence, still green. The IP is Merged with C, yielding a green C'. If *who* were to move to Spec-C at this point, that would be an instance of movement from a red phrase marker to a green phrase marker. The ungrammaticality of (24) suggests that (22) can be translated as follows:

(25)   No extraction is possible when its source and target are of different color

Let us assume, along with Nunes and Uriagereka (2000) and contra Larson (1988), that structural adjuncts are generated as modifiers of maximal projections. Then similar (even more straightforward) considerations apply in the case when the red phrase marker is an adjunct, accounting for Adjunct Condition violations (cf. ((3) above). On the other hand, consider (2), repeated here as (26):

(26)   Who$_i$ did you see [$_{DP}$ a picture of t$_i$]?

Suppose [$_{DP}$ *a picture of who*] is built and is assigned green color. Merger of this DP with V leads to creation of [$_{V'}$ *see a picture of who*], which is again green, since it is (part of) the same separately built phrase marker. It remains such until the end, when the C' is projected and *who* moves to Spec-C. Movement of *who* thus takes place within a singly colored phrase marker. The grammaticality of (26), along with the ungrammaticality of (24), suggests that (25) can be strengthened to (27):

(27)   *The Neo-Unified Approach (color version):*

Extraction is only possible when its source and target are of the same color

(27) effectively allows us to define an extraction domain in terms of the (same) color, and offers an informal paraphrase of (22) and (23). The 'red-green' distinction outlined above is implemented differently in different theories. We will concentrate on what we believe to be the most articulated theory of extraction domains in the minimalist literature, that in Nunes and Uriagereka (2000). We believe that this theory is most representative of the 'neo-unified' approach.[9]

2.2 Nunes and Uriagereka (2000)

Nunes and Uriagereka (2000) state the 'red-green' distinction as a distinction between a phrase marker that is linearized at PF prior to *wh*-extraction from it, and one that is not. According to Nunes and Uriagereka (2000), who adopt the Multiple Spell-Out theory of Uriagereka (1999), [$_{DP}$ *a picture of who*] in (24) must be linearized before Merger with [$_{V'}$ *hang on the wall*]. This early linearization is forced by a PF convergence requirement in the form of Kayne's Kayne (1994) Linear Correspondence Axiom (LCA), namely, that a linear order can only be established among syntactic objects which stand in asymmetric c-command relation. Informally, if [$_{DP}$ *a picture of who*] is not linearized before Merger with [$_{V'}$ *hang on the wall*], then the elements of the DP (*a, picture, of who*) will never c-command the elements of the V' (*hangs, on, the, wall*), and vice versa. Consequently, no

---

[9] Toyoshima (1997) introduces the notion of 'process' to mean, roughly, a separately built phrase marker. He proposes an *ad hoc* condition banning extraction across different processes, essentially a version of (27), as a minimalist counterpart of the CED.

linear order among the elements of the DP and the elements of V' can be established. On the other hand, if [DP *a picture of who*] is linearized and then Merged, then the label of the DP c-commands elements of [V' *hang on the wall*], hence, can be correctly linearized in relation to those. The linearized DP then is considered a morphological 'word' for the purposes of further syntactic processes. This means, for those authors, that the internal structure of that DP is 'inaccessible' for syntactic operations, such as extraction, capturing the intuition in (25).

Similar considerations apply in adjunct cases. Consider (3), repeated here as (28):

(28) ?*Who$_i$ did Mary cry [after Peter hit t$_i$]?

Here [vP *Mary cry*] and [CP *after Peter hit who*] are built independently. If [CP *after Peter hit t$_i$*] is not linearized before Merger with the vP, then the elements of the vP (*Mary, cry*) will never c-command the elements of the CP (*after, Peter, hit, who*); thus, no linear order can eventually be established. On the other hand, if the CP is linearized and then Merged with the vP, no such problem arises. The CP is then considered a gigantic 'word' that itself enters a c-command relation with the elements of vP, ensuring subsequent full linearization.

In contrast, [DP *a picture of who*] in (26) need not be linearized before it is Merged with *see*. This is so because this Merger establishes c-command relations between *see* and the elements of the DP. The same is true when the resulting phrase marker participates in further Mergers, up until the Merger with C, creating a C'. The PF linearization procedure need not apply before this point. Consequently, [DP *a picture of who*] is 'transparent' for

syntactic operations in the course of the derivation. In particular, *who* can be extracted from it.[10]

However, N&U's system entails an intrinsic redundancy. Consider the sentence in (29), involving a passivized subject:

(29)   ?*Who$_j$ was [$_{DP}$ a picture of t$_j$]$_i$ taken t$_i$ by Bill?

In a cyclic derivation of (29) [$_{DP}$ *a picture of who*] moves to Spec-TP before *who* is extracted. Now, N&U crucially assume a copy theory of movement, according to which t_i is a silent copy of [$_{DP}$ *a picture of who*] in the base generated position. N&U further assume that the higher copy is linearized before it is Merged in Spec-TP. This makes it opaque for the extraction. But in N&U's system, the silent copy of *who* is potentially available for extraction as well. This opens the door to a good derivation whereby *who* in (29) comes from the lower copy of [$_{DP}$ *a picture of who*]. To preclude this derivation, N&U appeal to a version of the condition on uniformity of chains (cf. Chomsky (1995c)), according to which the links of a chain must be uniform.[11]

---

[10] An account along similar lines can be envisioned in the strictly derivational system of Epstein (1999); cf. also Epstein et al. (1998).

[11] More specifically, N&U show that even in simple passives, chain uniformity must be involved:
(i)   Someone was arrested ~~someone~~
According to these authors, in the process of raising, a copy of *someone* is created, Spelled-Out at PF, and then inserted into the subject position. Movement of the underlying object to the subject position creates a non-uniform chain: the higher copy is Spelled-Out, and becomes a 'morphological word', the lower one is not. To explain the grammaticality of (i), N&U assume that *someone* is Spelled-Out <u>before</u> initial Merger with *arrested* (for reasons of convergence), and a copy of the Spelled-Out object Merges into the subject position, so the resulting chain is uniform. Similar considerations, presumably, apply in the case (29).

However, Takahashi (1994) shows that the chain uniformity condition suffices to rule out not only (29) but 'Subject Condition' effects everywhere, even in the baseline examples (1), under the standardly assumed predicate internal subject hypothesis. To illustrate, let us go back to (1a). In Takahashi's theory, movement proceeds by a series of successive adjunctions of the moving element to the maximal categories along its path (in accord with the condition Shortest Move, cf. Chomsky (1995c)). In (1a), [DP *a friend of who*] raises from Spec-*v*P to Spec-TP, forming a two-linked chain ([DP *a friend of who*], [DP *a friend of who*]). Then *who* is extracted from the higher link, and adjoined to it (on its way to the matrix Spec-CP), making a shortest move. This results in a chain ([DP *who* [DP *a friend of who*], [DP *a friend of who*]). This chain is not uniform because its links have different structure (similar considerations obtain in case of extraction from the lower link). Consequently, (1a) is ruled out. Similar considerations obtain for (29) and other cases of extraction from derived positions. We return to the chain uniformity analysis in Chapter 2.

Given this redundancy, N&U's account of non-extraction out of subject in (1) becomes superfluous, and may be dispensed with. But given the fact that subjects and adjuncts are treated in a unified manner in their theory, abandoning their account of subjects automatically means losing their theory of (non-)extraction out of adjuncts as well. This is a payoff for adopting the unification view in N&U's system.

In addition to this theoretical shortcoming, N&U's theory faces empirical problems. In fact, as we show immediately, these problems arise not just in the N&U's theory per se, but, rather, for the neo-unified approach as stated in (22) more generally.

2.3 Problems With the Neo-Unified Approach

As we saw above, (22) and (23) provide a new, derivational criterion distinguishing subjects and adjuncts ('non-complements'), as opposed to objects (complements). This criterion re-unifies subjects and adjuncts under one conceptual umbrella. However, by re-unifying subjects and adjuncts it inherits the major empirical problem of the pre-minimalist accounts of CED phenomena: it cannot distinguish between languages like English and the languages in Section 1.2, in which extraction out of subjects, but not out of adjuncts, is possible. The problem is undergeneration: examples involving extraction out of subject are wrongly predicted to be ungrammatical.

Consider, for instance, the Turkish example (12a), repeated here:

(12) a. [Op$_i$ [Ahmet-in t$_i$ git-me-si]-nin   ben-i   üz-dü-ğ-ü]   ev

        Ahmet-GEN go-inf-agr-GEN I-ACC   sadden-pst-comp-agr house

    lit. 'The house [which [that Ahmet went to _] saddened me]'

The clausal subject *Ahmet-in Op git-me-si* is a phrase marker with internal structure, and so is the V' (or *v'*) *ben-i üz-dü-ğ-ü*. Hence, by (23) they are assembled in parallel, and when they are Merged, the label of V (or *v*) projects. By (22), extraction from the clausal subject must be prohibited, contrary to fact. Similar considerations extend to other languages discussed in Section 1.2.

Naturally, the theory in N&U, implementing the approach in (22) and (23), inherits this problem as well. For N&U, both adjuncts and subjects are first linearized (which

makes them 'inaccessible') then Merged with the rest of the structure (cf the discussion in the previous section).

If one tries to accommodate the languages from Section 1.2, given N&U's implementation of the complement/non-complement distinction, one is virtually forced to say that subjects in those languages are phrase markers that need *not* be linearized at PF before Merging with the rest of the structure, just like complements in this theory. Under the usual assumptions about clause structure, specifically, concerning the point of insertion of subjects (at Spec-*v*P), we do not see any obvious way to state this accommodation in N&U's system. The reason for that is that N&U's analysis is too 'principled': it is couched in notions which do not seem parameterizable: phrase structure, LCA and linearization.

The neo-unified approach in the form of (22) and (23) also faces another empirical problem, concerning extraction from certain subjects even in English where the 'subject condition' otherwise holds. This is illustrated in (20), repeated here as (30), with the underlying structure in (31):

(30)   Who is there a picture of *t* on the wall?

(31)   there is [$_{SC}$ [$_{DP}$ a picture of who] [on the wall]]

The DP in (31) is Merged with the predicate of the small clause, both being phrase markers assembled in parallel, by (23). Whatever the label of the resulting constituent is, it is not that of the DP (the subject of the small clause), since small clauses do not have

the same distribution as DPs.[12] Since the label of the DP does not project upon Merger, (22) precludes extraction out of the subject DP.

Despite these problematic issues, suppose we insist that the complement/non-complement distinction is nevertheless correct, and only complements are in principle 'transparent' to extraction. Then we face two questions: one general, one specific. The general question is: how can the possibility that extractability out of subjects in languages discussed in Section 1.2, as well as out of subjects of small clauses in English, be reconciled within the neo-unified approach? Let us consider what seems to us to be the best possible extension of the neo-unified approach to the languages in which extraction out of subjects is possible.

2.4 An Extension of the Neo-Unified Approach

Let us repeat the basic idea behind the neo-unified approach:

(22)   A Neo-Unified Approach: If a phrase marker X was assembled in parallel with a

phrase marker Y, and then X and Y were Merged, whereupon Y projects, no

extraction is ever possible from X.

(23)   X is *assembled in parallel* with Y iff there exists a derivational point at which X

and Y co-exist in the derivational space, and are unconnected.

---

[12] For different proposals concerning the label of the resulting constituent, see Manzini (1983), Stowell (1978), (1981), Bowers (1993).

In effect, (22) can be taken as a definition of non-complement (for extraction purposes): a non-complement is a phrase marker built in parallel with another phrase marker, *and*, when the two Merge, it is the label of the other phrase marker that projects.

If we maintain the complement/non-complement distinction as the relevant distinction for extractability, we see only one conceivable way to account for extractability out of subjects in Section 2. We envision the following general scenario for implementing this idea.

(32) a. Instead of Merging at Spec-*v*P, subjects in those languages are Merged at the point where objects would Merge, that is, as a sister of V;[13]

   b. The subject stays there in the course of the derivation.

In a simple transitive clause, then, the subject and object would Merge simultaneously with V, resulting in a ternary structure:

(33)   [$_{VP}$ V DP$_1$ DP$_2$]

None of the DPs in (33) qualifies as a non-complement according to (22). Indeed, each DP is built in parallel with another DP (given that each DP has internal structure), but it is

---

[13] For instance, one may argue that there is no Spec-*v*P to Merge the subject into in these languages, because the category *v* is not utilized in these languages. This conjecture is in line with a recent suggestion in Lasnik (1999) that a functional category (AgrO for Lasnik) may or may not participate in the derivation, and the course of the derivation is adjusted accordingly in each case. This opens up a window for cross-linguistic variation: it is possible that certain languages may not use certain functional categories. On the other hand, in the system of Chomsky (1995b), (2000), the absence of *v* would imply completely different theta-relations (namely, no agents) in a language, and is unlikely to be a viable option. Some evidence that the subject theta-role may be assigned directly by V is given in Sabel (2001) for Malagasy.

not the case that when the two Merge, it is the label of the other DP that projects. Rather, in all cases, the label of the <u>head</u> V does. (The same is true if there are more than two DPs.) Hence, each DP in (33) is not a non-complement, which means that extraction out of it should be possible.[14] Note that this line of reasoning does not extend to adjuncts, which we continue to assume to be generated as modifiers of maximal projections. By (22), adjuncts still qualify as non-complements.

Note that (33) makes a claim only about the structure of VP, which is in some sense 'flat', or 'non-configurational'. This does not entail, in any way, that languages which would entertain (33) are also 'non-configurational' at other levels. Rather, what is at issue here is a clausal structure which is fully 'configurational' (hierarchical) except for the VP component. For example, this might be the structure that Kiss (1987) defends for Hungarian.

Let us now consider the second requirement, that the subject must stay in the base-generated position. This requirement is necessary, since extraction out of moved, or derived, subjects must be precluded independently. We know that because of the ungrammaticality of examples like (29) involving passive subjects in English:

(29)   ?*Who$_j$ was [$_{DP}$ a picture of t$_j$]$_i$ taken t$_i$ by Bill?

---

[14] The color metaphor illustrates this conclusion. Merging (colorless) V with two DPs of a different color, e.g. red and green, creates an object which assumes *both* these colors. The resulting VP is thus bi-colored, red and green. Consequently, Merging this VP with I creates a bi-colored IP; Merging this IP with C creates a bi-colored C'. Now notice that extraction out of either the red or green DP into Spec-C is possible, by (27). This is so because an element extracted out of either domain will land in a position of the same color as its departing site. (27) says nothing about additional color(s); hence those simply do not matter for extraction out of the respective domains.

Earlier, we criticized N&U's account of (29) in terms of chain uniformity, since it brings in substantial redundancy in their analysis. One might discover an alternative account of (29) in a theory like N&U, avoiding this redundancy. But whatever that alternative account is, it will automatically rule out extraction out of subjects in the languages under consideration, if those subjects undergo movement from their position inside VP in (33). We are then led to postulate that subjects (even passive) must, or at least can, stay in situ in those languages.[15]

Thus the neo-unified approach in principle allows for a possibility, in the form of (33), that subjects may behave as complements with respect to extraction. Before we consider the empirical motivation for the scenario in (32), let us address the specific concern as to whether and how (33) can be generated in an LCA-based system such as N&U.

The theory in N&U would allow both DPs not to be linearized before Merging with V (see Section 2.2), thus rendering them 'transparent' for extraction. However, generating the structure in (33) is not straightforward in N&U's system to begin with. The reason is that this structure is not binary branching and, as such, seems incompatible with the LCA, which is at the core of N&U's account.

Let us take a closer look at the situation. Recall that according to the LCA, a linear order can only be established among syntactic objects which stand in asymmetric c-command relation. Linear order mirrors asymmetric c-command: roughly, if *A*

---

[15] An alternative is that subjects raise, but only after the *wh*-extraction. This derivation is, however, counter-cyclic and thus should be precluded on independent grounds.

asymmetrically c-commands *B*, then (elements of) *A* precede (elements of) *B*, and conversely.

An *n*-ary structure as in (33) appears incompatible with LCA for the following reason. Suppose in (33) DP_1 and DP_2 are lexical items. Then V doesn't asymmetrically c-command DP_1 and DP_2; DP_1 doesn't asymmetrically c-command V and DP_2; and DP_2 doesn't asymmetrically c-command V and DP_1. Accordingly, no linear order between V, DP_1 and DP_2 can be established. Suppose DP_1 and DP_2 have internal structure. Again, DP_1 doesn't asymmetrically c-command DP_2, and DP_2 doesn't asymmetrically c-command DP_1. The ordering of elements of DP_1 and DP_2 cannot be established.

In fact, even a simple binary structure is prone to this effect of the LCA. Let us assume, following Chomsky (1995a) that LCA does not apply everywhere in the derivation, as in the original formulation of Kayne (1994), but, rather, at the PF component (after Morphology). Consider (34) involving a trivial (single-terminal) complement DP, assuming bare phrase structure of :

(34)   [$_{V'}$ saw it]

In (34), neither *saw* nor *it* asymmetrically c-commands the other; hence, cannot be ordered. To avoid this undesirable effect, Chomsky (1995a) suggests that *it* in this situation must raise, leaving in the original position a silent trace (or copy) which, arguably, does not need to be ordered, and so is ignored (or deleted) by the LCA.[16]

---

[16] As Chomsky points out (p. 437, fn. 35), raising of the verb does not exempt the complement from being raised. In our case, if *saw* is raised (and its trace ignored or deleted by the LCA), but *it* has not, then *it* still

But notice that once we have this option of avoiding the unwanted application of the LCA by raising, this same option is available in the case of the *n*-ary structure (33), even if it involves non-trivial DPs (i.e. with internal structure). That is, if both DPs in (33) have raised by the time LCA applies, then what will be left for the LCA to deal with is the verb and two traces, which, again, do not need to be ordered. In effect, this option in principle makes not only a binary, but also an *n*-ary structure compatible with the LCA.

Thus it seems in order to salvage the derivation involving the 'flat' structure in (33), one needs to postulate, in particular, that the subject must raise out of VP, for the LCA to be able to apply. A natural way of looking at this raising in a system like N&U is that such raising happens for usual syntactic reasons, such as feature checking, and LCA simply evaluates the output of that raising, submitted to the PF component.

But then we arrive at a contradiction. According to the reasoning so far, the subject must raise out of VP in (33). But the requirement (32b), which a system like N&U needs to even begin to approach the languages under consideration, states exactly the opposite, namely, that the subject must *not* raise. Thus, Chomsky's suggestion concerning circumventing the undesired LCA effect by syntactic raising is unlikely to help in this case.

Juan Uriagereka (p. c.) suggests that the contradiction can be avoided if the raising at issue does not take place in syntax, but at PF, and prior to the application of the LCA. In fact, this movement may happen exactly for the reason of satisfying the LCA, by breaking symmetric c-command. This suggestion is in line with a 'dynamic antisymmetry' approach

---

remains a single-terminal complement, either to V itself or to an intervening element, e.g. the next element up the tree.

recently pursued in Moro (2000). The possibility of PF movement may or may not be on the right track (for arguments against PF movement, see Bošković 2001; for arguments for it, see, e.g. Boeckx and Stjepanović (2001), Zubizarretta (1998)). The correlation between extractability out of subjects and subject raising at PF requires, of course, independent argumentation. Since we will ultimately arrive at a conclusion questioning the 'neo-unified' approach altogether, we will not pursue this possibility here.

We now turn to evaluating the empirical basis for the general scenario (32). The first part of the scenario (32) - the configuration in (33) - implies the absence of subject-object asymmetries in the languages from Section 2 due to the absence of structural hierarchy between subject and object.[17] Together with the second part of the scenario, a stronger prediction is made that the asymmetries must be absent in the *derived* structure. On the other hand, the existence of a subject-object asymmetry at *any* point would suggest that the subject has raised from its base-generated position, or, alternatively, that the structure of VP is hierarchical. Either one of these possible situations is sufficient to refute the scenario in (32).

During the past two decades or so, significant effort has been put, with considerable success, to establish the existence of hierarchical VP component in the languages listed in Section 1.2, on the basis of various subject-object asymmetries. Speas (1990) lists at least ten tests for configurationality, summarizing previous arguments from the literature and

---

[17] In addition, it implies the absence of structural hierarchy between direct and indirect objects, and other kinds of argument phrases, which are standardly assumed to exist in languages like English on the basis of certain binding asymmetries (cf., e.g. Barss and Lasnik (1986), Larson (1988)).

Another prediction made by the configuration in (33) is the absence of ergative/non-ergative distinction for intransitives, associated with the fact that in the latter, but not in the former, subjects must be 'external' to the argument structure of the verb (cf. Hale and Keyser (1993)). We know of no evidence corroborating this expectation for the languages under consideration.

adding several others (cf. Appendix). Most clear arguments for the subject-object hierarchy come from asymmetries in A-binding (subject can A-bind (into) the object, but not conversely), weak crossover effect (suggesting that the object does not c-command the subject), and licensing parasitic gaps by an A'-trace of the object (observing the anti-c-command requirement, see Chomsky (1986)), but not the subject. Some, or all of these, and further arguments, have been adduced in the literature on relevant languages. See, in particular, Saito (1985), Whitman (1986) for Japanese, Mahajan (1990) for Hindi, Georgopoulos and Ishihara (1991) for Palauan, Kural (1997) for Turkish (see also Merchant 1993 on weak crossover in Turkish), Horvath (1986), Marácz (1989) for Hungarian (but see Kiss (1987), (1994) for the opposite view), Pesetsky (1982), Bailyn (1995), King (1993) for Russian. These works demonstrate, in detail, that the 'flat' configuration in (33) cannot be the right structure for these languages.

To recap, the most natural extension of the 'neo-unified' approach to account for extractability out of subjects in the languages of Section 1.2, faces a number of problems, mostly empirical (unsupported expectations entailed by the 'flat' structure (33)). To these, we should add the problems with the 'neo-unified' approach itself, both empirical (an incorrect prediction concerning examples like *Who is there a picture of on the wall?*, cf. Section 2.3), and conceptual (concerning the redundancy with extraction out of derived subjects, cf. (29)). The range of problems considered casts doubt on the correctness of the 'neo-unified' approach in particular, and on the 'unified' approach more generally.

3. Conclusion

The traditional, pre-minimalist accounts of extractability explained the non-extractability out of subjects and adjuncts in languages like English by treating these two types of domains in a uniform manner. The idea behind the unification lies in the complement/non-complement distinction (cf. Cattell (1976)): both subjects and adjuncts are non-complements, whereas direct objects are. We argued that the pre-minimalist proposals implementing this 'unifying' idea, namely, via CED or Barriers, are problematic conceptually, since they stipulate a non-obvious connection between extractability out of a domain and external notions such as L-marking, defined via theta marking. More importantly, they are problematic empirically, since they lead to massive undergeneration for languages in which extraction out of subject is possible, as discussed in Section 1.2. A 'neo-unified' approach, explored in the recent minimalist theories, most prominently in N&U, inherits the empirical problems, since it seeks to retain the 'unified' spirit of 'CED'/Barriers. In order to incorporate languages in which extraction out of subjects is possible, one must postulate that subjects in those languages are 'complements', in the sense that a 'neo-unified' approach like N&U makes precise. However, we showed that the most natural extension of the 'neo-unified' approach to account for the extractability out of subjects in the languages under consideration, itself faces a number of empirical and conceptual problems. This points to the conclusion that the 'neo-unified' approach, at least in its current stage, is not capable of accounting for the entire cross-linguistic range of extractability patterns.

Thus, the previous unified theories, both pre-minimalist and minimalist, have not been successful in *stating* the 'unified' theory of extractability out of subjects and

adjuncts. Perhaps, the reason why it is so difficult to correctly state a 'unified' account so as to capture cross-linguistic variation in extractability lies precisely in the unification idea, that follows from the complement/non-complement distinction. As we pointed out in the beginning of this chapter, the distinction, and with it, the unification, is not imposed on us a priori, so it is at least possible that the distinction is simply immaterial for extractability.

Rejecting the idea of unifying subjects and adjuncts under one concept opens up a theoretical possibility that extractability out of subjects and adjuncts is regulated not by a single grammatical mechanism, but by two independent mechanisms of grammar. This possibility is certainly compatible with the empirical facts, namely, the cross-linguistic variation in extractability out of subjects as opposed to the apparently universal ban on (overt) extraction out of adjuncts (cf. Section 2). In the next two chapters we explore this possibility by developing an 'eclectic' approach, in which extractability out of different types of domains is regulated by different components of the grammar, in a non-overlapping manner.

APPENDIX: On (non-)applicability of certain 'configurationality' tests

In this appendix, we consider two tests that Speas (1990) devises on the basis of Marantz (1984). Speas claims that these tests distinguish the hierarchical structure and the 'flat' VP structure as in (33). We will explore the question whether this claim is indeed correct, and discuss some facts from Russian that the test bears on.

Marantz (1984) considers in detail mechanisms of formation of predicates and their representation at various linguistic levels. He argues that there exists an inherent

asymmetry in the way semantic roles (roughly, theta-roles, see Marantz's work for a more detailed discussion) are assigned to internal argument(s) of the verb, and logical subject, at the level of l-s structure - roughly, argument structure. In particular, the choice of the object in the argument structure of the verb affects the semantic role assigned to the logical subject. This can be demonstrated by the following:

(35)   a. kill a cockroach

         b. kill a conversation

         c. kill an evening watching TV

         d. kill a bottle

         e. kill an audience

The subject of the predicate in (35a) is a killer, but the subject of the predicate in (35b-e) is not a killer, even though the verb *kill* is, arguably, the same verb in all cases. Thus the meaning of the predicate strongly depends on the choice of object of the verb. On the other hand, the choice of subject argument does *not* influence the semantic role assigned to the object, as the following illustrates:

(36)   a. Harry killed NP

         b. The drunk refused to kill NP

         c. Cars kill NP

In (36), regardless of the choice of subject, the semantic role of the object is a thing

killed. Marantz brings these examples to support his claim that the argument structure of the verb provides a function from an (internal) argument to a predicate, and that the verb plus object form a predicate which assigns a semantic role to the subject, in a more or less compositional manner. Marantz further argues that the observed asymmetry is implemented at various levels of analysis, including syntax, provided the mapping between argument and syntactic structure is constrained on a principled basis (e.g. by the Projection Principle). Using Marantz's insight, Speas (1990) further argues that the asymmetry serves as a test for the syntactic structure in which verb and object form a constituent, to the exclusion of the subject.

What should happen in languages with 'flat' VP structures? Marantz points out, reasonably, that '... in a theory with symmetric argument structures, the choice of one argument can in no way affect the semantic role assigned to another argument of the predicate. All arguments are independent on par' (p. 23). On the assumption that the mapping between argument structure to syntactic structure is fairly straightforward, we expect, accordingly, that the asymmetry observed in (35) vs. (36) should not exist. The choice of the subject should not depend on the choice of the object, and vice versa. All else equal, the subject of predicates analogous to those in the analogues of (35) must always be a killer, and the object of the predicates analogous to those in (36) must be a killed object. More generally, in these languages a predicate would be formed when the verb is combined with all of its arguments simultaneously (cf. Jackendoff (1972)).

Speas (1990) sees a different prediction for the 'flat' VP languages from Marantz's account. She claims that in those languages 'S may influence theta role of O, and vice-versa' (p. 137). This clearly contradicts Marantz's vision of the relevant issues (see

above). It is true that if the asymmetry in 'influence' of the choice of one argument on the choice of the other somehow correlates with asymmetry in the hierarchical syntactic structure, the view which Marantz seems to endorse, then the absence of the latter should correspond to the absence of the former. But the absence of the asymmetry in the 'influence' of one argument on another in this particular case does not have to be understood to mean that subjects and objects should symmetrically influence each other's choice. Rather, it means the absence of 'influence' altogether. By the reasoning so far presented, any kind of 'influence' already suggests a hierarchical structure. It seems then that Speas' prediction amounts to a contradictory requirement that the verb and subject form a constituent, to the exclusion of the object, and also that the verb and object form a constituent, to the exclusion of the subject. The 'flat' structure cannot be at issue here.

Another one of Speas's configurationality tests is built around Marantz's discussion of idioms. An idiom, informally, is a 'chunk' of structure which carries a meaning largely independent of the meanings of its parts.[18] Marantz points out the preponderance of verb + object idioms in English, and the lack of idioms involving just the subject and the verb:

(37)   a. kick the bucket

         b. shake a leg

         c. chew the fat

---

[18] It is sometimes hard to distinguish idioms from 'metaphoric' phrases such as those in (35). For our present purposes, we will make a distinction along the lines of a suggestion of Howard Lasnik (p.c.) that a characteristic feature of idioms is the availability of a non-compositional reading, along with the compositional one. By this criterion, *take an advantage of NP*, for instance, is not an idiom (but see Chomsky (1995b) who considers it an idiom), since its meanings can, conceivably, be computed compositionally. In contrast, *kick the bucket* is an idiom since it has a meaning (in addition to the literal one) not predicted on the grounds of compositionality.

This, again, is consistent with the theory where the verb and the object form a constituent, to the exclusion of the subject.

If the preponderance of the verb+object idioms over verb+subject ones is indicative of a syntactic structure in which verb and object, but not verb and subject, form a constituent, then, by extension of reasoning, we expect the absence of preponderance of verb-object idioms in languages with the 'flat' VP. Whatever the exact mechanism of idiom formation, it should not make preference to targeting only one of the verbal arguments, but not the other. Do we expect any kind of 'preponderance' at all? The answer requires a more thorough study of idiom formation. Still maintaining that the map between argument and syntactic structure is straightforward, let us hazard the following:

(38)   The idiom formation process operates only on syntactic constituents

For (38) to be of any significant interest, it must be universal and not subject to parametric variation. The theory in (38) is consistent with the English facts observed so far and discussed in the literature. Given the hierarchical structure of English VP, it correctly predicts the existence of verb+object idioms (under V') with the open slot for the subject, as well as full propositional idioms, with all argument slots filled (cf. *the shit hit the fan*). It also correctly excludes subject+verb idioms in English. With respect to the 'flat' VP languages, we predict the existence of propositional idioms only. Neither verb+object, nor subject+verb idioms should be allowed, because none of those would form a constituent in syntactic structure.

Again, Speas (1990) believes that the prediction of a theory which accounts for the preponderance of verb+object idioms in English, for 'flat' VP languages, is different. She predicts 'S+V idioms, or only full sentential idioms' (p. 137) for those languages. Assuming that Speas' 'sentential idioms' more or less correspond to what we called propositional idioms, we fail to see why verb+subject idioms are predicted. The issue is not internal to a theory such as (38). As in the previous case, allowing verb+subject idioms amounts to allowing a *hierarchical* structure, in which verb+subject form a constituent, to the exclusion of object, but not a truly 'flat', or symmetrical structure.

Having laid out the predictions for the 'flat' VP languages from the domains of predicate formation and the distribution of idioms, let us now turn to examining the factual state of affairs in Russian. First, we will discuss non-idiomatic expressions and after that, we will turn to discussion of Russian idioms.

In Russian, as in English, the choice of the object of the verb may affect the semantic role assigned to the subject of the resulting predicate, as shown in the following examples (the examples below taken from Molotkov (1967)):

(39)   a. prolivat' sup

         spill   soup

         'to spill the soup'

    b. prolivat' krov'

         spill   blood

         'to defend smb. at war'

    c. prolivat' krokodilovy slezy

spill crocodile's tears

'to complain insincerely'

d. prolivat' svet na ...

spill   light on

'to bring evidence for...'

Thus a subject of (39a) would tend to spill liquids, whereas a subject of (39b) must be a war participant, and a subject of (39c) is an insincere complainer. In all cases, the verb *prolivat'* 'spill' is used. As in English, verbs that express a wide range of predicates depending on the choice of the object abound in Russian. However, it is also possible in Russian that the choice of subject affects the choice of the object. The following illustrates:

(40)   a. Ivan ne obidel NP

Ivan not offend

'John did not offend NP'

b. Bog ne obidel NP

God not offend

'NP is gifted, talented'

c. Boss ne obidel NP

Boss not offend

'NP's boss promoted NP (in salary, position, etc.)'

(41)   a. Sobaka ukusila NP

dog   bit

'A dog bit NP'

b. (Kakaja to) muxa ukusila NP

  some   fly   bit   NP

'NP behaves strangely'

(42)   a. Ivan placet po NP

  John cries at

'John mourns NP'

b. Verevka/viselica/gil'otina/ tjur'ma/... placet po NP

  Rope   gallows guillotine prison   cries   at

'NP deserves to be hung/decapitated/put in prison...'

(43)   a. Vrag napal na NP

  Enemy attacked on

'Enemy attacked NP'

b. (Kakoj-to) stix   napal   na NP

  certain   poem attacked on NP

'NP is in a certain mood'

(44)   a. Ivan   zaputal   NP

  'John confused NP'

b. Cert   poputal NP

  Devil confused NP

'NP succumbed into temptation'

The object of (40a) was simply not offended by the subject, whereas in (40b) the object is a gifted, talented person. In both cases, the same verb *obidet'* 'to offend' is used. Similarly for (41)-(44).

Marantz notes, in response to a criticism of J. Bresnan that in certain cases in English, the verb+subject combination appears to affect the semantic role of the object (cf. *the roof caved in on John*), that in such instances in English, the open slot is usually headed by a prepositional, or possessive phrase, and is, in fact, optional. Consequently, such instances may be analyzed as full propositional idioms (cf. [[*the roof caved in*] *on John*]. Notice that in none of the instances above is the object slot optional (with a possible, although not certain, exception in (42)). The prepositional phrases in (42) and (43) are headed by a preposition that is required by selectional properties of the verb (as in *rely on* in English), hence, cannot be dismissed as optional arguments.

Turning now to idioms, unsurprisingly, we find propositional idioms in Russian, cf:

(45)　　a. Dusa uxodit v pjatki (u NP)

　　　　　Soul goes   in heels

　　　　　'(Someone's) heart sunk'

　　　　b. Krov' udarila v golovy (NP)

　　　　　Blood hit   in head

　　　　　'Someone got a jolt of adrenalin, get anxious'

　　　　c. jabloku negde upast'

　　　　　apple nowhere to-fall

　　　　　'the place is crammed, overcrowded'

d. Glaza na lob   lezut (u NP)

eyes on forehead climb

'NP is getting overwhelmed'

However, Russian has verb+object idioms (see fn. 18):

(46)   a. dat' duba

give oak-acc

'to die'

b. sygrat' v   jashchik

play   into drawer

'to die'

c. metat' ikru

spawn caviar

'to hassle unnecessarily, over little things'

d. muxu   razdavit'

fly-acc to-squish

'to drink (alcohol)'

And Russian also has subject+verb idioms:

(47)   a. kondraska   stuknul NP

'Kondrashka (name) hit NP'

'NP died suddenly'

b. Bog pribral NP

God took

'NP died'

c. Kot naplakal NP

cat cried up

'very few, to a very small degree of NP'

The Russian data concerning the 'influence' of one argument on the choice of the other, and idiom formation, are not predicted under the 'hierarchical' structure, given Marantz's/Speas' original assumptions. In fact, they are not predicted under the 'flat' structure either. If Speas's predictions concerning the absence of the 'influence' asymmetry were correct, the data in (39)-(47) would actually suggest that Russian has a 'flat' VP. But we have already shown that Speas's predictions cannot go through, under Marantz's original assumptions. What the overall pattern of data seems to suggest then is that somehow the 'influence' asymmetry and idiom formation phenomena are not directly relevant to the issue of hierarchical/'flat' VP debate. Certain theoretical assumptions made so far must be questioned.

We believe the assumption that must be questioned concerns the mapping between the structure at which semantic roles are assigned (Marantz's l-s structure) - roughly, argument structure - and syntactic structure. At the time of Marantz (1984), it was natural to assume that the mapping was direct. Specifically, it was regulated by the Projection Principle (Chomsky (1981)), stating that the argument structure must be represented at

every syntactic level.[19] In particular, argument structure must be syntactically represented at D-structure, defined in Chomsky (1981) as a pure representation of grammatical function theta (GF-θ). As the theoretical thinking progressed, it became increasingly clear that the notion of D-structure is not a theoretical necessity and can be dispensed with, and furthermore, is not desirable on empirical grounds (cf. Bošković (1994) for empirical arguments against D-structure). The Minimalist program dispenses with the notion of D-structure entirely. Hence, it is no longer required that argument structure be represented at a particular level of representation. In particular, it becomes a theoretical possibility that an NP validates its argument status in a structural position different from that in which it was generated. Bošković and Takahashi (1998) explore this possibility in detail and suggest that theta-roles are, in fact, features that can (and must) be 'checked' in the course of the derivation (see also Bošković (1994), Hornstein (1999), Lasnik (1999), but see Chomsky (1995c) for the opposite view). An NP may in principle check theta-features in a derived position, by undergoing syntactic movement.

This opens a door to a different interpretation of the mapping between Marantz's l-s structure and syntax, which does not bear on the 'asymmetry in influence' and idiom formation phenomena. Suppose that the subject is Merged inside the VP along with the object, but checks its subject theta-role outside of it (e.g in the functional head v, Asp, or Agr). The verb itself has only one theta-role to assign, that of an object. Thus, in the *derived* structure, the argument structure will be represented as hierarchical. But at the

---

[19] In fact, as Marantz himself makes clear (p. 28), the assumption concerning the direct map between his l-s structure and syntactic structure is not trivial, and he explicitly restricts his claims concerning asymmetric 'influence' only to the former, not to the latter. Speas (1990), on the other hand, crucially assumes the direct map, as she uses Marantz's asymmetry as a test for a particular *syntactic* structure.

previous stage, the subject may have started out as a sister of V, forming a 'flat' VP structure, or as a sister of V' (specifier). Which of these possible structures, if any, should correspond to the one at which the 'asymmetric' influence of one argument on another, and idiom formation takes place (Marantz's l-s/argument structure)? Since there is no direct mapping, it is impossible to tell *a priori*. We conclude, therefore, that the 'influence' and idiom formation phenomena are orthogonal to the issue of the structure of VP. In addition, the Russian facts considered above provide an indirect argument for considering theta-roles as features.

CHAPTER 2

An 'Eclectic' Approach Part I: Chain Uniformity

1. Introduction

In Chapter 1 we considered the 'unified' approach in which extractability out of subjects and adjuncts is regulated by the same mechanism. We showed that this approach faces both conceptual and empirical problems. In this chapter we begin to explore an alternative approach under which extractability out of subjects and adjuncts is regulated by different mechanisms of the minimalist grammar. We begin by discussing the first part of the 'eclectic' theory, which deals with extraction out of subjects.

In Chapter 1 we pointed out the redundancy that arises under the neo-unified approach, in connection with examples of extraction out of derived subjects, illustrated in (1):

(1) ?*Who$_j$ was [a picture of t_j]$_i$ taken t_i by Bill

The mechanism on which the 'neo-unified' approach relies in order to derive the standard 'CED' effects in English is insufficient to account for cases of extraction out of derived subjects such as in (1). The additional mechanism invoked by N&U to supplement their 'neo-unified' theory, in order to account for (1), is based on a version of the chain uniformity condition (cf. Chomsky (1995c)). According to N&U, a copy of [*a picture of who*] in the derived position (Spec-IP) is not identical to the copy in situ, in a well-

69

defined sense (see Chapter 1 for discussion). Hence, (1) is ruled out as a violation of chain uniformity.

The redundancy arises, however, given Takahashi (1994) who showed that the chain uniformity condition alone is sufficient to account for all cases of non-extraction out of subjects in English, including (1) as well as standard 'subject condition' violations. In fact, the idea underlying a chain uniformity analysis potentially extends to the entire range of phenomena concerning extractability out of subjects, including the languages from Chapter 1 in which extraction out of subjects is possible. This line of analysis is pursued in detail in Takahashi (1994). In Section 2.2. of Chapter 1, we briefly reviewed the basic tenets of Takahashi's theory. In this chapter we explore a version of the chain uniformity analysis in more detail and strengthen its empirical and conceptual basis. In Section 2 we review Takahashi's original analysis. In Section 3, we restate this account in the current framework of assumptions. Section 4 discusses an extension of the chain uniformity analysis in the domain of extraction out of 'specific' NPs.

2. Takahashi (1994)

Wexler and Culicover (1981) formulate a 'Freezing Principle', which for our present purposes we interpret to encode the following descriptive generalization:[1]

---

[1] In effect, Wexler and Culicover's 'freezing principle' makes a much stronger claim, prohibiting extraction even out of the mother node dominating the moved domain. Their precise formulations (under the EST framework) are the following (p. 119):

(i) *Freezing Principle:* If a node A of a phrase marker is frozen, no node dominated by A may be analyzed by a transformation.

(ii) If the immediate structure of a node in a phrase marker A is nonbase, that node is *frozen*.

(iii) The *immediate structure* of A is the sub-phrase marker consisting of A, the nodes $A_1, A_2, ..., A_n$ that A immediately dominates, in order, and the connecting branches.

(iv) The immediate structure of A is a *base immediate structure* if A $\rightarrow$ $A_1, A_2, ..., A_n$ is a base rule. Otherwise, it is nonbase.

(2) No extraction is possible out of previously moved domains

Lasnik and Saito (1992), in discussing the absence of 'subject condition' effects in Japanese, offer an account which captures the essence of (2). They relate the absence of 'subject condition' effects to the status of the subject as a 'properly governed' position, which, for them, is a result of the structure in which subject is located in VP at 'S-structure', as opposed to the Spec-IP in languages in which 'subject condition' effects obtain (English). Combined with the internal subject hypothesis, under which subjects in all languages are generated within VP, the contrast between English and Japanese directly translates in terms of (2): in English, the subject moves to Spec-IP and 'subject condition' effects obtain, whereas in Japanese the subject stays in VP and 'subject condition' effects do not arise.

Takahashi (1994) pursues the direction of reasoning offered in Lasnik and Saito (1992) further and makes the 'subject condition' effect follow as a direct consequence of previous movement of the subject. Adopting the minimalist framework of Chomsky (1995c), Ch.1-3, Takahashi offers an account of subject condition effects in English and their absence in Japanese, based on two fundamental assumptions listed in (3) and (4):

(3)   *Chain Uniformity:* Chains must be uniform

Thus, for instance, if a subject DP moved from the Spec-vP to Spec-IP (under the internal subject hypothesis), the freezing principle prohibits extraction not only out of this DP, but out of the entire IP, including the DP itself. One potential problem that arises with respect to (i) is that, without additional provisos, it seems to rule out all successive cyclic movement (as the authors themselves note on p.583, fn. 43). On successive cyclic *wh*-movement, see Chapter 6 of the present work.

(4)   *Shortest Move:* Make the shortest move

The 'shortest move' condition (4) effectively forces an element undergoing A'-movement to reach its target site by a series of short successive adjunctions to the maximal projections along its path (see also Chomsky (1986a)). More generally, this condition requires movement to the closest asymmetrically c-commanding target, and that target is relativized according to the movement type (A'-movement to the closest A'-position, A-movement to the closest A-position). Furthermore, Takahashi interprets the chain uniformity condition (3) as a syntactic constraint which precludes derivational processes - such as movement - from 'disturbing' the uniform status of chains. Specifically, he proposes a Uniformity Corollary of Adjunction, a simplified version of which is shown in (5):

(5) *Uniformity Corollary on Adjunction (UCA):* Adjunction to a part of a non-trivial chain or coordination is not allowed

Putting aside for the moment the part of UCA that deals with coordinations, let us see how (5), in conjunction with (4), rules out standard 'subject condition' violations. Consider (6a), with a structure in (6b):

(6)   a. ?*Who does a picture of hang on the wall?

        b. Who does [$_{IP}$ [$_{DP}$ a picture of who] [$_{vP}$ [$_{DP}$ a picture of who] hang on the wall]]

In accordance with the internal subject hypothesis and 'shortest move' (4), in (6) *a picture of who* has moved from Spec-vP to Spec-IP, leaving behind a copy. The original and the copy then constitute an A-chain. *Who* is then extracted from the higher link of the chain. In accord with the shortest move (4) it first must adjoin to the maximal projection of the DP dominating it, creating the configuration in (7):

(7) Who does [$_{IP}$ [$_{DP}$ who [$_{DP}$ a picture of who]] [$_{vP}$ [$_{DP}$ a picture of who] hang on the wall]]

But (5) disallows this step, since it involves adjunction to a part of a non-trivial chain. *Who* must then skip the adjunction to DP, and proceed to the next available (intermediate) landing site, namely, adjoined to IP, creating the configuration in (8):

(8) Who does [$_{IP}$ who [$_{IP}$ [$_{DP}$ a picture of who]] [$_{vP}$ [$_{DP}$ a picture of who] hang on the wall]]

But this step of movement violates the 'shortest move' (4), since it skips one potential landing site. Thus *who* cannot reach its target without violating these derivational conditions. One might argue that at the stage (6b) the lower copy of *who* is also potentially available for the purposes of *wh*-movement. But even if this is so, similar considerations apply in this case as well: adjoining *who* to the DP is precluded by (5), whereas skipping this step and adjoining to vP would violate (4). The derivation then cannot converge.

There exist, in fact, two derivations of (6a) which circumvent (4) and (5). Consider a derivation in which *wh*-movement <u>precedes</u> movement of *a picture of who* to Spec-IP. That is, at the time of *wh*-movement, the subject DP is still located in Spec-vP. Suppose *who* is extracted from the DP and first adjoins to the DP, in accordance with (4). This results in (9):

(9) Who does [$_{IP}$ _ [$_{vP}$ [$_{DP}$ who [$_{DP}$ a picture of who]] hang on the wall]]

Since *a picture of who* has not moved, it constitutes a trivial chain, that is, a chain consisting of only one member, namely, the DP itself. Adjunction to a trivial chain is not precluded by (5). Then *who* continues its path, adjoining, successively, to vP, IP, and ultimately moving to Spec-CP. After that, the modified DP moves to Spec-IP, leaving behind a copy and creating an A-chain, as shown in (10):

(10) Who does [$_{IP}$ [$_{DP}$ who [$_{DP}$ a picture of who]] [$_{vP}$ [$_{DP}$ who [$_{DP}$ a picture of who]] hang

on the wall]]

The A-chain in (10) is uniform, in accord with the uniformity condition (3). Both (4) and (5) are observed at all stages of the derivation. The derivation, however, would violate the strict cycle, or principles which derive it (see Chomsky (1995c), Ch.4., Bošković and Lasnik (1999) and Chapter 3 of the present work for discussion).

The second scenario is the following. Suppose the derivation reached the stage (9), in which *who* adjoins to the DP in its base-generated position in Spec-vP. After that, this

modified DP moves to Spec-IP, resulting in (10). The chain resulting from this movement is uniform, in accordance with (5). Then *who* continues to move, adjoining to IP and ultimately moving to Spec-CP. Again, both (4) and (5) are observed at all stages of the derivation. Furthermore, the derivation observes strict cyclicity as well.

Ruling out this derivation is less straightforward under Takahashi's assumptions. Takahashi points out that the derivation involves 'chain interleaving': in effect the A-chain created by movement of the DP is formed 'in the middle of' formation of another chain, namely, created by movement of *who*. Following Collins (1994), Takahashi claims that interleaving chains in this manner violates a principle of economy of derivation. The reasoning behind this is the following. Chomsky (1995c), Ch.3, exploring the idea that derivational processes are subject to 'economy', points out that there are two natural ways in which economy might be understood: one, in (4), requiring shortest moves, and another one, requiring fewer steps in a derivation. These two formulations of economy seem to be in conflict: while the latter requires fewer steps of successive cyclic movement, thereby increasing the distance traversed by each step, the former prefers more steps, reducing the distance traversed by each step of movement. To resolve this paradox, Chomsky redefines the movement operation as Form Chain. As Chomsky and Lasnik (1993), Chomsky (1995c), Ch.3 point out, under Form Chain successive cyclic movement counts as one step, even though each part of the chain satisfies the 'shortest move' condition (4). Collins (1994) argues that the ban on chain interleaving follows directly from the conception of movement as Form Chain: if Form Chain takes place in one step, it cannot be the case that any other derivational action can intervene in the process of forming the chain. Chain interleaving thus results in adding derivational

step(s) to the process of chain formation, which, ultimately, violates the economy condition requiring fewer steps. The derivation is thus ruled out as well, but only under the conception of movement as Form Chain.

Takahashi's account of extractability out of subjects has several conceptual and empirical advantages. The main conceptual advantage is that it formalizes the intuition behind the descriptive generalization in (2). It does so in a rather elegant way, involving the independently justified notion of chain, but without involving additional notions of pre-minimalist constraint-based frameworks, such as government or barrier. The most important empirical advantage of Takahashi's account is that it seems to account for the entire range of the relevant phenomena, falling under the descriptive generalization in (2). Thus, consider extraction out of passive subjects in English:

(11)   ?*Who was [a friend of t] arrested?

The surface subject of (11) undergoes movement from the position of object of *arrested*, hence, by the time of *wh*-movement, is a moved domain. Hence, Takahashi's account correctly predicts extraction out of it to be ungrammatical.

It is well known (cf. Chomsky (1973), Kayne (1984)) that extraction out of subjects of Exceptionally Case Marked (ECM) verbs is also degraded:

(12)   ??Who do you believe [a picture of t] to be on sale?

Lasnik and Saito (1991), following Postal (1974), argue that an ECM subject actually undergoes raising to the matrix object position. This implies that by the time of *wh*-movement the matrix object is a moved domain. Again, Takahashi's account implementing the generalization in (2), straightforwardly rules out this derivation.[2]

Similar considerations obtain in cases of topicalization out of topicalized phrases (13a), and *wh*-extraction out of topicalized phrases (13b), which also results in degradation, as noted in Lasnik and Saito (1992). To these, we also add a case of extraction out of the *wh*-phrase in Spec-CP, as in (13c):[3]

(13)   a. ??[Vowel harmony]$_j$, I think that [articles about t$_j$]$_i$, you should read t$_i$

        b. ?*Who$_i$ do you think that [pictures of t$_i$]$_j$ John wanted t$_j$

        c. ??Who$_i$ do you wonder [which pictures of t$_i$]$_j$ Mary bought t$_j$

Now let us consider how Takahashi's account deals with languages in which extraction out of subjects is possible. Consider the Japanese example of comparative deletion in (14):

---

[2] For Lasnik and Saito (1991), this raising is obligatory. Lasnik (1999a), however, argues that the raising is 'optional', in the sense that the functional category driving raising may or may not participate in the derivation. Under either theory, however, [*a picture of who*] moves to the embedded Spec-IP from Spec-vP, given the internal subject hypothesis.

[3] The status of examples such as (13c) is a bit controversial. A clearer case provided in (i):
(i)   *Why do you wonder [how likely to fix the car t$_i$] John is?

(14) [OP [Mary-ga *t* yonda no]-ga   akirakana yorimo John-wa

Mary-Nom   read that-Nom is-obvious than   John-Top

takusan-no hon-o   yonda

many-Gen book-Acc read

'(*) John read more books than [that Mary read _] is obvious'

Following Lasnik and Saito (1992) (cf. also Kuroda (1988), Fukui (1986), Kitagawa (1986), cf. also Yatsushiro (1999)), Takahashi assumes that in (14) the subject *[Mary-ga Op yonda no]-ga* stays inside the VP in which it is base-generated, at least in overt syntax. As such, it constitutes a trivial chain (consisting of only one member). When the null operator is extracted, it first adjoins to the clause, in accord with (4). This adjunction is in fact allowed by (5), since it allows adjunction to trivial chains. Subsequently, it adjoins to higher projections until it reaches its destination, the matrix Spec-CP.

The step of the derivation in which the null operator adjoins to the subject is actually reminiscent of the one reflected in (9), which we considered as part of a possible derivation of an English sentence (6a). Recall that the derivation of which (9) was a step was ruled out because the subject would move to Spec-IP in overt syntax, after *wh-*movement takes place, thus violating the strict cycle. But the Japanese derivation does not violate the strict cycle if, as is assumed, the subject stays in vP, so that there is no acyclic movement to Spec-IP involved in this case. Note that, as the null hypothesis, this account of Japanese extends to the entire spectrum of languages in which it is possible to extract out of subjects, more concretely, languages of Chapter 1. This implies that subjects in those languages can stay in the positions in which they are base-generated.

It should be emphasized that Takahashi's theory does not postulate any language-specific differences that govern extractability out of subjects in a given language. Rather, it implements a descriptive generalization (2), taken as a universal claim that phrases that move become 'opaque' for extraction. We expect, then, that, in one and the same language, it would not be possible to extract out of the 'logical' subject when it clearly moved, whereas such extraction would be possible if one forces the subject to stay in situ.

In Chapter 1, we considered one such language: English. In contrast to standard cases of 'subject condition' violations (15a), in which the subject is standardly assumed to move from Spec-vP to Spec-IP, it is possible to force the subject to stay in situ in (15b) by inserting the expletive *there* in Spec-IP. Unsurprisingly, (15b) is grammatical.[4]

(15)   a. ?*Who$_j$ did [a picture of t$_j$]$_i$ [$_{vP}$ t$_i$ hang on the wall]?

         b. Who was there [$_{SC}$ [a picture of t] on the wall]?

German provides another example fulfilling this expectation. In this language, extraction out of subject seems to be precluded precisely in those contexts in which subject clearly undergoes overt movement. Thus, in simple matrix and embedded clauses, extraction out of an infinitival clause in the object as well as the subject position is possible (the German examples in this section are adapted from Haider (1997), Haider (2000) and Müller (1995) (see these authors, also Grewendorf (1989), among others, for discussion):

---

[4] Lasnik (1995a) who argues that the subject of the small clause in this case undergoes overt movement to Spec-AgrOP where it checks (partitive) Case against *be*. For the present purposes, we assume that this Case checking need not involve overt movement (cf. Lasnik (1999a)).

(16) [Welche Tür]ᵢ habt ihr [tᵢ damit aufzubrechen] vergeblich versucht?　　[object]

　　　Which door have you　to open with it　　in vain　　tried

　　　'Which door did you try in vain to open?'

(17) a. [Mit wem]ᵢ würde [tᵢ Schach spielen zu dürfen]　dich mehr freuen?　[Su]

　　　with whom would　chess　to-play to-be allowed you　more please

　　　'?*Who would [to be allowed to play chess with _] please you?'

　　b. [Welches Buch]ᵢ hat [tᵢ zu lesen] dir mehr Spaß gemacht?　　　　　　[Su]

　　　Which book　has　to read　you more fun-acc made

　　　'?*Which book did [to read _] was fun?'

　　c. Ich weiß nicht wenᵢ [PRO tᵢ gesehen zu haben] den Fritz beeindruckt hat [Su]

　　　I　know not who-acc　　seen　to have　　Fritz impressed has

　　　'?*I don't know who [to have seen _] impressed Fritz'

Haider (2000) gives an example of *wh*-extraction out of the subject of a V2 clause embedded under a 'bridge' verb in German. Interestingly, in this case the 'subject condition' effect obtains (see also the above authors for discussion), cf.:

(18) a. *[Mit wem]ᵢ hat sie gesagt [tᵢ Schach spielen zu dürfen]　würde sie sehr freuen?

　　　with whom has she said　　chess　to play to-be allowed would her much please?

　　　'*Who did she say that [to be allowed to play chess with_] would please her?'

b. *Welches Buch sagte sie [$_{CP}$ [t_i zu lesen] [habe [ihr Spaß gemacht]]]?

which book said she to read has her fun made

'*Which book did she say that [to read_] was fun?'

All else equal, a minimal account of the grammaticality contrast in (17) and (18) should take into consideration that the subject in (18) is in a clearly moved position. The contrast in grammaticality between (17) and (18) then can be attributed to the movement of the subject in (18), vs. non-movement of the subject in (17). Furthermore, Haider (op. cit.), Haider (1986), argues that not only the direct object, but all arguments and adjuncts are in a very local relation with the verb in overt syntax (which he defines in terms of government). It is a direct consequence of this claim that all verbal arguments, including the subjects, are inside the relevant verbal or functional projection in overt syntax. It is plausible, then, that German subjects can stay in situ in Spec-vP in overt syntax.[5]

We conclude that Takahashi's account correctly captures the entire range of facts concerning extraction out of subjects, as a subset of the empirical generalization in (2). Indeed, in Section 4 we show that this account also captures another relevant subset of data falling under this generalization, namely, extraction out of 'specific' phrases. However, there are two conceptual concerns arising with respect to Takahashi's analysis.

The first, and most important one in the light of the general goal of this work, is that it relies on narrow syntactic constraints in the form of (4) ('shortest move'), (5) ('Uniformity Corollary on Adjunction'), and a conception of movement as Form Chain.

---

[5] In a similar vein, Diesing (1992) argues that subjects in German appear either in Spec-vP (VP in her framework), or Spec-IP. As she shows, extraction out of subject is limited to the former context only.

(4) and (5) are prohibitive constraints: the former precludes longer moves, the latter adjunction to parts of non-trivial chains. Whereas (4) may be argued (as it was, by Chomsky) to follow on the grounds of 'economy', it is difficult to see how (5) can be independently justified.[6]

A similar concern arises with regard to chain uniformity itself. The chain uniformity condition as formulated in (3) is understood as a constraint on syntactic derivations. It is unclear, without additional qualifications, why the grammar needs this constraint. In other words, one may ask, reasonably, why is it that the non-uniformity of a chain results in a derivational crash.

The second concern is that in the framework that we adopt here, in which movement is seen as driven by the target ('Attract', Chomsky (1995c), Ch.4), but not by the moving element only (cf. 'shortest move'), Takahashi's chain uniformity account becomes unstatable. To see why this is so, consider (15a) again:

(15)   a. ?*Who$_j$ did [a picture of t$_j$]$_i$ [$_{vP}$ t$_i$ hang on the wall]?

As discussed above, under the conception of movement as subject to 'shortest move' (4), extraction out of the subject results in adjunction to one part of the A-chain ([*a picture of who*], [*a picture of who*]) at some point. This adjunction violates (5), and, more generally, chain uniformity (3) in that it creates a non-uniform chain ([*who*[*a picture of who*]], [*a picture of who*]). This is responsible for the ungrammaticality of (15a) (cf. (7)).

---

[6] Takahashi's (5) in fact recalls the ban against adjunction to arguments as in Chomsky (1986a); but that itself was a constraint the existence of which was difficult to justify.

Consider now the derivation of (15a) under the conception of movement as driven by the property of a target of movement, which Attracts the element that undergoes movement. On this derivation, a certain property of interrogative C Attracts *who* from the subject position. When *who* is extracted from the subject, nothing forces adjunction of *who* to the subject itself. Rather, *who* moves to Spec-CP in one 'fell swoop', resulting in the following configuration (we continue to assume the copy theory of movement):

(19) Who$_j$ C$_{[+Q]}$ does [$_{IP}$ [$_{DP}$ *a picture of who*$_j$]$_i$] [$_{vP}$[$_{DP}$ *a picture of who*]$_i$ hang on the wall]]
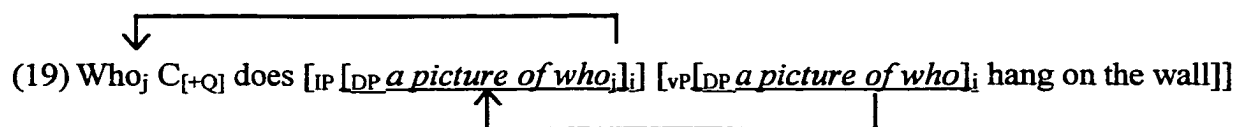
Note that at this point the chain (*a picture of who, a picture of who*), resulting from the (previous) movement of the subject from Spec-vP to Spec-IP is uniform, by all conceivable versions of chain uniformity understood as a syntactic condition. The chain formed by Attraction of *who* into Spec-CP is also uniform. This derivation then should converge, incorrectly predicting the grammaticality of (15a) (see, however, Ochi (1999b) for a proposal to reconcile Takahashi's analysis with the Attract framework).

In the following section, we restate Takahashi's account of extractability out of moved domains in terms of the framework that we are assuming. This restatement avoids the conceptual weaknesses above by exploring a possible motivation behind the chain uniformity requirement. It will be shown that this restatement also gives us further insight into the nature of chains.

3. Chain uniformity as a PF condition

Again, if chain uniformity (3) is understood as a condition on syntactic derivations, then it cannot rule out (19) since all chains formed at this point are uniform. The task then is to find a formulation of chain uniformity that can appropriately rule out the derivation.

In order to approach this task, let us first ask a more general question: why should the non-uniformity of a chain result in a derivational crash? Ochi (1999b) suggests that the answer lies in PF considerations. At PF, chains created by movement are subject to deletion of all copies except the highest (we will return to the question why the highest copy is retained and lower copies are deleted). Ochi hints that the key criterion for deletion of copies is identity. PF deletion is only possible if all copies are identical to each other. If copies are non-identical to each other, PF cannot delete copies, and an illegitimate (or uninterpretable) PF object results. Ochi's suggestion, in essence, reduces the chain uniformity condition to the following condition:

(20)   PF: delete copies under identity with each other.

Let us explore in more detail the intuition behind (20). At PF, syntactic constituents must be linearized. A crucial property of any linearization procedure is to establish a precedence relation between the terminal elements of the phrase marker. In fact, as Higginbotham (1983) points out, the precedence relation on terminals, or, more precisely, sound formatives, may be imposed by the physics of the human articulatory tract. In the present terms, that amounts to a rather plausible claim that the precedence relation is in

essence an interface requirement (of the PF interface). Note that the precedence relation is by definition a) asymmetric; and b) irreflexive.

Chains are also subject to linearization. Nunes (1999), (2001) argues that it is generally impossible to linearize more than one copy of a chain, since the precedence relation on elements including more than one copy of a chain cannot be established. To see why this is so, consider the following example (cf. Nunes 2001, p. 307):

(21)   *John* was [arrested *John*]

In (21) there are two non-distinct copies of *John*.[7] For Nunes, the linearization procedure is driven by the LCA: it establishes a (partial) linear order between two elements one of which asymmetrically c-commands the other. This procedure, for instance, will establish the precedence relation Π between the higher copy of *John* and the copula *was*, for instance. The same procedure will establish the precedence relation between *was* and the lower copy of *John*. Thus it must be the case at some point that Π = <*John*, *was*, *John*>. Since the two copies of *John* in (21) are non-distinct, it follows that *was* must precede and be preceded by the same element, namely, *John*. Π also implies that *John* precedes itself. Thus Π violates both defining characteristics of precedence: its asymmetry and irreflexivity. It follows that Π cannot be a linear order for (21).

In order to allow linearization, one copy of a chain must be deleted. This avoids the situation in which two non-distinct copies are evaluated for precedence. But deletion

---

[7] Chomsky (1995c), p. 227 proposes that two lexical items *l* and *l'* are marked as distinct for the computational system if they are formed by distinct applications of Select accessing the same lexical item in a numeration; thus copies of *John* in (21) count as non-distinct.

must be constrained. The usual criterion is recoverability, which is achieved if identity of copies is ensured. Thus the linearization procedure must involve scanning the structure submitted to PF for copies. It is a fairly uncontroversial view that Scanning includes establishing the identity of copies and deletion of non-higher copies (under identity). Thus (20) can be restated, more precisely:

(22)   SCAN:[8]

    a) establish identity of $\alpha_1,\ldots,\alpha_n$, where $\alpha_1,\ldots,\alpha_n$ are non-distinct copies, $n>1$

    b) delete $\alpha_2,\ldots,\alpha_n$

Let us now see how Scan (22), in this enhanced sense, may be responsible for the ungrammaticality of (15a), under the Attract conception of movement. Consider a simple transitive clause such as *pictures of John hang on the wall*. The subject *pictures of John* raises from Spec-vP to Spec-IP, forming an A-chain. This configuration is submitted at PF, where copies of *pictures of John* are identified. Since they are completely identical to each other, deletion of a lower copy is successful, and the resulting PF object is legitimate:

---

[8] Our Scan is very similar to Nunes (1999), (2001) Chain Reduction, but differs from it in that unlike the latter, it a) does not crucially resort to the LCA, b) includes establishing identity of copies. For Nunes, on the other hand, identity of copies follows from their non-distinctness. However, as Nunes himself acknowledges (his p. 306), non-distinctness is a syntactic notion (see fn. 7) and its transfer to PF constitutes a non-trivial assumption.

Bošković (2001) further argues that scanning for copies is directional: it proceeds from left to right. For Bošković, this is necessary to ensure that in the general case, the left-most/highest copy gets pronounced. Bošković shows that sometimes a lower copy can be pronounced, if the pronunciation of the highest copy leads to a PF non-convergence for independent reasons. Note that our Scan enforces deletion of non-first copies in general and in the form of (22) would not allow pronunciation of lower copies at all. This formulation suffices for our present purposes, however.

(23)   [IP [pictures of John] [VP [~~pictures of John~~] hang on the wall]]

Now consider the stage of the derivation of (15a), shown in (19). Here two chains are formed: ([*a picture of who*], [*a picture of who*]) and (*who, who*). Interestingly, these chains 'intertwine' in that a member of the chain (*who, who*) actually belongs to a member of the chain ([*a picture of who*], [*a picture of who*]). More generally, let us illustrate intertwining as follows:

(24)   CHAIN INTERTWINING:[9]

$[\alpha_1, \ldots \ldots [\ldots \alpha_k \ldots]_{\beta 1} \ldots [\ldots]_{\beta k} \ldots]$,

where $\alpha_1, \ldots, \alpha_n$ and $\beta_1, \ldots \beta_n$ are non-trivial chains (n>1)

Given this 'intertwining', it becomes a non-trivial question how these chains should be processed at PF; in particular, if there is a particular ordering in processing these chains.[10] Assume for the time being that a) the ordering exists, and b) this ordering is free: either chain can be processed first. Suppose the (*who, who*) chain is processed first. Both

---

[9] This is not to be confused with 'chain interleaving', discussed above in the sense of Collins (1994). The two are crucially different: chain interleaving is a sequence of derivational steps during which some derivational action intervenes in the chain formation process. Ours is a representational scheme of two completed chains, whereby a link of one chain is a member of the link of another one.

[10] An obvious example involving no 'intertwining' of chains is in (i).
(i)   [CP Who [IP ~~who~~ [VP ~~who~~ said that [IP John [VP ~~John~~ left]]]]]?
Here two chains are at issue, (*who, who, who*) and (*John, John*). Either chain can potentially be processed first at PF, without noticeable empirical differences.

members of this chain are identical, hence the lower copy of *who* is successfully deleted, resulting in the following configuration:[11]

(25)   Who_j did [a picture of ~~who~~] [_{vP} [a picture of who] hang on the wall]?

Now the remaining chain enters processing. But at this point this chain is non-uniform, according to Ochi's criterion: namely, the higher member is not identical to the lower. Consequently, deletion of the lower copy cannot take place, and an illegitimate PF object results. Thus (15a) is correctly ruled out.

Suppose now the ([*a picture of who*], [*a picture of who*]) chain is processed first, at the stage in (19). Both copies of the chain are identical, hence, by (20), deletion proceeds successfully, resulting in (26):

(26)   Who_j did [a picture of who] [_{vP} [~~a picture of who~~] hang on the wall]

The second chain is processed next, deleting the lower copy of *who* under identity:

(27)   Who_j did [a picture of ~~who~~] [_{vP} [~~a picture of who~~] hang on the wall]

---

[11] Note that this deletion does not extend to the copy of *who* in the lower copy of *a picture of who*, since that copy is not involved in creation of the chain (*who, who*). Here we follow the well known insight that chains reflect the history of movement, so one should be able to track down movement from the composition of the chain. We thank Željko Bošković for discussion of this point.

Deletion of all copies proceeds successfully, (20) is not violated, and the derivation converges, incorrectly. Clearly, something more must be said to preclude the second ordering.

In the reasoning above we assumed, implicitly, that Scan (22) applies to each chain in turn. Thus, in (19), in particular, there are two chains and Scan applies twice. More generally, if there are k chains in a sentence, then Scan must apply k times.

This, however, does not have to be the case. Suppose, instead, that Scan is a one-step procedure, much like Chomsky's Form Chain which was designed to encode the standard movement operation (see above). This conjecture actually includes two claims. One, crucially, is that establishing identity and deletion of all non-first copies count as one derivational step. The second is that Scan applies only once in a given sentence, establishing identity of all links of each chain, and deleting all non-first links of each chain. When there is only one chain in the sentence (cf. (23)), or when there are more than one non-intertwining chain (cf. fn. 10), this conception of Scan is indistinguishable from before. If a sentence involves intertwined chains in the sense of (24), the situation is different. In (19), again, links/members of the two chains ([*a picture of who*], [*a picture of who*]) and (*who, who*) need to be evaluated with respect to identity, and all non-first copies must be deleted. In the chain (*who, who*), identity of both links can be identified and the lower *who* can be deleted. But in the chain ([*a picture of who*], [*a picture of who*]), the identity of its higher and lower member must be established at the same time as deleting the copy of *who* in the higher member. Clearly, deletion of the higher copy of *who* would change the make-up of the higher member of the chain, making it non-identical to the lower member. However, we believe it is not the case that in this situation

Scan marks the higher copy of *a picture of who* as non-identical to a lower copy, and cannot delete the latter. Rather, we suggest that Scan is simply unable to evaluate the chain ([*a picture of who*], [*a picture of who*]) for identity, as this situation is indeterminate, and simply yields no output.[12]

There exists, in fact, a certain conceptual motivation of taking Scan to be a one-step operation. The argument comes from considerations of economy regarding the length of derivations. All things equal, a shorter derivation should block a longer derivation (cf. Chomsky (1995c), p. 314, 357), in terms of the number of applications of a given operation. Assuming that this reasoning extends to the PF portion of derivations as well, and given that Scan is an operation, it follows that the derivation in which Scan applies less times is more economical than the one in which Scan applies more times. Thus, in a sentence with k non-trivial chains, it is clearly more economical to apply Scan only once, rather than k times.

This argument is somewhat reminiscent to Chomsky's argument for Form Chain (note, incidentally, that under the present approach we no longer need to appeal to Form Chain, unlike Takahashi, since we no longer need the concept of 'shortest move'). Recall that this operation was designed to resolve the apparent paradox between two versions of the economy condition on (narrow syntactic) derivations: 'make the shortest move', and 'make the fewest steps'. Form Chain incorporated the properties of both: it allows the computational system to perform some derivational action in various portions of the structure (by successive cyclic adjunctions to maximal projections along the path of the

---

[12] In a metaphoric sense, this is similar to issuing two mutually exclusive instructions at the same time, which in our case would be 'identical' and 'not identical'.

moving element), whereas it counted as a single step (satisfying 'make the fewest steps'). With Scan, the situation is similar: under the present conception, it performs a certain action in various portions of the structure (namely, deleting copies under identity) whereas it counts as a single step, in accord with the economy spirit.

The account of ungrammaticality of the standard 'subject condition' cases (19) in terms of Attract, in syntax, and Scan as a PF operation related to linearization, carries over to all examples instantiating 'chain intertwining' (24), in particular, Lasnik and Saito's examples (13). Consider now a simple case of extraction out of object:

(28)   Who did you see [a picture of ~~who~~]?

Assume, contra Lasnik and Saito (1991), Johnson (1991), Koizumi (1995), in particular, that *a picture of who* in (28) does not undergo overt movement (see also Section 4). Then *a picture of who* constitutes a trivial chain. No deletion is required, hence, by definition, Scan does not 'notice' this phrase for the purposes of identification and deletion of copies (as, for instance, it does not 'notice' the TP for the same reason). On the other hand, movement of *who* has created a chain with two copies: Scan will identify both copies as identical and delete the lower one.

Consider again (14), involving extraction out of subject in Japanese:

(14)   [OP [Mary-ga   *t* yonda no]-ga   akirakana yorimo John-wa

            Mary-Nom   read   that-Nom is-obvious than   John-Top

takusan-no hon-o   yonda

many-Gen book-Acc read
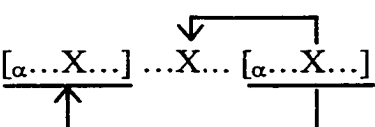
'(*) John read more books than [that Mary read _] is obvious'

This case is handled similarly to (28). The subject [*Mary-ga*   Op   *yonda no*]-*ga* does not move, hence, is a trivial chain, and no deletion is required. Extraction out of the subject then does not violate any restrictions on linearization at PF.

A question that should be addressed with respect to the Chain Intertwining analysis concerns 'remnant movement', postulated in a number of works (e.g. Webelhuth (1989), Müller (1998), (1999), Kayne (1998), Koopman and Szabolcsi (2000)), whereby the phrase undergoing movement contain trace(s) of previously 'evacuated' elements. An example is given in (29) (from Müller (1995)), and the general schema, assuming the copy theory, is shown in (30):

(29)   [t_i zu lesen]$_j$ hat keiner [das Buch]$_i$ t_j versucht

    to read   has no-one the book   tried

(30)   [$_\alpha$...X...] ...X... [$_\alpha$...X...]

Note that under the present assumptions, (30) appears to materialize the Chain Intertwining configuration (24), uninterpretable by Scan. Yet (29) is fully grammatical.

Similarly, Kayne (1998) and Koopman and Szabolcsi (2000) argue that at least certain instances of head movement should be reanalyzed as phrasal movement. In

particular, what apparent instances of X^0 are analyzed as movement of the XP after all of the material in that XP except the head itself has been moved out of it. This kind of phrasal movement patterns with the 'remnant' movement case above in that it also seems to materialize Chain Intertwining.

Čavar and Fanselow (1997) and Fanselow and Čavar (2000) propose an analysis of apparent 'remnant movement' examples as cases of full phrasal movement followed by 'scattered' deletion of different pieces of the resulting chain, at PF. This scattered deletion is forced by conditions on pronunciation, the exact formulation of which will not concern us here (see Fanselow and Čavar (2000) for stating these conditions in terms of Optimality Theory). Consider, for instance, (31) in German:

(31)   Bücher hat er keine gelesen

         books has he no   read

         'He didn't read any books'

According to Čavar and Fanselow (1997), the derivation of (31) involves (string-vacuous) movement of the object *keine Bücher* out of vP, and its subsequent topicalization. At PF, the laws of pronunciation force 'scattered' deletion of various pieces of the chain resulting from this movement, producing the observed word order, as illustrated below (see also Stjepanović (1999) for related discussion):

(32)   a. hat er [keine Bücher] [vP [keine Bücher] gelesen]

            has he no   books   no   books   read

b. [keine Bücher] hat er [keine Bücher] [$_{vP}$ [keine Bücher] gelesen]

    no   books   has he no   books   no   books   read

c. PF: [~~keine~~ Bücher] hat er [keine ~~Bücher~~] [$_{vP}$ [~~keine Bücher~~] gelesen]

Note that each instance of movement in (32) is full phrasal movement. Fanselow and Čavar (2000) argues for a similar kind of analysis of head movement in the sense of 'remnant' phrasal movement. Given the proposal by these authors, then, we tentatively assume the analysis of (29) as an instance of full phrasal movement followed by 'scattered' deletion, essentially as follows:

(33) a. hat keiner [das Buch zu lesen] [[das Buch zu lesen] versucht]

      has noone the book to read   the book to read   tried

b. [das Buch zu lesen]$_j$ hat keiner [das Buch zu lesen] [[das Buch zu lesen] versucht]

    the book to read   has no-one the book to read   the book to read tried

c. PF: [~~das Buch~~ zu lesen]$_j$ hat keiner [das Buch ~~zu lesen~~] [[~~das Buch zu lesen~~]

      versucht]

Clearly, our Scan (22) needs to be further modified, to accommodate instances of 'scattered' deletion in (32c) and (33c). The point, however, is that under this analysis (32) and (33) do not involve Chain Intertwining, as there is only one movement chain in each case, consisting of three copies of the moved XP. To the extent the scattered deletion approach can be maintained, instances of 'remnant' movement are compatible with the Chain Intertwining analysis.

Another question may arise as to why (34) is grammatical:

(34)   Mary, John likes, and Peter does too

(34) is a standard case of VP ellipsis, also involving topicalization out of the antecedent of the elided material. The structure of (34) is in (35):

(35)   Mary, John [likes ~~Mary~~], and Peter does [~~like Mary~~] too

Let us assume, along with Chomsky and Lasnik (1993), Lasnik (1997) and other works, that VP ellipsis involves deletion of the VP under identity, at PF (whereas it remains available for interpretation in the LF component). Under this assumption, the conditions on ellipsis are strikingly reminiscent to the conditions on deletion of copies: both involve deletion under identity. In fact, Chomsky (1995c), Ch.3. (p. 203) takes deletion of copies to be an obligatory variant of a more general operation of deletion, optionally applying in the case of ellipsis. In contrast, Chomsky (1995c), Ch. 4. (p. 252-253) suggests to reduce ellipsis to a variant of the copy deleting procedure. In view of this apparent parallel, the question arises as to whether the elided material and its antecedent constitute a 'chain' of some sort, which is subject to Scan at PF, similarly to regular chains. If so, (35) would materialize the 'chain intertwining' configuration (24) which, by hypothesis, Scan is unable to process, as in standard 'subject condition' cases. That would be an incorrect result.

There are, however, reasons to believe that deletion of copies is unrelated, or at least not equivalent, to ellipsis. Nunes (1999) points out three such reasons. First, it is well known that the elided VP (cf. [*like Mary*] in the second conjunct of (35)) can alternatively be realized with a characteristic low-flat intonation (cf. Chomsky and Lasnik (1993), Tancredi (1992)). For copies, no such option is available: a low-flat intonation on *Mary* in the first conjunct in (35), for instance, does not salvage *Mary* from deletion. Second, the identity relation is somewhat looser in the case of ellipsis, compared to the case of copy deletion. Deletion of copies operates with non-distinct elements, whereas ellipsis may operate with elements that are morphologically identical, but are distinct (in the sense of fn. 7). In fact, ellipsis may even operate on elements that are not morphologically identical at all. Consider the following Serbo-Croatian example from Stjepanović (1997b):

(36)   Ivan je pročitao   knjigu, a   Marija nije ~~pročitala   knjigu~~

Ivan is read-3sg.m. book   and Marija isn't read-3sg.f. book

'Ivan read a book, but Marija didn't'

Stjepanović (1997b) argues that the process deleting the crossed material in (36) is akin to VP ellipsis in English. In (36) the verb (or, rather, the past participle form) in the deleted phrase differs morphologically from its antecedent in the first conjunct in a φ-feature, namely, gender. Differences may also obtain in other inflectional features, such as tense

on the verb. Stjepanović (1997b) then proposes that the identity requirement should be relaxed so as to allow ellipsis in cases like (36).[13]

Third, under the unification of regular chains and ellipsis 'chains', it remains a mystery why the former, but not the latter, are subject to proper syntactic conditions as Last Resort and the c-command condition.

We conclude, accordingly, that cases involving ellipsis such as (34) are not problematic for the conception of Scan in (22) and the failure of its operation on intertwined chains of the kind in (24). If we are correct, it follows that deletion for the purposes of ellipsis is regulated by a mechanism different from Scan.

Let us now compare the Chain Intertwining analysis with Takahashi's original chain uniformity analysis with respect to the types of movement that are banned from taking place out of a moved domain. We begin with Takahashi's theory.

Recall that Takahashi's analysis crucially relies on the 'shortest move' condition (4), requiring movement to the closest asymmetrically c-commanding target, and that target is relativized according to the movement type. To recap briefly, the ultimate target of *wh*-movement in *Who does [a picture of t_{who}] hang on the wall?* (Spec-CP) is an A'-position, and so *wh*-movement is supposed to proceed by a series of adjunctions to the

---

[13] Nunes' own example of seeming non-identity of verbal forms is the pair *said* vs. *say* in the example of Chomsky (1995c), Ch.3:

(i) John said that he was looking for a cat, and so did Bill [~~say that he was looking for a cat~~]

The example is illustrative under the strictly 'lexicalist' view on verbal morphology adopted in Chomsky's minimalist work, according to which lexical items enter the structure fully inflected. Lasnik (1995b) proposes a 'hybrid' approach to verbal morphology, which he shows to have important advantages over the strictly lexicalist view. Under the 'hybrid' approach, main verbs are introduced into the structure as bare forms, and are later combined together with the affix by a version of Affix Hopping. Under Lasnik's approach, both main verbs in (i) will be introduced into the structure as *say*. Depending on the timing of application of Affix Hopping and ellipsis, *say* may be deleted deleted under strict identity with *say* in the first conjunct. Thus, under the hybrid approach, Nunes' example of non-identity becomes less obvious.

maximal projections along the path of movement of *who*, among which would be adjunction to the DP itself (the step that violates chain uniformity), since, by hypothesis, adjoined positions count as A'-positions.

Takahashi crucially assumes, however, that the position adjoined to DP (which, for him, is the specifier of DP) has a dual A/A' status. As such, it must also be an intermediate site for A-movement taking place out of the DP (in accord with the 'shortest move'). If the DP has moved before, it creates a non-trivial chain. Consequently, the intermediate step of adjunction to the link of the non-trivial chain makes that chain non-uniform, and so cannot take place. It follows that A-movement out of the moved DP is blocked, which accounts, in particular, for the ungrammaticality of (37a):

(37) a. *John seems that [pictures of t$_{John}$] are on sale

    b. John seems that [$_{IP}$ [$_{DP}$ John [$_{DP}$ pictures of John]] [pictures of John] are on sale]

The offending step of the derivation is shown in (37b): *John* adjoins to a link of a non-trivial chain, making the latter non-uniform.

Thus Takahashi's theory rules out instances of A-movement out of the moved DPs. However, this result depends on the assumption concerning the dual A/A' property of Spec-DP. This assumption, however, is not straightforward, as the status of Spec-DP is actually somewhat controversial. More importantly, the theory seems to preclude A-movement <u>only</u> out of DPs. It is well known, however, that DPs are not the only domains that move. For instance, clauses in the subject position, finite and infinitival, move as

well, in accord with the predicate internal subject hypothesis. Consider (38a), with the structure in (38b):

(38)   a. *This car seems that [for John to park t$_{thiscar}$ here] is illegal

       b. *This car* seems that [for John to park *this car* here] [for John to park *this car* here] is illegal

In order for (38) to be ruled out via chain uniformity under Takahashi's theory, A-movement of *this car* out of a link of the non-trivial chain formed by movement of the infinitival clause from Spec-vP to Spec-IP must be adjoining to the clause itself, as an intermediate step of movement. However, it is unlikely that a clause-adjoined position in (38) is an A-position. Clause-adjoined positions are most naturally regarded as A'-positions. Hence, the shortest move condition does not require *this car* to adjoin to the clause. Hence, this movement is impossible, by Last Resort. Rather, *this car* moves to the next available A-position (possibly, the matrix Spec-IP itself). For Takahashi, then, (38) must be ruled out for a reason unrelated to chain uniformity. But this conclusion is unappealing, given that intuitively, (38) and (37) call for a uniform analysis.

On the other hand, if the controversial assumption regarding the dual A/A' status of Spec-DP is dropped, and all adjoined positions are treated, rather naturally, as A'-positions, then neither (37) nor (38) are ruled out by the chain uniformity analysis. This is because A-movement does not need to adjoin to the link of a non-trivial chain in (37) and (38), and proceeds to the next available A-position (perhaps, Spec-IP). That is, under

most natural interpretation of Takahashi's theory, A-movement is actually allowed out of moved domains.

It should also be noted that Takahashi's theory allows head movement out of moved domains. This is so because head movement does not need to make an intermediate step of adjunction to a link of a non-trivial chain, formed by previous movement of a domain. The shortest move condition does not require this step, since this adjoined position is not the position of the relevant kind (head position) on the way of the moving head. The adjunction step is therefore precluded by Last Resort.

In contrast, it should be clear that the Chain Intertwining analysis precludes all instances of movement out of a moved domain: A-movement, A'-movement and head movement. This is so because any movement out of a link of a non-trivial chain leads to the Chain Intertwining configuration (24), which, by hypothesis, cannot be processed at PF. In this respect, we believe, the Chain Intertwining analysis more closely captures the spirit of the 'Freezing Principle' of Wexler and Culicover (1981) (cf. fn. 1). It is also more appealing in that it provides a uniform analysis of cases like (37) and (38).

In the following section, we discuss independent evidence for an account of 'subject condition' effects as 'moved domain' effects, in terms of a version of chain uniformity. The evidence comes from the distribution of questions involving *wh*-extraction out of 'specific' DPs.

4. Extraction from 'specific' DPs

4.1 Presuppositionality and movement

It is well known since Chomsky (1973) and Fiengo and Higginbotham (1981) that extraction out of DPs even in the direct object position results in degradation, if the DP is introduced by an determiners *the, that* or a 'strong' quantifier like *every* or *all*. These DPs are commonly referred to as 'specific':

(39)   a. ?*Who did you see every/all/most/the/that picture of?

   b. cf. Who did you see a picture/two pictures of?

(40)   a. *Who did John read every/all/most/the stori(es) about?

   b. cf. Who did John read stories about?

(41)   a. *Who did Mary make every/all/most/the movi(es) about?

   b. cf. Who did Mary make movies/two movies about?

In the above works, the ungrammaticality of the a. examples in (39)-(41), as opposed to grammatical b. examples, is attributed to the *Specificity Condition*, which bans extraction from 'specific' DPs. The full extent of the condition was always hard to evaluate, because it was not clear what notion of 'specificity' enters into consideration. According to Fiengo and Higginbotham (1981), 'specific' NPs are those 'having or purporting to have some definite reference'. Fiengo and Higginbotham (1981) deal mainly with definite NPs introduced by *the*, although they acknowledge that some quantifiers also induce 'specificity' in the indicated sense. See also Enc (1991) for a treatment of specificity as a semantic concept.

The Specificity Condition has long remained a red herring in the theories of locality, including most recent ones. The reason for that, we believe, is its uncertain nature: even though the condition was originally adduced to explain a syntactic effect concerning extraction, it appeals to a notion of specificity which is not, strictly speaking, a syntactic feature and remains rather poorly understood. A major conceptual problem with any formulation of the Specificity Condition as an independent constraint on extraction is that this condition is redundant. In its empirical coverage, it seems to overlap with principles deriving 'subject condition' effects. This becomes apparent given that in languages in which there is no 'subject condition' effect, extraction is possible out of 'specific' subjects. This can be seen in languages like Hungarian which have overt definiteness markers. Consider again the Hungarian example from Chapter 1:

(42)   Melyik színésznőnek$_i$ gondolja János, hogy t$_i$ a fényképe meglett?   [Su]

which actress's   thinks   Janos that   the picture-her turned up

'Which actress does John think that a picture of _ turned up?'

In (42), the subject from which *wh*-extraction takes place, is 'specific'/definite. If the Specificity Condition operated on top of whatever derives Subject Condition effects, then the grammaticality of (42) is a priori unexpected. The grammaticality of (42) thus suggests that 'specificity condition' must be restricted to objects. But relativizing the Specificity Condition in this manner makes it rather unappealing from the conceptual point of view. It will be our intent to show that the 'Specificity Condition' effects are epiphenomenal, and are fully explained by the chain uniformity theory.

Three kinds of approaches have been proposed in the literature to derive the Specificity Condition. A purely semantic account is suggested in Erteschik (1973) in terms of her notion of 'semantic dominance'. This notion is defined in terms of the presence or absence of some contextual reference for the NP. Thus NPs headed by quantifiers like *every, all* or determiners like *the* must be semantically dominant, whereas NPs headed by *a, many,* or *several* need not be.

A purely syntactic approach is advanced by Bowers (1988). Bowers assumes that quantifiers like *every, all,* and determiners like *the* are of category D, and so head a DP, whereas quantifiers *a* or *many* are attached within the NP. He then proposes an account of the contrast between a. and b. examples in (39)-(41) in terms of *Barriers*. Roughly, in the b. examples, the NP is L-marked by the verb, and so extraction can take place. In the a. examples, the NP is not L-marked by the verb, because the category DP intervenes between the two. Consequently, the NP serves as a 'blocking category' and a barrier (and the DP inherits barrierhood as well). The extraction from this NP results in a Subjacency violation.

Diesing (1992) proposes a somewhat mixed approach which draws from both syntactic and semantic insights. She argues that the relevant distinction behind the contrast in extraction in (39)-(41) has to do with the notion of presuppositionality. Since Milsark (1974), presuppositionality is considered a central notion behind the distinction between NPs introduced by various types of quantifiers. Quantifiers like *every, all* (called 'strong' quantifiers in Milsark (1974)) and determiners like *the*, introduce NP the existence of which is granted by all speakers in the discourse: these NPs are presuppositional. On the other hand, as Milsark (1974) and Fodor and Sag (1982) point

out, indefinites introduced by 'weak' quantifiers like *a, several, two* and others, are inherently ambiguous. On one interpretation, they are presuppositional, similarly to DPs headed by 'strong' quantifiers and determiners. On another interpretation, they have a existential, or 'weak', reading: they introduce a variable bound by existential closure. The two readings of indefinites are illustrated in (43):

(43)   a. There are two cars in my garage (unstressed *two*, asserts existence of cars)

           b. Two cars are in my garage (others are in the parking lot; presupposes the existence of cars)

Diesing adduces several contexts in which the existential reading seems to be singled out. Perhaps the most obvious one, known since Milsark (1974), involves existential *there*-context. Thus, *a man* in *there is a man* is most felicitously interpreted as existential, not presuppositional. We will be concerned, however, with Diesing's two other contexts. First, an existential reading is brought out in generic contexts involving a 'verb of creation' such as *write, paint, draw*. This reading may be reinforced by an adverb of quantification in conjunction with present tense:

(44)   a. I (often) write books about slugs

           b. I (usually) draw a map of New Zealand

(44a) does not permit a presupposed reading (cf. 'whenever there is a book about slugs, I write it'), but only one in which books about slugs are asserted into existence. Similarly

for (44b). With 'verbs of using' like *tell, buy, play, publish,* on the other hand, both readings of the indefinite object are possible:

(45)   a. I usually buy pictures of Picasso

       b. I usually play a sonata by Beethoven

However, as Diesing observes, extraction out of the indefinite object of the 'verb of using' results in the loss of the presuppositional reading:

(46)   a. Who do you usually buy pictures of?

       b. Who do you usually play sonatas by?

(46) is grammatical only on the existential reading of the object, similarly to (44).

The idea behind Diesing's version of the 'Specificity Constraint' is to ban extraction out of NPs which are, in this sense, presuppositional. Syntactically, Diesing argues that presuppositional NPs (or DPs) undergo obligatory movement outside the VP in which they are generated. According to Diesing, movement is driven by the need to escape the 'nuclear scope', which is the LF counterpart of the VP in her system. On the other hand, indefinite NPs in the non-presuppositional interpretation ('non-specific') do not move outside the VP.[14]

---

[14] Borer (1994) advances a version of this proposal, according to which non-presuppositional ('non-specific' referential, according to her) indefinites end up in a position lower than the position to which presuppositional ('specific') DPs overtly move.

One might, of course, reasonably object that Diesing's motivation behind syntactic movement of the presuppositional NP is actually semantic, hence questionable. But ignoring that for the time being, what is interesting for our present purposes is that the final version of Diesing's Specificity Condition restates the contrast at issue (cf. (39)-(41)) essentially as a 'moved domain' effect:

(47) *Revised Extraction Constraint:* Extraction cannot take place out of an NP that must raise out of VP before tree splitting. [p. 128]

Diesing points out that (47) recalls Wexler and Culicover's 'freezing principle'. Still adopting the *Barriers* framework, Diesing further argues that (47) is reducible to Subjacency: if an NP has moved, it ought to be in a non-L marked position, and would constitute a 'blocking category' and a barrier for extraction. By appealing to previous movement as a reason for ungrammaticality of extraction out of a domain, Diesing's (47) obviously develops the same idea we adopted for our present purposes.

According to Diesing, the contrast in (39), repeated here, receives the following account:

(39) a. ?*Who did you see every/all/most/the/that picture of?

b. Who did you see a picture/two pictures of?

In (39a), the object is presuppositional, and, as such, must move out of the VP where it is base-generated. In a cyclic derivation, *wh*-movement must take place after that. But (47)

precludes extraction out of the moved object. In (39b), on the other hand, the object *a picture/two pictures* only has the existential reading, under which a picture or two pictures are asserted into the domain of discourse. Consequently, the object does not raise, and (47) does not preclude extraction out of it.

We adopt Diesing's description of the relevant interpretation of indefinites: one presuppositional, which we assume to be of a generalized quantifier type $<<e,t>,t>$, and the other 'weak', or existential (of type $<e>$).[15] We further suggest that Diesing's account of the contrast in (39) in terms of movement is essentially correct (although we do not necessarily adopt her motivation for movement): presuppositional objects raise out of VP, but existential (non-presuppositional) ones stay in situ. Diesing draws support for this proposal mainly from German, on the basis of positioning the presuppositional/existential objects with respect to some sentential adverbs. The observation that overt movement of the object is necessarily associated with a 'specific', or non-presuppositional reading, has been made also on the basis of other languages in which overt object shift (sometimes termed 'scrambling') is visible on the surface, such as Dutch (De Haan (1979), De Hoop (1996), among others), and Icelandic Collins and Thráinsson (1996). It is more difficult to verify it in languages like English, which do not allow 'scrambling' of the object of the kind observed in these languages.

---

[15] For bare plurals, we follow the fairly standard view (cf., e.g. Link (1983)) that plural NPs denote 'group individuals' or 'pluralities', of type $<e>$.

See De Hoop (1996) for an alternative view on the typology of indefinites. She suggests that an NP in a weak reading can be either of individual type $<e>$ or 'predicate modifier' type $<<e,t>,<e,t>>$. 'Strong' readings are associated with the type $<<e,t>,t>$, which is obtain by type shifting of the NP initially of type $<e>$. We adopt Diesing's division, which itself is largely drawn on Milsark (1974), as it seems to us more restrictive.

Nevertheless, we believe English facts also support the distinction. Our argument consists in verifying that presuppositional objects are indeed higher in the structure than existential objects. Binding into a VP adverbial provides us with a clue in this regard. First, observe that existential objects behave just like other referential NPs in their ability to bind a reciprocal:[16]

(48)   Who do you usually see [$_{SC}$ [friends of t]$_i$ in each other$_i$'s clothes]?

(48) is parallel to (46) in that it involves a 'verb of using' and extraction, which removes the presuppositional reading of the object, leaving only the existential one. Now, we assume, following Lasnik (1999b), that a binding configuration can be established as a result of overt raising, but not covert 'feature movement' (see below). Observe now that object DPs headed by 'strong' quantifiers and determiners can bind into a VP adverbial:

(49)   I wrote all/these books$_i$ on each other$_i$'s advance

Significantly, an existential object in contexts like (44) cannot bind into a VP adverbial:

(50)   *I often write books$_i$ on each other$_i$'s advance

---

[16] We are grateful to Željko Bošković for suggesting this test, and William Snyder for his help in constructing (48)-(52).

For some speakers, such binding is nearly or almost impossible even in non-generic contexts, and the object is most felicitously understood as non-presuppositional (existential):

(51)   ?*I wrote two books$_i$ on each other$_i$'s advance

The same point is reinforced in (52):

(52)   #I often criticize two authors$_i$ during each other$_i$'s interviews

Informants report that *two authors* can only bind *each other* on the 'strong', presuppositional reading, in which the same two authors are criticized (e.g. on a radio show), e.g. every week. On the reading in which two different authors are being criticized every time, binding is impossible.

Aside from lending support to the account based on different structural positions of presuppositional and existential indefinites, (48)-(52) also provides an empirical argument against the suggestion in Larson (1988) that adverbials are projected in the innermost complement position, rather than VP adjoined. Under Larson's alternative, *on each other's advance* would be generated below the direct object in (49)-(52), hence, nothing would preclude binding a reciprocal and (50)-(52) are incorrectly predicted to be grammatical. At the same time, these data also support the claim advanced and defended in Lasnik (1999b), following in part Koizumi (1995), that objects in English undergo overt object shift, although we now suggest to restrict this raising to presuppositional

objects. Lasnik (1999b), and earlier, Lasnik and Saito (1991), show, in particular, that in ECM constructions an associate of the expletive *there* cannot bind into a matrix clause VP adverbial:

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

This is expected under Larson's suggestion since the adverbial modifying the matrix predicate could not have started lower than *two men*. However, the grammaticality of (54) is unexpected:

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

Again, the matrix adverbial could not possibly be generated lower than *two men* in the complement clause, hence, all else equal, it is not clear how binding obtains. Lasnik concludes that Larson's alternative therefore cannot be correct. Rather, he maintains, the adverbial is adjoined to the VP, in more standard, pre-Larsonian, fashion. (see also Stjepanović (1997a) for further arguments against the Larsonian structure).

Lasnik and Saito (1991), following in part Postal (1974), argue that *two men* in (54) raises to Spec-AgrO, the position in which it can bind into an adverbial adjoined to VP. This and Lasnik's subsequent work also presents a number of independent arguments that binding configurations of the sort in (54) are a result of overt raising to Spec-AgrO. In (53), on the other hand, there is no overt raising. As a result, no binding obtains.

Following Chomsky (1995c), Lasnik suggests that instead, formal features of *two men* raise to adjoin to *there* at LF. Departing from Chomsky, however, Lasnik claims that raising of formal features is not sufficient to establish a binding configuration. Lasnik then adopts for English the structure of VP proposed in Koizumi (1995) in which object overtly raises to Spec-AgrO and verb raises to a still higher head position, the head of a higher VP 'shell', resulting in the surface word order. Our binding test confirms this sort of analysis for English presuppositional objects.[17] Note, incidentally, that *two men* in (54) felicitously receives a presuppositional, but not existential, reading.

The chain uniformity theory that we adopted and further developed in the previous section readily formalizes the correlation between the impossibility of extraction out of 'presuppositional'/'specific' objects and their overt raising. Movement of the 'specific' object DP will create a chain, modifying which will lead to non-uniformity of that chain. Thus, the chain uniformity account offers a straightforward explanation of the Specificity condition on extraction in English.

---

[17] Željko Bošković (p.c.) points out to us that NPI licensing is apparently possible with object QPs introduced by the negative quantifier; cf.
(i)   The DA convicted no criminal during any of the trials
At the same time, presumably, the negative QP shows up in existential contexts, cf:
(ii)   There was no criminal in the room
If we assume on the basis of examples like (ii) that DPs headed by *no* are only non-presuppositional, then, given the possibility for *no criminal* to license an NPI in (i), our analysis would seem to predict that even non-presuppositional (existential) objects overtly raise, contrary to what we are claiming in the text. However, (i) is compatible with our proposal. First, according to speakers' intuitions, *no criminal* in (i) can have a reading actually presupposing the existence of a previously established set of criminals. According to Diesing (1992), Heim and Kratzer (1998), Reinhart (1995), and ultimately to Strawson's work (see also Barwise and Cooper (1981)), QPs headed by *no*, just as those headed by quantifiers like *every*, *all*, have a presuppositional reading. As a 'weak' determiner, however, *no* can have a non-presuppositional (existential) reading, as well.

4.2 Hindi vs. English: 'Specificity' and EPP

Mahajan (1992) represents, to our knowledge, the earliest attempt to unify 'specificity' effects with 'subject condition' effects. Mahajan makes an interesting suggestion that a 'specificity' effect in cases like (39a)-(41a) has to do (although indirectly) with the fact that the object in these cases is actually located in a specifier of a functional projection. In essence, he adopts the proposals of Johnson (1991) concerning overt object shift for reasons of checking structural (accusative) Case, also developed in Koizumi (1995) and Lasnik and Saito (1991), among other works, but restricts it to 'specific' objects (the result we, and earlier Diesing, also arrived at independently in the previous section). More precisely, he argues that the object moves to a spec position of some XP, in which it is governed by the AgrO, a sister of XP.

Mahajan then speculates that the impossibility of extraction out of specific DPs has to do with the fact that the derived position of the object is structurally similar to the position of the subject (also a specifier), namely, Spec-IP. Consequently, he attributes the ungrammaticality of (39a)-(41a) to the same condition which rules out extraction out of subject in English: according to him, extraction out of both kinds of domains results in a 'CED effect'.

Mahajan (1992) adduces strong evidence for the 'unification' approach to the two kinds of effects. He observes that it is possible to extract out of 'specific' DPs in Hindi. The relevant example is below:[18]

---

[18] Mahajan (1990) gives evidence that 'specific' objects in Hindi show object agreement with the verb, but non-specific ones do not, which we discuss below. Thus in (55) the agreement between *curaaii thii* and *kitaab* indicates that the phrase headed by *kitaab* is specific. Alternatively, 'specific' objects can be marked morphologically with the ending *-ko*.

(55)   Kiskii tum socte ho ki Mohan-ne   kitaab   curaaii   thii?

        whose you think   that Mohan-erg. book-f. stolen (f.) be (past.f.)

        lit. 'Of whom do you think that Mohan stole the book?'

At the same time, Hindi is a language in which extraction out of subject is possible:

(56)   a. Kiskii tum socte ho ki   kitaab corii ho gayii?

        whose you think   that book stolen got

        'Whose book do you think got stolen?'

        b. vo larka jise Mary čumti he tik   he vo bimar he

        the boy who Mary kisses   good is he sick is

        '*The boy who that Mary kisses is a good thing is sick'

The significance of (55) and (56) becomes clear when one considers those with their counterparts on English, which are, of course, ungrammatical. Taken together, the resulting paradigm points to a one-to-one correspondence between extraction out of subjects and extraction out of 'specific' DPs in a language: if one is allowed, the other is too (Hindi), if one is precluded, so is the other (English). The question then is how this correspondence can be accounted for.

Mahajan (1992) hints that an explanation of this parallel may be derived from considering both phenomena on a par as a 'CED effect'. Then the correlation in question is not coincidental: whatever principle is responsible for 'CED' phenomena, holds in

English, but not in Hindi. However, Mahajan remains explicitly silent concerning the nature of this principle (cf. his fn.7), pointing out simply that it might have something to do with the structural status of the position of the domain out of which extraction takes place (a specifier of a functional projection; see above).

The chain uniformity theory provides the beginning of a solution. First, we note that there is in fact some evidence that subjects in Hindi can stay in their base generated position. Hindi is a free word order language in which the unmarked surface word order is SOV. Mahajan (1990) (p. 32-33) discusses examples like the following:

(57)   a. ?mohan-ko$_i$ apne$_i$ baccoN-ne   ghar se nikaal diyaa

            Mohan-DO   self's children-SUB house from throw give-perf.

            lit. 'Self's children threw out Mohan from the house'

          b. cf. *apne$_i$ baccoN-ne   mohan-ko ghar se   nikaal diyaa

            self's children-SUB mohan-DO house from throw give

Mahajan argues that the ergative Case marker *-ne* marks an inherent Case on the subject. Consequently, the subject does not have to move to check structural Case, as standardly assumed for English. The object *mohan-ko*, on the other hand, is overtly located in some A-position from which it can license an anaphor inside the subject (Mahajan assumes that this position is Spec-TP). Assuming that *ghar se* 'from the house' is a sort of indirect argument of *nikaal diyaa* (or just *diyaa*), it is plausible that the subject *apne baccoN-ne*

is base generated in the highest VP, in accord with a structure that postulates more than one layer of arguments, inside a VP 'shell'.[19]

As discussed in the previous section, the chain uniformity provides an account of both the 'subject condition' effect, and the 'specificity' effect in English. In both cases, the domain out of which extraction takes place, has previously moved, creating a non-trivial chain. Extraction from a part of this chain then makes the latter non-uniform, which results in a PF violation. The chain uniformity account also naturally accounts for the correlation between the possibility of extraction out of subjects in Hindi (55) is possible and the fact that subjects can stay in Spec-VP: under chain uniformity, the non-moved subject constitutes a trivial chain, hence, extraction out of it does not lead to a violation of chain uniformity.

Let us now turn to (56). The chain uniformity theory, as a uniform account of 'subject condition' and 'specificity' effects, provides a natural explanation of its grammaticality as well. It follows that the 'specific' NPs in Hindi do not undergo movement, as in English; rather, they remain in situ, by the time extraction takes place. (Diesing's (47) has this consequence as well, unsurprisingly).

The claim that 'specific' objects do not move in Hindi is in flagrant contradiction with Mahajan (1990), who assumes the existence of overt object shift in Hindi, and further argues that NPs undergoing movement are interpreted as 'specific'. However, virtually all evidence that Mahajan adduces in support of overt object shift in Hindi pertains to the surface structural <u>position</u> of the 'shifted' object, which he claims to be Spec-AgrO, and

---

[19] The subject in Hindi does not have to bear the -*ne* marker. In this case, Mahajan argues that the subject raises to Spec-TP and obtains structural (nominative) Case, in the usual manner.

where the object, according to him, checks structural (accusative) Case. The motivation for a movement process to this position appears to be strictly conceptual, and is dictated by the pre-minimalist framework that he assumes, in which verbal arguments are generated at the level of 'D-structure' under VP representing the argument structure (or theta-grid) of the verb. Since AgrO is unrelated to the argument structure, the reasoning would seem to go, movement is necessitated to account for both the 'D-structure' position and the surface position.

Interestingly, the only empirical facts that Mahajan adduces in support of the 'argument shift' actually seem to argue <u>against</u> the movement to Spec-AgrO. The first argument concerns the absence of Condition A reconstruction effects with shifted NPs. Consider the following:

(58)   raam-ne$_i$ mohan-ko$_j$ apnii$_{i/j}$ kitaab   lOTaaii

         Ram -SUB Mohan-IO self's book-f(DO) return-perf.f.

         'Ram returned self's book$_{i/j}$ to Mohan$_j$'

In (58), the anaphor inside the direct object (DO) can be licensed by the NP in the indirect object (IO), or subject position. This is expected, under the rather plausible assumption that indirect objects are (or may be) generated higher than direct objects, and assuming the pronounced structure of VP with an Agr projection for each object, along the lines of Koizumi (1995). Since Hindi is a free word order language, the order of IO and DO may usually be reversed, without affecting grammaticality. However, doing so in (58) results in the change of binding possibilities:

(59)   raam-ne$_i$ apnii$_{i/*j}$ kitaab   mohan-ko$_j$ lOTaaii

Ram -SUB self's book-f(DO) Mohan-IO return-perf.f.

'Ram returned self's book$_{i/j}$ to Mohan$_j$'

In (59), the reflexive can no longer be bound by IO, but only by the subject.[20] But if DO had a chance to be generated below IO, as in (58), Condition A, an 'anywhere' condition (cf. Belletti and Rizzi (1988)), would presumably be satisfied, prior to movement. That Condition A is not satisfied suggests that DO has never been under IO, contra the movement analysis.

The second piece of evidence comes from the absence of Condition C reconstruction effects. (60) demonstrates a Condition C effect for DO:

(60)   *mE-ne use$_i$ raam$_i$ ki kitaab   dii

I-SUB him-IO Ram gen. book(f)-DO give-perf.(f.)

'I gave Ram's book to him'

However, if DO appears prior to IO, the condition C effect disappears:

---

[20] Note that the loss of binding by the indirect object cannot be attributed to a linear precedence requirement of the kind we proposed in Chapter 1 for Russian pronouns. As in Russian, in Hindi the anaphor can in principle linearly precede the binder, and still be licensed ((i) is from Mahajan 1990):
(i)   [apnii$_i$ raam$_j$ vaalii kitaab]$_k$ mE-ne$_i$ t$_k$ use$_j$ dii
      self's Ram's   book-DO I-SUB   him-IO give-perf.
      lit. 'My Ram's book, I gave to him'

(61)   mE-ne raam$_i$ ki kitaab$_i$ use$_i$ dii

         I-SUB Ram gen. book(f)-DO him-IO give-perf.(f.)

These facts find no direct explanation under the movement analysis. On the other hand, under the hypothesis that the 'shifted' position of the object is actually base-generated, nothing more needs to be said to ensure the absence of Condition A and C effects in (59) and (61).[21]

Neeleman (1994) discusses similar phenomena in Dutch and German. The latter, of course, is the language which Diesing originally used to support her claims concerning overt movement of presuppositional indefinites. Interestingly, extraction out of presuppositional DPs is in fact possible in German:

(62)   Über Chomsky$_i$ habe ich [den letzten Film t] leider   nicht gesehen

         about Chomsky have I   the last   film   unfortunately not   seen

         'Unfortunately, I have not seen the last film about Chomsky'

As Neeleman shows, in German (and Dutch), as in Hindi, it is possible for IO and DO to appear in either order, as shown below:

(63)   a. Ich habe dem Hans das Bild gezeigt

              I   have the Hans the picture shown

---

[21] Mahajan (1990) attempts to explain these facts by suggesting that theories of 'reconstruction', in particular, the view on Condition A proposed in Belletti and Rizzi (1988), must be reconsidered. Mahajan 1994, revisiting the same facts, leaves the issue open.

'I showed Hans the picture'

b. Ich habe das Bild dem Hans gezeigt

However, if DO contains an anaphor bound by IO, then it can only follow, not precede IO:

(64)   a. Ich habe [den Männern]$_i$ das Bild   von einander$_i$ gezeigt

I   have   the men   the picture of   each other shown

'I showed the men each other's picture'

b. *Ich habe das Bild von einander$_i$ [den Männern]$_i$ gezeigt

On the basis of this, and other sorts of evidence, Neeleman concludes that what looks like overt object shift is actually base-generation of the object (see also Neeleman and Reinhart (1998) for further discussion). In a similar vein, Miyagawa (1997) argues that the DO-IO and IO-DO positions are base-generated in Japanese (see also Bošković and Takahashi (1998) and references there). These results further support the idea that extraction out of presuppositional objects correlates with their non-movement.

So far, we have been able to keep to the chain uniformity theory to explain the presence and absence of 'specificity' effect in English and languages like Hindi, respectively. But the chain uniformity account does not completely suffice to explain Mahajan's observed correlation between extractability out of 'specific' DPs, on the one hand, and subjects, on the other. However, the chain uniformity theory establishes a common denominator between the two types of phenomena: movement. Recall that under

chain uniformity, extraction out of subjects is precluded exactly in those cases which involve previous movement of subjects. Similarly for cases of extraction out of 'specific' phrases. The question thus boils down to the following: what common mechanism underlies overt movement of the subject to Spec-IP and movement of a 'specific' phrase to Spec-AgrOP, and can this mechanism be parameterized?

We suggest that the answer can be formulated in the system of Chomsky (2000). Departing from the previous theories, Chomsky suggests that overt movement is triggered by the 'EPP requirement', which is essentially a requirement on functional projections to project and fill in a specifier. This is the only trigger for overt movement. Chomsky also postulates the operation Agree, which holds between two heads, 'probe' and 'goal', and consists in establishing a matching relation in their formal features, such as Case and φ-features.[22] Agree is an independent operation and does not by itself necessitate movement, unlike in the previous frameworks. Thus in English, Chomsky argues, subject undergoes overt movement due to the 'EPP requirement' of T, on top of the operation Agree that establishes a matching relation between the (uninterpretable) φ-features of T, and (interpretable) features of the subject DP.

For the case of object movement, Chomsky hypothesizes a feature that can optionally be present on the light verb v (or AgrO, in our terminology). This feature, which Chomsky terms P-feature, is essentially a generalized version of the 'EPP requirement': again, it requires a functional head to project a specifier. For ease of exposition, we refer to it as an 'EPP property'. If the EPP-property is present on v, it triggers overt 'object shift'

---

[22] This is a bit of oversimplification. For Chomsky, Case is not by itself a feature that needs to be checked, but, rather, a property that activates the 'goal' head for checking φ-features.

to the Spec-*v*. If the EPP-property is not present on *v*, 'object shift' does not occur, for reasons of economy allowing only motivated instances of movement. The semantics makes use of the difference between *v* with EPP-property and the one without EPP-property, although it cannot itself dictate one option or another. Chomsky further remarks that the relevant semantic considerations involve specificity. Following this suggestion in its essentials, we assume that objects that enter into the Agree relation with *v*, when the EPP property is chosen, are interpreted presuppositionally.[23]

Let us adopt this system in its essentials, adjusting for AgrO instead of *v* to ensure continuity with our previous discussion, and adopting the proliferate structure of VP along the lines of Koizumi (1995). Consider then the contrast between extraction out of presuppositional vs. non-presuppositional DPs in English (39)-(41), repeated below:

(39)   a. ?*Who did you see every/all/most/the/that picture of?

       b. cf. Who did you see a picture/two pictures of?

(40)   a. *Who did John read every/all/most/the stori(es) about?

       b. cf. Who did John read stories about?

(41)   a. *Who did Mary make every/all/most/the movi(es) about?

       b. cf. Who did Mary make movies/two movies about?

---

[23] We slightly depart from Chomsky (2000). According to Chomsky, it is the configuration involving an object raised to the "edge" of vP, that leads to a particular interpretational effect (viz. presuppositionality/'specificity') of the object. For us, the situation when *v* has the EPP property, and establishes an Agree relation with the object, is sufficient for the presuppositional reading of the object to arise. That is, for us, overt raising of the object is not directly related to its presuppositional interpretation. This distinction will become important when we consider the Hindi case, below.

In the a. examples, the EPP property is chosen on AgrO, triggering overt raising of the object, which in turn creates a non-trivial chain. Eventually, *wh*-extraction takes place from part of the chain. At the interface with semantics (LF), the raised object is interpreted presuppositionally. At PF, the chain is evaluated along the lines of Section 3, and is found illegitimate. Consequently, the examples are ruled out as PF violations.

In the b. examples, the EPP property is not chosen, and no overt raising takes place. No non-trivial chain is therefore created. Later in the derivation, *wh*-extraction then takes place. Consequently, at LF no presuppositional interpretation is assigned to the object. At PF, no illegitimate objects are produced. The examples are therefore grammatical.

Thus, following Chomsky, we assume that a common mechanism underlying overt movement of both subjects and 'specific' objects in English is the EPP property.[24]

Consider now extraction out of 'specific' objects and subjects in Hindi, illustrated in (55) and (56), repeated (in part) below:

(55)   Kiskii tum socte ho ki Mohan-ne   kitaab   curaaii   thii?

         whose you think   that Mohan-erg. book-f. stolen (f.) be (past.f.)

         lit. 'Of whom do you think that Mohan stole the book?'

(56)   a. Kiskii tum socte ho ki   kitaab corii ho gayii?

         whose you think   that book   stolen   got

         'Whose book do you think got stolen?'

---

[24] It remains an open question in Chomsky (2000) why the EPP feature on *v* is allowed to be optional, but the EPP feature on T is obligatory in English.

As we argued above, in both cases overt movement is not triggered, unlike in English. We are now in a position to make a proposal as to where the parametric difference between English and Hindi lies.

Consider first (55), involving extraction out of a presuppositional DP. Mahajan observes (cf. fn. 18) that presuppositional DPs show obligatory morphological object agreement with the verb, also shown in (55). If the DP is not 'specific', it does not show morphological object agreement with the verb. This is illustrated in simple cases in (65) from Mahajan (1990):

(65)   a. siitaa-ne laRkaa dekhaa   (object agreement)

        Sita-erg boy-m saw-m

        'Sita saw the boy'

    b. siitaa-ne laRkaa dekh rahii hE   (no object agreement)

        Sita-erg boy-m see-prog.-be-f.

        'Sita is looking for a (suitable) boy (to marry)

According to Chomsky, subject agreement (with the verb) in English involves establishing the relation Agree between the functional head T and the NP in φ-features, and subsequent raising of the subject to Spec-TP, in satisfaction of the EPP-property. Similarly, we suggest that in Hindi AgrO (or v) establishes the Agree relation with the object. We propose that just like in English, AgrO has an optional EPP property. However, this EPP property is of a different kind. Chomsky (2000) suggests, indirectly, that the EPP property can in principle be satisfied either by XP elements (phrases), or by

X^0 elements (heads). Following in part Alexiadou and Anagnostopoulou (1998)'s discussion of the EPP property of Tense, we want to suggest that this distinction in the nature of the EPP property is realized parametrically. In languages like English, it is satisfied by XP elements, whereas in Hindi, it is satisfied by X^0 elements. In the present case, AgrO needs an X^0 to be satisfied, and the only element that qualifies for this purposes is the V head. V then raises to AgrO, in satisfaction of the EPP property, and this results in 'rich' morphological object agreement (this captures the intuition, also explored in Alexiadou and Anagnostopoulou (1998), that overt V-movement is more likely to have a morphological reflex). At the interface of LF, as usual, the agreeing object is interpreted presuppositionally. At PF, no illegitimate objects are produced. (55) is therefore grammatical.

Consider now the derivation in which the EPP property is not present. Again, the Agree relation is established between AgrO and the object (necessary for checking Case and φ-features). However, V does not raise to AgrO in this case. Consequently, no 'rich' morphological agreement ensues. At LF, the object is interpreted non-presuppositionally. At PF, no illegitimate objects are produced. Hence, extraction out of non-presuppositional objects is correctly predicted to be grammatical.

Let us now turn to (56). Importantly, nothing more needs to be said in order to ensure the absence of 'subject condition' effect in this example. Here, T establishes the Agree relation with the subject, and has the EPP property, which needs to be satisfied.[25] Since the EPP property is satisfiable only by X^0 elements, T will seek a suitable head to attract

---

[25] Unlike the case of v, we do not assume that the EPP property T is necessarily associated with presuppositionality (see also Chomsky (2000)).

for the EPP purposes. The available head in this case is the V+AgrO complex, which then raises to adjoin to T. No subject movement is required. Consequently, the subject does not create a non-trivial chain, and extraction out of it does not affect grammaticality.

The proposed account recalls to a large extent the argument in Alexiadou and Anagnostopoulou (1998), (1999) made for the EPP in the domain of Tense. Alexiadou and Anagnostopoulou argue that the 'rich' verbal morphology of null-subject languages contains a nominal element bearing the features found on English subject pronouns. Consequently, they argue, EPP in those languages is satisfied by verb movement to the subject agreement head, producing a head-adjunction structure. This contrasts to languages like English, in which verbal morphology is sufficiently impoverished so that this option is not available. It should be noted that while our present proposal parameterizes the choice of the EPP property to XP and X^0 elements, Alexiadou and Anagnostopoulou (1999) try to derive this distinction on more principled grounds. They argue, in particular, that head-adjunction structure is preferable to projecting a specifier, for reasons of economy of projection (see also Bošković (1997c) and Pesetsky and Torrego (2000) for related discussion). Hence, when both options are available, the head-adjunction option is chosen, rendering the projection of a specifier unnecessary. We believe that something along these lines may be on the right track, although at this point we prefer to leave the distinction at issue as a parameter.

Our proposal has a number of attractive properties. First, it provides a principled 'unified' account of Mahajan's observed one-to-one correspondence between extraction out of subjects and out of presuppositional DPs. Second, it provides strong empirical support to the framework of Chomsky (2000) in which overt movement in natural

language is attributed to a single property of the computational system, namely the EPP property, arguably a conceptually desirable and elegant move. Third, it provides considerable support to the chain uniformity theory, extending its empirical coverage beyond the original spectrum of data (extractability out of subjects). Fourth, it is minimalist in spirit, as it needs no additionally postulated mechanisms over and above the minimal set of assumptions concerning the theory of movement in Chomsky (2000).

5. Conclusion

The chain uniformity theory, explored, modified and extended in this chapter, has a number of important consequences that need to be further explored. For one thing, the proposal unifying 'subject condition' and 'specificity' effects in terms of the parametric nature of the EPP property obviously extends beyond Hindi, to other languages, including those from Chapter 1, in which 'subject condition' effects do not obtain: Turkish, Japanese, Hungarian, Palauan, Russian, German, Basque, and Navajo. It follows that even in those languages, the EPP property is an X^0 property. By the criteria alluded to in Alexiadou and Anagnostopoulou (1998), this should correlate with the *pro*-drop status of the language. This seems to be the case for most languages in the list, although questions arise in the case of German (which is argued to be at least a partial pro-drop language in Grewendorf (1989), but that is not an uncontroversial assumption) and Russian (argued to be only a non-thematic pro-drop language in Franks (1995), but see Benedicto (1993) who argues that Russian is a full pro-drop language). Similarly, the results achieved in this chapter on the basis of Hindi (and German) imply that in all the languages above it is possible to extract out of 'specific' DPs. Unfortunately, it is difficult to test this prediction

in most languages except Hungarian, where extraction out of NPs is generally very difficult or impossible, to begin with. For Hungarian, the prediction is borne out (see (42) above).

To the extent the results we arrived at in this chapter hold, the chain uniformity theory represents a viable part of the 'eclectic' approach to 'CED'-type extractability that we are in the process of developing. In the next chapter, we turn to developing the second part of the 'eclectic' approach, which addresses in detail the syntactic behavior of adjuncts and issues related to extraction out of them.

CHAPTER 3

An 'Eclectic' Approach Part II: Late Adjunction

In this chapter we propose a theory of 'adjunct condition' effects, and their apparently universal character. The theory that we propose is based on the hypothesis that adjuncts enter the structure (by Merge) postcyclically, that is, after the cyclic portion of the derivation has been completed. The proposal itself is outlined in Section 1. The rest of the chapter is devoted to building an argument for the postcyclic Merger of adjuncts (cf. also Stepanov (2000a), (2001)). It will be shown, in particular, that postcyclic Merger of adjuncts follows from the bare phrase structure theory of Chomsky (1995a) given a version of the condition on structure building along the lines of 'Least Tampering' in Chomsky (2000), namely, that the set of syntactic relations must be preserved in the existing structure.

1. 'Adjunct Condition' effects: A proposal

A textbook case of an 'Adjunct Condition' effect is exemplified below:

(1)   ?*What did John [$_{VP}$ [$_{VP}$ go to bed] [after Peter fixed t$_i$]]?

As mentioned in Chapter 1, unlike 'subject condition' violations, 'adjunct condition' effects appear to be universal: they hold consistently from language to language. Any theory explaining 'adjunct condition' effects must address their universality.

In the current framework, one may think of at least three conceivable possibilities for explaining the universal character of 'adjunct condition' effects in (1). One may attribute them to a constraint at the PF interface, which derives, for instance, from the impossibility to linearize a chain one part of which is located inside an adjunct. Alternatively, one may attribute them to a constraint at the LF interface, which has to do with the impossibility of interpreting constructions in which the *wh*-phrase overtly moves out of a structural adjunct. The third possibility is that the ungrammaticality of extraction out of adjuncts has something to do with the nature and syntactic behavior of adjuncts, as opposed to non-adjuncts (viz. complements). Adjuncthood itself is a phrase structure notion, and is, presumably, universal. For the present purposes, we discard the first two possibilities, as we do not see how either of those could instantiated. We believe the third possibility is correct. In this, we partially agree with Nunes and Uriagereka (2000) who also derived the 'adjunct condition' effect based on the phrase structure status of adjuncts (although, of course, we disagree with them in the issue of whether this approach should be extended to 'subject condition' effects; see Chapter 1). The latter possibility is the one that we will explore in this chapter.

We assume the standard minimalist theory of structure building in terms of Merge and Attract, along the lines of Chomsky (1995c) (as far as we can see, our proposal is also restatable under the assumptions concerning Merge and movement in the system of Chomsky (2000)). We suggest that the syntactic behavior of adjuncts that plays a role in inducing 'adjunct condition' effects is captured in the following hypothesis:

(2) *The Late Adjunction Hypothesis (LAH):* Adjuncts must be Merged into the structure postcyclically.

As we show in Section 2, (2) is not an additional constraint imposed on the computational system, but a straightforward consequence of the structure building algorithm in the minimalist bare phrase structure theory, and needs no additional justification. In Section 2.2.4, we also propose a definition of adjuncthood that delimits the spectrum of application of LAH. For now, let us adopt (2) and see how it derives the ungrammaticality of (1).

The phrase *after Peter fixed what* is an adjunct. By LAH, this phrase must be Merged with the rest of the structure (cf. the vP [$_{vP}$ ... *go to bed*]) postcyclically. In particular, it cannot be Merged by the time the interrogative feature Q of the matrix complementizer is Merged with the IP *John go to bed*. In other words, there exists a point in the derivation at which the two phrase markers are unconnected to each other (irrelevant details omitted):

(3)   [$_{CP}$ Q [$_{IP}$ John go to bed]]

        [$_{Adj}$ after Peter fixed what]

Under the fairly standard minimalist assumptions concerning feature checking, the Q feature of the matrix C in (1) must be checked at the point of insertion, at the next step of the derivation.[1] (In terms of Chomsky (2000), this feature of C has the EPP property that

---

[1] See Lasnik (1999c) for detailed evaluation of the recent versions of the checking theory, and Chomsky (2000) for its further development.

must be satisfied by displacing *what* into its specifier). At this point it is appropriate to clarify our assumptions concerning the movement operations. We follow Chomsky (1995c), and later works in adopting the c-command condition on movement chains, namely, that α must c-command its trace. This view confines all movements within a single phrase marker (see Nunes (1995) and Bobaljik and Brown (1997), among others, for alternative views).

Under the view confining movement to a single phrase marker, the relevant feature of the matrix C cannot be satisfied by *wh*-movement in (1), since the only available candidate for such movement, namely, *what*, is not part of the same phrase marker. As a result, this feature remains unchecked, which is responsible for the ungrammaticality of (1).[2]

The analysis extends to all instances of complex structural adjuncts. Thus, for instance, assuming that restrictive and non-restrictive relative clauses are structural adjuncts, the so called Complex NP island, exemplified below, straightforwardly falls under the 'adjunct paradigm':[3]

(4)   a. ?*What did John meet the man who was wearing?

       b. *What did John meet Peter, who was wearing?

---

[2] The well known argument/adjunct asymmetry with regard to extraction, traditionally stated in terms of Subjacency/ECP (cf. Lasnik and Saito (1984)), requires further attention.

[3] As Pesetsky (1987), fn. 14 notes, adjuncts in Japanese seem to be Complex NPs. For instance, *mae-ni* 'before' is presumably marked by the (dative) Case marker *ni*. Hence, one would not expect a distinction between Complex NP effects and 'Adjunct Condition' effects.

Several possibilities arise concerning the third case, involving noun complements:

(5)   ?(?)What did John hear [a claim [that Peter bought t]]?

If Stowell (1981) is correct, the embedded clause in this case is an adjunct as well, and the case reduces to the first two. However, examples like these sound generally much better than examples involving relative clauses. This suggests that perhaps there is no Subjacency effect at all in these cases. See, however, Chomsky (1986a).

Beside complex NPs, the proposed analysis of 'Adjunct Condition' extends to other kinds of 'unextractable' domains that can be (re-)cast as structural adjuncts (e.g. *if*-clauses; cf. Cattell (1976), among others).

2. The Late Adjunction Hypothesis

2.1 Previous arguments

2.1.1 Adjuncts can *be merged late*

Perhaps the best known argument that adjuncts <u>can</u> be Merged counter-cyclically is due to Lebeaux (1988), (1991), (see also Riemsdijk (1981), Freidin (1986)). It involves certain 'anti-reconstruction' effects with respect to Condition C. Consider the following examples:

(6) a. ?* Which argument that John$_i$ is a genius did he$_i$ believe?

  b. Which argument that John$_i$ made did he$_i$ believe?

  c. *He$_i$ believed the argument that John$_i$ is a genius

d. *He$_i$ believed the argument that John$_i$ made

The R-expression inside a clause which is an argument of an NP induces a Condition C violation when a coreferent NP c-commands the extraction site, as shown in (6a) vs. (6c). However, if the clause modifies the NP (that is, is a structural adjunct), the Condition C effect disappears ((6b) vs. (6d)). Assuming that Condition C is an 'everywhere' condition (that is, must be satisfied at every point in the derivation), Lebeaux concludes that the structural adjunct-modifier in (6d) has an option of not being Merged at D-structure at all. This option is not available for the structural complement as the argument clause in (6a) vs. (6c).[4]

Lebeaux accounts for these facts by developing a theory that allows adjuncts to be inserted acyclically. He argues that arguments are required by the Projection Principle, whereas adjuncts are not. Arguments are thus required to be present at 'D-structure' but adjuncts can be added later in the (overt part of the) derivation.[5]

Kuno (1997), Postal (1997) and Lasnik (1998) put to a closer scrutiny the part of Lebeaux's paradigm concerning the noun-complement constructions. In particular, Lasnik (1998) argues at length that most of his examples involving noun complements can be plausibly ruled out on independent grounds, namely, because of their pragmatic oddity.

---

[4] There is a conceptual reason why Lebeaux does not adopt a stronger version of his claim. He takes the adjunction to be an instance of an operation Adjoin-α. Working in the late GB framework, he designs this operation after Move-α, a process that applies freely. Thus, for Lebeaux, claiming that adjuncts must be merged late would amount to depriving Adjoin-α of its free character, a non-minimal assumption in his terms.

[5] Lebeaux's system can be recast in the minimalist framework which restores the mechanism of generalized transformations as a principal method of structure building. See Chomsky (1995c), Ch.3), Kitahara (1995).

Once the pragmatic factor is controlled for, such examples become quite acceptable, as the following shows (examples from the authors mentioned above):

(7)   a. Which piece of evidence that John$_i$ was guilty did he$_i$ successfully refute?

b. How many arguments that John$_i$'s theory was correct did he$_i$ publish?

c. Whose claim that the Senator$_i$ had violated the campaign finance regulations did he$_i$ dismiss as politically motivated?

d. The claim that the director$_i$ was corrupt, he$_i$ was unwilling to discuss

e. The widespread belief that John$_i$ is incompetent, he$_i$ deeply resents

Although this type of counterexample does not directly affect the late adjunction hypothesis, it casts some doubt on whether the paradigm in (6) can motivate it. There are, however, other pieces of evidence that support the late adjunction idea. One argument comes from Bošković and Lasnik (1999) who discuss the Pseudo-opacity effects that obtain with adjunct *wh*-extraction, but not at all with argument extraction (cf. Rizzi (1990)), as in the following examples:

(8) a. [Combien de livres]$_i$ a-t-il beaucoup consultés t$_i$

'How many of books did he a lot consult?'

b. *Combien$_i$ a-t-il beaucoup consultés [t$_i$ de livres]?

'How many did he a lot consult of books?'

cf. c. Combien$_i$ a-t-il consultés [t$_i$ de livres]?

Here, in both cases the extracted *wh*-phrase crosses the modifier *beaucoup*. Boškovic and Lasnik's concern is why the extraction of argument in (2a) is fully grammatical and not even a Subjacency violation, which is normally the case with argument extraction from *wh*-islands (cf. *??how many books do you wonder whether John read*). Assuming the Lasnik and Saito (1984), (1992) theory of locality of movement, according to which argument traces must be checked for locality restrictions ('γ-marked') in overt syntax, and adjunct traces at LF, Bošković and Lasnik (1999) argue that the grammaticality of (8a) is expected if the modifier *beaucoup* (presumably, an adjunct) enters the structure non-cyclically, after the *wh*-movement has taken place.[6] In particular, *beaucoup* may enter the structure <u>after</u> the locality restrictions on the (argument) trace are checked. (Note that *beaucoup* still must enter the structure in overt syntax, given that it has phonological content.) (8b) is ungrammatical because in this case, the locality restrictions on the adjunct trace must be checked at LF, and at that level *beaucoup* has invariably become part of the structure. On similar grounds, Bošković and Lasnik (1999) argue for non-cyclic insertion of certain heads, e.g., certain complementizers.

Another argument is due to Nissenbaum (1998) who offers an analysis of parasitic gap (PG) constructions based on the idea advanced in Heim and Kratzer (1998) that syntactic movement creates derived predicates (λ-abstracts) at LF. Consider the example in (9):

---

[6] Bošković and Lasnik (1999) are concerned with the empirical effects of Chomsky's (1995) definition of 'strong' features, as follows:
(i) 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. (p.233-234)
This definition, they argue, allows the acyclic merger of *beaucoup* in (8) if it does not have any strong features (under the view that there is no independent cyclic principle). See also fn. (21).

(9)   Which book did John [$_{VP}$ [$_{VP}$ file _] [$_{Adj}$ Op without PRO reading _]] ?

According to Nissenbaum, the complex adjunct denotes a two place predicate (type <e, <E, t>>), E being the event type, e the individual type, and t the truth value. If so, its (cyclic) Merger with the VP *[John file which paper]* leads to a compositionally uninterpretable outcome because of type mismatch, since the (segment of the) VP which becomes a sister of the adjunct is taken to be a one place predicate (type <E, t>). Nissenbaum suggests a possible way around this problem. He argues that the *wh*-phrase first moves within the VP leaving an (intermediate) trace when it undergoes *wh*-movement. This movement creates a λ-abstract which transforms the VP into a predicate of the matching type (<e, <E, t>>). Crucially, as Nissenbaum notes, this movement has to take place <u>before</u> the adjunct is Merged with the VP. If the adjunct is Merged prior to the movement, then the latter will create a λ-abstract over the <u>entire</u> VP + adjunct constituent, which gives rise to the type mismatch problem. For further arguments, see Fox (1999), Nissenbaum (2000).

2.1.2   *Adjuncts* must *be merged late*

Ochi (1999a) explores the PF merger analysis of English verbal morphology of the type in Bobaljik (1995a) and Lasnik (1995b) (cf. also Halle and Marantz (1993)). According to these authors, Infl in English is affixal, hence must merge with a V, a PF process that requires adjacency. This PF merger can proceed in (10a), but not in (10b) where the

adjacency is disrupted by the negative head *not* (the working assumption here is that 'do-support' applies whenever PF merger fails):

(10)   a.   John Infl [vp leave]   (cf. John left)

        b.   John Infl *not* [vp leave]   (cf. John did not leave)

Bobaljik (1995a) observes that adverbs (presumably, adjuncts) apparently do not interfere with the adjacency requirement:

(11)   John Infl *quickly* leave   (cf. John quickly left)

Ochi suggests that this state of affairs arises because the PF merger of Infl and V can take place <u>prior</u> to Merging the adverb, *quickly* in (11). He adopts the Multiple Spell-Out model of Uriagereka (1999) and others to implement this idea. The derivation of (11), according to Ochi, involves the following steps (assuming the VP-internal subject hypothesis):

(12) a.   Create: [vp John leave]

     b.   Merge Infl (followed by subject movement):

           John Infl [vp John leave]   → *Spell-Out and PF merger of* Infl *and* leave

     c.   Merge *quickly*:

           John quickly left   → *Spell-Out*

The next question is, what blocks the derivation of (13):

(13)   *John did quickly leave

Here the adjunct must have been merged cyclically and the PF merger fails; as a result, *do*-support applies (cf. (10)). It is not clear, then, why (13) is ungrammatical. As Ochi points out, however, the explanation is readily available if the option of cyclic insertion is effectively excluded for adjuncts: they must be merged non-cyclically.

The next piece of evidence involves certain facts concerning multiple *wh*-fronting in Slavic. Bošković (1997a) discusses the somewhat unusual behavior of multiple *wh*-questions in Serbo-Croatian with respect to Superiority (Chomsky (1973)). In particular, he notes that while in short distance questions the order of *wh*-phrases is free, in embedded contexts it is fixed as predicted by Superiority, as shown in the following examples:

(14)   a. Ko   je koga vidio?

Who is whom seen

'Who saw whom?'

b. Koga je ko   vidio?

Whom is who seen

(15)   a. Jovan i   Marko ne znaju   ko   je koga istukao

Jovan and Marko not know   who is whom beaten

'Jovan and Marko do not know who beat whom'

b. *Jovan i   Marko ne znaju   koga   je ko   istukao

Jovan and Marko not know   whom   is who beaten

Under Chomsky (1995c) notion of Attract,[7] Superiority follows from economy considerations, given that features must be checked in the most economical way (cf. also Oka (1993)). In particular, Attract should pick the highest *wh*-phrase in the structure for the purposes of *wh*-movement (cf. the Minimal Link Condition). Details aside, Boškovic argues that in (14) the matrix C with a strong Q feature is not projected overtly. Consequently, no Attraction by C takes place, that is, there is no *wh*-movement in the sense of moving to Spec-CP. Hence, Superiority is irrelevant. Furthermore, *ko* and *koga* adjoin to a functional projection below C, call it FP, in an instance of what Boškovic terms focus movement. According to Bošković (1999), the category F hosting the *wh*-phrases in Serbo-Croatian is of Attract-All type, that is, it is satisfied by Merging all available *wh*-phrases. F can Attract the *wh*-phrases in any order, since, regardless of the order of movement of *wh*-phrases from their base-generated position (e.g. from within the VP), the same number of (full) nodes is crossed to satisfy the Attract-All property of F. This accounts for the lack of Superiority with focus movement.

In (15), on the other hand, the embedded C must be projected before it is embedded in a larger structure. Hence, under Attract, *wh*-movement must take place. Here, movement of only one of the *wh*-phrases suffices to check the feature of the interrogative

---

[7] "K *Attracts* F if F is the closest feature that can enter into a checking relation with a sublabel of K".

C (i.e. we are dealing with an Attract 1F head, which induces Superiority effects). According to the Minimal Link Condition, this has to be *ko*, which is generated the highest. The other *wh*-phrase *koga* again moves to adjoin to FP. This situation is shown in (16):

$$\overbrace{\quad\textit{focus-movement}\quad}$$

(16)   znaju [$_{CP[+Q]}$ ko [$_{FP}$ koga [$_{FP}$ ...   ... t   t]]]
         $\underbrace{\wedge\quad\textit{wh-movement}\quad}$
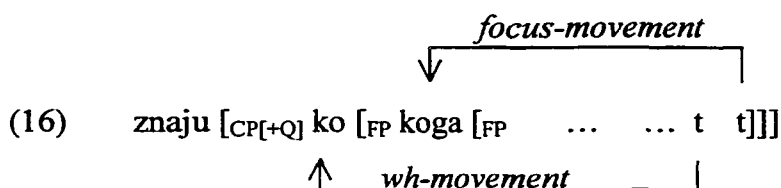
Notice now that the timing of *wh*-movement and focus-movement in (15) is crucial in order to derive the well-formed (15a). Considerations of cyclicity would suggest that focus movement takes place before *wh*-movement. But if so, then both *ko* and *koga* will move to adjoin to FP. Since Superiority is irrelevant for the focus movement, either of these two *wh*-phrases can be the highest *wh*-phrase in the structure, or the closest one to the interrogative C. If *koga* happens to be the highest, C will Attract it for the purposes of *wh*-movement resulting in the ungrammatical (15b), a Superiority violation. On the other hand, if *wh*-movement applies first, this ensures that the highest *wh*-phrase (*ko* in (15)) gets attracted, which accounts for Superiority in (15a). This entails that focus movement by adjunction to FP must take place non-cyclically. The ungrammaticality of (15b) thus suggests that the option of cyclic adjunction must effectively be excluded, leaving non-cyclic adjunction as the only possibility.[8]

---

[8] Boškovic's analysis is developed so as to force acyclic movement of adjuncts in the Superiority contexts, but on somewhat stipulative grounds. See Boškov007 (1997a) for details.

Finally, consider the following constructions involving the reciprocal *each other* ((18b) is from Johnson (1987), attributed to R. Kayne):

(17)a. ?*What evidence that each other$_i$'s friends brought up at court did the lawyers$_i$ refuse to talk about?

cf. b. The lawyers$_i$ refused to talk about the evidence that each other$_i$'s friends brought up at court

(18) a. ?*Which proofs that demonstrated that pictures of each other$_i$ had been forged did they$_i$ read?

cf. b. They$_i$ read proofs that demonstrated that pictures of each other$_i$ had been forged

In (17) and (18), *each other* is inside a restrictive relative clause modifying (*what*) *evidence* and (*which*) *proofs*, respectively. We take the relative clause to be a structural adjunct. The reciprocal in question is subject to some version of Condition A of the binding theory, which we take, for the present purposes, to state that *each other* must have an antecedent c-commanding it within a certain domain at some point in the derivation (cf. Belletti and Rizzi (1988), Lebeaux (1991), Chomsky (1995c)). In the good b. sentences, *the lawyers* and *they* c-command the reciprocal. The a. sentences, however, are ungrammatical, which is expected under the late adjunction hypothesis. Note that if the adjunct must Merge with the *wh*-phrase postcyclically, that is, after *wh*-movement has

taken place, there is never a point in the derivation where the antecedent c-commands *each other*. This leads to a Condition A violation in (17a) and (18a).[9]

2.2 Formal proposal

2.2.1 'Least Tampering'

As a starting point, we adopt the minimalist bare phrase structure system (Chomsky (1995a), (1995c)b, (2000)). This system involves the basic structure building operation Merge, which combines syntactic objects taken from the lexicon as well as those already formed in the course of the derivation, forming legitimate syntactic objects of the set-theoretic form $\{\gamma, \{\alpha, \beta\}\}$, γ a predictable label.

We follow the idea expressed in various forms in Chomsky (1995c), (2000), Watanabe (1995) that operations of the computational system - in particular, Merge - tend to preserve existing structure, rather than 'tamper' with it, where 'tampering' means making certain changes in the structural make-up of the phrase marker. Chomsky (2000) further argues that what should be preserved is a set of basic relations in a given syntactic object (phrase marker), in particular, c-command.[10] We adopt here the definition of c-command that goes back to Reinhart (1976):

(19) α *c-commands* β iff neither α nor β dominates the other and the first branching node that dominates α dominates β.

---

[9] See also Ishii (1998).

[10] Chomsky also includes sisterhood in this set. For us, it does not matter if sisterhood is included as a separate relation. On the other hand, if c-command is defined in terms of sisterhood and containment, as Chomsky proposes in the same work, then our present definitions can be correspondingly reformulated.

The definition of dominance (20) encodes the segment/category distinction, along the lines of May (1985), Chomsky (1986a), (1995a), (1995c).

(20) α *is dominated* by β only if it is dominated by every segment of β.

Thus the notion "the first branching node that dominates α" in (19) is restricted to categories. Dominance is taken to be irreflexive. Adopting the segment/category distinction allows us to distinguish adjuncts and non-adjuncts (specifiers, complements). In Section 2.2.4 we formalize this distinction.

Chomsky (2000) defines the 'least tampering' property of the computation only in connection to basic relations involving the label that projects (cf. his (59), p. 137). Suppose the property is in fact more general and holds of any syntactic object (phrase marker) in the process of structure building. Taking c-command to be the relevant relation, let us assume the following version of this property. Given a syntactic object labeled X, structure building operations applying to it should neither subtract, nor add new c-command relations inside this object.[11] More formally, define 'Least Tampering' as follows:

---

[11] An idea similar to a version of 'Least Tampering' is pursued in Frank and Vijay-Shanker (1998) who argue that operations allow changes in the set of c-command relations as long as new relations are added, but nothing is subtracted (their 'Derivational Monotonicity' requirement). For them, this also holds for operations applying inside an object. For instance, suppose that D adjoined counter-cyclically to C in (i), resulting in (ii):

(i)　　　　A　　　　　　(ii)　　　　A
　　　　／＼　　　　　　　　　　／＼
　　　B　　C　　→　　　　B　　C
　　　　　　　　　　　　　　　　／＼
　　　　　　　　　　　　　　　D　　C
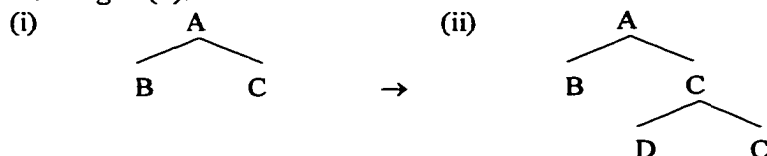
(21) *Least Tampering (modified from Chomsky's (59))*

Given a choice of operations applying to a syntactic object labeled α, select one that does not change @(α).

@(X) – a set of c-command relations in a syntactic object labeled X.

(21) can be naturally conceived of as an Economy condition, in the sense that it chooses between (continuations of) derivations that do and do not change the c-command relations inside an existing phrase marker, in favor of those that do not. As we show below, it is this condition that, in conjunction with the structure building procedure, effectively forces adjuncts to be Merged after the cyclic portion of the derivation.

2.2.2 Merger by Substitution

Chomsky (1995c) assumes that the operation Merge applies at the 'root' of tree/phrase marker. As he points out,

(22)   We assumed ... that Merge applies at the root only. In the bare phrase structure system, it is easy to see why this is expected. Suppose that the derivation has reached stage Σ, with objects α and β. Then Merger may eliminate α and β in favor of the new object K $= \{\gamma, \{\alpha, \beta\}\}$, with label γ. That is the simplest kind of merger. (p. 248)

Then the set of c-command relations inside the object labeled A increases (it now includes those involving D), but no existing relations are subtracted. In contrast, we posit a stronger condition that disfavors introduction of new relations inside the object (in particular, by introducing new elements into it), in addition to requiring the existing ones to be kept.

The idea behind Chomsky's conceptual argument is that it is in some sense more difficult for Merge to 'look inside' a phrase marker for its target, rather than simply target the root. Chomsky (1995c); (2000) suggests that structure building operations only 'see' the label of their target (see also Section 2.2.4). Thus, in terms of (22), Merge would not be able to 'look inside' β, but only at the label of β itself. But this is not entirely obvious. Indeed, what can in principle prevent an operation from looking inside the (complex) phrase marker? In addition, the label visibility condition, if inviolable, is so strong that it effectively excludes all instances of acyclic Merger, including adjunction, even as an option. The facts in Section 2.1, then, would have to be explained in some other way.

The 'Least Tampering' property in general, and its specific formulation in (21), offers a way to restate the idea concerning Merger at the root in a more principled manner. First, let us define 'root' as follows:

(23)   A *root* is a category c-commanded by no other category

(23) is based on the definition of c-command as in (19) (cf. also 'attachment root' of Frank et al. (1999)). Suppose now we apply 'substitution' Merge to build a syntactic object $L=\{\beta,\{\alpha, K\}\}$ (the label β is either α or K depending on which one projects), where $K=\{X, \{X, Y\}\}$, as shown in (24a). The set of c-command relations in L is shown partially in (24b):

(24) a.   L =   β
              /\
            α   X^max
                /\
               X   Y
                   /\
                  ... ...

b. @ (β) = {<α, X^max>, <X^max, α>, <α, X>, <α, Y>, <X, Y>, <Y, X>, ...}
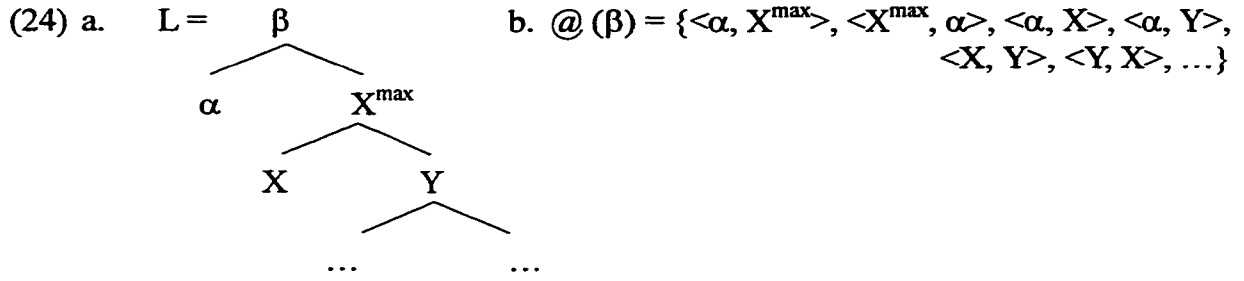
By (23), the category labeled β in (24) is a root, since it is c-commanded by no other category. Suppose the next step in the derivation is Merging γ with L by substitution. Merger may apply either at the root (that is, with β), or not, for instance to X^max. If Merger applies at the root, a new object M={δ, {γ, L}} is created, as shown in (25a). The set of c-command relations in M is shown partially in (25b).

(25) a.   M =   δ
              /\
            γ   β
                /\
               α   X^max
                   /\
                  X   ...

b. @ (δ) = {<γ, β>, <β, γ>, <γ, α>, <α, X^max>, <X^max, α>, <α, X>, ...}

c. @ (δ) = {<γ, β>, <β, γ>, ...} ∪ @ (β)
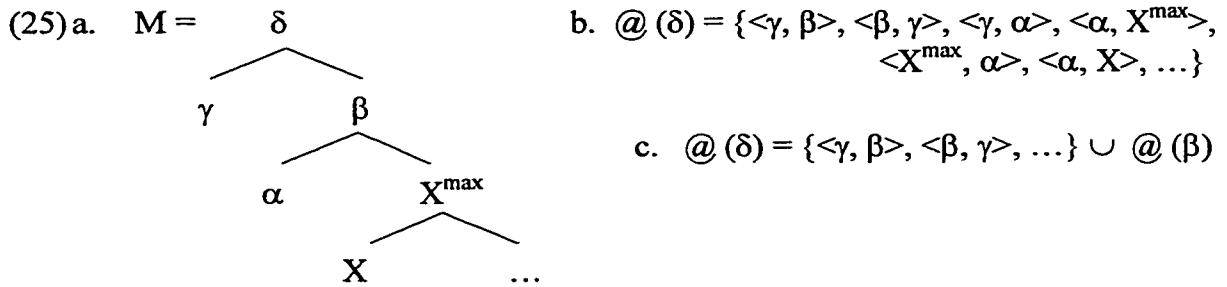
Note, importantly, that upon the Merger of γ with L, the set of c-command relations inside L does not change: no existing relations are subtracted, no new relations are added. This Merger simply established additional relations in the new object M, as (25c) shows (cf. Epstein et al. (1998), Epstein (1999)). Suppose now, instead, that Merge applies to γ and X^max inside L, as shown in (26a) (for ease of exposition, take γ to be a head). The set of c-command relations now is shown in (26b):

(26) a.
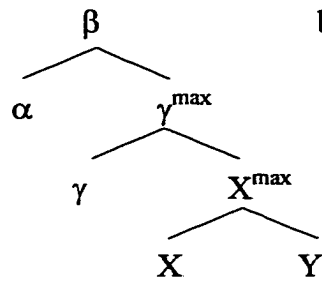
b. @ (β) = {<α, γ^{max}>, <γ^{max}, α>, <α, γ>, <α, X^{max}>, <γ, X^{max}>, <γ, X>, ...}

The object in (26a) is still labeled β, but the set of c-command relations in it has changed, as can be seen from comparison of (26b) with (24b). In particular, new c-command relations - those involving γ - have been added in the set. Thus, given (21), the derivation reaching the stage in (24) can (locally) compare the option of Merging something at the root, or not at the root, and determine that the option of Merging at the root leads to preservation of the existing set of c-command relations.[12]

The 'Least Tampering' condition in (21) can thus be formulated in the following manner:

(27)   Merge at the root when possible (where 'root' is defined as in (23)).

2.2.3 Merger by Adjunction

Consider now how the proposed structure building algorithm works in case of adjunction Merge.

---

[12] Again, similar argumentation can be constructed if sisterhood is included in the set of basic relations inside a phrase marker; see fn. 10. It should be stressed that comparison between the option of Merging at the root or Merging not at the root does not have to involve one and the same γ: different objects may be considered for Merger.

Suppose γ Merges with L in (24) by adjunction. If γ is Merged somewhere inside L, this leads to a change in @(β), the set of c-command relations inside L (still labeled by β): new relations involving γ are now included, similarly to the substitution case in (26). If, however, γ is Merged at the root (β), this creates a segmented object, as in (28).

(28)

$$
\begin{array}{c}
\beta \\
\diagup\diagdown \\
\gamma \qquad \beta \\
\diagup\diagdown \\
\alpha \qquad X^{max}
\end{array}
$$
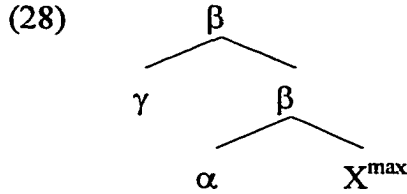
Note that this Merger of γ does not result in a change in @(β). This is so because the "first branching node..." clause in the definition of c-command (19) does not apply (there is no such node), consequently, c-command between γ and the elements of L is undefined. (27) then forces γ to be Merged with β (at the root), as a more economical choice.

Furthermore, according to the definition in (23) (and (19), (20)) in (28) there are <u>two roots</u>, namely, γ and the segmented category labeled <β,β> (recall, β is α or X, depending on which one projects). This is so again because the "first branching node..." clause in the definition of c-command (19) does not apply, and c-command between γ and <β,β> is undefined. It follows that neither c-commands the other. In other words, @(<β,β>), the set of c-command relations inside (28), does not include pairs involving γ (see also Frank et al. (1999)):

(29) $@(<\beta, \beta>) = \{<\alpha, X^{max}>, <X^{max}, \alpha>...\}$

Since γ and <β, β> are both roots the null hypothesis appears to be that further Merger is potentially possible with either. Note, however, that (cyclic) Merger of some ρ to γ results in a 'structure of the following type, in which γ is dominated by more than one node:

(30)
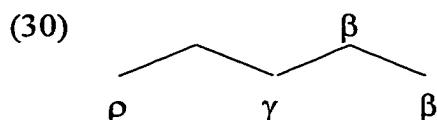
We assume, on fairly plausible grounds, that (30) is not a legitimate syntactic object. There are several ways to exclude (30), proposed in the literature. One possibility might be to rule it out as a violation of dominance relations, along the lines of Lasnik and Kupin (1977), whose formalization of phrase marker excludes non single-rooted trees ('forests'). In particular, their definition of phrase marker involves a requirement that there be a (non-terminal) element that dominates the entire (terminal) string. There is no such an element in (30), hence (30) is not a legitimate phrase marker. A similar result can be achieved by the Single Root Condition of Partee et al. (1993).[13] See also Collins (1997) who rules out (30) by appealing to the LCA (Kayne (1994)).

Consider what happens if ρ is Merged with the other root, <β, β>. If this Merger is adjunction, this creates a multi-segmented structure:

---

[13] We do not discuss here the question of translating these alternatives into the bare phrase structure theory. See Frank and Vijay-Shanker (1998) for relevant discussion.

A *Syntax* reviewer notes, correctly, that the Single Root Condition must hold at all stages of the derivation, in order to preclude unwanted derivations involving certain types of multi-domination structures (see, in particular, Bobaljik (1995b) and Gärtner (1999) and references therein).

(31)

$$
\begin{array}{c}
\beta \\
\diagdown \\
\rho \quad \beta \\
\diagdown \\
\gamma \quad \beta \\
\diagdown \\
\alpha \quad X^{max}
\end{array}
$$
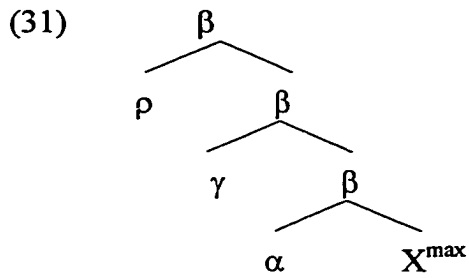
Here the set of c-command relations in the existing structure is not changed; this Merger is therefore legitimate.

If the Merger is substitution, a new category is created, resulting in the following structure:

(32)

$$
\begin{array}{c}
\rho^{max} \\
\diagdown \\
\rho \quad \beta \\
\diagdown \\
\gamma \quad \beta \\
\diagdown \\
\alpha \quad X^{max}
\end{array}
$$
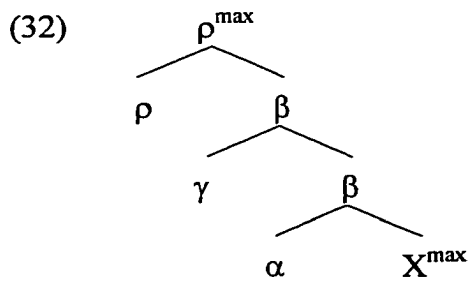
But now note that this Merger results in the change in the set of c-command relations inside the existing structure, labeled $<\beta, \beta>$. Specifically, now γ and $<\beta, \beta>$ c-command each other, since the "first branching node..." clause of the definition (19) applies. Thus the pairs involving γ are added in the set:

(33) $@(<\beta, \beta>) = \{<\gamma, <\beta, \beta>>, <<\beta, \beta>, \gamma>, <\gamma,\alpha>, <\gamma, X^{max}>, <\alpha, X^{max}>,...\}$ (cf. (29))

This Merger is excluded in favor of a derivation when ρ is Merged directly with β when γ is not yet in the structure.[14]

Thus, the object in (28) can only represent the last cyclic Merger in the derivation.[15] If at this point the derivation contains no unused items (either lexical, in the Numeration, or assembled in the derivational 'workspace') to be Merged, the overt (structure-building) part of the derivation ends here. Specifically, (28) may illustrate an adjunction of some XP to matrix CP (e.g. a topic). The next question that arises is what happens if there are more syntactic objects available for (substitution) Merger. Let us now turn to a concrete example for clarity.

Consider the sentence in (34a), the Numeration for which is given in (34b):

(34)   a. Adeola fixed the car with a hammer

        b. {Adeola, fixed, the, car, with, a hammer, v, T, C}

---

[14] The derivation thus may involve a two-step look-ahead. We then depart from the radically local view on derivational economy (cf. Chomsky (2000), Collins (1997)). It is possible to restate the proposal in the text as to not involve look-ahead at all. Suppose instead of deriving (27) from 'least tampering' considerations, we take it as a primitive condition in the grammar, and, furthermore, modify the definition of c-command in (19) as follows:

(i)   α c-commands β iff α does not dominate β and every category that dominates α dominates β.
                                                                            (cf. Chomsky (1986a))

For the case of substitution Merge, the results follow as before. For the case of adjunction, suppose the derivation reached the stage (24), and γ is Merged, creating (28). Under (23) and (i), this structure involves no root at all (both γ and <β, β> now c-command each other). Hence the next cyclic Merger cannot apply, and (32) is never derived.

We thus face a dilemma: to allow a slight amount of look-ahead or stipulate (27) as a primitive. We adopt the former as a more principled choice. In addition, it is not clear if the two-step look-ahead of the kind suggested above raises any problem at all, as the issue of just how much look-ahead is allowed in the computation is still under debate (see, e.g., Chomsky (2000), Nakamura (1998)).

[15] This also extends to adjunction by movement. Cyclic adjunction is thus excluded, including successive cyclic adjunction. See, in particular, Lasnik and Saito (1992) who argue against certain cases of successive cyclic IP adjunction.

Assume that the adverbial *with a hammer* is a true vP adjunct (see Section 2.2.4 for more discussion). Take the point where the vP is constructed. Suppose that *with a hammer* is Merged cyclically to vP, forming the segmented constituent, as shown below:

(35)   [vP [vP Adeola fixed the car] [with a hammer]]

At this point the Numeration contains the unused items T and C. But neither of these items can be merged cyclically with the vP, since that would lead to a change in the set of c-command relations inside the vP (cf. (32)). This derivation is thus canceled, since there is in fact a more 'economical' alternative (see below). Now notice that (27a) gives an option for counter-cyclic Merger. In cases like (34a), however, this Merger is impossible for independent reasons, such as selection. For example, one cannot Merge T inside vP simply because the selectional requirements of T would thereby be violated. Similar reasoning applies with regard to insertion of C. This derivation is then canceled as well, because of the non-exhausted Numeration (cf. (Chomsky (1995c), Ch. 4).

Consider now another continuation of the derivation of (34a). In particular, when the [vP *Adeola fixed the car*] is completed, the next step is a ('substitution') Merger of T (triggering the displacement of *Adeola* to its Spec, as standardly assumed), thereby creating a T'/TP. This Merger does not lead to a change of c-command relations inside the vP (and is therefore preferred to Merging of *with a hammer* instead at this point). At this point, two objects remain to be Merged: the PP *with a hammer* and the complementizer C. If the PP is now Merged with the vP (non-cyclically), this changes the set of c-

command relations inside the TP (in particular, new relations are added). On the other hand, if C is Merged cyclically (at the root), the set of c-command relations inside the TP is not changed. Merger of C is thus a more economical choice.

After C is Merged with the TP, *with a hammer* is the only remaining object in the numeration. It can then Merge postcyclically. This Merger is allowed insofar as it ensures convergence, by (27).

In cases when several modifiers modify different categories, each of the adjuncts Merges postcyclically to their corresponding category, as in (36):

(36)   John said *after class* that he couldn't come *because he was sick*

Similarly, in cases when an adjunct is inserted inside an adjunct, each of those Merged postcyclically to the corresponding host:

(37)   John left *because he couldn't stay after class*

We have demonstrated that a cyclic Merger of an adjunct prevents further cyclic structure building. If at this point there remain more elements in the derivational 'workspace' (including the Numeration) that cannot be Merged acyclically, the derivation is canceled. The present theory thus forces all non-adjuncts to be Merged cyclically, and adjuncts postcyclically. The notions 'cyclic' and 'created by substitution' are equivalent, just as the terms 'postcyclic' and 'created by adjunction' are.

2.2.4 The dichotomy of Merge

In order to assess the entire empirical coverage of the proposed structure building algorithm it is necessary to identify the class of syntactic objects that are Merged by adjunction. The question can be stated more generally: what regulates the choice of substitution Merge and adjunction Merge in each case? The topic continues to be under debate in the literature (cf. e.g. Chomsky (1995c), Chametzky (1996), (2000)). We believe that, in the system of Chomsky (2000), an answer may be formulated in terms of featural makeup of objects that undergo Merge. Chomsky (1995c), (2000) distinguishes two fundamentally different types of features: features that are interpretable at the PF and LF interfaces, and features that are uninterpretable at the interfaces, but are satisfied in the course of syntactic computation. Chomsky (2000) assumes that uninterpretable features make the phrase they reside in an active candidate for movement.

Note that there are two dichotomies in Chomsky's system: the uninterpretable/interpretable dichotomy, and the substitution/adjunction dichotomy. Driven by the minimalist spirit, one can thus take a reductionist view and establish a systematic correlation between the two dichotomies. Let us make this step. We suggest that in addition to being instrumental in triggering movement, uninterpretable features also have a different kind of property: they trigger projection of a full category. We then define adjunction as follows:

(38) A non-projecting syntactic object α is Merged with a syntactic object β by *adjunction* iff the label of α contains no active ('unchecked') uninterpretable feature(s).

Note that (38) allows uninterpretable features to be present in the object labeled α, as long as they are not part of the label, that is, as long as they are 'buried' inside a complement of the head α.[16]

The definition of substitution follows as a mirror image of (38):[17]

(39) A non-projecting syntactic object α is Merged with a syntactic object β by *substitution* iff the label of α contains active ('unchecked') uninterpretable feature(s).

It should be kept in mind that the definition in (39) does not necessarily imply checking uninterpretable feature(s) immediately upon Merger. Rather, both (38) and (39) imply a more general claim that projecting a separate category is a property of uninterpretable features, regardless of when they are actually checked.

Given (38) and (39), the late adjunction algorithm forces objects that contain uninterpretable features in their label to be Merged cyclically, and those that do not contain such features in their label - postcyclically.[18] This holds for pure Merger (base-

---

[16] This has relevance for 'adjunct condition' cases as in (1), repeated here:
(i) ?*What did John [$_{VP}$ [$_{VP}$ go to bed] [after Peter fixed t$_i$]]?
In the *after*-clause, the [+*wh*] feature of *what* (uninterpretable in the system of Chomsky (2000)), is 'buried' sufficiently deep inside the clause as to not figure in the label, and assuming that *after* has no other uninterpretable features, the *after*-clause is Merged by adjunction, not substitution.

[17] The 'non-projecting' restriction is needed to exclude indeterminate situations of the following sort: if α has uninterpretable feature(s), and β does not, then by (39) α is Merged with β by substitution, becoming a specifier of β (β projects), but, at the same time, by (38), β adjoins to α, becoming an adjunct to it (α projects). The 'non-projecting' restriction resolves this indeterminacy, by appealing to the basic assumption that only one of the two elements combined by (the binary) Merge actually projects, whereas the other one does not. As a result, both participants in the Merger are never considered simultaneously with regard to determination of their phrase structure status; rather, only the one that does not project, is so considered.

[18] The idea that only elements containing uninterpretable features must Merge cyclically, and others postcyclically, is also pursued in Ishii (1997) and (1998). Ishii (1997) implements this idea by making it follow from two stipulated conditions (the 'Immediate Checking Principle' and 'Earliness Principle on

generation) as well as Merger as part of movement. For the latter, we are thus driven to the view that movement resulting in adjunction must be triggered by some property of the target of movement only (e.g. 'EPP', or 'strength'). In contrast, movement resulting in substitution must involve an uninterpretable feature of the element that moves, along with a property of the target of movement (cf. Chomsky (2000)).

One consequence of (38) and (39) under the late adjunction algorithm concerns wh-items *how, why* and the like. These items have traditionally been called 'wh-adjuncts', by analogy with their non wh-versions (cf. *with a hammer* or *because John was sick*). However, in the system in Chomsky (2000) (although not in Chomsky (1995c)) the wh-feature is an uninterpretable feature.[19] It follows, by (39), that these items must enter the structure by substitution, just like argument wh-phrases such as *who* or *what*. Possibly, these items are generated in a specifier position of a functional projection close to VP. Under the late adjunction hypothesis, then, they will be generated cyclically. For instance, in (40) *how* is Merged by substitution, and later raises to the embedded Spec-CP where it checks its uninterpretable wh-feature.

(40)   a. John wonders [cp how Peter fixed the car t]

        b. How did John say that Peter fixed the car?

Select') plus certain additional provisos. Ishii (1998) posits, instead, a single condition stated in terms of selection. In contrast, our goal is to formalize this idea within the standardly adopted notions, without introducing new principles into the grammar. Note that for us, uninterpretability itself does not enforce a particular timing of Merger; rather, independent considerations do; see the discussion above.

[19] This follows given Chomsky's proposal that uninterpretable features make the phrase they reside in an active candidate for movement.

The result conforms with the common understanding of behavior of *how* which is standardly assumed to be subject to cyclic structure building. Note that without (38) and (39), the derivation of (40) under the late adjunction hypothesis is less straightforward. In particular, the late adjunction algorithm forces *how* in (40) to be Merged into the structure postcyclically. It is not clear, then, how postcyclic Merger of *how* can be reconciled with its requirement for *wh*-movement which is a cyclic operation. One can speculate that items like *how* and *why* do not undergo *wh*-movement at all; rather, they are inserted directly into the Spec-CP (cf., e.g. Hegarty (1991), cf. also Law (1993), Murasugi (1992)). However, under this line of analysis it is difficult to account for the fact that in (40b), for instance, *how* can modify the lower clause, although it is in the higher Spec-CP. In addition, it does not explain why in *wh*-in situ languages *why* and *how* are located in-situ. If (38) and (39) are adopted, these problems do not arise.[20] The structural distinction between *wh*- and non-*wh*-versions of circumstantial adverbials is captured by the proposed distinction in terms of uninterpretable features: *how* and *why* have the uninterpretable *wh*-feature, whereas *with a hammer* or *because John was sick*, arguably, have no uninterpretable features in their label.

---

[20] A possible alternative analysis that makes *wh*-items like *how* on a par with regular arguments such as *who* or *what* could be formulated if one adopts the proposals of Hagstrom (1998) and Bošković (1998b). According to these authors the interrogative Q feature of C is generated inside a clause over which interrogation takes place. These authors give evidence (the former for *wh*-in situ, the latter for overt *wh*-movement languages) that the Q feature can be Merged with the *wh*-phrase. This suggests that the Q feature can select the *wh*-phrase. Being subject to selection (regardless of whether selection involves checking uninterpretable features), *wh*-phrases like *how* are then not 'adjuncts', but rather, arguments, hence, by the late adjunction mechanism, are Merged cyclically. This analysis, however, raises a number of issues, in particular: what selects the Q morpheme itself, that is, why is the complex [Q [*how*]] not an adjunct? Note that the approach in the text does not face this difficulty.

3. Further Consequences of the Late Adjunction Hypothesis

A conceptual consequence of our proposal is that it derives Chomsky's (1995), Ch.3., exemption of adjuncts from his Extension requirement. In fact, the Extension requirement itself is also derivable from the 'least tampering' condition (27).[21] Below we discuss several empirical consequences of the late adjunction hypothesis. We begin with the discussion of raising constructions involving experiencers.

3.1 Raising constructions and the Minimal Link Condition

Examples like (41) involving raising verbs present a well known problem for the minimalist theory, pointed out in Chomsky (1995c), Ch. 4):

(41)   John$_i$ seems to Mary [t$_i$ to be smart]

In (41) it is the lower subject that raises to the matrix subject position, even though there is a closer candidate for raising, namely, the experiencer *Mary* which c-commands *John* at the time of raising (we take 'closeness' to be defined in terms of c-command, cf.

---

[21] Note that (38) and (39), in conjunction with the late adjunction hypothesis, derive the generalization (i), discussed in Bošković and Lasnik (1999) as a consequence of Chomsky's (1995b, p. 234) 'strength' conception of the cycle in (ii) (see fn. 6).

(i)   Acyclic Merger of an element α is in principle possible if α contains no strong features in its label.

(ii) 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.

On the assumption that 'strength' has to do with uninterpretability (but see Chomsky (2000)), the late adjunction hypothesis and (38) and (39) formalize the strongest version of (i): not only is acyclic Merger of α possible, it is to be forced, in a postcyclic manner. See also Ochi (1999b) for relevant discussion.

It is worth noting that our approach is a sort of the opposite of the one in Bošković and Lasnik (1999). These authors point out that Chomsky's (ii) overlaps with his Extension condition in capturing cyclicity effects, and propose to abandon the Extension Condition, leaving (ii). In contrast, our approach makes use of the principle behind the Extension Condition ("Merge at the root"), and dispenses with (ii). The full extent of compatibility between the two approaches remains to be determined.

Chomsky (1995c)). That the experiencer c-commands into the lower clause can be seen from Condition C effects:[22]

(42)   They seem to him$_i$ to like John$_{j/*i}$

The grammaticality of (41) is thus problematic, given the Minimal Link Condition or Attract Closest (MLC/AC) (Chomsky (1995c), Ch.4). Specifically, under the MLC/AC the strong EPP feature of matrix T should not be checked by the D-feature of *John* since there is a closer D-feature, namely, that of the experiencer. Furthermore, all else equal, the MLC/AC predicts that (43) should be grammatical:

(43)   *Mary$_i$ seems to t$_i$ [John to be smart]

In (43) Mary raises to the matrix Spec-T, checking the EPP feature of T (perhaps after entering a Case checking configuration with *to*, similarly to 'quirky' Case instances). The formal features of *John* may raise to adjoin to the matrix T at LF, checking Case and, perhaps, φ-features. (Note that we assume here, following Chomsky (1995c), Ch. 4; (2000), that (A-) traces are in some sense 'invisible' to computation, hence do not block raising across them). Thus under the MLC/AC nothing blocks this derivation.

Torrego (1996) argues that the experiencer in raising to subject languages including

---

[22] The fact that the experiencer can bind (cf. (42)) suggests that it is an A-position, which at the first glance does not fit easily with its structural status as adjunct. There is a way of reconciling these two properties of experiencers. One can suppose, reasonably, that the A-position in question is not the entire phrase containing the experiencer (*to him* in (42)) but, rather, just the experiencer (*him*). Being a complement of *to*, the experiencer qualifies as an A-position. See also Kitahara (1997).

English is an adjunct to the lower clause (cf. the bracketed part in (41)). We want to modify her proposal somewhat and suggest that *to Mary* in (41) is an adjunct in the matrix clause, possibly to the lower VP projection in the extended argument structure of *seem*, in which *seem* undergoes short movement to a higher functional head (e.g. little v).[23] Given this proposal, it is now easy to see how the raising facts can be reconciled with the MLC/AC. Under the present theory, if the experiencer is an adjunct, it does not enter the structure until after the raising took place. That is, at the time of raising, there is no closer candidate to be Attracted to the matrix T than the lower subject. The derivation of (41) should then proceed as follows:

(44)   1) Create: T seems [$_{CP/IP}$ John to be smart]]

        2) Attract *John* (closest):   John T seems [$_{CP/IP}$ t$_i$ to be smart]

        3) Insert *to Mary*:   John$_i$ seems [to Mary] [$_{CP/IP}$ t$_i$ to be smart]]

Notice that we do not have to require *to Mary* to always be Merged late in this case: having it as an option would suffice. Thus, apparently, a theory that simply allows adjuncts to be Merged postcyclically (as an option), a weaker version of the present algorithm, is capable of making this result follow. However, any version of such a theory is of no avail with respect to ruling out (43), if the MLC/AC is adopted. As long as the option of cyclic adjunction is available, it remains unclear why (43) is ungrammatical. On

---

[23] Another possibility, suggested to us by Howard Lasnik (p. c.), is that *to Mary* is right adjoined to the (main) vP, or to some kind of 'point of view' projection within the main clause, and the embedded clause undergoes obligatory extraposition. All possibilities derive the correct word order; we leave the ultimate choice between them open.

the other hand, under the present theory (43) will never be generated. The experiencer

will not be Attracted as a (closest) candidate for checking the features of T, for the trivial

reason: at that point the experiencer is not yet part of the structure. This line of reasoning

allows us to maintain the MLC/AC intact.

Consider now the following pair of interrogative sentences:

(45)   a. ?[To who(m)]$_j$ does John$_i$ seem t$_j$ [t$_i$ to be smart]?

         b. *Who(m)$_j$ does John$_i$ seem [to t$_j$] [t$_i$ to be smart]?

(45a), in which the option of pied-piping the PP containing the experiencer is chosen, is

significantly better than (45b), in which the experiencer undergoes *wh*-movement,

stranding the preposition. Under the late adjunction hypothesis, this contrast receives the

following analysis. We adopt the proposal in Chomsky (1973) that in PPs containing a

wh-phrase, in particular, *to whom*, the *wh*-feature may reside either in the label of the *wh*-

phrase, or in the label of the PP. If the former, *wh*-movement targets the *wh*-phrase only,

stranding the preposition. If the latter, then *wh*-movement involves pied-piping of the

entire PP. As Chomsky points out, this accounts, in particular, for the grammaticality of

examples like (46) (without positing optionality of pied-piping):

(46)   a. Who(m) did I give the book to?

         b. To who(m) did I give the book?

Thus in (45a), where the pied-piping option is realized, the *wh*-feature resides in the label

of the PP (cf. (46b)).[24] Recall now that we are assuming, following Chomsky (2000), that

the *wh*-feature belongs in the class of uninterpretable features (see Section 2.2.4). But

under the definition in (39), it follows that *to whom* in (45a) enters the structure by

substitution, rather than adjunction. (In particular, it may be a specifier of the lower VP, or

of some 'point of view' phrase, see fn. 23). Hence, it is predicted to undergo normal *wh*-

movement in a cyclic fashion.[25]

In (45b), stranding of *to* signals that the *wh*-feature resides in the label of the *wh*-

phrase, that is, *who(m)* itself (cf. (46a)). The label of the PP *to whom* now does not

contain the uninterpretable feature(s). Consequently, by (38), *to whom* is an adjunct and

must Merge post-cyclically. As a result, *whom* cannot participate in the (cyclic) *wh*-

movement. (45b) is thus ruled out.[26]

---

[24] Similar considerations apply for 'complex' *wh*-phrases such as that in (i), except that here the preposition stranding option is not available:

(i)   a. [After which class] do you usually have a large break?

[25] A potential question may arise here as to why *to whom* in (45a) does not create an intervention effect, given that it is now a specifier, and c-commands into the lower clause thus qualifying as a closer candidate. We assume that as part of *wh*-movement, *to whom* undergoes a short movement to the 'edge' of vP first (cf. Chomsky (2000), also Bošković (1997b)), at the stage of the derivation depicted in (i):

(i)   T [$_{vP}$ [to whom]$_j$ [seem t$_j$] [John to be smart]]

It is sometimes observed that once the PP has moved, the DP inside it can no longer be a c-commander, as demonstrated in (ii):

(ii)   a. Which men$_i$ did you talk to t_{DP} about each other$_i$?

        b. *?To which men$_i$ did you talk t_{PP} about each other$_i$?

Thus in (i) the D-feature of *whom* does not c-command *John*, hence is not a closer candidate. Still within the realm of issues related to the MLC/AC, one may wonder why (iii) is not allowed, if *to whom* is a specifier:

(iii)   *[To who(m)]$_j$ t$_j$ seems t$_j$ [John to be smart]?

In (iii) *to whom* raises to the matrix Spec-IP first, and then on to Spec-CP by *wh*-movement. To rule out (iii), it suffices to maintain, on fairly natural grounds, that PPs are not appropriate checkers of the EPP feature of Infl, at least in English. See Bresnan (1994) for arguments that PPs that seem to be able to check the EPP feature of Infl in so called Locative Inversion constructions (cf. *Under the bed is a good place to hide*) are actually NPs with an elided N head.

[26] The same analysis applies to the contrast in (i), which mirrors the contrast in (45):

(i)   a. ?*Who did it seem to *t* that John is smart?

        b. To who(m) did it seem *t* that John is smart?

Our treatment of experiencers leads to certain cross-linguistic consequences. Some languages, in fact, do allow sentences of the type in (43) (modulo preposition stranding). Icelandic is a case at hand. In this language, as in English, the ('quirky') Dative experiencer is optional in raising constructions (cf. (47a) vs. (47b)). The Case of the experiencer, presumably, is inherent, and comes from the verb (via theta-marking). In addition, it has been suggested that the experiencer also has a structural Case feature on top of the inherent Case (cf. Chomsky (1995c), Frampton and Gutmann (1999)) Thus the experiencer can check the (structural) Case of matrix T, whereas the subject of the lower clause stays in situ, as shown in (47a) (examples from Boeckx (1998):

(47) a. þeim   hafði   virst   Ólafur   veragáfaður

     them-dat   has   seemed Olaf-nom   be intelligent

     'They regard Olaf as intelligent'

  b. cf. Ólafur   hafði   virst   veragáfaður

     Olaf-nom   has   seemed be intelligent

     'Olaf seemed to be intelligent'

In the present terms, these results are expected if a) the MLC is observed, b) the experiencer in Icelandic is not an adjunct, in contrast to English. That is, the experiencer is Merged cyclically in Icelandic. We thus expect, all else equal, that in Icelandic raising across the experiencer should be blocked. This is indeed what we find:

(48)   *Ólafur hafði virst þeim   [t vera gáfaður]

        Olaf-nom has   seemed them-dat. be   intelligent

        'Olaf seemed to them to be intelligent'   ·

What is the relevant distinction between Icelandic and English that makes experiencers adjuncts in the latter, but not in the former? Our approach to adjunction/substitution in terms of (38)/(39), repeated here, may provide an answer:

(38)   A non-projecting syntactic object α is Merged with a syntactic object β by *adjunction* iff the label of α contains no active ('unchecked') uninterpretable feature(s).

(39)   A non-projecting syntactic object α is Merged with a syntactic object β by *substitution* iff the label of α contains active ('unchecked') uninterpretable feature(s).

If the experiencer in Icelandic is not an adjunct, that is, enters the structure by substitution Merge, then (39) implies that, unlike in English, it has an uninterpretable feature in the label. In fact, there is a good candidate for such a feature, namely, Case. Note that in English, but not in Icelandic, the experiencer is a prepositional phrase (PP). It is well known that prepositions (in English, at least) are Case-checkers. It is then conceivable that in English, the Case of the experiencer is checked internal to the PP, so that its label contains no unchecked uninterpretable features (unless the PP is [+*wh*], see above). In

contrast, in Icelandic, the experiencer is a DP which needs to check structural Case (on top of the inherent Case, see above). Hence, it has to check its (structural) Case by something else in the structure, specifically, by matrix T, as in (47) (see also Sigurðsson (1992), Freidin and Sprouse (1991), Chomsky (2000), among others).

If the present analysis is on the right track, we expect that languages that allow subject raising across the experiencer will not allow raising of the experiencer, and vice versa. In other words, all else equal, languages display either the English pattern or Icelandic pattern, depending on the morpho-syntactic makeup of the experiencer. Various cross-linguistic studies of the raising constructions confirm that this state of affairs obtains (see, in particular, Torrego (1996) and Boeckx (1998); the latter including discussion of apparent counterexamples).

3.2 Approximative Inversion in Russian[27]

The next consequence we consider here with regard to postcyclic Merger of adjuncts concerns adjunction by movement, and involves the phenomenon of approximative inversion in Russian. In Russian numeral expressions, nouns typically follow the numeral as in English, as shown in (49).

(49)   sem'   dnej

         seven   days-gen.

---

[27] The argument in this section is largely drawn from Bošković (UConn class lectures, 2000).

Franks (1994), (1995) argues extensively that numeral phrases in Russian (cf. (49)) instantiate either a 'bare' Quantifier Phrase (QP) or a quantified DP. When the DP option is taken, the numeral phrase appears in usual DP positions. In particular, it can occupy Spec-IP, triggering agreement with the predicate and being capable of binding a reflexive or control, thus exhibiting typical subject properties. Under the QP option, none of this behavior is possible, suggesting that the QP stays in its base generated position within the predicate phrase (in conformity with the internal subject hypothesis). Thus, in (50a), the optional agreement pattern suggests that *pjat' zenšcin* can be either a DP in Spec-IP or a VP internal QP, whereas in (50b) the inability to bind a reflexive in the absence of agreement suggests that only the DP option is available (examples from Franks (1995), p. 121):

(50)   a.   Pjat' zenšcin smotreli/smotrelo na Ivana

                Five women   looked-pl./neut.   at Ivan

        b.   Pjat' zenšcin smotreli/*smotrelo na sebja

                Five women   looked-pl./neut.   at self

In Russian numeral expressions, the noun may optionally precede the numeral, which results in approximative reading, as in (51):

(51)   My proveli v Milane nedeli dve

        We spent in Milan weeks two

        'We spent about two weeks in Milan'

Franks shows that phrases involving approximative inversion are not DPs, but QPs, since they do not display typical subject properties: agreement, binding a reflexive or reciprocal, or control, as (52) demonstrates:

(52)   a. Zenŝcin pjat' smotrelo/??smotreli na Ivana

        Women five looked-neut/pl.   at Ivan

    b. Zenŝcin pjat' ?*smotrelo/?*smotreli na sebja

        Women five looked-neut/pl   at self

    c. Zenŝcin pjat' *staralos'/*staralis' vojti   v   dom

        Women five tried-neut/pl   to-enter into house

According to Franks, approximative inversion in (51)/(52) involves adjunction of *nedeli/zenŝcin* to the QP (or, alternatively, to some category inside the QP). Furthermore, Franks proposes that approximative inversion cannot take place in DPs because adjunction to DPs is impossible. Arguably, this follows from Chomsky's (1986) ban on adjunction to arguments. (52b, c) and the agreement version of (52a), all involving the DP, are then ruled out. But in the present form the analysis faces a difficulty given that the numeral DP itself contains a QP (Franks assumes that the numeral itself is located in the specifier of QP, or, alternatively, in the head Q). Given that Franks relies on adjunction to QP in standard cases, the question then arises as to why *nedeli/zenŝcin* in (51)/(52) cannot adjoin to this QP inside the DP. Franks' analysis incorrectly allows for such a possibility.

The late adjunction hypothesis, combined with the chain uniformity theory explored in Chapter 2, offers an alternative way of accounting for the approximative inversion data, while maintaining Franks' structural assumptions. Note that under the internal subject hypothesis movement of a DP from inside the VP (or vP) to Spec-IP creates a non-trivial chain (DP, t_{DP}), consisting of two copies of the DP. According to the theory of Chapter 2, both copies must be identical in order for the chain to be successfully linearized at PF. The formal identity requirement can be naturally extended as to exclude any structural modifications <u>within</u> the DP (e.g. adjunction to a category contained in the DP): any such alteration will make the head of the chain formally different from other member(s), and chain uniformity will be violated.[28] Thus, a quantified DP in (50b) moves from Spec-vP to Spec-IP. Consequently, this DP forms a non-trivial chain, and adjunction to the head of the chain - the DP itself or within the DP - is not allowed. Hence the impossibility of approximative inversion in (52b,c) and the agreement version of (52a).

The prohibition on modification does not extend to chains consisting of a single member - trivial chains. This is so because a modification of (a head of) a trivial chain creates an object which is identical to itself, as it is the only link in the chain. Thus, if a DP/QP does not move, it is a trivial chain, and its modification is allowed.

Recall that a bare QP in Franks' system remains in its base-generated position within the VP (or vP). Thus, the QP forms a trivial chain, and adjunction to it is allowed. This accounts for the standard cases of approximative inversion in (51) and the non-agreement version of (52a).

---

[28] In particular, the procedure Scan, discussed in Chapter 2, will mark the copies of the chain as non-identical and would not be able to perform deletion of non-highest copies.

This alternative analysis, however, also allows a loophole. An unwanted option arises under which *zenščin* in the relevant cases of (52) adjoins to the DP (or within it) <u>before</u> the latter moves out of the VP, while this DP still is a trivial chain. Then the resulted segmented object moves from Spec-VP to Spec-IP, creating a chain which consists of identical (modified) copies, hence is uniform. This again wrongly predicts approximative inversion.

The late adjunction theory closes this loophole. Suppose the derivation has built a *v*P labeled *v* and the numeral QP (or DP) *pjat' zenščin* is inserted in Spec-*v*P. There exists an option of adjoining *zenščin* to the QP/DP or within it (e.g. Q), or cyclically insert the next element, e.g. T(ense). Merger of T (by substitution) does not lead to a change of c-command relations inside the *v*P. In contrast, adjunction of *zenščin* to the QP/DP or within this phrase is not allowed, as it leads to the change in the set of c-command relations inside the *v*P (still labeled *v*): relations involving the phrase *zenščin* are now included in this set (cf. Section 2.2.3). Hence, the cyclic option will be chosen.

Thus the solution lies not in the distinction between the categorial status of numeral phrases which do and do not allow adjunction to them, as Franks proposes, but, rather, in the distinction between those numeral phrases that move and those that do not, the distinction needed in Franks' system in any case.

3.3 Questions with wh-*in situ in an adjunct*

Recall that according to the view pursued in this chapter, Adjunct Condition effects with *wh*-movement arise because, at the time when the movement has to take place, the adjunct containing the moving element is unconnected to the phrase marker containing

target of movement, due to the particular timing of structure building as encoded in the LAH. This implies that after the adjunct becomes connected (last thing in overt syntax, see Section 2.2.3), Adjunct Condition effects should no longer arise. In this section we investigate a number of phenomena that relate to this expectation.

First, consider the English multiple *wh*-questions in which one *wh*-phrase is in an adjunct:

(53) Who left after Bill read what?

As is well known, in English questions of this sort only the first *wh*-phrase moves to the matrix Spec-CP. Importantly, however, *what* in (53) has matrix reading, which means that it establishes a dependency with the matrix C across the adjunct boundary.

Consider also a representative example of *wh*-in situ inside an adjunct in Japanese (from Ochi (1999b)):

(54) John-wa [Mary-ga nani-o katta ato] kaetta no?

     John-top Mary-nom what-acc bought after left Q

     '*lit.* What did John leave after Mary bought?'

As seen from the translation, *nani-o* in (54) has the matrix reading as well. Generally, examples like these are known to be well-formed in *wh*-in situ languages (see Huang (1982)).

Three major approaches to locality of A'-movement have been developed in the literature that address, in particular, the well-formedness of the example type in (53) and (54), as opposed to usual Adjunct Condition violations. Let us briefly recap them here. The traditional, most influential, approach is due to Huang (1982), who states it in terms of Subjacency. This approach presupposes that *what* in (53) and (54) moves at LF ('covertly') to Spec of the matrix interrogative C, across the adjunct boundary. The well-formedness of these examples, on this approach, is accounted for by stipulating that Subjacency holds at 'S-Structure', but does not hold at LF. Covert movement, on this approach, is not subject to island effects.

The second approach, originally due to Nishigauchi (1986), (1990), is recently revived in somewhat different terms in a number of works including Richards (2000), and, to a various extent, in Watanabe (1992b), (1992a) and Hagstrom (1998). Like Huang's, this approach presupposes the existence of LF (covert) movement, but, unlike Huang's, does not exclude covert movement from island effects. Rather, it maintains that these effects, in particular Adjunct Condition effects in (53) and (54), may be circumvented in covert movement if the entire island, viz. the adjunct in (53) and (54), moves ('pied-pipes') to the matrix CP domain at LF, rather than just the *wh*-phrase. The 'pied-piping' approach allows one to keep Subjacency as a condition on both S-Structure and LF movement, *a priori* a more natural theory.

In contrast to the previous two, the third approach, due to Tsai (1994) and Reinhart (1998), presupposes that (argument) *wh*-in situ like *what* in (53) and (54) do not move at LF at all. Rather, they are licensed in situ by a special mechanism of unselective binding.

Subjacency is irrelevant for this licensing, since unselective binding, not involving movement, is insensitive to syntactic island boundaries.

Regardless of which of these three approaches is adopted, the LAH makes the right predictions concerning the absence of Adjunct Condition effects in instances like (53) and (54). Suppose that, in accord with Huang's approach, *wh*-phrases may move at LF. The LAH correctly predicts that (53) and (54) will be derived, since at LF the adjunct and the matrix part have already joined together and nothing prevents LF movement of *what* in these examples across the adjunct boundary. Significantly, under the LAH, we no longer need to state any differences for LF movement as opposed to overt movement, in terms of locality, in particular, Huang's original stipulation that Subjacency does not hold at LF. In fact, the notion 'Subjacency' is (or, rather, continues to be) superfluous here.

The LAH is also compatible with the 'pied-piping' as well as unselective binding approaches. If we adopt the 'pied-piping' approach, then the LAH predicts, correctly, that at LF the adjunct will be part of the structure, hence, able to participate in movement operations. Similar considerations hold under the unselective binding approach, under the most natural assumption that this mechanism is operative at LF.

3.4 The LAH and Cyclicity

It might appear at first that the LAH is incompatible with the spirit of the strict cyclicity.[29] As we saw in the previous discussion, the LAH virtually forces all instances of adjunction

---

[29] The issue concerning the cyclic property of derivations goes back at least to Fillmore (1963) and Chomsky (1965), and is stated explicitly as the Strict Cycle Condition of Chomsky (1973):

(i) *The Strict Cycle Condition (SCC)*: No rule can apply to a domain dominated by a cyclic node A in such a way as to affect solely a proper subdomain of A dominated by a node B which is also a cyclic node.

(by Merge or Move) to target the constituents of the phrase marker which by the time of application are already 'embedded' in larger domains (unless it applies to a matrix node, e. g. CP), hence, one might think that it runs afoul of strict cyclicity. The goal of this section is twofold. First, we strengthen the argument for the LAH by considering an alternative which does not invoke postcyclic operations, and showing that, given a larger empirical realm, this alternative has a questionable status at best. Second, we argue that, under a closer scrutiny, the LAH may as well be compatible with the strict cyclicity.

3.4.1 A multi-planar alternative

Noam Chomsky (p. c.) suggests to us that it might be possible to avoid postcyclic operations, in particular, for the purposes of accounting for the standard 'adjunct condition' effects, if one makes use of a third dimension of syntactic 'space' (see, among others, Goodall (1987), Moltmann (1992), Muadz (1991), Stepanov (1997) for various implementations of this theoretical possibility). Informally speaking, the idea is that adjuncts are attached cyclically to their correspondent hosts, but in a different 'plane', in some sense 'orthogonal' to the one in which the rest of the derivation proceeds. This multi-planar view can account for the Adjunct Condition effects, and, possibly, for other empirical data discussed above, if one assumes that wh-extraction cannot take place established across planes. Thus who in (1), repeated here as (55), cannot move to the matrix Spec-CP since the two are located in different planes:

(55)   ?*What/??Which car did John [VP [VP go to bed] [after Peter fixed tᵢ]]?

The empirical basis for a multi-planar account of 'Adjunct Condition' effects and, in particular, for the absence of syntactic dependencies across planes, may be drawn from a range of phenomena discussed in Safir (1986). In particular, Safir is concerned with the absence of weak crossover effects in non-restrictive relative clauses, as shown in (56):

(56)   John$_i$, who$_i$ his$_i$ wife loves t$_i$, arrived early   [Safir's (16)]

Details aside, Safir proposes that non-restrictive relatives enter the structure at the level LF' (cf. also Chomsky (1982)), the level past LF, where, according to Safir, certain conditions inducing cross-over effects apply (namely, his 'Parallelism Constraint on Operator Binding'). Thus, he argues that coindexation of *John* and *who* in (51) takes place at LF'. Seeking independent support for LF', Safir further argues that it is the same level at which parentheticals enter the structure. This explains an observation made in McCawley (1982) that both non-restrictive relatives and parentheticals behave in a similar manner, as if they are not in the structure, with respect to VP ellipsis, presumably a pre-LF' phenomenon:

(57) John sold Mary, *who had offered him $600 an ounce*, a pound of gold, but Arthur refused to (=refused to sell Mary a pound of gold, ≠refused to sell Mary, who had offered him $600 an ounce, a pound of gold, ≠refused to sell Mary)

     [McCawley's (6a)]

(58) John talked, *of course*, about politics, and Mary did too (= Mary talked about politics too, ≠Mary talked too, ≠Mary talked, of course, about politics too)

[McCawley's (5a)]

Safir's LF' theory does not translate easily into the traditional T-model of grammar. One obvious reason is postulation of an additional level of representation which runs counter usual intuitions about language design (see Chomsky 1995 for a general discussion). Furthermore, given the standard architecture of grammar, in which the levels of representation that receive (parts of) the product of a syntactic derivation, PF and LF, do not 'talk' to each other, this theory also raises another, more immediate, concern. Namely, if a non-restrictive relative, or other piece of structure, enters the derivation after it branched off to the PF interface, then how can this structure have a phonological reflex, and, moreover, be linearized at PF with respect to other constituents? It appears that we are practically forced to say that LF' is a level that has access to the PF component, and at the same time is similar to LF in terms of the relevance of binding relations (cf. coindexation of *John* and *who* in (51)). Since this is clearly a high price to pay, in light of available evidence, it would then be more desirable to reformulate Safir's account so as to avoid the reference to this level.

A multi-planar framework provides a natural way of restating Safir's analysis. Suppose non-restrictive relatives (as well as parentheticals) enter the derivation along with their correspondent relative heads, cyclically. Suppose also that they are attached to their heads in a plane distinct from the one in which the rest of the derivation takes place. The absence of weak cross-over effects in (51) then follows immediately since the relevant binding relations (e.g. between *John* and *who*) cannot be established across different planes. Similar considerations obtain with respect to the identity relation on VP

ellipsis in (57). Furthermore, if no syntactic dependencies can be established across planes in general, we expect that extraction out of non-restrictive relatives is also impossible, which is indeed the case, as shown in (4b) repeated here as (59):

(59)   ?*What did John meet Peter, who was wearing?

♦

Under the assumption that non-restrictive relatives are adjuncts (see Section 1), the ungrammaticality of (59) can be attributed to an Adjunct Condition effect. Thus, no additional levels of representations need to be postulated in the multi-planar analysis,[30] and no 'postcyclic' operations are necessary to account for the Adjunct Condition effects.

Now the question is whether a multi-planar version of Safir's approach can be extended to a more general theory of extraction out of adjuncts, replacing the LAH. We believe the answer is negative. First, as Safir points out, weak cross-over effects obtain with restrictive relatives, as opposed to non-restrictive relatives (cf. (56))

(60)   *?A man$_i$ who$_i$ his$_i$ wife loves t$_i$ arrived early   [Safir's (15)]

Similarly, VP ellipsis targets restrictive relatives, recognizing them as part of the structure:

---

[30] Although, of course, some additions to the technical machinery have to be made, particularly, concerning the procedure linearizing the material in different planes, at PF.

(61) John sold a lady *who had offered him $600 an ounce*, a pound of gold, but Arthur

refused to (=refused to sell the lady who had offered him $600 an ounce, a pound of

gold, ≠refused to sell a lady a pound of gold, ≠refused to sell the lady)

Safir accounts for the contrast between restrictive and non-restrictive relatives by

stipulating that only the latter enter the derivation at the level of LF'. In the present terms,

restrictive relatives must enter the derivation in the same plane as the rest of the structure,

in order for the relevant relations (binding in (60), identity in (61)) to be established. But

that cannot be true in the multi-planar framework, since, crucially, extraction out of

restrictive relatives is still ungrammatical, as shown in (4a) repeated here as (62) (cf.

(59)):

(62)   ?*What did John meet a man who was wearing?

If restrictive relatives enter the derivation in the same plane as the rest of the structure,

then it is unclear what prevents establishing a *wh*-movement dependency in (62). In other

words, the ungrammaticality of (62) must be due to something other than an Adjunct

Condition effect, and remains a mystery under this approach.

There is another, potentially more serious, problem with the multi-planar approach,

involving constructions where the Adjunct Condition effects are obviated. For instance,

let us return to (53) repeated here as (63):

(63)   Who left after Bill read what?

If the adjunct and the rest of the structure are in different planes in the ungrammatical (55), as the multi-planar approach assumes, then, all else equal, the grammaticality of (63) would suggest that the adjunct there somehow ends up in the *same* plane as the rest of the structure (similar considerations apply to the Japanese (54)). A proponent of the multi-planar approach might try to derive the contrast in a principled way by claiming that in (63) we are dealing with *wh*-movement after Spell-Out (namely, LF *wh*-movement) and Spell-out forces the two planes to 'collapse' into one, perhaps for reasons related to linearization at PF. But let us remember that we are dealing here with the same level of representation at which, as Safir observes, weak cross-over effects do not obtain (cf. (59)). By allowing the adjunct plane to 'collapse' with the other plane(s) we are predicting cross-over effects, incorrectly, voiding Safir's insight that constitutes the basis for our multi-planar approach. Hence, the adjunct in 53 can*not* be in the same plane as the rest of the structure. It is extremely difficult to reconcile both the absence of weak cross-over effects in (59) and the absence of Adjunct Condition effects in (63), under the multi-planar approach based on Safir's theory. The difficulty is particularly striking if we assume a theory of unselective binding along the lines of Tsai (1994) and Reinhart (1998) according to which the dependency between the matrix C and the *wh*-in situ in (63) is not that of movement, but of binding, the same sort of relation as the one involved in weak cross-over effects. Thus, it appears that the multi-planar theory which has a chance to be

best articulated at the moment, given Safir's proposals, is not capable of replacing the LAH.[31]

In contrast, the LAH does not face the problems outlined above. First, it subsumes extraction out of both restrictive and non-restrictive relatives under one conceptual umbrella, providing a uniform account of Adjunct Condition effects in both. Second, as we saw in the previous section, it straightforwardly accounts for the absence of Adjunct Condition effects in (63) and the like (e.g. (54)). In addition, note that under the LAH it is not necessary to view the absence the Adjunct Condition effects in (63) and the absence of weak cross-over effects with non-restrictive relatives, as in (59), along the same lines. The LAH explains the former, whereas some other condition, perhaps even the multi-planar theory of non-restrictive relatives itself, may be correct for the latter. Some evidence that this suggestion may be on the right track comes from questions containing *wh*-in situ inside a non-restrictive relative clause (which, recall, we consider as adjuncts):

(64)   ?*Who met Peter, who was wearing what?

The LAH treats (64) on a par with (63) (see the discussion in the previous section). Hence, it cannot be responsible for the ungrammaticality of (64). On the other hand, the degradation in (64) is consistent with the multi-planar approach, which precludes establishing any syntactic relation between the elements in the matrix and relative clause,

---

[31] It is also not entirely clear, under the multi-planar approach, how simple cases like (i) are derived:
(i)   How did John fix the car?
Presumably, some kind of 'cross-planar' movement ought to be allowed. We will not pursue this possibility here.

specifically, that pertaining to the *wh*-in situ. Thus, the multi-planar approach may in fact complement the LAH in a general theory of grammar.

3.4.2 Postcyclic operations

Before we conclude the discussion of the LAH let us remark on the issue of how this algorithm fares with respect to the Strict Cycle Condition in (49). In order to evaluate this issue, one should be clear on the theoretical status of the SCC itself. The early pre-minimalist view, going back to Fillmore (1963) and Chomsky (1965) assumed that the SCC exists as a principle of grammar. Later studies, beginning at least with Freidin (1978), argued that the SCC is epiphenomenal and deduced its effects from independent principles of grammar. Continuing in the same vein, in the minimalist framework Chomsky (1995) encodes cyclicity into the independent properties of the computational system, such as his Extension condition and the definition of 'strong features' (in fact, with certain redundancy, as Bošković and Lasnik (1999) point out; see fn. 6 and 21; see also Epstein (1999) for a slightly different way of incorporating cyclicity effects).

Regardless of whether the SCC is real and has its place, in some form, in the theory of grammar, or, rather, is fully derivable from other independent principles, we believe that the LAH is not inconsistent with it. The key point here is that under the LAH, adjunction operations may be naturally regarded as *postcyclic* (in the true sense of the term) that is, take place after all cyclic transformations have applied. This means that the SCC simply does not extend to adjunction transformations.

Further, it may turn out that the entire empirical domain of the SCC, or whatever principle(s) derive it, include only the phenomena involving substitution operations, but

not adjunction (by Merge or by Move). Chomsky (1995c), p. 191 in his discussion of the Extension requirement, suggests that this is indeed the case. Some of the relevant constructions Chomsky discusses involve superraising and *wh*-island violations (with regard to the latter, see also Kawashima and Kitahara (1996)):

(65)   a. *[TP John seems [TP it is certain to be here]]

       b. *[CP How did John wonder [CP what Mary fixed]]?

Without violating any known condition on locality of movement, each of the examples in (65) may have a good derivation, which involves countercyclic operations. Thus in the raising construction (65a) it is possible to raise John directly to the matrix Spec-TP, and then insert *it* in the intermediate Spec. Similarly, in (65a) *how* may move directly to the matrix Spec-CP, and *what* can then (acyclically) raise to the embedded Spec-CP. But, as Chomsky notes, these examples crucially involve substitution operations, not adjunction ("Note that the very strong empirical motivation for the strict cycle just given does not apply in these cases" [p. 191]). In the absence of evidence to the contrary, it seems reasonable to maintain that only substitution operations are relevant for the SCC. We then conclude that, regardless of the status of SCC, the LAH is (perhaps trivially) consistent with its spirit.

4. Final Remarks on the 'Eclectic' Approach

In this and preceding chapter we have developed an 'eclectic' approach to the extractability out of non-complements. We hope to have shown that the 'eclectic'

approach is conceptually attractive, avoids conceptual and empirical problems faced by the more traditional 'unified' approach, and most importantly, has a greater empirical coverage than the latter.

From our 'eclectic' account, it follows that subject (more generally, 'derived position') islands have a PF-related nature, whereas adjunct islands are a consequence of the 'narrow syntactic' properties of grammar.

A similar conclusion was reached independently by Merchant (1999) in his detailed study of Sluicing. Merchant argues at length that 'derived position' islands are indeed 'PF islands', that is, islands that arise as a violation of some PF principle. Consider (66) involving extraction out of subject in English, and the corresponding Sluicing example (67):

(66)   *Which artist did she say that [a biography of_] appeared last year?

(67)   A biography of some artist appeared this year - guess which!

It has been known since Ross (1969) that Sluicing can 'repair' island violations. Now, Merchant argues independently that Sluicing involves deletion of an IP at PF. It follows that the violation caused in a subject island structure can be 'repaired' at PF by deletion of the entire IP including the island. Thus, Merchant reasons, in order for a violation to be repairable at PF, the violation itself must have a PF nature. Of course, Sluicing also ameliorates adjunct islands (cf. *If Ben talks to someone, Abby will be mad, but I don't remember who*). But here, Merchant argues, the amelioration is only apparent. According

to him, there exists an alternative analysis of the sluice which does not involve an adjunct island at all.[32]

One further piece of empirical support for the 'eclectic' approach comes from the well known difference in grammaticality between 'subject condition' and 'adjunct condition' violations. In general, 'subject condition' violations seem to have lower acceptability than 'adjunct condition' violations. This sort of difference is unexpected under the 'neo-unified' approach such as Nunes and Uriagereka's, based on the complement/non-complement distinction. On the other hand, under the 'eclectic' approach in which the two types of violations are handled by different mechanisms, such difference in acceptability is not surprising.

Another piece of evidence comes from recent studies of a curious psycholinguistic phenomenon known as syntactic satiation. It has been known anecdotally that speakers who initially judge sentences involving certain grammatical violations as unacceptable, evaluate them as increasingly acceptable after repeated exposure. Snyder (2000) and Hiramatsu (1999) induced satiation effects experimentally and discovered that English speakers show significant satiation effects with regard to sentences involving extraction out of subjects ($p < .01$ by paired t-test, Hiramatsu (1999)). The same speakers, however, do not show a satiation effect with respect to sentences involving extraction out of adjuncts. This divergence is unexpected, if extractability out of subjects and adjuncts is regulated by the same mechanism of grammar, as is the case under the traditional unified approach, or the neo-unified one. On the other hand, if extractability out of subjects and

---

[32] For Merchant, the identity requirement for sluicing is not structural, but, rather, semantic. See also Lasnik (to appear), who questions Merchant's empirical arguments for his conclusion.

adjuncts are regulated by different mechanisms, this divergence is not surprising. Furthermore, under the 'eclectic' theory outlined above, in which extractability out of subjects is a matter of the PF interface, Snyder's and Hiramatsu's results suggest that satiation is a process that targets the PF side of the grammar. It remains for future research to determine the full extent of this conclusion.

CHAPTER 4

'NP-shells'

1. Introduction

In the preceding chapters we have shown that the cyclic domain approach implemented in the minimalist framework successfully handles a wide range of phenomena associated with extraction out of non-complements, viz. subjects and adjuncts, correctly capturing the cross-linguistic variation, in cases where it is attested. In this and next chapters we begin to look at various kinds of movement dependencies that involve finite clauses in a complement position. Particular attention will be devoted to investigation of theoretical issues associated with *wh*-extraction out of finite complements.[1]

It is well known that in English and many other languages a *wh*-element can be extracted out of a finite clause:

(1)   a. What did John believe [(that) Peter bought *t*]?

         b. How did Bill think [(that) Molly fixed the car *t*]?

Chomsky (1973) proposes that *wh*-extraction is impossible across a tensed (=finite) clause boundary, unless the moving item goes through a local COMP, which serves as a sort of an 'escape hatch'. This idea was at the heart of the standard view that 'long-

---

[1] With regard to complement NPs/DPs, see an account of contrast in extraction arising with respect to their (non-)presuppositionality in Chapter 2. We leave aside conditions regulating extractability out of complement NPs, many of which seems to have lexical nature. See Chomsky (1973), (1977), Bach and Horn (1976), and, more recently, Müller (1995), Müller and Sternefeld (1995), Davies and Dubinsky (2001) for discussion.

distance' *wh*-movement, as this kind of extraction is normally referred to, is *successive cyclic*, that is, proceeds as a series of local movements through intermediate COMPs.

As has been noted early on (already in Chomsky 1973), however, not all languages freely allow extraction out of finite clauses. Below is a list of degraded counterparts of (1) from Russian (2), Polish (3) (Giejgo (1981)), Finnish (4) (Ross (1967)), and Georgian (5) (Harris (1981)) which usually allow/require *wh*-fronting in simple *wh*-questions:

(2) ?*Kogo Ivan skazal, čto ljubit Maria?

        whom Ivan said   that likes Maria

        'Who did John say that Mary likes?'

(3) ?*Kogo ty wiesz [że Janek lubi t]?

        whom you know   that John loves

        'Who do you know that John loves?'

(4) *Mitä hattua uskoit   ettei   hän koskaan käyttänyt?

        Which hat you believed that not she ever   used

        'Which hat did you believe that she never wore?

(5) *sad tkva gelam, (rom) mama çavida?

        where said Gela that Father went

        'Where did Gela say that Father went?

A similar situation obtains in certain *wh*-in-situ languages. In particular, Dayal (1996a) notes that a *wh*-phrase inside an embedded question cannot have matrix scope/construal in Hindi, unlike in Chinese and Japanese. This suggests that the relevant dependency cannot be established at any point, including LF:[2]

(6) a. Anu jaannaa   caahtii   hai [ki   kis-ne kyaa khariidaa]

     Any know-Inf want-pres   that who   what buy-past

     'Anu wants to know who bought what'

     NOT 'What is such that Anu wants to know who bought it?'

     NOT 'Who is such that Anu wants to know what s/he bought?'

b. cf. ni   xiang-zhidao [shei mai-le   sheme]

     you wonder   who buy-asp what

     'Who is such that you wonder what s/he bought?'

     'What is such that you wonder who bought it?'

Under the assumption that *wh*-movement is 'unbound', the lack of long-distance dependencies in these cases is puzzling. Several attempts have been made in the literature to explain the restriction on long-distance movement at issue, (see Bailyn (1995), Zaenen (1983), Stepanov and Georgopoulos (1997) for Russian, Giejgo (1981) for Polish, Dayal

---

[2] It is well known that French allows *wh*-in situ as well as *wh*-movement in short distance matrix clauses. With respect to long-distance interrogation, French instantiates a mixed pattern whereby overt *wh*-movement is possible (as in (1)), but *wh*-in situ is not (as in (6)). See Bošković (1998a) for more discussion and an account of this mixed pattern.

(1996a) for Hindi) with some illuminating results, but the general aspect of the problem remains far from fully resolved.

In this chapter, we argue that *wh*-extraction out of finite complements in Russian and Polish is precluded because finite complements in these languages are structurally realized as structural adjuncts to a clausal pro-form, in a configuration we term 'NP-shell', following the term coined in Müller (1995), Müller and Sternefeld (1995). Extraction out of finite complements thus amounts to some form of a Complex NP island violation, and is accounted for along the lines of Chapter 3. We then argue that this form of complementation is basic, rather than the standardly assumed structure in which the finite clause is a complement of the subordinating predicate. In chapter 6, we develop a general theory of forming *wh*-dependencies in finite clauses, a theory that reconciles the 'shell' hypothesis with the availability of long-distance questions of the type (1) in those languages that allow them.

This chapter consists of two main parts. In the first part, we outline an approach to finite complementation in terms of "NP-shells", on the basis of Russian and Polish. In the second part, we discuss the strategy of *wh*-scope marking in Slavic in detail and show how it provides further evidence for the NP-shell hypothesis.

2. NP-Shells

2.1 The Paradigm

Extraction out of finite complements in many varieties of Russian (noted to various extent in Bailyn (1995), Comrie (1972), Müller (1995), Pesetsky (1982), Zaenen (1983)) consistently leads to a mild to serious degradation in most, if not all, environments

involving *wh*-movement, including *wh*-questions (7) (cf. (2)), relative clauses (8), bcomparative constructions (9) (cf. Chomsky (1977)) and cases of Sluicing involving 'sprouting' (10) (cf. Ross (1969), Chung et al. (1995)). Furthermore, extraction of adjuncts are generally more degraded than of arguments, suggesting the familiar asymmetry involving 'Subjacency/ECP' effects with islands (cf. Lasnik and Saito (1984)):

(7) a. ?*Kogo Ivan polagaet, čto Maria ljubit _?

   Whom Ivan supposes that Maria likes

   'Who does John think (that) Mary likes?'

   b. *Kak Ivan dumaet, čto Petr počinil mašinu _?

   How Ivan thinks that Petr fixes   car

   'How does John think that Peter fixed the car?'

(8) a. ??Čelovek, kotorogo ja sčital   čto Petr znaet _

   Person   whom   I thought that Petr knows

   'The person who I thought Peter knew'

   b. ?*Gorod, v kotoryj my dumali čto etot pevec nikogda ne priedet _?

   City in which we thought that this singer   never   not will-come

   'The city where we thought that this singer will never come'

(9) ??Ivan prines bol'še fruktov čem Marina utverždala čto on prineset _

   Ivan brought more fruits   than Marina claimed   that he will-bring

'John brought more fruit than Marina claimed that he will bring'

(10) ??Marija skazala čto Petr razbil okno/porezal palec tol'ko ja ne mogu vspomnit' čem

Maria said   that Petr broke window/cut finger but   I not can   remember what-inst.

'Mary said that Peter broke a window/cut his finger but I can't remember with what'

It is sometimes noted that without the complementizer *čto* 'that', at least some of the

relevant sentences become more acceptable:

(11) a. (?) Kogo, Ivan polagaet, Maria ljubit _?

Whom Ivan thinks   Maria likes

'Who does John think Mary likes?'

b. (?) Gorod, v kotoryj, my dumali, etot pevec nikogda ne priedet _?

City   in which we thought this singer   never   not will-come

'The city where we thought this singer will never come'

However, it is unclear if the examples in (11) represent instances of genuine long distance

extraction. There is a possibility that the clauses *Ivan polagaet* and *my dumali* in (11) is

some sort of a parenthetical constituent inserted inside a root question (*Kogo Maria*

ljubit? 'who does Maryt like?', and ...*v kotoryj etot pevec nikogda ne priedet* 'in which

this singer will never come', respectively). For one thing, it is possible to set off these

clauses with a comma intonation, suggesting that this possibility is real. Since we are not

aware of a clear test that distinguishes the parenthetical from the true matrix clause in

(11), we do not consider (11) as part of the relevant paradigm and put it aside for the purposes of this discussion.

The sentences in (7) through (10) improve when the complement of a matrix verb is not a finite, but a subjunctive or infinitival clause. Subjunctive clauses in Russian are introduced by a limited set of predicates inducing irrealis mood, specifically volitional predicates of wish, desire or command, such as *xotet'* "want", *predpočitat'* "prefer", *prikazat'* "order", etc.[3] The complementizer of a subjunctive clause also changes its morphological form:

(12)   Čto Ivan xočet kupit' _?

   What Ivan wants to-buy

   'What does John want to buy?'

(13)   Čto Ivan xočet čtoby   Petr kupil _?

   What Ivan wants that-sbj Petr bought

   'What does John want Peter to buy?'

Certain adverbial 'reason' clauses also utilize the subjunctive complementizer *čtoby*. In these cases extraction is generally degraded:

(14)   a. Ivan priexal v Moskvu čtoby kupit' dom

         Ivan came   in Moscow that to-buy house

---

[3] See Brecht (1972) for a detailed examination of those predicates.

'Ivan came to Moscow in order to buy a house'

b. ??Čto   Ivan priexal v Moskvu čtoby kupit'?

what Ivan came   in Moscow that to-buy

lit. 'What did Ivan come to Moscow in order to buy?

We assume that *čtoby*-clauses in cases like (14b) fall into the class of adverbial adjuncts, similarly to English and the ungrammaticality is due to their adjunct status.

A similar picture also obtains in Polish (see Giejgo (1981)). As in Russian, in this language extraction out of finite clause is degraded, while extraction out of subjunctive/infinitives is acceptable:

(15) a. *Kogo ty wiesz   że   Janek lubi t   [Pol]

whom you know   that John loves

'Who do you know that John loves?'

b. Co   chcesz żebym ja zrobił?

What want that   I did

'What do you want me to do?'

c. Kogo on chce spotkać?

Who he wants to-meet

'Who does he want to meet?

In the next section we offer an account of the ungrammaticality of (7) through (10) in terms of a Complex NP island violation (see our approach to deducing the relevant

locality constraint in Chapter 3). Our account is largely based on earlier insights in Comrie (1972) concerning the structure of finite complements in Russian and Giejgo (1981) for Polish, and supports an approach to finite complementation based on an extension of their proposals.

2.2 Correlatives and finite complements in Russian

A finite complement clause in Russian can be optionally preceded by an overt pro-form *to*, as shown below:

(16)   Petr utverždal   (to) čto Ivan ljubit Mašu

        Petr claimed   to   that Ivan loves Masa

        'Peter claimed   that John loves Mary'

For the purposes of discussion, let us refer to this pro-form as a *correlative*, borrowing the term from Comrie (1972) and Giejgo (1981) who discusses a similar element in Polish (see below). In his detailed investigation of the distribution of *to*, Comrie (1972), following descriptive literature, identifies the following descriptive characteristics of the latter:

(17) a) it literally means "that";

   b) it functions as a demonstrative referring to the clause that follows it;

   c) it has no detectable effect on the sentence meaning;

   d) if present, may induce an intonational break between it and the clause that follows

Property (17c) must be emphasized. The correlative *to* has no contribution to the meaning of the sentence. In this respect, it crucially differs from the English *it* in contexts similar to (16), illustrated in (18b) (cf. Rothstein (1995))

(18)   a. John and Mary announced that they got married

         b. John and Mary announced it that they got married

(18a) and (18b) cannot be both uttered in the same situation. (18a) is a simple assertion about John and Mary's announcement. In (18b), on the contrary, John and Mary's marriage is taken by the speaker as a point of fact.[4] The difference between (18a) and (18b) is likely to be presuppositional in nature. This intuition is further supported by the following implicature cancellation test: a continuation of (18a) and b) with an explicit denial of a state of affairs reported in the embedded clause leads to a contradiction, but only when *it* is present. Thus, while (19a) is not contradictory, (19b) is:

(19)   a. John and Mary announced that they got married, but, in fact, they never did.

         b. #John and Mary announced it that they got married, but, in fact, they never did.

The pronoun *it* in (18b), arguably, contributes to the meaning of the sentence. In contrast, a version of (16) with or without the correlative *to* can be uttered essentially in the same

---

[4] In this sense, *announce* patterns with factive verbs in the sense of Kiparsky and Kiparsky (1970). See Lasnik and Saito (1991) for a related claim.

set of situations, as an assertion about Peter's particular claim. Furthermore, sentences involving *to* with a non-factive verb like *utverždat'* 'to claim' pass the implicature cancellation test:

(20)   Ivan vsegda utverždal to čto on ljubit boršč, no na samom dele

Ivan always claimed   *to* that he likes borscht but in fact

eto ne tak - on ego terpet' ne možet

this not so   he it tolerate not can

'John always claimed that he liked borscht but in fact this is not true - he can't stand it'

(20) is not contradictory. This supports the conclusion that the correlative *to* does not contribute to the meaning of the sentence.

The question of defining natural classes of predicates that may and may not occur with *to* has been a subject of debate in the traditional literature (see Comrie 1972 for an overview). In point of fact, the set of such verbs varies somewhat from speaker to speaker but generally includes predicates such as *utverždat'* "claim", *zabyt'* "forget", *znat'* "know", *govorit'* "tell", *objasnit'* "explain", *podtverdit'* "confirm", *ponjat'* "understand", but not verbs like *otvetit'* "reply", *dumat'* "think", *(pred)polagat'* "suppose" or *sčitat'* "believe".[5] Thus the following is not acceptable:

---

[5] Some variation indeed exists. For instance, Pesetsky (1982), p. 244 reports an example involving *to* occurring with *polagat'*, as acceptable:
(i)   Ja polagaju to čto Maša pjana
       I suppose   that Maša is drunk
To our ear, (i) sounds somewhat awkward. On the other hand, Comrie (1972) reports similar examples involving *to* with *skazat'* "to say" as unacceptable, whereas we do not discern any degradation in these cases, especially in contexts involving comparing alternatives, cf.

(21)   *Petr sčital   to čto   Ivan ljubit Mašu

Petr believed   that Ivan loves Maša

'Peter believed that John loves Mary'

Giejgo (1981) was, to our knowledge, the first to suggest that the correlative is responsible for the impossibility of extraction out of finite clauses in Polish. Specifically, she observes that in Polish, like in Russian, finite complements can be introduced by *to*, and proposes that the two form a constituent (NP) (cf. (16)):

(22) Wiesz   [NP (to) [CP że   Janek lubi Maria]]

know-you   that John loves Mary

'You know that John loves Mary'

Giejgo argues that the same structure is involved in the extraction example (15). She postulates that *to* is uniformly present in the syntax of finite complementation, but can optionally delete at PF. This leads Giejgo to a natural conclusion that the deviance of (15) is due to a locality violation as a result of extraction out of a 'complex NP island'. Let us call the NP complement of the matrix verb an 'NP-shell', borrowing the term from Müller (1995).

(ii) Ja skazal to čto Petr ne ljubit Mariju - šepotom,   a   to čto Marija ne ljubit Petra - gromko
      I said   that Petr not loves Mary   whisper-instr and   that Mary   not likes Petr -outloud
      'I said that Peter doesn't love Mary - in whisper, and that Mary doesn't love Peter - outloud'

As one might expect, extraction out of a complement preceded by overt *to* is ungrammatical as well:

(23)   ?*Kogo Petr skazal   to   čto Ivan ljubit?   [Rus] (cf. (21))

Who Petr said   that Ivan likes

'Who did Peter say that John likes?'

However, in its present form Giejgo's account cannot be maintained. The reason is that in both Polish and Russian the extraction is ungrammatical even with verbs like *utverždat'* "claim", where, as (21) demonstrates, *to* is not allowed. The relevant examples follow (see also (7)-(10)):

(24)   ??Kogo Petr sčital   čto Ivan ljubit?   [Rus] (cf. (21))

Who Petr believed that Ivan likes

(25)   a. Mysli (*to) że   Janek lubi Marie   [Pol]

thinks-he   that John loves Mary

'He thinks that John loves Mary'

b. ??Kogo ty   myslisz że   Janek lubi?

Whom you think   that John loves

'Who do you think that John loves?'

(21) and (25a) would appear to suggest that the structure of complements of *sčitat'/mysleć* are not NPs. Hence, all else equal, extraction should not be precluded by

the NP boundary (complex NP island effect). If so, the degraded status of (24) and (25a) is unexplained.

Despite this problem for Giejgo's account, we wish to claim that her intuition concerning the NP-shell complementation is nevertheless correct. We believe the reason why this problem arose in the first place was the failure to provide a formal criterion for distinguishing the classes of verbs that may and may not be followed by the overt correlative *to*. In what follows, we devise such a criterion and extend Giejgo's insight to cover the entire range of Slavic constructions disallowing long-distance *wh*-extraction.

2.3 Finite Complements, Correlatives and Case Theory

Recall that for Giejgo, the correlative *to* heading an NP-shell is necessarily base-generated as a complement of certain verbs, such as *wiesz*, and then deleted at PF. Suppose that finite complements of *sčitat'/mysleć* are also 'encapsulated' inside an NP-shell where *to* is a 'correlative' pro-form, exactly as complements of *wiesz* (cf. (22)). Suppose that alongside the overt *to*, there exists a correlative TO which is exactly like *to* in terms of its properties (cf. (17)) except it has no phonological matrix. TO necessarily occurs with verbs like *sčitat'/mysleć*, so that the structure of (21) and (25a) is the following:

(26) a. Petr sčital [NP TO [CP čto Ivan ljubit Mašu]]

 Petr believed that Ivan loves Maša

 b. Mysli [NP TO [CP że Janek lubi Marie]]

 thinks-he that John loves Mary

Positing the silent correlative TO allows us to dispense with Giejgo's assumption that overt *to* in cases like (16) and (22) undergoes optional PF deletion. Instead, we now say that these cases involve either the overt *to* or silent TO.

The obvious question that arises at this point concerns the distribution of both correlatives. We propose to account for this distribution via Case theory. Note that the verbs that allow the overt correlative *to* with finite complements are generally verbs that take direct objects, that is, are structural Case-assigners/checkers, as shown below:

(27)   a. Ja utverždal /ponjal   /pomnju/ podtverdil/   (to) čto Ivan ušel

    I claimed   understood remember confirmed   that Ivan left

    b. Ja utverždal /ponjal   /pomnju/ podtverdil/   *(éto)

    I claimed   understood remember confirmed   this

On the other hand, verbs that do not allow overt *to* generally do not assign/check Case:

(28)   a. ?*Ja dumaju/sčitaju/ (pred)polagaju/   to čto Ivan ušel

    I think   consider   suppose   that Ivan left

    b. *?Ja dumaju/ sčitaju/ (pred)polagaju eto

    I   think   consider suppose   something

These correlations can be accounted for if we assume, adopting the checking approach to the Case theory (see Chomsky (1995c)), that NPs headed by overt *to* necessarily bear the

Case feature, whereas phrases headed by silent TO are allowed not to have it (even though they may). In other words, *to* must be drawn from the lexicon with a Case feature, whereas TO can be drawn from the lexicon either with or without a Case feature.

Note that this account extends to cases where the Case assigning element is not necessarily a verb. The following demonstrates that either *to* or TO may occur with a finite clause in subject position, where Nominative Case is assigned/checked by Infl ((29b) is adapted from Comrie (1972), p.1):

(29)   a. (To) čto Petr ljubit Mašu, bylo dostatočno očevidno

        that Petr likes Maša   was enough   obvious

        'That Peter likes Mary was obvious enough'

      b. (To) čto Kolja sdal ekzamen, udivilo   nas vsex

        that Kolja passed exam surprised us all

        'That Kolja passed the exam surprised us all'

Evidence that phonologically silent elements may or may not have a Case feature is given in Franks (1995), Uriagereka (1988) and Raposo and Uriagereka (1990), among others.[6] If a correlative has a Case feature, it must be checked against a matching feature

---

[6] For instance, Raposo and Uriagereka (1990) discuss the following examples in Galician Portuguese where the expletive can be either null (ia) or overt (ib):

(i)   a. En Beirute, [e] ficaron onte   [sc [os soldados] sen   armas]
      In Beirut   became yesterday those soldiers without guns
    b. *En Beirute, *el* ficaron onte   [sc [os soldados] sen   armas]
      In Beirut   it became yesterday   those soldiers without guns

As they show, the matrix Infl assigns Nominative to *os soldados* in (ia), and the null [e] remains Caseless. In (ib), both NPs (*el* and *os soldados*) need Case, but only one (Nominative) can be assigned by Infl. On the other hand, the same null [e] can be assigned Case, as shown below:

of some functional head. In cases like those above, that means that the matrix verb must be a Case assigning verb. On the other hand, if a correlative does not have a Case feature, then the matrix verb must not have this feature either, since if it does, it will remain unchecked in the course of the derivation.

Under the view outlined above, it follows that *to* preceding a finite complement may occur only with Case assigning verbs, and TO either with Case-assigning or non Case assigning verbs. In all cases, the presence of an NP-shell blocks extraction out of a finite complement.

We are now in a position to offer an analysis of ungrammaticality of sentences involving extraction out of finite clauses in Russian and Polish. We suggest an analysis in terms of our 'late adjunction' theory in Chapter 3. Recall that for us, complex NPs containing finite clausal complements actually involve adjunction of the clause (see also Stowell (1981)). According to the late adjunction theory, this adjunction is postcyclic. That means, in particular, that in the derivation of (2a), reproduced below, at the point of insertion of the interrogative C into the matrix clause, the finite clause embedded under the matrix verb is actually unconnected with the matrix clause. This is shown in (30), where K1 stands for the matrix clause, K2 for the embedded clause:

(2) a. ?*Kogo Ivan polagaet, čto Maria ljubit _?

(ii) a. [e] chegou un estudiante tarde
          arrived a student   late
      b. El chegou un estudiante tarde
          arrived a student   late
Raposo and Uriagereka argue that in (ii) *un estudiante* receives (partitive) Case from the verb, while the null [e] in (iia) or overt *el* receives Nominative from Infl.

Whom Ivan supposes that Maria likes

'Who does John think (that) Mary likes?'

(30)   K1: Q Ivan polagaet   TO

              Q Ivan supposes

        K2: čto Maria ljubit kogo

             that Maria loves who

This results in a timing mismatch resulting in the inability of the uninterpretable feature of matrix interrogative C to be checked at the point of its insertion in K1, similarly to other cases involving extraction out of adjuncts (see Chapter 3 for details).

It should be noted that under the late adjunction analysis, adjunction of K2 or any finite complement will be literally to *to*/TO, which is a non-projecting term, hence, is an maximal phrase (XP) in accordance with the bare phrase structure. The representation of the basic structure of the finite complementation in Slavic should then be adjusted slightly, as in (31):

(31)   Ivan polagaet   [$_{TO}$ [$_{TO}$ TO][$_{CP}$ čto Maria ljubit Petra]]

          Ivan presupposes   that Mary loves Peter

The analysis in (30) recalls an original analysis of clausal complementation in terms of generalized transformations developed in Chomsky (1955). In that system, sentences like *John knew that Peter left* would be generated by combining two phrase markers,

John knew it and *Peter left,* into a single phrase marker *John knew it # Peter left* (#

marking the sentence boundary). The generalized transformation would then delete #,

remove *it,* and insert *that* before the complement clause. Thus before the application of

the generalized transformation, a situation similar to (30) would arise. In Chomsky's

system, the difference between Russian/Polish and English would have to be stated in

terms of the applicability of the (portion of the) transformation rule deleting the

'correlative' in the matrix object position in English (cf. *it*), but not in Russian (cf. *to*/TO).

But while Chomsky's analysis is based on the notion of 'kernel' sentences, our present

analysis need not rely on it. The fact that the K2 is a clause is epiphenomenal: what

matters is that it is a structural adjunct. The Russian case in (30) and especially similar

cases involving overt *to* also demonstrate that we do not need to complicate the

transformational process with deletion of TO/to and insertion of *čto* as would be

necessary under the 'kernel' analysis. It remains to be seen, of course, how to state the

difference between Russian/Polish and English, in our system. We postpone this task

until Chapter 6.

Consider again subjunctive and infinitive complements (of volitional predicates). As

we know, extraction is possible out of these complements (cf. (12), (13) and (15b, c)).

Following the reasoning so far, we conclude that NP-shell must not be involved in these

cases, so that no Complex NP island ensues. Independent evidence supports this

conclusion. In Polish, a nominal clitic like *go* "him" can climb out of an infinitive and

subjunctive clause, but not out of a finite one:

(32)   a. On go_i chce   przywitać t_i

He him wants to-greet

'He wants to greet him'

b.   On go$_i$ chce [żeby ona chwaliła t$_i$]

He him wants that she praised

'He wants her to praise him'

c. cf. ?*On go$_i$ wie   że   ona chwaliła t$_i$

He him knows   that she praised

'He knows that she praised him'

Given that clitic climbing is subject to locality (see Giejgo (1981), Progovac (1993) for further discussion), this suggests that movement in (32a, b) does not cross a complex NP boundary (violating the 'Complex NP island').

Another piece of evidence comes from the well-known 'obviation effect', whereby the subject of the subjunctive clause must be disjoint in reference with the subject of the matrix clause:

(33)   Ivan$_i$ xočet [čtoby on*$_{i/j}$ poexal v Ameriku]

Ivan$_i$ wants that   he   went   in America

'Ivan wants him to go to America'

This suggests that the pronoun in the subjunctive clause is subject to Condition B, which operates in sufficiently local domains. If (33) involved an NP-shell, the latter could serve as a local domain ('governing category', cf. Chomsky (1981)) for the pronoun *on*,

allowing its coreference with the matrix subject (for reasons why the subjunctive clause in Russian itself does not serve as a local domain, see Avrutin and Babyonyshev (1994))

Under the assumption that volitional predicates do not select an NP-shell, one predicts that these predicates (if Case-assigning) would not be able to occur with the correlative *to*. This is indeed the case, as shown below:

(34) ?*On xočet to čtoby Ivan vyjgral

He wants   that   Ivan won

'He wants John to win'

If *xotet'* selected an NP shell, the ungrammaticality of (34) is unexpected given that *xotet'* is a Case assigning verb, and can occur with direct object NPs:

(35)   On xočet kofe

'He wants coffee'

However, if *xotet'* does not take an NP-shell as a complement, the contrast between (34) and (35) does not raise a problem.

Of course, the ultimate plausibility of the account of non-extractability out of finite clauses in Slavic in terms of NP-shells depends to a large extent on the independent evidence for the existence of the silent correlative TO. There is indeed strong evidence to this end, which involves the correlation in distribution of the correlative TO in declarative sentences and the so called '*wh*-scope marker' *kak* in *wh*-scope marking interrogatives,

which suggests that both elements employ the same structural position. We present this evidence in Chapter 5, where we consider the *wh*-scope marking strategy in detail.

2.4 Alternatives

Even though the NP-shell approach to non-extractability out of finite clauses in Slavic based on Giejgo's idea may appear somewhat simplistic, it does not suffer from various shortcomings of the alternative accounts proposed in the literature.

Zaenen (1983), working in the framework of Lexical Functional Grammar, introduces a morphological feature [+/- BND] on a complementizer to allow or prohibit creation of long-distance *wh*-dependencies. She suggests that in Russian, finite indicative complementizers are specified for the negative value of the BND feature (disallowing a dependency across them), whereas subjunctive complementizers are specified for the positive value, allowing such a dependency. This accounts for the contrast between unacceptable (7) and acceptable (13) in Russian. However, without independent motivation for the [+/- BND] feature, the analysis remains stipulative. More importantly, Zaenen's account implies that the indicative/subjunctive distinction in Russian is morphosyntactic in nature, since the value of this feature affects a syntactic process, namely, creation of a long-distance *wh*-dependency. That, however, is incorrect: subjunctive mood in Russian is not triggered by syntactic factors, but determined entirely by the lexical semantics of predicates selecting subjunctive clauses. As pointed out above, subjunctives are generally complements of irrealis predicates of various sorts. Thus, the

proposal attributing the possibility of creating long-distance dependencies in Russian to a morphological feature does not seem to be revealing.[7]

Greenberg (1988), adopting the theory of barriers in Chomsky (1986a), accounts for the non-extractability out of finite clauses by assuming that the complementizer čto introducing finite clauses does not support the intermediate step of (successive cyclic) movement, hence movement crosses two bounding nodes (taken to be IP in Russian). Bailyn (1995) (see also Pesetsky (1982)) attributes the non-extractability to Stowell's (1981) Case Resistance Principle (CRP), by virtue of which the finite embedded clauses in Russian obligatorily 'extrapose' from their base generated position of complement of the verb, to escape Case assignment by the verb. In their 'extraposed' position, they become structural adjuncts. Consequently, extraction out of those amounts to a 'CED' violation. This latter view is quite similar in spirit to our 'correlative' analysis. However, Greenberg's and Bailyn's proposals are stipulative: it is not clear how they distinguish between Russian, on the one hand, and languages in which extraction from finite clauses is possible, on the other (cf. (1)). In addition, Bailyn's account is based on the CRP, a construct the plausibility of which has been questioned by many (see, e.g. Bošković (1995) for arguments and criticism).

A different sort of analysis is proposed in Stepanov and Georgopoulos (1997). They attribute the impossibility of extraction to a parameter on phrase structure regulating the order of application of the structure building operations Merge and Move in well-defined cases. In English, they argue, *wh*-extraction out of finite clauses in English takes place

---

[7] The situation is different in languages like Icelandic, Romance or Austronesian in which subjunctivity may in fact be triggered by (morpho-)syntactic factors. We take no stand with regard to the plausibility of Zaenen's proposal for those languages.

after the embedded clause became a part of the phrase marker containing the target of *wh*-movement (matrix interrogative C), as always assumed. In Russian, *wh*-extraction (by Move) within the embedded clause must take place prior to Merging the embedded clause with the matrix one, hence, by hypothesis, cannot ever take place long-distance. This proposal involves complications of the phrase structure component, concerning the way two clausal syntactic objects can be combined, as well as parametrization in structure building, which are not easy to accommodate into the current versions of the theory. In addition, the proposal concerning the Merger after *wh*-extraction within the embedded clause does not fit easily with the strict cyclicity. Also, these authors' approach did not leave much room for explaining either the contrast in grammaticality between extraction of arguments vs. adjuncts, or the distribution of the correlatives *to* and TO, that may precede the finite clause in Russian/Polish. Nevertheless, Stepanov and Georgopoulos's analysis shares with the present one an important idea, which we believe to be on the right track, namely, that the finite complement is assembled as a separate phrase marker, along with the matrix clause, the insight which, as we noted, goes back to Chomsky (1955). Our present account is better than Stepanov and Georgoupolos's, since, as we showed, this idea no longer has to be stipulated bringing in additional complications. Rather, it follows from the minimalist theory of structure building discussed in Chapter 3.

3. NP-shells cross-linguistically

The question that arises at this point is why languages like Russian and Polish employ this, seemingly complicated, form of subordination, whereas in languages like English, on the other hand, clausal subordination seems to be realized by

generating a finite clause, headed by a complementizer C, as a complement of the subordinating predicate. At this point we begin to work towards an account of the relevant differences between the two types of languages.

We observe that the NP-shell form of finite clause complementation employed in Russian and Polish finite complementation is not accidental or language specific. According to Kiss (1987), the NP-shell is also utilized in Hungarian. In Hungarian, embedded complements in declarative sentences can optionally be introduced by an overt pro-form, as shown below (see also Kiss (1987)):

(36)   a. (Azt)   mondtad   [hogy eljönnek   a gyerekek]

             it-acc said-2sg-def.DO   that away-come-3pl the kids-nom

             'You said that the kids would come'

    b. Számítunk rá   hogy eljönnek   a gyerekek

             count-1pl it-al that away-come-3pl the kids-nom

             'We expect that the kids will come'

The matrix verb displays the definite object agreement, suggesting that the object that the verb agrees with has a nominal nature. The pronoun *az*, as described by Kiss, is optional, means 'it' and does not contribute to the meaning of the sentence. This makes the distribution of *az* look very similar to the distribution of Slavic *to* (cf. (17), Section 2.2).

Extraction out of finite complements in Hungarian is blocked if the finite complement is preceded by a clausal pro-form such as *azt*, as Kiss shows:

(37)　　*János pulóvert_i　　szeretné　azt　hogy kössek neki　t_i

John pullover-acc　would-like　it-acc　that　knit-I　for-him

'As for John, it is a pullover that he would like it that I knit for him'

Based on facts like this, Kiss suggests that that the clausal proform in Hungarian forms a constituent (NP) with the embedded finite complement in Hungarian. Hence, extraction out of the finite complement amounts to a Complex NP violation.

At least for some speakers consulted extraction out of a finite complement is degraded (to a various degree from ?? to ?*) even when there is no overt pro-form, unless the matrix verb is *akar* 'want' taking a subjunctive complement. This is, again, in striking similarity with the corresponding Russian facts (cf. Section 2.2)

(38)　　a. ???/?* Kit　　mondott János hogy Peter　meglátogatott?

whom said　John　　that　Peter　visited

'Who did John say that Peter visited?'

b. János mikor_i akarja hogy induljunk t_i?

John　when　wants　that start-we

'When does John want that we start?'

It seems possible, then to extend our NP-shell proposal made on the basis of Slavic: in addition to overt pro-form, Hungarian also has a silent pro-form (e.g. AZ), and finite complements in Hungarian are uniformly NP-shells. In fact, Kiss (1987)

acknowledges such a possibility in a footnote (p. 139, fn. 5). Horvath (1997) also considers overt *az* and silent *pro* on a par, as expletive elements associated with the finite CPs. This possibility is supported, in particular, by the pattern of object agreement observed, for instance, in (36a): even if there is no overt pro-form (e.g. *az*) in a sentence, the definite object agreement obtains, still indicating that the verb agrees with something that has a nominal nature. We will return to the discussion of the Hungarian facts in Chapter 5, where the NP shell hypothesis is confirmed with evidence from *wh*-scope marking.[8]

Müller (1995) and Müller and Sternefeld (1995) explore in detail the NP-shell approach to finite complementation for German. They argue that finite complements in German are headed by a phonologically silent pronominal N. The N may undergo incorporation into the subordinating V, in certain well-defined contexts. In particular, N undergoes abstract incorporation into propositional V when the V selects, in some sense, for a nominal. For these authors, selection for N defines the 'bridgehood' property of the verb. If V does not select for N, no incorporation takes place. In this manner, these authors account for the fact that extraction out of finite clausal complements in German is allowed only in certain well defined contexts, which are

---

[8] Kiss, however, dismisses this possibility, suggesting, instead, that in sentences without overt pro-form, the finite complement is not an NP-shell, but, rather, a CP, as always assumed. This suggestion is made in light of the following grammatical examples:

(i) János két dolgot$_i$ ígért meg Máriának hogy meg tesz t$_i$
    John two things promised Mary-dat that does
    'As for John, it was two things that he promised Mary that he would do'

Kiss argues that these examples, as well as (37), involve long operator movement. Kiss does not discuss *wh*-extraction out of finite clauses as in (38a) in detail. The ungrammaticality of (38a) and the grammaticality of (i) apparently constitutes conflicting evidence, if one assumes that they involve the same movement process. We will restrict ourselves with the cases of *wh*-extraction, and remain agnostic with regard to the nature of examples like (i).

determined by the lexical properties of the subordinating verb (cf. also Svenonius (1994) who argues that selection involves feature checking).

A survey of diachronic literature suggests that finite complementation via NP-shells is employed to a great extent in Early Indo-European languages, and even in the early stages of languages which currently seem to utilize the 'normal' V-CP structure. English is one case at hand. In modern English, extraction out of finite clauses is possible (cf. (1)), which suggests that at least by the time of extraction, there is no NP-shell separating the complement clause and the subordinating verb, which would block extraction. On the other hand, Old English (<11th c.) and Middle English (11-15th c.) had productive syntactic configurations which descriptive grammars of these languages usually place under the rubric 'recapitulation and anticipation' (Mitchell and Robinson (1992), p. 66., Burrow and Turville-Petre (1992), p.55-56). One such configuration involves the use of a pronoun *þæt* introducing a finite clause, as shown below:

(39) þa þæt Offan mæg ærest onfunde, þæt se eorl nolde yrhðo geþolian

     lit. 'Then the kinsman of Offa first learned that thing (the first *þæt*), that the leader

     would not   tolerate slackness'   [Mitchell and Robinson (1992)]

The first *þæt* in (39) is an object of the matrix verb. The second *þæt* seems to play the same role as the complementizer *that* in Modern English. This is parallel to its modern Russian analogues, except for a slight difference in the morphological make-ups of the corresponding elements: *to* functions as the object of the matrix verb, and *cto* is

standardly taken to be a complementizer. In fact, the description of *þæt* by historical grammarians suggests that it bears striking similarities with modern Russian *to*, listed in (17). In particular, similarly to Russian *to*, the pronoun *þæt* a) literally means 'that' (or 'that thing'); b) functions as a demonstrative referring to the clause that follows it; c) has no detectable effect on the meaning of the sentence. The latter property clearly sets *þæt* apart from the Modern English *it* in contexts like (18b), which, as discussed in Section 2.2, does affect the meaning of the sentence.

Anticipating the discussion in the next section, we believe it is appropriate at this point to quote the following note of Mitchell and Robinson (1992) which seems to us interesting in at least two respects. The first part offers a quite insightful historical perspective on the 'NP-shell' structure in Modern English, which is in fact taken up in the modern philosophy of language (Davidson (1969), see the next section for details). The second part of the note may serve anecdotally as illustrating the point that a historical view always needs to be checked against the synchronic perspective (in this case, languages like modern Russian).

(40) 'It is possible that in the sentence 'He said that he was ill', 'that' was originally a demonstrative - 'He said that: he was ill' - which gradually became a part of the noun clause. If so, the introduction of the second *þæt*... illustrates clearly the difficulty our ancestors seem to have had in collecting and expressing complicated thoughts'. (p. 67)

In addition to *þæt*, other elements can be used in Old English to 'introduce' a finite complement, in particular, *þæs* and *hit*. The following example illustrates the point with *þæs*, which seems no different from *þæt* in this context:

(41) <u>þæs</u> ic gewilnige and gewysce mid mode, þæt ic ana ne belife æfter minum leofum

þegnum

lit. 'That thing I desire and wish in my mind, that I should not remain alone after my

beloved thanes'

(42) shows that these elements can co-occur in sentences involving several clausal embeddings:

(42) <u>þæt</u> is micel wundor þæt <u>hit</u> ece God æfre wolde þeoden þolian, þæt wurde þegn swa

monig forlædd be þam lygenum...

lit. 'That is a great wonder that eternal God the Lord would ever permit it, that so

many a   thane should be deceived by those lies'

It seems plausible, then, that finite complements introduced by or *þæt, þæs*, or *hit* as in (40) and (41) have the same structure as Russian *to*-complements, that involving an NP-shell.

The correlative complementation of this kind may in fact be an Early Indo-European feature. Holland (1984) cites an example of Homeric Greek which, arguably, involves a similar phenomenon:

(43) leússete gàr <u>tò</u> pántes, hó moi géras érkhetai állēi

 you see for this all that my prize goes elsewhere

 'For you all see this, that my prize goes elsewhere'

Holland gives similar examples for Vedic Sanskrit, Avestan, early Latin and Hittite.[9] Interestingly, Holland argues, independently, that markers like -*tò* in (43) are semantically empty artifacts in Early Indo-European, and their meaning, if any, is recoverable from the meaning of the subordinated clause of which they are pro-forms. This again, suggests that the distribution of *tò* and similar pronouns can be described along the lines of (17), similarly to Russian/Polish *to*. It follows, then, the NP-shell structure may have historical roots, and it is likely that its current realization in Russian and Polish simply preserves an Indo-European feature.

4. Davidson's conjecture

The NP-shell hypothesis also has some philosophical justification. Davidson (1969) considers sentences like the following:

(44) Galileo said that the earth is round

---

[9] In fact, Holland demonstrates that the phenomenon of using 'subordinating' pronouns like *tò* in (43) is much more widespread across constructions, including various sorts of adverbial modification and relative clauses, among other types. Similarly, Old and Middle English seem to use the 'recapitulation and anticipation' strategy quite productively, for different types of subordinated clauses.

(44) represents a class of sentences involving a finite complement of an 'indirect speech' predicate. Investigating the logical properties of these sentences, Davidson points out that the truth value of the entire sentence does not depend on the truth of the utterance following *that*. In other words, the semantics (logical form) of the sentence is computed independently of the logical form of the 'embedded' clause. In this respect, Davidson proposes that these complements are not clausal in nature, despite the appearance. Rather, the complement of these verbs is a demonstrative pronoun *that* to which the sentence that follows it is related cataphorically. According to Davidson, (44) is interpreted as (45), under what is sometimes referred to as a 'paratactic' analysis:[10]

(45)   Galileo said that. The earth is round

Adopting and elaborating on Davidson's proposal, LePore and Loewer (1989) extend the class of sentences analyzable along the lines of (45) to other 'propositional attitude' verbs, including *believe, expect* etc. Thus under Davidson's conjecture, (44) is really two independent sentences. However, even though Davidson's argument about a certain degree of autonomy of the embedded clause is sound, it most likely cannot be the case that the two clauses never constitute a unit. After all, there exists well known types of logical dependencies between the element in the matrix and

---

[10] It is beyond the scope of this chapter to survey the large amount of productive philosophical debate generated by Davidson's conjecture. See, e.g., Hand (1991), LePore and Loewer (1989), Rumfitt (1993), Segal and Speas (1986) and references cited therein. Rumfitt (1993) also suggests an implementation of Davidson's conjecture at the LF interface.

embedded clause, which could only be accounted for if the two constitute a unit at a certain point. One such dependency is the possibility of variable binding into the embedded clause:

(46)   Everyone$_i$ says that he$_i$ admires Picasso

Under the standard version of the syntax-semantics map, when the composed syntactic structure is submitted to the semantics interpretation at the level of LF, it follows that by LF the two must constitute a unit.

Torrego and Uriagereka (1993), assuming that Davidson's conjecture extends to all propositional verbs, suggest a syntactic interpretation of Davidson's paratactic analysis, which is essentially a version of the NP-shell proposal. According to these authors, the embedded clause is adjoined to a (possibly null) NP, which is the real complement of an epistemic verb. This is contrasted with subjunctive complements of volitional verbs which, as Torrego and Uriagereka argue, are true complements of the verb. The following are schematic representations of both types of complements, according to these authors:

(47)   [V$_{epistemic}$ [[(possibly null) NP]$_i$ [embedded clause]$_i$]]

(48)   [V$_{volitional}$ [embedded clause]]

We believe Torrego and Uriagereka's version of Davidson's conjecture can be rather naturally represented in the system that we have developed so far. Consider a version of (44), involving the correlative *to*, in Russian:

(49)   Galilej skazal to cto zemlja kruglaja

Galileo said   that earth round

'Galileo said that the earth is round'

Under our approach, *cto zemlja kruglaja* is a structural adjunct. Recall that in Chapter 3 we argued that structural adjuncts enter the structure postcyclically. In other words, there exists the derivational point right before the adjunct Merges with the matrix clause:

(50)   a. Galilej skazal to

Galileo said *to*

b. Cto zemlja kruglaja

That earth round

The embedded clause in instances like (44) will be adjoined last in overt syntax, before submitting the structure to LF. Note that the situation in (50) is strikingly reminiscent of Davidson's analysis in (45). It remains an open question, of course, what the connection is between the apparent 'truth-value' independence of the embedded clause observed by Davidson, and its late insertion into the structure.

It should also be noted that, even though we have so far collected evidence that the NP-shell complementation is entertained in Slavic, but not in English, Davidson's conjecture (elaborated along the lines above), as well as Chomsky's original analysis of finite complements, was illustrated precisely with English, the language in which syntactic *wh*-extraction facts offer no evidence for the NP-shell (at least, at the relevant point in the syntactic derivation). We interpret this state of affairs as stronger support for the NP-shell hypothesis in Slavic, but until Chapter 6 we remain agnostic with regard to the manner in which the syntax of 'paratactic' analysis is realized in English.

5. Conclusion

Having established the fact that long-distance extraction in regular questions (formed by regular *wh*-movement) is blocked in Russian and Polish by virtue of entertaining the NP-shell complementation type, the question remains as to whether the relevant difference between Slavic and languages like English lies indeed in different types of finite complementation, or, rather, both types of languages realize their finite complements as NP-shells, which by default precludes extraction. If the latter, then the next question, of course, is how good long-distance *wh*-extraction sentences are derived in languages like English.. Keeping these questions in the back of our mind, we now turn to discussing the strategy of *wh*-scope marking, which in most, if not all, cases serves a means alternative to asking a 'long-distance' question in Slavic and other languages. The relevance of this strategy becomes apparent when we find that it exploits exactly the type of complementation we are considering for Slavic, namely, the 'shell' complementation.

The desideratum for a fruitful inquiry at this point is to explore the intimate connection between the structure of finite complementation (answering the question above) and the structure of *wh*-scope marking questions. It is this connection, we claim, that can shed light onto the most important question of this part of our study: deriving the locality of *wh*-movement in finite clause domains.

CHAPTER 5

Wh-Scope Marking and the 'Shell' Complementation

1. Introduction

The phenomenon of *wh*-scope marking has recently become an issue lively and fruitfully debated in the syntactic and semantic literature (cf. Lutz and Müller (1996) for a general discussion), particularly after the seminal work by McDaniel (1989) (who has also coined the term 'partial *wh*-movement' for this type of construction). It has been found to be a productive question-forming strategy in a wide variety of languages including German, Romani, Hindi, Iraqi Arabic, Malay and Hungarian (see McDaniel (1989), Dayal (1994), Wahba (1991), Horvath (1997), among others). The following are representative examples of *wh*-scope marking questions from German and Hindi.

(1) **Was** glaubst Du, **wen** sie liebt?   [German]

     what believe you who she loves

     'Who do you believe that she loves?'

(2) siitaa-ne **kyaa** socaa   ki   ravii-ne **kis-ko** dekhaa   [Hindi]

     Sita-erg what thought that Ravi-erg who   saw

     'What did Sita think, who Ravi saw?'

A number of influential analyses of the phenomenon have been proposed (cf. Lutz and Müller (1996) for a general discussion).

221

What has not, in our view, been sufficiently acknowledged in the previous literature is the existence of a fundamental connection between the availability of *wh*-scope marking, long-distance *wh*-movement, and the structure of finite complementation. This connection will be explored in this and the next chapter. Recall that we argued, in Chapter 4, that finite complements (CPs) are attached in the course of the derivation to a sentential pro-form in the matrix clause, forming a so called NP-shell. In this chapter, we discuss the strategy of *wh*-scope marking in detail and argue that it provides further evidence for the NP-shell approach. In Chapter 6, we take the next step and argue that both the *wh*-scope marking and the long-distance strategies utilize a version of the 'NP-shell' structure, and complement each other in a way that can be formally defined.

In Sections 2-7, we investigate the *wh*-scope marking strategy in Russian and Polish (cf. also Stepanov (2000b)). We choose these languages for two main reasons. First, unlike the languages mentioned above, the phenomenon of *wh*-scope marking in Slavic has not been systematically investigated in the previous literature. Second, because of their syntactic and semantic properties, the Slavic data can be shown to provide important grounds for testing several competing syntactic and semantic theories of *wh*-scope marking, and shed new light on the issue of what the correct syntactic and logical structure of *wh*-scope marking questions should be. In particular, we will argue that a version of the Indirect Dependency Approach proposed in Dayal (1994), (1996b) and adopted and modified in this chapter, is the correct analysis of the *wh*-scope marking in Russian and Polish. Section 8 addresses the *wh*-scope marking strategy in other languages, utilizing the empirical results of previous studies, and argues that our version

of the Indirect Dependency straightforwardly accounts for most properties of *wh*-scope marking questions in those languages.

2. The characteristics of Slavic *wh*-scope marking

Consider the following examples:

(3)   **Jak** myślisz,   **kogo** Janek lubi?   [Pol]

how think-you whom John loves

'What do you think, who does John love?'

(4)   **Kak** vy dumaete, **kogo** ljubit Ivan?   [Rus]

how you think   whom loves John

'What do you think, who does John love?'

First, we observe that questions of this type elicit an answer that does not involve supplying the value to the *wh*-phrase *jak/kak* 'how' in the higher clause. Rather, a felicitous answer to (3) and (4) involves supplying the value for the *wh*-phrase *kogo* 'whom' in the embedded clause, as the following shows:

(5)   Myśle   że Janek lubi Marię   [Pol]

Think-I that John loves Mary

(6)   Ja dumaju čto Ivan ljubit Mariju   [Rus]

        I think   that John loves Mary

        'I think that John loves Mary'

In this connection, note that (5) and (6) represent the same kind of answer as would be given to a long-distance question in languages which allow long-distance extraction, cf:

(7)   Who do you think (that) John loves?

In effect, this feature is characteristic of the class of questions which Riemsdijk (1982) termed *wh*-scope marking (WSM) questions, in the sense that the first *wh*-phrase *how* 'marks' the scope of the second *wh*-phrase in overt syntax. In the following discussion we refer to the first *wh*-phrase as the *wh*-scope marker, simply as a mnemonic term.

Another distinctive feature of WSM questions is that despite the superficial appearance of them being composed of a sequence of unrelated *wh*-questions, usually reflected in the English translation, they actually represent a single sentential unit. That this is also true for Slavic (3) and (4) can be demonstrated by the bound variable test (cf. Dayal (1996b), Horvath (1997)). Consider the following sentences:

(8)   **Jak** [kazdy student]$_i$ myśli, **gdzie** go$_i$ posla?   [Pol]

        How every student   thinks where him send-they

(9)   **Kak** sčitaet [každyj iz   studentov]$_i$, **kuda** ego$_i$ mogut otpravit'?   [Rus]

How thinks   every from students   where him can   send-they

'Where does every student think that they can send him'

In (8) and (9), the bound variable reading obtains, which suggests that quantifier in the matrix clause c-commands into the lower clause. The presence of c-command indicates that (8) and (9) represent a single phrase marker and should be treated by means of sentential grammar.[1]

Of course, in languages which do not have *wh*-scope marking it is still possible to construct a discourse mimicking the word order in (3) and (4). Crucially, though, in those languages the bound variable reading does not obtain. Here is an illustration from English (see fn. 1):

(10)   What does [every student]$_i$ think? Where will they send him $*_{i/j}$?

Other properties of Slavic *wh*-scope marking questions are common among *wh*-scope marking constructions found in other well-studied languages like German or Hindi. This is true, in particular, of the following, non-exhaustive, list of properties.

---

[1] The bound variable test, of course, does not rule out the possibility that (3) and (4) may also have a second source, namely, as two distinct questions. This possibility seems to be universal. In languages like German, the two possibilities are easy to tease apart because each is reflected in the word order (no V2 in the embedded clause in WSM questions, yes in question sequences), but in Slavic, there is no tangible reflex, except, perhaps, subtle intonational differences. Thus while the WSM option tends to be realized in Polish and Russian with a single intonational contour, the sequential option induces two more or less separate contours. In (8) and (9), inducing two intonational contours appears to result in blocking the bound variable reading, although the relevant judgments are somewhat volatile. Here we put aside the sequential option, which clearly belongs outside the sentence syntax (see Dayal (1996b) for a unified (discourse) semantic treatment of question sequences and WSM questions).

I. Any *wh*-phrase can appear in the embedded clause of a *wh*-scope marking question:

(11)   **Jak** myślisz,   **gdzie/kiedy/dlaczego/jak** Maria tańczyła?   [Pol]

How think-you where/when/why/how   Mary danced

(12)   **Kak** vy   dumaete, **gde/kogda/zachem/kak** tancevala Marija?   [Rus]

How you think   where/when/why/how   danced Mary

'What do you think, where/when/why/how did Mary dance?'

II. The embedded clause can have more than one *wh*-phrase (in other words, the lower clause can be a multiple *wh*-question):

(13)   **Jak** myślisz,   **kto co**   czytają?   [Pol]

How think-you who what read

(14)   **Kak** vy   dumaete, **kto čto**   čitaet?   [Rus]

How you think   who what read

'Who do you think read what?'

III. The *wh*-scope marker and the associated *wh*-word must not be in the same clause (the requirement Müller (1996) termed 'antilocality'):

(15)   *****Jak co**   studenci przeczytali?   [Pol]

How what students read

(16)   *****Kak čto** vy čitaete?   [Rus]

How what you read

intended: 'What are you reading?'

IV. The predicate in the first clause must not (s-)select a question (that is, not be *wonder*-type), despite the fact that the complement itself appears to be:[2]

(17)   *__Jak__ sie zastanawiales, **co**   studenci przeczytali?   [Pol]

how wonder-you,   what students read

'??What do you wonder that the students read?'

(18)   *__Kak__ vy sprosili, **čto** studenty čitajut?   [Rus]

How you asked what students read

'What did you ask that the students read?'

V. *Wh*-scope marking questions in which clausal negation separates the *wh*-scope marker and the 'real' *wh*-phrase are ungrammatical (cf. Rizzi (1992)):

(19)   *__Jak__ nie myślisz,   **co**   studenci czytają?   [Pol]

how not think-2pl what students read

---

[2] Horvath (1997) and Dayal (1996a) notice, for Hungarian and Hindi, respectively, that it is possible to have a *wh*-scope marking question with verbs that do select a question, if the embedded clause contains two *wh*-phrases, as in (i) (Dayal (1996a), p.85):
(i)   jaun kyaa puuchh rahaa thaa ki   merii kis-se baat karegii yaa nahiiN
       John what ask   -PROG-P   that Mary who talk-do   or   not
       'What was John asking, whether Mary will talk to who or not?'
In Slavic, however, questions like (i) are unacceptable, as the following example illustrates:
(ii)   *Kak vy sprosili čto budut li studenty čitat' ?   [Rus]
       How you asked what will Q the students read
       'What did you ask whether the students will read?'

(20)   *Kak vy ne   dumaete, čto čitajut studenty?   [Rus]

    how you not think   what read   students

    'What don't you think that the students read?'

VI. The *wh*-scope marker cannot be separated from the true *wh*-phrase by more than one clause:

(21) *Jak Janek   sądzi że   Piotrek myśli, co   studenci przeczytali?   [Pol]

    How Janek judge that Piotr   thinks what students read

    'What did John say that Peter think that the students read?'

(22) *Kak vy dumaete, čto Ivan sčitaet, čto pročitali studenty?   [Rus]

    How you think   that Ivan believes what read   students

    'What do you think that John believes that the students read?'

VII. Slavic allows the second clause of a *wh*-scope marking interrogative to be a yes/no question rather than a *wh*-question (similarly to Hindi, but not German):

(23) **Jak** myślisz,   czy   Piotrek przyszedl?   [Pol]

    How think-pl. whether Piotrek came

    'What do you think about whether Peter came?'

(24) **Kak** vy sčitaete, budet li zavtra   dožd'?   [Rus]

    How you think, will Q tomorrow rain

    'What do you think about whether it will rain tomorrow?'

VIII. Iterating the clause containing the *wh*-scope marker generally leads to degradation (unlike in German or Hindi):

(25) ??**Jak** Janek   sądzi, **jak** Piotrek myśli,   **co**   studenci przeczytali?   [Pol]

How Janek judge how Piotr   thinks   what students read

'What did John say that Peter think that the students read?'

(26) ??**Kak** vy dumaete, **kak** Ivan sčitaet,   **čto**   pročitali studenty?   [Rus]

How you think   how Ivan believes what read   students

'What do you think that John believes that the students read?'

Note that the questions listed above involve matrix verbs that (roughly) mean *think* or *believe*. In point of fact, very few verbs can license *wh*-scope marking questions in Polish and Russian beside *think*, and that is why *wh*-scope marking is a very limited phenomenon in these languages, as opposed to, say, German or Hindi.[3] For instance, if we replace *think* in (3) and (4) with a verb meaning *say* or *understand*, the sentences become much worse in the relevant reading - as true questions:

(27) *****Jak** Janek powiedział, **kogo** Piotrek lubi?   [Pol]

how John said   whom Peter   loves

(28) *****Kak** Ivan skazal, **kogo** ljubit Petr?   [Rus]

---

[3] In German, *wh*-scope marking is possible with "bridge" verbs like *think* or *say*. In Hindi, effectively any verb not s-selecting questions (non *wonder*-type), including factives, can participate in *wh*-scope marking See Lutz and Müller (1996) for more discussion.

how John said   whom loves Peter

'Who did John say Peter loves?'

In Section 4 we will determine precisely the class of verbs that can participate in *wh*-scope marking in Slavic and show that verbs like *say* do not belong to this class, which is responsible for the ungrammaticality of (27) and (28).

The relevant interpretation of (3) and (4) may be salvaged by replacing the initial *wh*-phrase *how* with *what*. However, when pronounced, the relevant sentences necessarily involve two separate intonational contours, suggesting that what we are dealing with is two separate questions, rather than a single *wh*-scope marking sentence (cf. (10)):

(29) **Co**   Janek powiedział? **Kogo** Piotrek lubi?   [Pol]

   what John said   whom Peter loves

(30) **Čto** skazal Ivan? **Kogo** ljubit Petr?   [Rus]

   what said John   whom loves Peter

   'What do you think? Who does John love?'

That indeed two different sentences are involved is supported by the bound variable test: as expected, in such questions the bound variable reading is not available.

(31) *Co   [kazdy student]ᵢ powiedzial? **Gdzie** goᵢ   posla?   [Pol]

   What every student said   where him send-they

(32) *Čto skazal [každyj iz studentov]ᵢ? **Kuda** egoᵢ otpravjat?   [Rus]

  What said every from students   where him send-they

  'Where does every student think that they send him'

To conclude the paradigm, *what* cannot be used with verbs like *think* in *wh*-scope

marking questions (although it can in sequential questions):

(33) *Co   myślisz,   **kogo** Janek lubi?   [Pol]

  What think-you whom John loves

(34) ?*Čto vy dumaete, **kogo** ljubit Ivan?   [Rus]

  What you think   whom loves John

  'Who do you think John loves?'

The above examples show that the *wh*-scope marker in Slavic is necessarily a *wh*-phrase literally meaning *how*. This contrasts with German and other well studied *wh*-scope marking languages in which the *wh*-scope marker is usually a *wh*-phrase meaning *what* (see, e.g., Fanselow (1997) for a cross-linguistic overview of *wh*-scope marking).

We now turn to discussing one major approach to *wh*-scope marking proposed in the literature, the Direct Dependency Approach. It will be demonstrated that this approach cannot be the right analysis of Slavic *wh*-scope marking.

3. The Direct Dependency Approach

The Direct Dependency Approach (DDA) goes back to Riemsdijk (1982) and is spelled out in detail in McDaniel (1989) on the basis of German and Romani. The fundamental assumption of the DDA is that *wh*-scope marking questions such as German (35a) receive an interpretation similar to the one of the corresponding long-distance questions (cf. (35b)):

(35) a. **Was** glaubst du **wo**   Maria getanzt hatte?

What think you where Maria danced has

'What do you think, where did Maria dance?'

b. **Wo** glaubst du dass Maria getanzt hatte?

What think you that Maria danced has

'Where do you think that Maria danced?'

As Riemsdijk (1982) and others point out, in the case of several embeddings the *wh*-scope marker can appear not only clause-initially but in every intermediate clause that dominates the one containing the true *wh*-phrase in German, as illustrated below:

(36) **Was** glaubst du, **was** Peter meint, **mit wem** Maria gesprochen hat?

What believe you what Peter thinks with whom Mary talked   has

'Who do you believe Peter thinks Mary talked with?'

Given the alleged similarity with long-distance questions, the proponents of the DDA assume that the structure of a *wh*-scope marking question is in the relevant respect similar. Specifically, they assume that the embedded clause is headed by a [-*wh*] C; that is, the embedded clause is not a question. Thus the lower ('true') *wh*-phrase is located in the specifier of [-*wh*] CP in the overt part of the derivation. Under the DDA, the *wh*-scope marker plays the role of an expletive-like element which has the formal feature [+*wh*] but does not contribute to the meaning of the sentence. The *wh*-scope marker is base-generated in the matrix, [+*wh*] Comp (note, in passing, that the matrix Comp is the only [+*wh*] Comp in a *wh*-scope marking question.)

McDaniel (1989) proposes an account of *wh*-scope marking in terms of '*wh*-chains'. In particular, according to McDaniel, the sequence of scope marker(s), the true *wh*-phrase and its trace form a '*wh*-chain' at S-structure, which is subject to locality restrictions.[4] A *wh*-chain parallels expletive-associate A-chains in the sense of Chomsky (1986b) (e.g. a chain <*there, a man*> in *there is a man in the garden*). By analogy with expletive-associate chains, some researchers claim that at LF the true *wh*-phrase raises to the higher clause and 'replaces' (in the relevant sense) the *wh*-scope marker, in accord with the principle of Full Interpretation. Thus at LF the true *wh*-phrase has scope over the whole sentence, capturing the alleged interpretational similarity with the correspondent long-distance examples.

---

[4] McDaniel's formal definition of a *wh*-chain is the following:
(i)   A chain $C = (a_1, a_2, ..., a_n)$ is a <u>*wh*-chain</u> iff:
        a. $\forall a_i$, $1 \leq i < n$, a_i locally A'-binds a_{i+1}
        b. $\forall a_i$, $1 \leq i < n$, a_i is a *wh*-element
        c. a_n is a variable in IP internal position, and
        d. for any scope marker a_i, $1 \leq i < n$, $(a_1, a_2, ..., a_n)$ contains a true *wh*-phrase.

Let us see how the DDA fares with regard to Slavic (3) and (4), repeated here as (37) and (38):

(37) **Jak** myślisz,   **kogo** Janek lubi?   [Pol]

how think-you whom John loves

(38) **Kak** vy dumaete, **kogo** ljubit Ivan?   [Rus]

how you think   whom loves John

'Who do you think John loves?'

Suppose, in accordance with the DDA, that the *wh*-scope marker *jak/kak* in Polish and Russian is an expletive-like element that only has the (syntactic) *wh*-feature. *Kogo* 'whom' would then have to raise at LF and replace *jak/kak* obtaining the wide scope over the sentence.

One empirical argument against adopting the DDA for Slavic comes from the fact that in both Polish and Russian, the second clause of a *wh*-scope marking interrogative may be a yes/no question rather than a *wh*-question, as shown in (23) and (24), repeated as (39) and (40):

(39)   **Jak** myślisz,   czy   Piotrek przyszedl?   [Pol]

How think-pl.   whether Piotrek came

What do you think, did Peter come?

(40)   **Kak** vy sčitaete, budet li zavtra   dožd'?   [Rus]

How you think, will Q tomorrow rain

'What do you think, will it rain tomorrow?'

In (39) and (40) there is no element that can possibly raise at LF and replace the 'expletive' scope marker. One might argue that *czy* 'whether' in Polish undergoes LF raising. However, Beck and Berman (1996) on the basis of Hindi (in which questions like (39) and (40) are also possible) argue against raising of *whether* in *wh*-scope marking questions altogether, demonstrating that such raising always leads to a non-existent interpretation.[5] Under the 'expletive replacement' analysis of the DDA, the non-replacement of the *wh*-scope marker (recall that it has an expletive status for McDaniel) must lead to a violation of Chomsky (1986b) principle of Full Interpretation, which does not allow expletive elements to enter LF. Thus, (39) and (40) cannot be accounted for within the classic DDA analysis.[6]

The same problem arises for a version of the DDA developed by Cheng (1997). Adopting the minimalist framework of Chomsky (1995c), Cheng suggests that the *wh*-scope marker is simply a morphological spell-out of the *wh*-feature of the 'true' *wh*-

---

[5] Beck and Berman)'s type of argument goes roughly as follows. After the raising of *whether*, the denotation of Polish (39) would be schematically represented as (i):

(i) ||WHETHER|| (|| you think (that) Peter came||)

The compositional interpretation assigned to (i) amounts to the question "Do you think that Peter came?" and has the following set as its answers:

(ii) {I think Peter came, I don't think Peter came}

However, the answers in (ii) are not appropriate for the original question (39), rather, the set in (iii) is:

(iii) {I think Peter came, I think Peter didn't come}

Raising of *whether* then renders (39) the wrong interpretation (ii) and at the same time deprives it of its correct interpretation (iii). (Notice, incidentally, that the second members of the sets in (ii) and (iii) are hard to distinguish in English because of the 'negative transportation' effect. In Slavic, however, these have clearly distinct interpretations.)

[6] The same is true for Hindi counterpart of (39) and (40). In the literature on *wh*-scope marking, examples like these constitute strong evidence against adopting the DDA for languages like Hindi.

phrase. That means, for example, that in examples (39) and (40) *jak/kak* is a morphological spell-out of the *wh*-feature of *kogo*. Cheng maintains that the possibility of this kind of 'dissociation' of the *wh*-feature from the rest of the *wh*-phrase is due to the morphological makeup of *wh*-phrases in certain languages. One can immediately see that Cheng's proposal cannot account for the grammatical (39) and (40) either, since these examples involve no *wh*-phrase at all which would possibly contribute the *wh*-feature to be spelled out as *jak/kak*. Any DDA-based account that capitalizes on the 'association' between the *wh*-scope marker and the 'true' *wh*-phrase will face this problem with respect to Slavic, as well as other languages in which questions like (39) and (40) are possible.

Another potential argument against adopting the DDA for Slavic concerns the issue of raising of the true *wh*-phrase, e.g., *kogo* 'whom' at LF in (39) and (40) above. Recall that under the DDA the true *wh*-phrase has to raise to replace the *wh*-scope marker at LF. The problem arises if one adopts the minimalist framework of Chomsky (1995c)) in which all movement is essentially feature-driven. Bošković (1998a) shows that LF movement, which involves pure feature movement in this framework, is, in fact, more restricted than overt movement, the opposite of the standard claim (cf. Huang (1982)).[7] With this in mind, recall that under the DDA *wh*-scope marking questions and the

---

[7] One piece of evidence discussed in Bošković (1998a) comes from the domain of long-distance *wh*-questions in French. French is an overt *wh*-movement language. However, it is also possible to leave a *wh*-phrase in-situ in short-distance questions, but not in long-distance questions:
(i)   a. Tu   as   vu   qui?
          you have seen whom
          'Who did you see?'
       b. Qui   as-tu   vu?
(ii)   a. *Jean et   Pierre croient que Marie a   vu   qui?
          Jean and Pierre believe that Marie has seen whom
          'Whom do Jean and Pierre believe that Marie saw?'
       b. Qui Jean et Pierre croient-ils que Marie a vu?

correspondent long-distance questions are simply the alternatives (cf. (35) above). The *wh*-phrase can potentially raise either overtly or covertly. The overt raising of the true *wh*-phrase would result in the usual long-distance question. However, as discussed in the preceding chapter, long-distance extraction out of a finite clause results in degradation in Russian and Polish.

(41) ?***Kogo** myślisz że Janek lubi?   [Pol]

     Who think-you that John loves

(42) ?***Kogo** vy dymaete čto Ivan ljubit?   [Rus]

     Who you think   that John loves

     'Who do you think John loves?'

If Bošković's (1998a) conjecture that LF movement is more restricted than overt movement is correct, then, given the already degraded status of (41) and (42), it is very unlikely that *kogo* moves at LF in (39) and (40).

The above considerations suggest that the DDA is unlikely to be the right approach to the Slavic *wh*-scope marking.[8]

A DDA-type analysis faces a more general problem. Recall that under the DDA the embedded clause in a *wh*-scope marking question is headed by a [-*wh*] Comp; it is not a

---

[8] This conclusion, of course, does not undermine the possibility of the DDA for languages like German which does not present the same kind of empirical problems as Slavic, although it speaks against adopting the DDA as a uniform account of *wh*-scope marking. See Beck and Berman (1996) for more discussion on this issue. Several researchers, notably Herburger (1994) and Dayal (1996a) argue against the DDA on the whole, in particular, against its assumption that a *wh*-scope marking question receives the identical LF interpretation with the correspondent long-distance question. See above works for the arguments.

question. Thus the lower ('true') *wh*-phrase is in the Specifier of CP marked [*-wh*] in overt part of the derivation. It is unclear, under this approach, why the *wh*-phrase moves into this position in languages with overt *wh*-movement. This problem becomes even more acute in the minimalist framework where all movement is subject to the economy principle of Last Resort: it does not happen unless triggered by a formal inadequacy of the target of movement (under the conception of Movement as Attract). But there is no apparent inadequacy that would be satisfied by moving a *wh*-phrase into this position. In point of fact, the question is reminiscent of the one posed by the successive cyclic property of *wh*-movement, for which, too, no obvious explanation is available (see Chapter 6).

The next section argues for a syntactic structure of Slavic *wh*-scope marking questions along the lines of the 'sentential expletive' approach proposed in various forms for Hindi (Mahajan (1990), (1996)) and Hungarian (Horvath (1997)). Our structural proposal concerning *wh*-scope marking in Russian and Polish provides an independent confirmation for the NP-shell structure discussed in Chapter 3.

4. The Syntactic Structure of Slavic *wh*-scope marking questions

Dayal (1994), (1996a), Fanselow and Mahajan (1996), Mahajan (1990), (1996), Srivastav (1991), and, more recently, Horvath (1997) suggest a structural parallel between the *wh*-scope marker and various 'sentential expletives' in declarative sentences. In particular, Dayal and Mahajan consider the Hindi scope marker *kyaa* 'what' on a par with the sentential expletive *yeh* 'it', based on the fact that either of those may occur in the same position - object of the matrix verb, as illustrated by the following examples:

(43) a. siitaa-ne **kyaa** socaa   ki   ravii-ne **kis-ko** dekhaa

   Sita-erg what thought that Ravi-erg who   saw

   'What did Sita think, who Ravi saw?'

b. siitaa-ne yeh jaantaa hai ki   ravii-ne **kis-ko** dekhaa

   Sita-erg it   know   that Ravi-erg who   saw

   'Sita knows (it) who Ravi saw'

Importantly, Mahajan suggests that at 'D-structure', the *wh*-scope marker *kyaa* forms a constituent with the embedded finite clause. (Similar considerations apply in the case of the sentential expletive *yeh*). A very similar proposal in advanced in Herburger (1994).

Following these authors, we propose that in Russian and Polish (37) and (38), too, the *wh*-scope marker *jak/kak* forms a constituent with the second clause at some point in the structure building, as shown below:

(44) a. myślisz   [xP **jak** [cP **kogo** Janek lubi]]   [Pol]

   b. vy dumaete [xP **kak** [cP **kogo**   Ivan ljubit]]   [Rus]

      you think   how   whom John loves

We further suggest that the *wh*-scope marker *jak/kak* is a head that takes the finite clause as a complement. (The reasons for taking *jak/kak* as a head in *wh*-scope making questions is discussed below.) The linear order in Slavic *wh*-scope marking questions is then derived by overtly moving *jak/kak* to the matrix Comp:

(45) [$_{C'}$ *jak/kak*$_i$ [$_{IP}$ ... [$_{XP}$ t$_i$ [$_{CP}$ ...]]]]
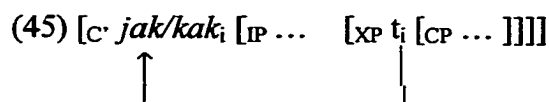
In what follows, we discuss the motivation behind the proposed syntactic structure.

Note that the structure in (44), in which the *wh*-scope marker heads a projection which 'encapsulates' a finite clause, mirrors the structure of finite complementation in Slavic in the form of NP-shells, along the lines of Chapter 3. Let us call this projection 'XP-shell', circumventing the issue of exactly what category the *wh*-scope marker *jak/kak* belongs to. The difference between the XP-shell and NP-shell is that in the former, the *wh*-scope marker takes the CP as a complement, as noted above, whereas in an NP-shell, the finite clause is attached to the N pro-form as an adjunct (cf. Chapter 4). This structural difference can possibly be attributed to a relevant difference between the 'correlative' pro-form, in the NP-shell, and the *wh*-scope marker, in the XP shell, which may be stated in selectional terms: the *wh*-scope marker s-selects a question, whereas the 'correlative' pro-form has no selectional properties.

Recall now, as the discussion in Chapter 4 makes clear, there is more than one correlative in Slavic, namely, overt *to* and silent TO, as exemplified below:

(46)   a. Petr utverždal [$_{NP}$ to/TO čto Ivan ljubit Mašu]

         Petr claimed   that Ivan loves Masa

         'Peter claimed   that John loves Mary'

b. Petr dumal   [$_{NP}$ TO/*to [$_{CP}$ čto   Ivan ljubit Mašu]]

Petr believed   that Ivan loves Maša

The proposed structural similarity between NP-shells and XP-shells thus suggests that there must be common aspects of formal distribution of the *wh*-scope marker and at least one of these correlatives. Such commonality indeed exists between *jak/kak* and silent TO, and can be captured by Case theory. Recall that the distribution of TO in finite complements is determined by its Case properties; namely, that it can be drawn from the lexicon either with or without the Case feature. If it is drawn with the Case feature, Case theory predicts that it can occur with the class of matrix predicates that are necessarily Case-assigners; those include verbs like *utveržžat'* or *znat'* (cf. (46a)). If TO is drawn without the Case feature, it can occur with the class of matrix predicates that are not Case-assigners; those include verbs like *dumat'* (cf. (46b)). Now, it is natural to assume that *jak/kak* does not bear the Case feature. This is suggested by the absence of any Case-related morphological reflex in this element which would be expected otherwise in morphologically 'rich' languages like Polish or Russian. Consequently, the XP-shell headed by *jak/kak* in (44) does not have the Case feature. If that much is on the right track, this establishes a common property of the XP in (44) and the NP headed by the correlative TO in (46b): both constituents are Caseless. On the basis of this common property, we suggest that the Slavic scope marker *jak/kak* is an (overt) *wh*-counterpart of the Caseless TO in *wh*-scope marking questions.

This proposal makes an important prediction concerning the distribution of *wh*-scope marking questions in Slavic. We have seen above that TO-complements are selected only

by *think*-type verbs, which cannot have the Case feature, but not *know*-type verbs that have it. Since the complements headed by the *wh*-scope marker *jak/kak* are also Caseless, the prediction is that these complements will be selected by exactly the same type of verbs, namely, *think*-type verbs, but not *know*-type verbs.

The prediction is borne out. The following shows that *wh*-scope marking questions in both Polish and Russian are acceptable precisely with *think*-type verbs, but not *know*-type verbs:

(47) a. **Jak** Piotr   myśli/ przypuszcza/sądzi/ uważa,   **kogo** Janek lubi?   [Pol]

How Peter thinks/ supposes/   judges/considers whom John loves

b. ***Jak** Janek powiedział/mówił/wie/rozumie   /pamięta, **co** studenci przeczytali?

How John said   /said / knows/ understands/remembers what students read?

(48) a. **Kak** vy   dumaete/polagaete/sčitaete/ predpolagaete, **kogo** ljubit Ivan?   [Rus]

How you think/   suppose/ consider/ suppose/   whom loves John

b. ***Kak** vy skazali/govorili/znaete/zabyli, **kogo**   ljubit Ivan?

How you said /told   /know /forgot whom loves John

Thus, Case theory makes the correct distinction concerning the class of matrix predicates that support *wh*-scope marking in Slavic.[9] We take this as an indication that our proposal

---

[9] Note that if the proposal is correct, it shows that Slavic *wh*-scope marking questions, despite being a limited phenomenon, form an open class directly corresponding to the class of non-Case marking verbs that take clausal complements.

concerning the underlying position of the *wh*-scope marker is on the right track. Furthermore, the NP-shell hypothesis proposed for Slavic in Chapter 3 receives strong independent support.[10]

A note should be made concerning the mode of overt movement of the *wh*-scope marker in Russian and Polish *wh*-scope marking questions. We postulated that the *wh*-scope marker is a head and so undergoes head movement to the matrix C, as illustrated in (45). The question then arises, why, instead of the [+*wh*] head, the entire XP headed by it does not undergo movement in (37) and (38), resulting in the ungrammatical (49) and (50) (in the next section we indeed argue that these are in fact the LF structures of (37) and (38)):

(49)   *[Jak [kogo Janek lubi]] myślisz,?   [Pol]

             how think-you whom John loves

(50)   *[Kak [kogo ljubit Ivan]] vy dumaete?   [Rus]

             how you think   whom loves John

             'Who do you think John loves?'

We suggest that the answer to the above question lies in the particular view concerning an optimal way of feature checking. In the system of Chomsky (1995c), there are two types of configurations involved in the definition of checking domain, namely, head-to-head (or

---

[10] The only difference between the *jak/kak* and Caseless TO is that the latter is structurally realized as a non-projecting element, hence a maximal phrase (see Chapter 4), whereas the former is taken to be a head, which must project into a phrase when combined with the (embedded) question in a *wh*-scope marking construction. However, this difference is immaterial in light of the crucial similarity in their distribution, captured by Case theory: both phrases must occur in a Caseless position.

feature to head) adjunction and specifier-head. Several authors (Alexiadou and Anagnostopoulou (1999), Nunes (1998), Pesetsky and Torrego (2000), Bošković (1997c), cf. also Stateva (to appear) for related discussion) have argued that when there is an option of choosing whether to move a head (or a feature) of the phrase, or the phrase itself, the computational system in fact prefers establishing a checking relation among features via head adjunction, and implements the second mode of checking, by projecting a specifier, only when head adjunction violates independent principles. Thus Nunes (1998) points out that elements adjoined to a head H a priori form the most natural configuration for the checking domain of H: a given element moves to enter into a checking relation with the features of H, not with the projection formed by H and its complement; he further suggests that adjunction to head is the unmarked option. Thus, adjunction to head H is the most optimal way of checking features. Non-minimal (including more than one lexical item) maximal projections, on the other hand, cannot adjoin to heads (X^0 elements), presumably, because the morphological component would rule out the structure in which a phrase is 'buried' inside an X^0. For the latter case, the computational system resorts to the option of projecting a specifier, to accommodate the moving phrase. Movement in this case, according to Nunes, satisfies Morphology. Nunes's theory provides the following explanation for the ungrammaticality of (49) and (50): even though there was an option of checking the [+wh] feature in the most optimal way (by head adjunction), the system chose the less optimal way of checking this feature by projecting a specifier of C.

A similar results obtains in the system of Alexiadou and Anagnostopoulou (1999). These authors argue that feature checking by adjunction to head H is preferred over

feature checking by moving to the specifier of H, for reasons of economy of projection, understood, roughly, as the tendency to extend the structure (by projection) as less as possible. Thus, if there is an option of feature checking by head adjunction, this option is chosen as more economical, since the other option (moving to a specifier) involves an extra step of projection.

Further support for the structure of *wh*-scope marking proposed for Russian/Polish comes from considering the semantic aspect of *wh*-scope marking. It can be shown that the structure (44) provides for a straightforward compositional analysis of *wh*-scope marking, in the framework of Hamblin (1973) augmented by Dayal's (1994, 1996) proposal that the clause containing the 'real' *wh*-phrase is a restriction on the existentially quantifying *wh*-scope marker.

5. The Semantics of Slavic *wh*-scope marking

5.1 The Indirect Dependency Approach

In a series of studies, Dayal (1994), (1996a), (1996b) develops an Indirect Dependency Approach (IDA). From the syntactic point of view, the IDA presupposes that each clause in a *wh*-scope marking question forms a separate question, or a local *wh*-dependency:

(51) jaun **kyaa**$_i$ soctaa hai [$_{CP}$ ki   merii **kis-se** baat karegii]$_i$

   John what think-PR   that Mary who   talk do

   'What does John think, who will Mary talk to?'

Under this view, the *wh*-scope marker *kyaa* is not an expletive-like element, but a regular *wh*-phrase. Note, in particular, that Dayal crucially assumes that *kyaa* is an XP which, just like in the regular questions, must originate in the complement position of the matrix predicate (and, under the view that LF movement exists, moves to Spec-CP at LF in Hindi).

The semantic part of the IDA is couched in the framework of Hamblin (1973), in which the denotation of a question is a set of its propositional answers. A *wh*-expression in this framework is interpreted as an existential quantifier. For instance, the denotation of the question in (52a) is (52b) which may yield sets such as (52c):

(52) a. Who will Mary talk to?

b. $\lambda p' \exists x$ [p'= $^\wedge$will-talk' (m, x)]

c. {$^\wedge$Mary will talk to Bill, $^\wedge$Mary will talk to Sue, ...}

According to Dayal, the embedded local *wh*-dependency in (51) receives an interpretation similar to its English counterpart (52a), namely, (52b). The first local *wh*-dependency, Dayal claims, is identical to the one in questions like (53a), which receives an answer as in (53b).

(53) a. jaun kyaa soctaa hai

John what think-PR

'What does John think?'

b. jaun soctaa hai ki vo tez hai

John thinks   that he smart be

'John thinks that he is smart'

Based on (53), Dayal proposes that the *wh*-scope marker (*kyaa* in Hindi) quantifies over propositions. In Hamblin's terms, the denotation of (53) would be (54a) resulting in sets like (54b):

(54) a. λp∃q [p= ^think' (j, q)]

    b. {^John thinks that he is smart, ^John thinks that Mary will talk to Bill, ^John thinks that Mary will talk to Sue}

The denotation of (51) is a result of combination of the denotations in (54a) and (52b), as shown below:

(55)   λp∃q [λp'∃x [p'= ^will-talk' (m, x)] (q) & p= ^think' (j, q)]

    ⇒ λp∃q [∃x [q= ^will-talk' (m, x)] & p= ^think' (j, q)]

Informally, a *wh*-scope marking question as in (51) denotes a set of propositions of the form *John thinks q*, where the value of *q* is limited to the set of answers to the question *who will Mary talk to*. The limitation correctly excludes propositions like *John thinks he is smart* as a possible answer, from the set in (54b). Under this view, (51) may be paraphrased with something like the following:

(56) For which proposition q, q the answer to "Who will Mary talk to?", John thinks q?

Dayal's main proposal is that the second *wh*-question (cf. *who will Mary talk to* in (51)) provides a restriction on the matrix existential quantifier in the form of the *wh*-scope marker (*kyaa* in (51)). This proposal is in line with the traditional view that natural language does not allow unrestricted quantification. The situation with *wh*-scope marking is thus reduced to usual properties of interrogation understood as existential quantification.

The idea has obvious natural elegance and appeal. It suggests that a compositional analysis of *wh*-scope marking should essentially be parallel to the analysis of other, more familiar types of overtly restricted quantification. The latter is standardly assumed to involve the structural relation of sisterhood between a quantifier and its restrictor. For example, in the interpretation of expressions like *every girl*, in which *girl* is a restrictor on the universal quantifier, the two must be structural sisters in the syntax, in order for a compositional analysis to go through.

However, Dayal does not adopt this direction and proposes a compositional analysis which is less straightforward. First, in Dayal's analysis, there is effectively no syntactic relation between the *wh*-scope marker and the finite clause-restrictor. For Dayal, the *wh*-scope marker originates in the argument position of the verb, and the clause-restrictor is adjoined to the matrix clause in an 'appositive' manner. Semantically, Dayal suggests building the restriction on the *wh*-scope marker into the lexical meaning of the latter. To do so, she introduces a variable T (Topic) bound by the *wh*-scope marker whose type is $\ll s, t>, t>$, that is, a set of propositions. The denotation of *kyaa* is $\lambda Q \exists q[T(q)\ \&\ Q(q)]$.

Given this definition, Dayal (1996a) proposes the following compositional analysis of (51):

(57)

$$\text{CP}$$

$\lambda T_i [\lambda p \exists q \ [T_i(q) \ \& \ [p={}^\wedge\text{think'}(j,q)]](\lambda p' \exists x \ [p'={}^\wedge\text{will-talk'} \ (m,x)])$
$\Rightarrow \lambda p \exists q \ [\lambda p' \exists x \ [p'={}^\wedge\text{will-talk'} \ (m,x)](q) \ \& \ [p={}^\wedge\text{think'}(j,q)]]$
$\Rightarrow \lambda p \exists q \ [\exists x \ [q={}^\wedge\text{will-talk'} \ (m,x)] \ \& \ [p={}^\wedge\text{think'}(j,q)]]$

$$\text{CP} \qquad\qquad \text{CP}_i$$

$\lambda Q \exists q \ [T_i(q) \ \& \ Q(q)] \ (\lambda q \ [p={}^\wedge\text{think'}(j,q)]) \qquad \lambda p' \exists x \ [p'={}^\wedge\text{will-talk'} \ (m,x)]$
$\Rightarrow \exists q \ [T_i(q) \ \& \ \lambda q \ [p={}^\wedge\text{think'}(j,q)](q)]$
$\Rightarrow \exists q \ [T_i(q) \ \& \ [p={}^\wedge\text{think'}(j,q)]]$
$\Rightarrow \lambda p \exists q \ [T_i(q) \ \& \ [p={}^\wedge\text{think'}(j,q)]]$

Spec $\qquad \lambda q \qquad$ C'$\lambda p'$ [p=p'] $({}^\wedge\text{think'}(j,q))$
$\lambda Q \exists q \ [T_i(q) \ \& \ Q(q)] \qquad\qquad \Rightarrow p={}^\wedge\text{think'}(j,q)$

kyaa $\qquad\qquad$ C^0 $\lambda p'$[p=p'] $\qquad$ IP $^\wedge\text{think'}(j,q)$
what$_i$ $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ kis-se $\quad$ meri t baat karegii

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ jaun t soctaa hai $\quad$ who-with Mary $\quad$ will talk
$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ John thinks
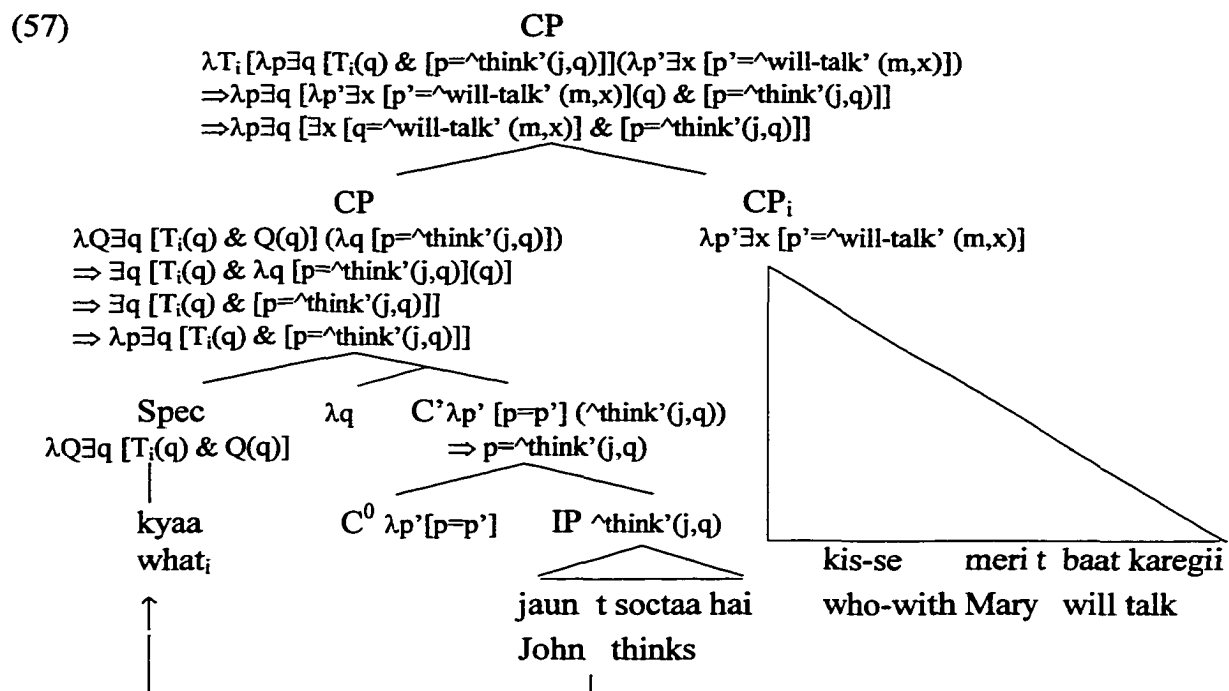
The *wh*-scope marker in the Spec of the leftmost CP takes the set of propositions of the type *John thinks q* as its argument, yielding the denotation of that CP. The denotation of the rightmost CP is basically the one in (52b). The denotation of the entire *wh*-scope marking question (51) is obtained by combining the denotations of both CPs by function application.

Arguably, introducing the Topic variable T deprives the system of its natural elegance, presumably because of the somewhat stipulative character of T.[11,12] As noted

---

[11] See, however, Dayal (1996b) where she proposes to employ the Topic variable T in the analysis of discourses of the kind *What do you think? What does Peter like?*, attempting to give it some independent justification.

above, the most natural way of implementing Dayal's idea would be to pursue the analogy with other types of quantification (cf. *every girl*), in which restricting a quantifier takes place under the structural relation of sisterhood. More specifically, one expects that the *wh*-scope marker and its restrictor clause must be sisters, for the purposes of compositional analysis.[13] Such is not the case in Dayal's analysis in (52): there, the *wh*-scope marker *kyaa* and the second clause restriction are different constituents at every point of the syntactic derivation.[14]

We would like to suggest an alternative compositional analysis of *wh*-scope marking which, we believe, implements Dayal's original idea about the second clause being a restriction on the *wh*-scope marker in a more straightforward and natural manner than Dayal's own analysis. In particular, we show in the following sections that the syntactic structure we proposed in (44) for Slavic *wh*-scope marking constructions (cf. Section 4)

---

[12] In addition, potential technical objections arises with respect to incorporating T into the compositional analysis. Recall that T provides a restriction on *kyaa* and this restriction must be the rightmost CP in (51). Given the denotations of the 'left' ($\lambda p \exists q$ $[T_i(q)$ & $[p=^\land think'(j,q)]]$) and 'right' ($\lambda p' \exists x$ $[p'=^\land will\text{-}talk'$ $(m,x)]$) CPs, however, it does not seem possible to replace T with the denotation of the rightmost CP simply as a result of combining the denotations of both CPs. This is because neither of the denotations is in the domain of the other. To solve this problem, Dayal suggests binding the variable T_i by abstracting over it, yielding a function from question denotations to question denotations $\lambda T_i \lambda p \exists q$ $[T_i(q)$ & $[p=^\land think'(j,q)]]$ (e.g., Dayal (1996a), pp.63-64). Then the denotation of the rightmost CP is in the domain of this resulting function. This gets the result. Similar considerations obtain concerning abstracting over the propositional variable p in calculating the meaning of the leftmost CP. However, free λ-abstraction is an operation that would not be allowed in more restrictive frameworks, such as Heim and Kratzer (1998) in which each instance of λ-abstraction is independently motivated (e.g. as an interpretive effect of syntactic movement).

[13] Here we largely disregard certain unorthodox views which imply separating the restriction from its operator at LF. One such view is Chomsky (1993) stipulated Preference Principle, formulated so as to 'minimize the restriction on the operator if possible', favoring interpreting *wh*-operators separately from their restrictions.

[14] Dayal (1996b) modifies this analysis, assuming, that at least in German, that the two do in fact form a constituent at LF (p.123). However, this assumption remains somewhat stipulative. In addition, Dayal (1996b) still makes use of the topic variable T, hence, arguably, inherits the above mentioned inadequacies.

provides for a semantic analysis which straightforwardly fulfills the expectation concerning the sisterhood of the *wh*-scope marker and the finite clause.

5.2 Slavic wh-scope marking: The Logical Form

In Section 4, we established that the *wh*-scope marker forms a constituent with the second clause at 'D-structure' in Russian and Polish (cf. (44)). The surface word order is derived by movement of the *wh*-scope marker in overt syntax, in accord with the general pattern of Slavic *wh*-fronting.

We propose the following way of deriving the LF of *wh*-scope marking quesitons. After *wh*-movement applies to the structure in (44) moving the *wh*-head into C, the surface word order obtains:

(58) [$_{C'}$ **jak** myślisz   [$_{XP}$ t_{jak} [**kogo** Janek lubi t_{kogo}]]]?   [Pol]

      how think-you   whom John loves who

(59) [$_{C'}$ **kak** vy dumaete [$_{XP}$ t_{jak} [**kogo** ljubit Ivan t_{kogo}]]]?   [Rus]

      how you think   whom loves John whom

We assume that it is the trace (lower copy) of the *wh*-scope marker that matters for the purposes of semantic computation, rather than the moved item itself. Following Heim and Kratzer (1998), we assume that the trace of *jak/kak* is of the same semantic type as the moved item (see below). The trace (lower copy) of the *wh*-scope marker in (58) and (59) forms an XP constituent with the second clause.

Incorporating Dayal's insight, we propose that the XP constituent is a kind of Quantifier Phrase. As such, it undergoes QR at LF, resulting in (60) and (61):

(60) LF: [$_{CP}$ [$_{XP}$ t_{jak} [$_{CP}$ **kogo** Janek lubi]]$_i$ (jak) [$_{IP}$ myślisz   t$_i$]   [Pol]

        whom John loves   how   think-you

(61) LF: [$_{CP}$ [$_{XP}$ t_{kak} [$_{CP}$ **kogo** ljubit Ivan]]$_i$ (kak) [$_{IP}$ vy dumaete t$_i$]   [Rus]

        whom loves John   how   you think

Note, crucially, that in the LFs (60) and (61), the *wh*-scope marker forms a constituent with the second clause in a *wh*-scope marking question.[15]

In the next section, we return to the issue of what motivates this LF movement, as well as the quantifier status of the *wh*-scope marker.

5.3 A compositional analysis of Slavic wh-scope marking

We suggest the following compositional analysis of *wh*-scope marking, using the LF in (61):

---

[15] Dayal (1996b) proposes a similar LF for German. Her proposal, however, differs from ours in that she assumes that the second clause undergoes LF movement and attaches to the wh-scope marker in Spec-CP. See fn. 14.

(62)

$$\mathbf{CP}$$

$\lambda Q\lambda p''\exists q''[\exists x[q''=^\wedge loves\ (i,\ x)]\ \&\ Q(q'')(p'')]\ (\lambda r\lambda q\ [q=\ ^\wedge think'\ (you,\ r)])$

$\Rightarrow \lambda p''\exists q''[\exists x[q''=^\wedge loves\ (i,\ x)]\ \&\ [\lambda r\lambda q\ [q=\ ^\wedge think'\ (you,\ r)]](q'')(p'')]$

$\Rightarrow \lambda p''\exists q''[\exists x[q''=^\wedge loves\ (i,\ x)]\ \&\ [p''=\ ^\wedge think'\ (you,\ q'')]]$

$\lambda r\lambda q\ [q=\ ^\wedge think'\ (you,\ r)]$

$\mathbf{XP}$

$\lambda P\lambda Q\lambda p''\exists q''[P(q'')\ \&\ Q(q'')(p'')](\lambda q'\exists x[q'=^\wedge loves\ (i,\ x)])\quad \lambda r$

$\Rightarrow \lambda Q\lambda p''\exists q''[[\lambda q'\exists x[q'=^\wedge loves\ (i,\ x)]](q'')\ \&\ Q(q'')(p'')]$

$\Rightarrow \lambda Q\lambda p''\exists q''[\exists x[q''=^\wedge loves\ (i,\ x)]\ \&\ Q(q'')(p'')]$

$\mathbf{C'}$

$\lambda p\lambda q\ [q=p]\ (^\wedge think'\ (you,\ r))$

$\Rightarrow \lambda q\ [q=\ ^\wedge think'\ (you,\ r)]$

kak
how

$\lambda P\lambda Q\lambda p''\exists q''[P(q'')\ \&\ Q(q'')(p'')]$

$\mathbf{CP}\ \lambda q'\exists x[q'=^\wedge loves\ (i,\ x)]$

$\mathbf{C}$

$\lambda p\lambda q\ [q=p]$

$\mathbf{IP}\ ^\wedge think'\ (you,\ r)$

vy dumaete t$_{XP}$
you think

kogo$_j$ ljubit Ivan t$_j$
whom loves John

Both the matrix and the lower CP denote a question, that is, are of type $<<s,t>,t>$. The matrix IP contains a trace of the moved XP which is interpreted as a propositional variable r of the type $<s,t>$. As in Dayal's analysis, the IP denotes an open sentence $^\wedge think'\ (i,\ r)$, and C^0 serves as a function turning propositions into sets of propositions, that is, questions (of the $<<s,t>,t>$ type), so that the denotation of C' is $\lambda q\ [q=\ ^\wedge think'\ (you,\ r)]$ (type $<<s,t>,<<s,t>,t>>$). Movement of the XP creates a λ-abstract (cf. Heim and Kratzer (1998)), the denotation of which is $\lambda r\lambda q\ [q=\ ^\wedge think'\ (you,\ r)]$. This denotation is submitted as an argument of the function denoting the moved XP, which we discuss next.

We define a lexical entry for *kak* in its *wh*-scope marking meaning as an existential quantifier over propositions:

(63) $\| kak \| = \lambda P_{<<s,t>,t>}\lambda Q_{<<s,t>,<<s,t>,t>>}\lambda p''\exists q''$ [P(q'') & Q(q'')(p'')]

Informally, the *wh*-scope marker takes a question of the form P as its input and maps it onto the function Q which, when given the value of two propositional arguments of the type <s,t>, yields a question. Note that our restricting function P here is similar to Dayal's T in that it serves as a built-in restriction on the existentially quantifying *kak*. In our case, however, no *ad hoc* measures are required in order to combine this function with the denotation of the actual set of propositions (of the form *who does John love*). This is so because the CP node denoting this actual set of propositions is a sister of the *wh*-scope marker, and, as a result, can combine with it via functional application, resulting in the denotation of the XP as $\lambda Q\lambda p''\exists q''[\exists x[q''=^{\wedge}loves\ (i,\ x)]\ \&\ Q(q'')(p'')]$. The type of the XP is $<<<s,t>,<<s,t>,t>>,<<s,t>,t>>$.[16] The denotation of the XP combines with the λ-abstract, and yields the denotation of the matrix CP. One can easily check that the resulting denotation is in conformity with Dayal's desired interpretation in (55).

The motivation behind the QR of the XP becomes straightforward. Since the matrix verb *dumaete* "think" needs a propositional complement (type <s,t>), but, instead, has the XP of the above type as its sister in the overt syntax (cf. (44)), this results in a type mismatch at LF. The type mismatch can be resolved by movement of the XP, which

---

[16] Sigrid Beck (p.c.) pointed out to us an interesting type parallelism between this type of the XP and that of regular QPs like *every man*, namely, $<<e,t>,t>$, whereby the individual type e corresponds to the propositional type <s,t> and the truth-value to values of questions, in our Hamblin-type semantics. This indirectly supports our claim in the previous section that the XP is a kind of Quantifier Phrase.

leaves a trace of the right type $(<s, t>)$.[17] Overall, the proposed analysis preserves the spirit of the classical Hamblin/Karttunen-type analysis of wh-questions (see von Stechow (1996) and Beck (1996) for recent arguments for Hamblin/Karttunen's framework).[18]

6. Addressing the properties

The proposed analysis of *wh*-scope marking entails an account for most of the properties of Slavic *wh*-scope marking outlined in Section 2. Certain properties seem to be best explained as syntactic consequences, while others receive explanation in semantic terms.

Property **I** (the possibility of any *wh*-phrase in the lower clause of a *wh*-scope marking question), Property **II** (the lower clause can be a multiple *wh*-question), as well as Property **VII** (the lower clause can be a yes/no question) follow naturally since, under the IDA, a *wh*-scope marking question involves two interrogative syntactic dependencies. It does not matter whether the lower dependency involves a *wh*-question (single or multiple) or a yes/no question; as long as it is a question type (set of propositions), the entire *wh*-scope marking sentence is interpretable, and its interpretation is computed along the lines above.

Property **III** ('antilocality') can now be seen from the point of view that root sentences involving both the *wh*-scope marker and the *wh*-phrase are not questions. This follows from the compositional analysis of the XP in (62), given the lexical entry for the

---

[17] Such LF movement is in fact the only option in a framework in which no type shifting operations are available. See, however, Heim and Kratzer (1998), among others, for an alternative. Note also that this movement should involve pied-piping, otherwise it does not resolve type mismatch.

[18] See, in particular, Hagstrom (1998), Reinhart (1998), Rullmann and Beck (1998) for alternative treatments of *wh*-questions.

wh-scope marker in (63). As mentioned above, the type of this XP is <<<s,t>,<<s,t>,t>>,<<s,t>,t>>. This is not a question type (<<<s,t>,t>), hence, the XP cannot be a question.

Consider Property **IV** (the matrix verb may not (s-)select a question), illustrated in (17), repeated below:

(17) *__Jak__ sie zastanawiales, **co** studenci przeczytali?   [Pol]

       how wonder-you,   what students read

     'What do you wonder, what did the students read?'

We agree with Dayal that the ungrammaticality of (17) is a result of semantic type mismatch. Recall that the variable *r* in (62) is of propositional type <s,t>. As Dayal points out, in questions like (17) the matrix verb of the *wonder* type allows only for sets of propositions in its answer. The variable *r* must therefore be of type <<s,t>,t>. But then the function Q defined as being of type <<s,t>,<<s,t>,t>> cannot combine with it via functional application because of the type mismatch. If Q were to combine with this variable it would have to be of a more complex type <<<s,t>,t>,<<s,t>,t>>, contra our conventions.

Property **V**, the impossibility of negating the matrix predicate (cf. (19) and (20)) follows from the head status of the *wh*-scope marker *jak/kak*. Assuming that negation is realized as a head as well, moving of the *wh*-scope marker across the negation amounts to a violation of the Minimal Link Condition/Attract Closest, essentially a Relativized Minimality violation (see also Herburger (1994). Alternatively, the impossibility of

negated matrix predicates may be attributed to a pragmatic constraint of the sort argued in Dayal (1996b). Either of these two possibilities suffices for our purposes.

Consider now property VI, the impossibility of moving the *wh*-scope marker long-distance. The relevant Polish example (21) reproduced below with some structure:

(21) ??**Jak**$_i$ Janek sądzi [$_{NP}$ TO [$_{CP}$ że Piotrek myśli [$_{XP}$ t$_i$ [$_{CP}$ **co** studenci przeczytali]]]]?

How Janek judge   that Piotr   thinks   what students read

'What did John say that Peter think that the students read?'

As (21) shows, movement of the *wh*-scope marker is clause (CP-) bound. Recall from Chapter 4 that any embedded indicative CP in Polish (and, by extension, in Russian) is 'encapsulated' in an NP-shell headed by the correlative. Extraction out of such constructions therefore creates a 'Complex NP island' effect.

Finally, Property VIII involving sentences with iteration of the *wh*-scope marker remains somewhat of a puzzle under the present analysis. The relevant Russian example (25) is reproduced below indicating the structure assigned to it:

(25) ??**Kak**$_j$ vy dumaete, [$_{XP1}$ t$_j$ [$_{CP}$ **kak**$_i$ Ivan sčitaet, [$_{XP2}$ t$_i$ [$_{CP}$ **čto** pročitali studenty]]]]

How you think   how Ivan believes   what read   students

Counterparts of (25) are acceptable in other *wh*-scope marking languages like German and Hindi. If what we said so far is correct, the apparent deviance cannot be attributed to semantic factors. Given our lexical entry for *jak/kak* in (63), it can be shown that the *wh*-

scope marker can successfully combine via functional application with questions denoted by the intermediate CP in (62) (see also Dayal (1994)). This is so because there is no semantic distinction between *wh*-scope marking questions and regular *wh*-interrogatives: both are of the same semantic type $<<s,t>, t>>$, hence, either can become an argument to the function denoted by the *wh*-scope marker. There is also no obvious reason why (25) would involve a syntactic violation: the derivation of (25) involves a series of short movements by *wh*-scope markers *kak*, each within its own clause (aside from the 'true' short *wh*-movement in the lowest clause).

An apparently related fact is that embedded *wh*-scope marking questions are also somewhat deviant in Slavic:

(64)   ??Janek sie zastanawial **jak** Piotrek sadzi, **co**   studenci przeczytali   [Pol]

        John   wonders   how Peter   judges what students read

(65)   ??Ivan   sprosil **kak** Petr sčitaet, **čto** pročitali studenty   [Rus]

        John   asked how Peter believes what read   students

One possibility suggests itself in light of the important observation of Reis (1996) that *wh*-scope marking constructions are intimately related to corresponding sequential/parenthetical questions. Reis shows that in German, the properties of the matrix clause in *wh*-scope marking questions closely resemble the properties of what she calls an Integrated Parenthetical, in constructions such as (66):

(66)   **Was** glaubst Du, **wen** liebt sie?

what believe you who loves she

'What do you think, who does she love?'

Reis (1996) argues convincingly that the *wh*-scope marking question is a historic descendant of the Integrated Parenthetical construction, a result of diachronic transition from the domain of discourse into the domain of sentential grammar (an idea further pursued in Dayal (1996b)). Suppose this is indeed the case. Now, if this diachronic transition is a more or less general process, we expect it to take place in other languages as well. That said, it is plausible that it is an ongoing and as yet non-completed process in languages like Russian and Polish. As a result, it may be that (sentential) *wh*-scope marking, while already existing as such in these languages, has not yet established itself as a fully productive strategy in certain contexts, so that instances like (25), (64) and (65) are not yet part of the sentential grammar.

7. Interim Summary

In this part of our discussion, we have considered the strategy of *wh*-scope marking in Russian and Polish, in the framework of Dayal's IDA. On the basis of these languages, we have seen that the structure of *wh*-scope marking questions is closely related to the structure of declarative sentences, under the hypothesis that finite complements in declarative sentences are NP-shells. Slavic data thus provide an important empirical argument for 1) the NP-shell hypothesis, 2) the analysis of *wh*-scope marking in terms of the IDA, and 3) the existence of a fundamental structural connection between declarative sentences with finite embeddings, on the one hand, and *wh*-scope marking questions, on

the other. The connection is rooted in Case theory: the Case properties of the matrix predicate must match the Case properties of the correlative, or the *wh*-scope marker. Thus our analysis explains why the constituent headed by the *wh*-scope marker and including the finite-clause restrictor is generated in the domain of the matrix verb. In contrast, Dayal's related proposal that the *wh*-scope marker must originate in the argument position of the verb is stipulated, and left unexplained.

In the following sections we strengthen the case for the syntactic and semantic analysis of *wh*-scope marking in the general framework of IDA and modified along the lines above.

8. IDA in the 'Shell'-Based Syntax: Cross-Linguistic Arguments

Given the present understanding of the nature of syntax-semantics interface, it is a common assumption that in all languages that entertain *wh*-scope marking questions their interpretation is derived from the same or similar LF. This implies that the syntactic structure of *wh*-scope marking questions, and syntactic processes leading to the appropriate LF, must be more or less similar as well. In other words, the structure we proposed in (44)/(62) for Slavic must be operative in all languages that entertain the *wh*-scope marking strategy. In this section, we extend our analysis for Hungarian, Hindi and German. We also address a number of apparent problems for the IDA that have been discussed in the literature, and show that most of them are straightforwardly accounted for under the modified version of the IDA proposed in this chapter.

8.1 Hungarian

The existence of *wh*-scope marking as a sentential phenomenon in Hungarian has been observed and discussed in Horvath (1997), (1998), Kiss (1991). Horvath (1997) provides strong arguments for the analysis of *wh*-scope marking constructions in Hungarian in which the *wh*-scope marker originates in the complement position of the matrix verb. The regular *wh*-scope marker in Hungarian *mit* means *what*, presumably, a Case-marked nominal. Hence, we expect that *wh*-scope marking verbs must be Case-marking verbs. Horvath (1997) argues extensively that this is indeed so. The facts concerning morphological Case and object agreement demonstrate that the *wh*-scope marker and the matrix verb must be (at some point) in a local structural relation:

(67) a. Mit   mondtál,   hogy kinek   vett   János   színházjegyet?

     What-acc said-2sg-indef.DO that who-dat bought John-nom theatre-ticket-acc

     'What did you say, for whom John did buy a theater ticket?'

   b. Mire   számítasz, hogy melyik fiúval   fog Mari   beszélni?

     On-what-allative count-2sg that which boy-with   will Mary-nom speak-inf

     'On what do you count, with which boy will Mary will speak?'

   c. *Mire mondtál,   hogy kinek   vett   János   színházjegyet?

     What-allat. said-2sg-indef.DO that who-dat bought John-nom theatre-ticket-acc

(68) a. Tudják   hogy melyik fiút   szereted

       know-3pl-def.DO   that which boy-acc like-2sg-def.DO

       'They know which boy you like'

b. Mit   tudnak/*tudják   hogy melyik fiút   szereted?

what-acc know-3pl-indef.DO/def.DO   that   which   boy-acc   like 2sg-def.DO

'What do they know, which boy you like?'

<div align="right">(adapted from Horvath (1997), p. 527)</div>

(67) demonstrates that the Case of the *wh*-scope marker is determined by the choice of the verb, in conformity with the general Case assigning pattern by that verb. It also shows that the matrix verb displays indefinite object agreement when it occurs with the *wh*-scope marker, similarly to regular *wh*-questions, cf:

(69)   Mit   mondott   Mari?

What-acc   said-3sg.-indef.DO Mary

'What did Mary say?'

(68a) shows that finite clausal complements trigger definite object agreement on the verb in Hungarian. (68b) demonstrates, crucially, that in a *wh*-scope marking question the verb agrees with the *wh*-scope marker (by showing the indefinite agreement), not the clausal complement.

As we argued so far, *wh*-scope marking questions in Russian and Polish utilize essentially the same syntactic structure as that assigned to regular declarative sentences with embedded finite complements - the structure. of an NP/XP-shell. Interestingly, Kiss (1987) and especially Horvath (1997) note a similar parallel for Hungarian. As discussed in Chapter 4, Kiss suggests that the clausal proform (cf. *az*)

in Hungarian declarative sentences forms a constituent with the embedded clause, essentially an NP-shell. Horvath (1997) suggests that the *wh*-scope marker forms a constituent with the embedded clause in *wh*-scope marking questions in Hungarian, at least at LF. Horvath (1997) further brings in the same motivation for this parallel, in terms of the Case theory: the morphological Case of the *wh*-scope marker (in *wh*-scope marking questions), and of the clausal pro-form (in declarative sentences with embeddings), and is fully determined by the Case assigning properties of the matrix (subordinating) verb.

Our analysis extends to Hungarian straightforwardly. Consider (68b), for example. Under our analysis, *mit* in *wh*-scope marking questions is a head, and forms a constituent with the embedded clause prior to *wh*-movement. The surface word order is derived after *wh*-movement, as shown in (70a). At LF, the *mit*+CP constituent undergoes QR, as shown in (70b), and receives an interpretation along the lines of (62):[19]

(70) a. Mit tudnak   [t_{mit} [hogy melyik fiút   szereted]]?

        what know-3pl-indef.DO   that which boy-acc   like 2sg-def.DO

     b. LF: [t_{mit} [hogy melyik fiút   szereted]]   (mit) tudnak

           that which boy-acc like 2sg-def.DO what know-3pl-indef.DO

---

[19] Horvath in fact proposes an LF for Hungarian *wh*-scope marking question that is virtually identical to ours (via a somewhat different syntactic derivation). However, she does not adopt Dayal's IDA semantics of *wh*-scope marking, but rather, derives the interpretation via a mechanism of 'operator-feature percolation'.

8.2 Hindi

Hindi, for which the IDA was originally designed by Dayal, is easily amenable to our syntactic analysis based on Case theory. Consider again the question in (51):

(51)   jaun **kyaa**$_i$ soctaa hai [$_{CP}$ ki   merii **kis-se** baat karegii]$_i$

John what think-PR   that Mary who   talk do

'What does John think, who will Mary talk to?'

Let us adopt Mahajan (1990), (1996) proposal that *kyaa* and the second clause in a *wh*-scope marking question form a constituent at an underlying level, generated in the argument position of the verb. Mahajan assumes that *kyaa* is an XP. We, instead, take *kyaa* in *wh*-scope marking questions to be a head, taking the finite clause-restrictor as its complement, much like *jak/kak* in Slavic.[20]

The Case theoretic considerations support the view that the constituent headed by *kyaa* is generated in the domain of the matrix verb. Given that *kyaa* is clearly a nominal (literally meaning "what"), it is reasonable to assume that it must have a Case feature (recall that under our assumptions, overt nominals must have Case). This makes a prediction that *wh*-scope marking verbs in Hindi must be Case-assigners. The prediction is borne out. For instance, *soc* 'think' and *kah* 'say' are commonly used verbs that support

---

[20] See fn 10.

wh-scope marking when inflected for habitual aspect.[21] Importantly, these verb forms can take direct object DPs as well:[22]

(71) a. (ruko,) mE kuch   soctaa   huun   (Srija Sinha, p. c.)

    (wait,) I   something think-hab. be-pres.

    '(Wait), I will think of something'

b. mE ek baat kahtaa huun   (dhyaan-se   suno.)

    I   one fact say-hab. be-pres. (care/caution with listen-2sg.)

    'I will tell you something (Listen to me carefully)'

At the same time, we were not able to find verbs that participate in *wh*-scope marking in Hindi and are not Case assigners.

In line with our previous discussion of Russian/Polish, we now claim that the *wh*-scope marker and the finite clause in Hindi are interpreted as a constituent in the semantic component, by undergoing QR and yielding a logical form similar to (60) and (61):

(72) LF: [t_{kyaa} [_{CP} ki merii **kis-se** baat karegii]] jaun (kyaa) soctaa hai

         that Mary who talk do   John what   think-PR

    'What does John think, who will Mary talk to?'

---

[21] Habitual aspect marking seems to be the necessary condition on predicates participating in *wh*-scope marking. This possibly relates to the point of view orientation of *wh*-scope marking questions. See Reis (1996), Dayal (1996) for more discussion.

[22] This confirms Dayal's and Mahajan's suggestion that *kyaa* is a *wh*-correlate of the sentential pro-form *yeh* (cf. (43) above), which implies that verbs that take *yeh* as a direct object (that is, Case-assigning verbs) and the *wh*-verbs supporting *wh*-scope marking are the same set.

There are at least two possibilities to make (72) compatible with the surface word order as in (51). One possibility is that the *kyaa*+CP constituent is base-generated preverbally, and the finite clause is obligatorily extraposed in overt syntax (on the extraposition analysis of finite clausal complements, see Dayal (1996a)). At LF, what gets interpreted is the trace of the extraposed clause (given that it has the same semantic type, see Section 5.2) in the position of complement of *kyaa*. This possibility presupposes that SOV is the basic word order in Hindi. A second possibility is that the *kyaa*+CP constituent is base-generated postverbally, and in the course of the derivation *kyaa* moves into a preverbal position by an overt operation (perhaps, for reasons related to Case-checking, or incorporation; for the latter possibility, see Chapter 6). This possibility is in line with Kayne (1994) Linear Correspondent Axiom (LCA) which enforces the SVO word order in Hindi (in fact, universally). A version of this possibility is explored in Mahajan (1996). Under our analysis, then, it would be the trace of the *wh*-scope marker gets interpreted in Hindi, exactly as in Polish and Russian. This latter possibility is more conceptually appealing since there is no need to posit obligatory extraposition and explain why it has to occur. Later, we present reasons to adopt this view. For now, it is not necessary for us to decide which of these possibilities is realized. Crucially, our Case theoretic proposal concerning the syntax of *wh*-scope marking coupled with our semantic analysis in the IDA framework incorporates the original Hindi paradigm with no additional stipulations.

8.3 German

Consider again (1):

(1) **Was** glaubst Du, **wen** sie liebt?

what believe you who she loves

'Who do you believe that she loves?'

The German *wh*-scope marker *was* literally means 'what', hence is, arguably, a nominal element that has a Case feature. Consequently, our analysis predicts that in German the set of matrix verbs supporting *wh*-scope marking must be verbs that are Case-assigners. This prediction is correct. Höhle (1996) observes that matrix verbs in *wh*-scope marking questions can also take a direct object like *das* 'that', or (regular) *was* 'what', meaning that they can check Case:

(73)   Was   sagt/denkt/glaubt   Hanna wen Peter mag?

What says/thinks/believes Hanna who Peter likes

(74)   a. das sagt/denkt/ glaubt   Hanna   (adapted from Höhle (1996), pp. 42-43)

that says/thinks/believes Hanna

b. was   denkt Hanna?

what thinks Hanna

In fact, the reverse is also true: propositional verbs that do not support *wh*-scope marking, do not take direct objects either, that is, they are not Case-assigners. This can be illustrated with the verb *bemerken* which is ambiguous between 'notice' and 'remark'. Crucially, in its 'remark' meaning, *bemerken* neither supports *wh*-scope marking, nor

takes direct objects, as shown in (75). (75a) and (75b) are acceptable on the 'notice' meaning only. A similar observation holds for *zugestimmt* 'agreed', as (76) and (77) demonstrate:

(75)   a. (*)Was hat Fritz bemerkt wen Peter mag?   (Sigrid Beck, p. c.)

          what has Fritz remarked who Peter likes

        b. (*)Ich bemerke das

          I remark that

(76)   *Was hast du zugestimmt wen wir einladen sollen?

      What have you agreed whom we invite should

                              (Stechow and Sternefeld (1988))

(77)   *Du zugestimmt das

      you agreed this

Also, not allowed in *wh*-scope marking are complex object-verb predicates and predicates like *es scheint* "it seems", none of which take accusative objects:

(78)   a. *Was ist Peter des Glaubens/der Meinung, wohin Petra gefahren ist?

      What is Peter (of) the belief the opinion where-to Petra gone is

      b. *Du bist das des Glaubens

      you are that (of) the belief

(79)   *Was scheint es (dir), womit könnte man ihm helfen?

      What seems it you-dat where-with could one him help

Our analysis accounts for the correlations observed in examples (73) through (76) straightforwardly, if a) *was* in *wh*-scope marking constructions is a head (cf. Herburger (1994)) b) the constituent headed by *was* and including the finite clause in *wh*-scope marking questions is generated close to the matrix verb, so that *was* can check its structural Case in a local relation with the verb. At LF, the *was*+CP constituent in cases like (1) undergoes QR, resulting in (80):[23]

(80)   LF: [CP [XP t_was [CP **wen** sie   liebt]] (was) [IP glaubst Du]]

                   who she loves   what   think   you

Müller and Sternefeld (1996) questions the applicability of the IDA to German and argues, instead, that German *wh*-scope marking should be analyzed along the lines of the DDA (Section 3). Recall that according to the IDA, the *wh*-scope marker

---

[23] There remain two puzzles which need to be resolved, if the IDA approach is adopted for German, as discussed, in particular, in (Höhle (1996), Müller and Sternefeld (1996), Beck and Berman (1996)) . The first puzzle is that German disallows *wh*-scope marking questions with factive verbs, whereas in Hindi such questions are acceptable:

(i)   *Was   hast du   vergessen, wen wir einladen sollen   (Beck and Berman (1996))
     what have you forgotten whom we invite   should
     'What did you forget, whom should we invite?'

(ii) jaun kyaa jaantaa hai Mery kis-se baat karegii?   (Dayal (1994))
     John what knows   Mary who with will-talk
     'What does John know, who will Mary talk to?'

The second puzzle concerns the fact that in German, *wh*-scope marking questions are not allowed if the embedded CP is a yes/no questions. In contrast, in Hindi such questions are allowed:

(iii) *Was glaubst du ob   die Maria mit dem Hans gesprochen hat (Beck and Berman (1996))
     what think you whether   the Maria with the Hans talked has
     'What do you think, did Maria talk with Hans?'

(iv) ravi-ne   kyaa kahaa ki   anu aayegii yaa naahiiN   (Dayal (1996b))
     Ravi-erg   what say- that Anu come-fut or not
     'What did Ravi say, will Anu come or not?'

was is a 'meaningful' *wh*-quantifier, and should behave similarly to other *wh*-phrases. In particular, it should be able to occur with another *wh*-phrase in the same clause of the *wh*-scope marking question, as in the cases of multiple *wh*-interrogation. This is indeed the case in a number of dialects of German, as shown in (81):

(81) Was meint wer [CP [mit wem]ᵢ sie tᵢ gesprochen hat]?

　　　what thinks who　　with whom she　　talked　　has

(81), in which *was* is moved to the clause-initial position is not necessarily problematic for the IDA. However, (82) seems to be problematic:

(82)a. *Wer meint was [CP [mit wem]ᵢ sie tᵢ gesprochen hat]?

　　　　who thinks what　　with whom she　　talked　　has

　b. cf. Wer meint was?

　　　Who thinks what

Given that *was* can generally occur in a non-fronted position, as indicated in (82b), the ungrammaticality of (82a) is unexpected. On the other hand, under the DDA, in which the *wh*-scope marker *was* is an expletive inserted into Spec-CP, the ungrammaticality of (82a) can be directly accounted for since *was* in (82a) can never be in a position other than Spec-CP (that is, *was* in (82a) and (82b) are different, albeit homonymous, items).

Fanselow (1997), however, offers an interesting possibility which allows one to reconcile the ungrammaticality of (82a) with the IDA. He shows independently that whenever *was* appears together with an overt restriction (which he terms 'associate'), it must raise to Spec-CP. First, observe that *wh*-phrases in German are ambiguous between the *wh*-quantifier reading and the indefinite reading: thus *was* in certain contexts can mean either *what* or *something*, as exemplified below:

(83)   a. Wer hat dir was gesagt?

         who has you what said

         "Who said something to you?" OR

         "Who said what to you?"

Fanselow notes that *was*, on either reading, may be associated with a restriction in the form of a syntactically independent adjective, as shown below:

(84)   a. Er hat mir was schönes gesagt

         he has me what nice   said

         'He said something nice to me"

       b. Was hat er dir denn schönes gesagt?

         what has he you ptc. nice   said?

         'Which nice things did he say to you?'

Observe now that for the *wh*-quantifier reading of *was* plus the restriction in the form of an AP (sch?nes) to obtain, the latter must be in Spec-CP; it cannot be in situ:

(85)   Wer hat dir denn was schönes gesagt

who has you ptc what nice   said

'Who said something nice to you?'

NOT: 'who said which nice things to you?'

Now recall that under the IDA, especially in the form that we develop in this chapter, the embedded CP in *wh*-scope marking question provides a restriction on *was*, similarly to the AP above. Thus the unavailability of the *wh*-reading of *was* in (85) is on a par with the ungrammaticality of (82a): in both cases, *was* with a restriction fails to move to Spec-CP.[24] At the same time, the insight that in *wh*-scope marking questions, *was* which takes the embedded CP as a restriction must be in Spec-CP, is preserved, correctly.

If the two cases are parallel, however, the question is why there is at least an indefinite reading available in (85), but apparently not in (82a) (given it's total ungrammaticality). Fanselow speculates that the *wh*-scope marker *was* differs from regular *wh*-phrases in German in that it only has a *wh*-quantifier reading. In this respect, it behaves more like English *wh*-phrases, which are unambiguously *wh*-quantifiers.

---

[24] Under this argument, it must be assumed that the *wh*-scope marker *was* does not have an indefinite reading, unlike the regular *was*.

8.4 Wh/quantifier interaction

Several authors (Pafel (1996), von Stechow 1996) argued against the IDA on the whole, on the basis of certain facts concerning scope interaction between quantifiers like *everyone* and the *wh*-scope marker. Consider the following pair:

(86)   a. Wo$_i$ glaubt   jeder   daß sie gerne   t$_i$ leben würde?

          Where believes everyone that she readily   live   would

          'Where does everyone believe that she would like to live?'

      b. Was   glaubt   jeder   wo$_i$ sie gerne   t$_i$ leben würde?

          What believes everyone where she readily   live   would

(86a), involving long-distance *wh*-movement, is ambiguous: either *jeder* has wide scope with respect to *wo* or vice versa. Interestingly, the *wh*-scope marking (86b) only has the reading in which *jeder* scopes over *wo*. Put differently, (86a) has the reading in (87a) and (87b), whereas (86b) has only the reading in (87b):

(87) a. For which place x, does everyone believe that she would like to live at x?

    b. For every person y, for which place x, does y believe that she would like to live at x?

But under Dayal's version of the IDA (cf. Section 5.1), in which *was* is associated with the set of propositions denoted by the restrictor clause, nothing prevents scope interaction

similar to that in (86a): namely, either *jeder* should scope over *was*, or vice versa. Arguably, then, (86b) should have the following readings (cf. (56) above):

(88) a. For which proposition p, p the answer to "Where would she like to live?", everyone believes p?

b. For every person y, for which proposition p, p the answer to "Where would she like to live?", y believes p?

It is very difficult, if not impossible, to exclude the reading (88a) in Dayal's original IDA analysis (cf. (57)). In her analysis, the *wh*-scope marker binding the Topic variable T which stands for the restrictor clause, would scope over *jeder* in the matrix clause, at LF. Von Stechow (1996) proposes a solution within the DDA (cf. Section 3). Recall that under the DDA, the *wh*-phrase from the lower clause raises at LF to the matrix Spec-CP, possibly 'replacing' the *wh*-scope marker, where the latter functions as a semantically void element. The solution is based on the ban on quantifier crossing at LF, proposed in Beck (1996):

(89)`   *Beck's filter:*

$*[\alpha_i...negation/quantifier...t_i^{LF}...]$

Beck's filter bans the reading in (87a) since it involves LF movement of *wo* across the quantifier *jeder*, resulting in the configuration of the type in (89). The reading (87b), under this proposal, is derived by QR-ing *jeder* to take the widest scope, and subsequent

LF movement of *wo*. In this case *wo* crosses only the trace of *jeder*, and a configuration as in (89) does not arise. Both the illegitimate and legitimate LFs for (86b) are shown below:

(90)   a. *[CP wo_i   glaubt   jeder [t_i^LF sie gerne t_i lebem würde]

            where believes everyone she readily live   would

        b. jeder_j [CP wo_i   glaubt t_j [t_i^LF sie gerne t_i lebem würde]]

            everyone where believe   she readily live   would

The capability to account for the missing readings in (86b) under the DDA, assuming Beck's filter, is then taken by many as an independent argument for the DDA. However, this solution is not forced upon us. In fact, once Beck's filter is adopted, it becomes possible to account for the missing readings <u>without</u> giving up the IDA framework. This possibility arises if one adopts the LF and compositional analysis for *wh*-scope marking that we developed for Slavic in Sections 5.2 and 5.3 (cf. (62)). Recall that in our analysis, the constituent headed by the *wh*-scope marker undergoes LF movement to Spec-CP. Thus the illegitimate LF for (86b) should actually look more like the following:

(91)   *[NP was [CP wo   sie gerne t_i lebem würde]]_j glaubt   jeder   t_j

            what   where she readily live   would   believes everyone

But note that when the *was*-constituent undergoes LF movement, it crosses the quantifier *jeder*, creating the configuration in (89), which violates Beck's filter. The situation is

parallel to von Stechow's (90a), but now this LF is excluded strictly within the IDA framework. Similarly, the good reading for (86b) is derived similarly to von Stechow's (90b), except for the LF movement of the *was*-constituent:[25]

(92) jeder$_j$ [$_{CP}$ [$_{NP}$ was [$_{CP}$wo   sie gerne t$_i$ lebem würde]]$_k$ glaubt   t$_j$ t$_k$]

       everyone   what   where she readily live   would   believe

The argument involving *wh*/quantifier interaction makes an important empirical distinction between Dayal's and our proposals concerning the structure and LF of *wh*-scope marking questions, again demonstrating the empirical advantage of the analysis of *wh*-scope marking we proposed on the basis of Polish and Russian.

9. Conclusion

In this chapter we have considered the *wh*-scope marking strategy. We argued that theories of *wh*-scope marking based on the IDA are on the right track. The most important syntactic characteristic of the IDA is that each clause in a *wh*-scope marking question forms a local *wh*-dependency. In other words, the complementizer heading each clause is interrogative ([+Q]). We adopted and modified Dayal's original analysis of *wh*-scope marking in the IDA framework, including both the syntactic and semantic aspects. In particular, we proposed that the *wh*-scope marker

---

[25] The hypothesis that the *was*-constituent (or, for that matter, *wo* under the DDA account above) moves after QR of *jeder*, seemingly violating cyclicity, can be reconciled with our late adjunction theory of Chapter 3, given that both movements involve adjunction at LF, hence must be postcyclic. See Chapter 3 for details.

forms a QP-like constituent with the embedded clause, at all derivational points, in parallel with other cases of overtly restricted quantification. We provided empirical evidence that the QP is generated as an object of the matrix verb in Russian and Polish. In the course of the empirical argument, it became clear that *wh*-scope marking questions in these languages utilize the same general form of finite complementation as in simple declarative sentences, namely, NP-shells. Furthermore, we argued that at LF, the QP undergoes QR, leaving behind the trace of the propositional type. The proposed syntactic structure was shown to receive a straightforward compositional analysis. We then discussed our version of the IDA analysis with respect to Hindi, Hungarian and German. Finally, we addressed a number of apparent problems noted in the literature with regard to the IDA framework, and showed that most of those receive a straightforward explanation under the theory proposed in this chapter.

CHAPTER 6

Long-Distance Extraction as Residual Wh-Scope Marking

1. Introduction

In this chapter we address the issue of locality of *wh*-movement in examples like (1):

(1) Who does Peter believe [_CP_ that John likes *t*]?

Since Chomsky (1973) it has been standardly assumed that long-distance *wh*-movement in (1) proceeds in a successive cyclic fashion, stopping by the intermediate Spec-CP(s) on its way to the matrix Spec-CP. The literature contains discussion of numerous pieces of empirical evidence supporting this view of long-distance movement (see Torrego (1983), (1984), McCloskey (1991), Kayne and Pollock (1978), Collins (1994), Zaenen (1983) to mention a few). The common theme underlying these studies is that in some languages, the long movement of a *wh*-phrase leaves morphological or syntactic reflexes within the clauses that it passes through, on its way to the matrix clause.

On the intuitive level, it is clear that movement in natural language is sufficiently local, and potentially indefinitely long 'fell swoop' steps of movement are not allowed. The question then arises, how to most adequately state the property of grammar responsible for this sort of locality.

278

Beginning with Ross (1967) and Chomsky (1973), the literature aimed at explaining the local property of movement (including successive cyclic movement) by adopting a fundamentally constraint-based approach. The fundamental assumption underlying that approach is that movement (Move α) can in principle proceed indefinitely long, but independent grammatical constraints on movement or representations preclude steps of movement that are 'too long', such as movement across a clausal boundary. The effort then has been put to find and motivate the best formulation of these constraints. The usual line of inquiry was to formulate the constraints in such a way as to not only account for the successive cyclic character of long movement, but also extend to other cases of local movement, exhibited, in particular, in 'island' configurations. The well known pre-minimalist proposals restricting successive cyclic movement so as to proceed in a local fashion includes Chomsky's Subjacency, formulated either in terms of Bounding Nodes (Chomsky (1973), (1977)) or Barriers (Chomsky (1986a)), Lasnik and Saito (1984)'s γ-marking procedure involving a locality condition on traces, Locality Condition and Bounding Condition of Koster (1978), Koster (1987). For more recent proposals, see Chomsky (2000), (2001).

However, incorporating the successive cyclic property of *wh*-movement into the theory has always been somewhat problematic. Under standard assumptions, only the ultimate landing site of the moving *wh*-phrase - matrix C - is interrogative, that is, marked [+Q] or [+*wh*]. The intermediate C of the clausal complement of the regular propositional attitude verb that supports long-distance extraction, is not marked [+*wh*], as it cannot itself host a *wh*-phrase (2a), simply because propositional attitude

verbs do not select a question. Rather, it is the same C that occurs in an embedded clause of a declarative sentence (2c):

(2) a. *Peter believes [CP what John likes]

   b. cf. Peter wonders [what John likes]

   c. Peter believes [CP that John left]

What is then the purpose of a step of *wh*-movement to a Spec of [-*wh*]? In the systems of Lasnik and Saito (1984) and Chomsky (1986a) the intermediate traces of long-distance *wh*-movement are necessary in order to license the original trace of the moving *wh*-adjunct, in cases like (3):

(3) How do you believe [*t'* that Peter [fixed the car] *t*]?

In (3), the original trace is licensed by the antecedent trace via 'antecedent-government' in a sufficiently local configuration (no 'barriers' intervene between *t'* an *t*, in the sense of Chomsky (1986a)).

In the minimalist framework, where movement takes place only when necessitated (principle of 'Last Resort'), for reasons of economy of derivation, the problem of legitimizing intermediate steps of successive cyclic movement becomes acute. If, as standardly assumed in minimalism, *wh*-movement serves to satisfy some formal property of the matrix [+*wh*] C, it follows that only the final step of successive cyclic movement is justified. Movement to the intermediate Spec-CPs

then appear to be in flagrant contradiction with the Last Resort principle, since it takes place without any obvious driving force.

In most versions of the minimalist framework, then, intermediate stages of successive cyclic movement were assumed to be driven by a feature in intermediate C which needs to be checked by overt movement (see, in particular, Chomsky (1995c), Fanselow and Mahajan (1996)). However, such a feature is extremely hard to motivate. A priori, the C found in intermediate stages of wh-movement is the same C found in embedded declarative clauses (cf. (2c)). However, no wh-movement is required in embedded declarative clauses. The existence of such a feature, therefore, is problematic.

In Chomsky (2000), non-final stages of successive cyclic wh-movement are 'indirectly feature-driven', in the sense that they are needed to guarantee eventual convergence, in compliance with the 'phase impenetrability condition'. By analogy with 'defective' T, found in ECM constructions, Chomsky suggests, indirectly, that each intermediate C may have a defective P feature which needs deletion but is unable to delete the set of features of the 'goal', that is, moving wh-phrase.

In this chapter, we explore a theory of long-distance wh-movement under the cyclic domain approach, which, in contrast to the traditional approaches to successive cyclicity mentioned above, is non-constraint-based. Pursuing such a theory is justified in several respects. We will demonstrate that a theory based on the cyclic domain approach dispenses with the need to impose constraints and mechanisms on grammar which enforce locality in long distance movement, thus simplifying the design of the grammar significantly. This is in conformity with the minimalist thrust

in favor of theories that reduce the amount of complexity of the computational system of language. Second, we show that it makes the successive cyclic *wh*-movement in complete conformity with the principle last resort, removing the most glaring peculiarity of long-distance *wh*-constructions. Finally, our theory is empirically and conceptually advantageous in that it serves as a unifying minimalist theory of long-distance *wh*-movement, *wh*-scope marking phenomena, and finite complementation.

2. Wh-scope marking and long-distance *wh*-questions

Since the seminal work of Riemsdijk (1982), long-distance *wh*-movement and *wh*-scope marking have long been thought to be related. For one thing, WSM questions trigger the kind of answer similar to that for long-distance questions. This can be illustrated for those (not all) dialects of German in which long-distance *wh*-extraction is possible. Thus (5) is a felicitous answer to both (4a) and (4b) (see also Chapter 5, Section 1):

(4) a. Was glaubst du   wen sie liebt?

   What think you who she loves

   'What do you think, who does she love?'

   b. Wen   glaubst du daß sie liebt?

   Whom think   you that she loves

   'Who do you think that she loves?'

(5)   Ich glaubst daß sie Peter liebt

I think   that she Peter   loves

'I think that she loves Peter'

The Direct Dependency Approach of McDaniel (1989) (cf. Chapter 5, Section 3), incorporates this similarity between the two types of questions in the notion of *wh*-chain: for McDaniel, a *wh*-chain, roughly, is a sequence of one or more *wh*-scope markers (if any), contentful *wh*-phrase and traces of *wh*-phrases, in that order. As noted in Chapter 5, under the Direct Dependency the *wh*-scope marker is a kind of expletive, associated with the contentful *wh*-phrase. Thus, under the DDA, both the *wh*-scope marking strategy and the long-distance strategy are special cases of a more general process of *wh*-chain formation. Both (4a) and (4b) have the same LF, stated in terms of a *wh*-chain.

In contrast, in the literature on the Indirect Dependency Approach (IDA, Dayal (1994), (1996a), (1996b)), the parallel between *wh*-scope marking questions and long-distance questions is not explored. Recall that under the IDA, the *wh*-scope marker is not an expletive, but, rather, a contentful *wh*-phrase, as in regular *wh*-questions. Nothing in the IDA so far suggests anything in common between *wh*-scope marking and long-distance questions in (4a) and (4b): in (4a), the fronted *wh*-scope marker quantifies over propositions, whereas in (4b) the fronted *wh*-phrase quantifies over individuals. Under the IDA, the LFs for (4a) and (4b) are different.

Consequently, if one were to adopt the IDA for German, the fact that (4a) and (4b) trigger the same kind of answer would remain a mere coincidence.[1]

In a larger cross-linguistic picture, it is Hindi, not the German dialects permitting both (4a) and (4b), that seems to represent the 'unmarked' case. In fact, the descriptive generalization that holds with the exception of these dialects, is that in the languages that entertain the *wh*-scope marking strategy in the sense of Chapter 5, including Hungarian, Russian, Polish and a number of dialects of German, overt long-distance *wh*-extraction out of finite clauses results in a degradation of a varying degree. Conversely, languages in which overt long-distance *wh*-extraction out of finite clauses is possible (e.g. English, French, Scandinavian), arguably, do not entertain the (sentential) *wh*-scope marking strategy. In other words, with the exception of the German dialects that permit (4a) and (4b), the *wh*-scope marking (in the sense of Chapter 5) and long-distance *wh*-extraction strategies are in complementary distribution.

This complementary distribution confirms the existence of a fundamental connection between both types of questions, and strongly suggests some common denominator that unifies the two strategies. It is our impression that this connection has been overlooked in most theories of *wh*-scope marking and locality of long *wh*-movement.

---

[1] The state of affairs implying lack of parallelism between the two strategies is perhaps due to the historical reasons, namely that the IDA was originally developed for Hindi, in which extraction out of finite clause is impossible (see, e.g. Dayal (1996a)).

In this respect, McDaniel's (stipulated) notion of *wh*-chain came closest to recognizing the connection, by providing a uniform account of *wh*-dependencies in long-distance and *wh*-scope marking questions. However, even the notion of *wh*-chain cannot account for the complementary distribution: it misses the fact that languages tend to utilize either one, or the other strategy, but not both. In a more general perspective, McDaniel's theory, and the DDA on the whole, suffer from a number of conceptual and empirical drawbacks, as discussed in Chapter 5. In particular, it does not explain why in long-distance *wh*-questions the *wh*-phrase must pass through the intermediate Spec-CP(s), but simply encodes the requirement into the notion of *wh*-chain. Furthermore, it does not explain why the *wh*-phrase in (4a) can move to a specifier of a seemingly [-Q] complementizer. Overall, we concluded, along with a number of previous works, that the DDA cannot be adopted as a unified theory for *wh*-scope marking.

It is also important to realize that in the previous accounts, the connection was explored largely in one direction only: the guiding intuition in those accounts seems to be that the *wh*-scope marking strategy is some sort of derivative, or a version of, the long-distance *wh*-strategy. That is, whatever syntactic structure and processes are utilized in long-distance *wh*-questions, must be present in *wh*-scope marking, but not the other way around. We want to argue that the method of inquiry concerning these two types of constructions must be reversed. We take the connection to mean, rather, that the structure of *wh*-scope marking indeed underlies the syntax of long-distance *wh*-questions in which successive cyclic effects are observed.

Bayer (1996) makes a suggestion which is in fact in line with this method of inquiry. He adopts essentially a DDA type approach to *wh*-scope marking, according to which the *wh*-scope marker is base-generated in the 'scope marking' position, Spec-CP. He then extends the base-generation account to *wh*-phrases in general. To implement that, Bayer resorts to the mechanism of generalized transformations in the sense of Chomsky (1995c), Ch.3. Chomsky assumes that a generalized transformation (GT) extends a phrase marker by first creating a 'placeholder', and then inserting another phrase marker into it. Bayer capitalizes on the insertion option. He argues that insertion of the *wh*-scope marker into Spec-CP proceeds via GT. In the case of long-distance *wh*-movement, he suggests, the only 'real' movement of the *wh*-phrase is from its base generated position to the most local Spec-CP. Instead of moving the *wh*-phrase to the next Spec-CP, as the standard successive cyclic analysis would maintain, a copy of the *wh*-phrase is inserted in that Spec-CP by GT, similarly to generation of a *wh*-scope marker. The chain formed by the previous step is 'extended' via assigning the same index on the inserted copy of the *wh* as on the original (locally moved) copy. The same procedure applies in each Comp, including the most matrix one. All copies except the highest are then deleted, creating the effect of long-distance *wh*-movement.

The goal of Bayer's analysis is a respectable one. Bayer's analysis attempts to place the two types of constructions under one concept, thus reducing properties of specific constructions to interaction of general principles. In addition, Bayer aims at explaining perhaps the hardest problem in the modern theory of A'-locality, namely, the successive cyclic effect. However, the analysis is based on a number of

stipulations and remains counterintuitive. First, Bayer has to stipulate that insertion of a *wh*-phrase in the Spec-CP by GT (including movement) marks the complementizer as [+*wh*], to explain why the *wh*-phrase moves even to the most local Spec-CP. Second, it is not clear how the index assignment enforces the identity of copies in Bayer's analysis. Third, Bayer has to assume that the chain formation process is independent of movement. These assumptions are rather difficult to maintain in the current framework.[2]

Let us now turn to our proposal. In Chapter 5, we have established a common factor that explains the correlation between the availability of *wh*-scope marking and the absence of long-distance *wh*-movement in Russian and Polish: both types of constructions utilize the structure of finite complementation in the form of NP-shells. On the one hand, the NP shell renders a 'Complex NP island' for long-distance extraction. On the other hand, an NP-shell (or, rather XP-shell) is the underlying structure of *wh*-scope marking questions. This common factor becomes transparent under the IDA approach to *wh*-scope marking questions. However, at that point we left open the question what is responsible for the grammaticality of long-distance extraction examples in languages like English, given the generality of the NP-shell hypothesis.

In this chapter, we want to take the next step and argue that the NP-shell hypothesis is at the core of the other side of the correlation, too, accounting for languages in which long-distance extraction, but not *wh*-scope marking *per se*, is

---

[2] See also Müller (1996) who proposes an account of *wh*-scope marking questions and long-distance *wh*-questions stated in Optimality Theory.

available. It will be shown that once the second part of the correlation is established, the main property of long-distance movement - its successive cyclic character - follows directly as a residual effect of forming a *wh*-scope marking question, in the sense of the IDA.

3. Head movement of the *wh*-scope marker

In this section we consider several aspects of *wh*-scope marking that will become instrumental in our subsequent discussion of long-distance *wh*-movement.

In Chapter 5 we developed a syntactic version of the IDA whereby the *wh*-scope marker is Merged with the embedded clause of a *wh*-scope marking question; as schematized in (6):

(6) $[_{CP[+Q]} \ldots [_{VP} V \, [_{WP} W_{SM} \, [_{CP[+Q]} \ldots]]]]$ (W_{SM} = *wh*-scope marker)

W_{SM} is a head in our analysis. There are reasons to believe that in a number of *wh*-scope marking languages, W_{SM} undergoes head movement to the higher V, and the $V+W_{SM}$ complex behaves as one structural unit. Hindi is one such language. As observed by Mahajan (1996), adverbs like *abhii* "just now" can generally appear preverbally (among other positions) in Hindi. However, in *wh*-scope marking questions, *abhii* cannot separate the *wh*-scope marker and the verb:

(7)   siitaa-ne kyaa (*abhii) socaa ki   ravii-ne kis-ko dekhaa

       Sita-erg what just now thought that Ravi-erg who   saw

'What did Sita think just now, who did Ravi see?'

Interestingly, in declarative sentences involving sentential pro-form *yeh* preceding the embedded clause, *abhii* can appear between *yeh* and the verb.

(8) siita-ne yeh abhii socaa ki ravii-ne tumhe dekhaa

Sita-erg it just now thought that Ravi-erg you saw

'Sita thought it just now that Ravi saw you'

This suggests that *yeh* and *socaa* do not form a constituent. The simplest account of the difference between (7) and (8) seems to be that *kyaa* is affixal, which suggests that it is a head, but *yeh* is a phrase.[3] This confirms the structural difference between NP-shell in declaratives, and the XP-shell in *wh*-scope marking questions that we postulated for Russian: according to our analysis, the embedded clause is adjoined (postcyclically) to *yeh*.

In Hindi, *kyaa* also has a second meaning, as a yes/no question marker. As Mahajan observes, in this meaning, an adverb can also intervene between it and the verb:

(9) siitaa-ne kyaa kal tumhe dekhaa thaa

Sita-erg Q yesterday you-dat saw be

---

[3] Bayer (1996) who discusses the *wh*-scope marking strategy in Bangla (Bengali) also argues that the *wh*-scope marker in that language is a head.

'Did Sita see you yesterday?'

The V+W$_{SM}$ complex can participate in further movement processes. Wahba (1991) describes the *wh*-scope marking strategy in Iraqi Arabic, as exemplified in (10):

(10) Š-   tsawwarit Mona Ali ištara šeno?

W$_{SM}$ thought   Mona Ali bought what

'What did Mona think Ali buy?'

Wahba points out that in (10) the *wh*-scope marker *š*- must occur in the beginning of the sentence. Crucially, it cannot be separated from the verb, e.g. by the subject:

(11) *Š- Mona tsawwarit Ali ištara   šeno

W$_{SM}$ Mona thought   Ali bought what

Note that subjects can otherwise precede the verb in questions:

(12) Mona šaafat meno?[4]

Mona saw   whom

'Who did Mona see?'

---

[4] According to Wahba (1991), *wh*-phrases can stay in situ, or move in Iraqi Arabic in matrix questions.

A similar situation obtains in Albanian, the language in which the *wh*-scope marking strategy is discussed in Turano (1998). In this language, as in Hindi, the marker *a* can ambiguously be used as a yes/no question particle, or a *wh*-scope marker. In the *wh*-scope marking reading, again, *a* must occur clause-initially, and cannot be separated from the verb by the subject (subjects can usually appear before verb):

(13) a. A   mendon (ti) se   kë   ka takuar Maria?

W$_{SM}$ think-2s you that who has met   Mary

'Who do you think Mary met?'

b. *A   ti mendon se   kë   ka takuar Maria?

Turano (1998) reports that the version with the *wh*-scope marker in the embedded clause is ungrammatical in Albanian:

(14)   *A   mendon a   Maria thotë se kush ka lexuar librin?

W$_{SM}$ think   W$_{SM}$ Mary says that who has read book

'Who do you think that Mary says read the book?

Turano argues that (14) is ungrammatical because the *wh*-scope marker, a [+*wh*] element, is inserted into a Spec-CP which is not interrogative. Note, however, that the sentence improves significantly if the order of subject and verb is reversed in the embedded sentence:

(15) ?A mendon a   thotë Maria se kush ka lexuar librin?   [Dalina Kalluli, p.c.]

W$_{SM}$ think W$_{SM}$ says Mary that who has read book

This suggests that Turano's explanation cannot be correct: rather, (14) is ungrammatical because the embedded subject *Maria* illegally intervenes between *a* and the verb. At the same time, this strengthens the point that *a* and verb form an unbreakable *a*+Verb complex. As Dalina Kallulli (p. c.) points out to us, indeed certain items can intervene between *a* and the verb, but those are crucially restricted to a) clitics; b) aspectual markers, in other words, other heads. This is illustrated below:

(16)   A   e   imagjinon (dot) (ti) se   kë   ka takuar Maria?

W$_{SM}$ cl-acc imagine   (at all) (you) that who-acc. has met Mary

'Who do you imagine/think Mary has met?'

(17)   A   po   mendon (ti) se   kë   ka takuar Maria?

W$_{SM}$ prog think you that who-acc. has met Mary

'Who are you thinking that Mary has met?'

In (16), the clitic doubling the embedded question intervenes between *a* and the verb; in (17), it is the progressive marker. This pattern is consistent with the head adjunction analysis, in which W$_{SM}$ adjoins to V, along with other relevant head(s).

It is not the case, however, that the *wh*-scope marker adjoins to V in overt syntax in all *wh*-scope marking languages. In Russian (and Polish), in particular, the *wh*-scope marker can be separated from the verb, as can be seen in standard cases:

(18)   **Kak** vy dumaete, **kogo** ljubit Ivan?

how you think   whom loves John

'What do you think, who does John love?'

In German, too, the *wh*-scope marker does not form a unit with the verb, as can be seen from an example of an embedded *wh*-scope marking question (from Höhle (1996)):[5]

(19)   Heinz möchte wissen, was du   glaubst, wer Recht hat

Heinz wants know   what you think   who right was

'Heinz wants to know who do you think was right'

The question then arises, what is responsible for this parametric difference in the behavior of the *wh*-scope marker.

Let us consider why the *wh*-scope marker raises in languages in which it does. The possibility we would like to explore is that this raising correlates with the <u>affixal</u> nature of the *wh*-scope marker. In other words, we propose the parametric difference

---

[5] In matrix *wh*-scope marking questions, the *wh*-scope marker is located next to the verb, and nothing can be inserted in between, but this is simply because of the matrix V2 effect in German.

in question is reduced to the lexical nature of the *wh*-scope marker: Hindi *kyaa*, Iraqi Arabic *š-* and Albanian *a* in the *wh*-scope marking meaning are in fact affixes, while Russian *kak*, Polish *jak*, and German *was* are not affixes, but independent lexical items. The latter also include, apparently, Hungarian *mit* and Romani *so* (cf. Chapter 5).

The characteristic descriptive property of affixes is that they cannot be morphologically independent, but, rather, need a host to attach to. Recognizing the existence of affixes as separate heads participating in syntactic computation is not in the spirit of the strictly lexicalist view, which goes back to Chomsky (1970) and is revived in the Minimalist framework in Chomsky (1995c), Ch.3. According to the lexicalist view, lexical items are inserted from the lexicon into the derivation fully inflected, and their inflectional features are 'checked' in the course of a syntactic derivation. However, several authors, notably Lasnik (1995b), Bobaljik (1995a) adduce strong empirical arguments which suggest that the lexicalist view should be relaxed, so as to allow certain affixes to participate in the syntactic derivation. In postulating *wh*-scope markers in the above languages as verbal affixes, we follow these authors, for whom verbal affixes such as English past tense *-ed* occupy a head position.

How does the affixal property of a *wh*-scope marker correlate with its raising?[6] Following Bobaljik (1995a) and Bošković (2001) we assume that the

---

[6] In the preceding discussion, we implicitly assumed that the raising in question is to V, so as to form a head adjunction complex. In fact, this does not have to be so. The data above are compatible with an analysis in which the raising at issue is not to V, but to a functional category above it (e.g. *v*), to which V also adjoins. Most importantly, though, we want to maintain that a segment of V and the *wh*-scope marker are sisters.

morphological/PF component serves as a filter for syntactic derivations; that is, syntactic derivations that yield an output that cannot be interpreted by morphology/PF, are ruled out. There must be a formal feature driving raising of the *wh*-scope marker in syntax (call it F), but this raising feeds satisfaction of its affixal requirement at PF (cf. also 'merger' in the sense of Halle and Marantz, another operation that joins terminal nodes under a head level category; cf. also (syntactic) 'incorporation', in the sense of Baker (1988)).

Note that raising of the *wh*-scope marker in syntax does not change the compositional way the semantics of a *wh*-scope marking question is computed (cf. (60) of Chapter 5). Recall that there we argued that it is the trace of the *wh*-scope marker that is interpreted at LF in Russian and Polish, rather than the (copy of the) *wh*-scope marker in the moved position. The only adjustment that needs to be made is that even in *wh*-in situ languages like Hindi, it is the trace of the *wh*-scope marker that is interpreted at LF, rather than the scope marker itself, in order to get the semantics of the *wh*-scope marking.

4. 'Long-distance' movement: A proposal

Under our analysis in Chapter 4, it follows that finite complements even in English are realized as NP-shells. More precisely, we assume that the object of *believe* is a silent correlative, to which the embedded finite clause is adjoined (postcyclically, cf. Chapter 3): in other words, finite complements have the structure as shown below:[7]

---

[7] To avoid unnecessary complications at this point, we consider an example not involving overt C *that*.

(20)   John believes [NP [NP ∅] [CP Peter likes Mary]]

The question now is, how the standard long-distance *wh*-question (21) is derived:

(21)   Who do you believe John likes?

Given (20), (21) must be ruled out for the same reason as in Russian, as a violation of a 'Complex NP island'. This is expected under the assumption that the structure of the finite complement by the time of *wh*-extraction is the same in (20) and (21). We would like to maintain that this is indeed the case, and the derivation in which the *that*-complement has the NP-shell structure in (21) is indeed canceled for violating of principles responsible for a 'Complex NP island' effect (see Chapter 3). We suggest, rather, that there is another derivation of (21) which converges successfully, and in which the 'complex NP effect' is voided. Let us consider this alternative derivation.

We argued in Chapter 5 that for Russian and Polish, the NP-shell structure is a 'prerequisite' for the *wh*-scope marking structure, under the IDA approach. Specifically, the derivation of a *wh*-scope marking question (22) in overt syntax in Russian includes the steps in (23):

(22)   **Kak** vy dumaete, **kogo** ljubit Ivan?

how you think   whom loves John

'What do you think, who does John love?'

(23) a. Build   [$_{C'}$ Ivan ljubit **kogo**]

                                                                 I.   loves who

b. *Wh*-move *kogo*   [$_{CP}$ **kogo** Ivan ljubit t$_{kogo}$]

                                                                 who I.   loves

c. Merge a *wh*-scope marker   [$_{XP}$ **kak** [$_{CP}$ **kogo** Ivan ljubit t$_{kogo}$]

                                                         how   who I.   loves

d. Merge matrix V   dumaete [$_{XP}$ **kak** [$_{CP}$ **kogo** Ivan ljubit t$_{kogo}$]]

                                               think   how   who I   loves

e. Continue building   [$_{C'}$ vy dumaete [$_{XP}$ **kak** [$_{CP}$ **kogo** Ivan ljubit t$_{kogo}$]]

                                               you think   how   who I.   loves

f. *Wh*-move the *wh*-scope marker [$_{CP}$ **kak** vy dumaete [$_{XP}$ t$_{kak}$ [**kogo** Ivan ljubit t$_{kogo}$]]

                                               how you think   who I.   loves

Given that finite complements are NP-shells in English as well, suppose that the *wh*-scope marking structure, in the sense of the IDA, is utilized in English too. Assume that there exists a *wh*-scope marker in English and it is a phonologically silent head, as in Basque (see Chapter 5). Let us refer to the *wh*-scope marker as WHAT. In accordance with the general schema in (6), the *wh*-scope marker forms a constituent with the embedded clause in a *wh*-scope marking question. It follows that the derivation of a 'long-distance' *wh*-question in English begins essentially as a derivation of a *wh*-scope marking question, shown in (23a)-(23d). The structure of (21) at the point of Merging the matrix verb is thus:

(24) believe [$_{XP}$ WHAT [$_{CP[+Q]}$ who John likes t]]

We suggest that $\varnothing_{[+wh]}$ in English is affixal, similarly to Hindi *kyaa*, Iraqi Arabic *š-* and Albanian *a*. Consequently, the *wh*-scope marker raises to adjoin to the matrix V, at the point illustrated in (24). Here we follow to a large extent Müller (1995), Müller and Sternefeld (1995) (reviewed briefly in Chapter 4), who, like us, also adopt the NP-shell hypothesis even for English and further argue that that the head of the finite 'shell' can 'incorporate' into the matrix verb. We depart from these authors, however, in two important respects. First, for us the element that undergoes incorporation is the silent *wh*-scope marker WHAT, while in these works, it is the empty pronominal of an NP-shell, of the kind found in declarative sentences (cf. (20)). Second, in the above works, the incorporation takes place at LF, while we claim that incorporation takes place in overt syntax.

The 'incorporation' of the *wh*-scope marker in (24) into V results in (25):

(25) [$_{VP}$ WHAT+believe [$_{XP}$ t$_{WHAT}$ [$_{CP}$ **who** John likes t$_{who}$]]]

(26) shows the derivational stage at which the matrix interrogative C$_{[+Q]}$ is Merged:[8]

(26) [$_{C'}$ C$_{[+Q]}$ do you [$_{vP}$ [$_{VP}$ WHAT+*believe* [$_{XP}$ t$_{WHAT}$ [$_{CP}$ **who** John likes t$_{who}$]]

---

[8] The complex WHAT+*believe* can further adjoin to *v*, in the course of the derivation. As far as we can see, nothing in our analysis depends on whether it does. For simplicity, though, we assume that the complex does not raise.

$C_{[+Q]}$ seeks a matching [+wh] feature in order to attract it to its specifier (or establish the Agree relation, in terms of Chomsky (2000)). WHAT must bear the wh-feature (it is clear, from the previous discussion, that wh-scope markers in overt wh-movement languages typically behave like regular wh-phrases with respect to wh-movement). Hence, WHAT+believe is a candidate for attraction.

Note that in (26) who has already moved to a specifier of the intermediate CP, satisfying the property of the interrogative C, with its [+wh] feature. Chomsky (1995c) suggests that certain kinds of features can be checked more than once. Suppose that this is true for the [+wh] feature. That is, suppose that who in the intermediate Spec-CP can potentially participate in further checking operations. Observe also that in the moved position, WHAT does not c-command who (under the definition of c-command in terms of the first branching node; cf. Chapter 3). Hence, by Closeness defined in terms of c-command (Chomsky (1995c)), WHAT is not closer to $C_{[+Q]}$ than who. Furthermore, we assume that head movement traces, similarly to traces of A-movement (Chomsky (1995c)), do not count as potential intervenors. In other words, head movement traces cannot participate in attraction/Agree. Thus, a trace of WHAT in the original position does not intervene between $C_{[+Q]}$ and who. That is, who is another candidate for attraction into the matrix Spec-CP.

One continuation of this derivation thus involves attraction (or Agree) between $C_{[+Q]}$ and the [+wh] of who. In the framework of Chomsky (2000), the [+Q] feature of C has the 'EPP property', requiring to project an overt specifier. Who therefore raises to the matrix Spec-CP, satisfying the 'EPP property' of C. Movement of who from the embedded Spec-CP

results in (27), which corresponds to the word order in (21). This is a usual long-distance *wh*-question:

(27)   [c· **Who** C[+Q] do you [vP [vP WHAT+*believe* [XP tWHAT [CP twho John likes twho]]

Now suppose C[+Q] establishes Agree relation with the [+*wh*] feature of the WHAT+*believe* complex. If the complex then raises to adjoin to C[+Q], this results in the ungrammatical (28a), with the structure roughly in (28b).

(28)   a. *believe you who Peter likes?

        b. [CP [C WHAT+believe] you [XP tWHAT [CP who Peter likes twho]]]

Arguably, a similar derivation converges in Iraqi Arabic and Albanian (cf. Section 3). Some parametric property must then distinguish between (28a) and their grammatical counterparts in these languages. Given that English is an overt *wh*-movement language, it seems safe to assume that the 'EPP property' of the interrogative C in English *wh*-questions is satisfiable only by XPs.[9] (28b) then is ruled out since the (XP-type) EPP property is not satisfied in this case. We tentatively suggest that C in Iraqi Arabic and Albanian is of the type satisfiable by X[0] elements (see also Chapter 2 for discussion of the possibility of 'EPP' to be satisfied by heads, and references there.).

---

[9] See also Bošković (2001) who suggests that heads may be lexically specified concerning how their features are checked: in the Spec-head configuration, or via head adjunction.

Thus the absence of what is standardly called the *wh*-scope marking strategy (of the 'in situ' type found in Hindi and Russian) in languages like English is now explained simply by the fact that English C has the 'EPP property' that is satisfiable only by XPs. At the same time, the presence of the long-distance *wh*-movement strategy is due to a) the fact that the silent *wh*-scope marker undergoes incorporation to V, and b) the fact that the *wh*-phrase in the embedded clause is a suitable candidate for checking the EPP property of the matrix C.

Let us now consider why languages like Russian have overt *wh*-scope marking, but do not have long-distance *wh*-questions. Consider (29) ((29a) = (18)):

(29)   a. **Kak** vy dumaete, **kogo** ljubit Ivan?

          how you think   whom loves John

          'What do you think, who does John love?'

          b. ?*Kogo vy dumaete cto Ivan ljubit?

          Who you think   that John loves

          'Who do you think that John loves)

Recall that *kak* is not an affix, hence, does not raise to V. Thus at the point the matrix interrogative C is inserted, the structure is as follows (cf. also Chapter 5):

(30)   [$_{CP}$ C$_{[+Q]}$ vy dumaete [$_{XP}$ **kak** [**kogo** ljubit Ivan]]]

          you think   how whom loves John

Since *kak* did not incorporate, it c-commands *kogo*; hence, the [+*wh*] feature of *kak* serves as an intervenor for the [+wh] feature of *kogo*. *Kak* then is the closest candidate for *wh*-fronting.

Thus, syntactic *wh*-scope marking in place of the long-distance *wh*-movement is due to the fact that the *wh*-scope marker is not affixal. Consequently, no incorporation takes place, and the *wh*-phrase in the embedded clause is 'invisible' for further attraction.

Our proposal concerning deriving the long-distance structure from the *wh*-scope marking structure does not make new predictions about the interpretation of long-distance *wh*-questions.[10] The standard Hamblin-type LF and interpretation of (31), for instance, is in (32a) and (32b), respectively:

(31) Who did John say that Peter likes?

(32) a. LF: $\lambda p \exists x$ [person(x)(w) & p=λw'.John said that Peter likes x in w'], where p is

    a   proposition, and w and w' are variables over possible worlds

    b. {John said that Peter likes Mary, John said that Peter likes Sue, John said that

        Peter likes Molly, etc.}

Under the standard view, (32) is derived on the assumption that the embedded clause *that Peter likes x*, where *x* is the trace of *who*, is an open sentence, a set of possible worlds in which Peter likes *x*. The compositional denotation of the matrix VP is

---

[10] We thank Penka Stateva for very helpful discussions on this point.

derived by functional application: the matrix verb takes the proposition denoted by the embedded clause as an argument, and then combines with the subject, to derive the interpretation of the IP. The IP node denotes another proposition. The complementizer C then converts it into a set of propositions, i.e. a question of type $<<s,t>,t>$ (see Chapter 5). The *wh*-phrase *who* which has moved to Spec-CP has created an index which binds the free variable x in the embedded clause (see Heim and Kratzer (1998)). Finally, *who* combines with the C' to yield the interpretation in (32), which is a question of type $<<s,t>,t>$.

Now, under the present analysis, the embedded clause is in fact a question, that is, a set of propositions (type $<<s,t>,t>$). Obviously, combining the question denotation with the denotation of the matrix verb would lead to a type mismatch. This reflects the empirical fact that matrix verbs in long-distance (and *wh*-scope marking) questions are exactly those that do not select questions. But note that in our analysis, it is not the verb that must combine with the embedded clause. The incorporation of the *wh*-scope marker into the matrix verb has a semantic effect, such that the *wh*-scope marker and the verb form a complex predicate, and it is this complex predicate that combines with a question in the embedded clause, as seen in (27). The semantic component combines the denotation of this complex predicate and the question denotation of the embedded clause, so as to yield the usual denotation of the VP node. From this point, the computation proceeds in the manner characteristic of the standard view, described above. Thus we do not have to re-evaluate the standard semantics of the long-distance *wh*-constructions, but, rather, we only require a slight adjustment of its compositional analysis.

One may now notice an apparent discrepancy concerning which copy of the incorporated *wh*-scope marker actually matters for interpretation. Recall again that in *wh*-scope marking questions in which the *wh*-scope marker incorporates (Russian, Hindi, Iraqi Arabic, Albanian), it is <u>not</u> the *wh*-scope marker itself, but, rather, its trace that is interpreted at LF. In particular, the denotation of the trace combines with the denotation of the embedded question, to yield the denotation of the XP, which later undergoes QR (cf. (60) of Chapter 5). For long-distance *wh*-questions, we are now claiming that the *wh*-scope marker <u>itself</u> is interpreted at LF. Specifically, its denotation is combined with the denotation of the matrix verb with which the *wh*-scope marker has incorporated, to yield the interpretation of the complex predicate. Is there a principled distinction concerning the issue of which copy of *wh*-scope marker is interpreted at LF?

We suggest that there is no need to regulate which copy gets interpreted in each case. Either copy of the *wh*-scope marker can be interpreted in both long-distance, and *wh*-scope marking questions. Consider the relevant cases.

Consider the Hindi example again:

(33)   siitaa-ne **kyaa** socaa   ki   ravii-ne **kis-ko** dekhaa

Sita-erg what thought that Ravi-erg who   saw

'What did Sita think, who Ravi saw?'

In Hindi *wh*-scope marking questions, as we showed above, the *wh*-scope marker incorporates into the verb. If the trace of the incorporated *wh*-scope marker is

interpreted, we arrive at the semantics of *wh*-scope marking, along the lines of Chapter 5. Suppose the *wh*-scope marker itself is interpreted. Then it forms a complex predicate with the matrix verb. This complex predicates combines with the embedded question. So far the computation proceeds as in English. If *kis-ko* moves, overtly or covertly to Spec-CP in (33), we would end up with the structure (27) for (33). In other words, we would interpret (33) as a long-distance question. But recall from Chapter 4 that *wh*-phrases cannot be extracted from Hindi, either overtly or covertly (Dayal (1996a)). Hence, *kis-ko* does not move to the matrix Spec-CP in (33). Consequently, the structure is uninterpretable.

Consider now (21) in English, with the structure in (27), repeated here:

(21) Who do you believe John likes?

(27) [c′ **Who** C$_{[+Q]}$ do you [vP [vP WHAT+*believe* [xp twHAT [cp t$_{who}$ John likes t$_{who}$]]

As suggested before, the higher copy of WHAT, rather than its trace, is interpreted in (27), in order to get the long-distance interpretation. Suppose now that the trace of the *wh*-scope marker in the position of head of the XP is interpreted (cf. (24)). Suppose also we interpret the copy of the *wh*-phrase in the Spec of the intermediate CP, rather than the highest one. Thus, instead of formation of the complex predicate, the denotation of the trace combines with the denotation of the embedded question, to yield the denotation of the XP, which later undergoes QR (cf. (60) of Chapter 5). In other words, we will interpret (21), or any English long-distance *wh*-question, as a *wh*-scope marking question. That is, we are now claiming that English long-distance

wh-questions have two different LFs, one corresponding to the 'usual' long-distance structure (cf. (32a)), and the other corresponding to the *wh*-scope marking structure. This claim has little empirical repercussions, however. Empirically, the semantics of long-distance questions, and the semantics of *wh*-scope marking questions are impossible to distinguish, so this outcome is in fact consistent with the speakers' intuitions.

5. Consequences

From the resulting picture, the complementary distribution of the (syntactic) *wh*-scope marking and long-distance strategies mentioned in Section 2 is reduced to a simple morpho-lexical parameter, namely, whether the *wh*-scope marker is affixal in a given language.

As mentioned in Section 2, in certain varieties of German both long-distance and *wh*-scope marking questions are available. Let us refer to that dialectal group as German$_1$. We repeat examples in (4) below:

(4) a. **Was** glaubst du **wen** sie liebt?

   What think you who she loves

   'What do you think, who does she love?'

   b. **Wen**   glaubst du daß sie liebt?

   Whom think   you that she loves

   'Who do you think that she loves?'

As (4) shows, German₁ appears to violate the complementarity of the long-distance and *wh*-scope marking strategies. In particular, it seems to allow both the Russian option, and the English option. From what we said so far, it follows that the German₁ *wh*-scope marker *was* is both affixal and non-affixal, a seemingly paradoxical situation.

The paradox can easily be resolved, however, if we postulate that German₁ in fact entertains <u>two</u> *wh*-scope markers in the lexicon. Conceptually, nothing precludes this possibility: it is well known that languages allow overt and null counterpart of the same token: pronouns in *pro*-drop languages are one example. Suppose that in addition to *was*, there exists another *wh*-scope marker in German, which is exactly like *was*, but differs from it in that it is phonologically silent; we dub it WAS. In contrast to the non-affix *was*, utilized in *wh*-scope marking languages, WAS is in fact an affix. Consequently, it undergoes raising to V. The derivation of the long-distance example (4) is then the same as in English, and proceeds along the lines outlined above.

This proposal makes two predictions. First, if WAS is exactly like *was*, we expect it to occur exactly with the same matrix verbs as *was*. In other words, the set of verbs supporting *wh*-scope marking and those supporting long-distance movement in German₁ must be co-extensive. This prediction is borne out. A number of contributions in (Lutz and Müller (1996)) note that *wh*-scope marking in German is possible only with 'bridge' verbs, which also license long-distance *wh*-extraction (see also Chapter 5 on German)

The second prediction is that it should be possible to combine both *wh*-scope markers in the same sentence in German₁. That this is indeed the case is suggested by grammaticality of examples like (34a), which, in our analysis, has the structure as in (34b):[11]

(34)   a. **Was** glaubst du **mit wem** Hans meint daß Jakob gesprochen hat?

           What think you with whom Hans thinks that Jakob talked   has

           What do you think, with whom does Hans think that Jakob talked?

    b. [$_{CP1}$ **Was** glaubst du [t$_{was}$ [$_{CP2}$ **mit wem** Hans WAS+meint [t$_{WAS}$ [$_{CP3}$ t'$_{mitwem}$ daß Jakob t$_{mit\ wem}$ gesprochen hat]]]]]

The *wh*-phrase *mit wem* originates in the most embedded CP3 as an argument of *gesprochen* and subsequently *wh*-moves to local Spec-CP3, in accord with the general pattern of *wh*-movement. The silent WAS is Merged with CP3 and subsequently adjoins to V, giving way to further *wh*-movement of *mit wem* from Spec-CP3 to Spec-CP2. Overt *was* is then Merged with CP2, and later *wh*-moves to the interrogative C. The verb *glaubst* is also in C, but it followed its own path, due to a matrix V2 requirement in German.

In other varieties of German, long-distance *wh*-extraction of the sort in (4b) above is not allowed, even though the *wh*-scope marking structure (4a) still is. Let us

---

[11] As several authors in Müller (1996) report, some speakers also accept the following:
(i)   **Was** meinst du daß sie gesagt hat **wann** sie kommen würde?
       what think you that she said has when she come   would
       'What do you think that she said, when she would come?'
To the extent (i) is acceptable, we suspect that both *wh*-scope markers are involved here as well.

refer to this variety as German$_2$. German$_2$ thus patterns with languages like Russian and Polish in the relevant respect. Under our analysis, the only difference between German$_2$ and German$_1$ is that there is only one *wh*-scope marker *was*. There is no *wh*-scope marker like WAS. Interestingly, most of those speakers that do not accept long-distance *wh*-movement in general, also do not accept instances of *wh*-scope marking as in (34a). This is expected, since nothing would support long-distance *wh*-extraction from Spec-CP3 to Spec-CP2 in those dialects.

A similar dialectal split seems to obtain in Hungarian: while all dialects apparently have the *wh*-scope marker strategy, only some dialects allow long-distance *wh*-extraction (cf. Chapter 4). Grammatical examples of long-distance *wh*-extraction are reported, in particular, in Marácz (1989), Kiss (1987), p.c., and in reference to J. Horvath. Some of our informants, however, do not accept long-distance *wh*-extraction. Consider the following:[12]

(35)   a. Mit   gondolsz   hogy ki   látta Jánost

        what-acc. you-think   that who saw   John-acc.

        'What do you think, who saw John'

      b. #Kit   gondolsz   hogy látta Jánost?

        who-acc you-think   that saw John-acc

        'Who do you think that John saw?'

---

[12] Note that the extracted *wh*-subject in (35b) bears Accusative case. This situation is typical for subject extraction in Hungarian, and will not concern us here.

(35b) demonstrates that, similarly to German, certain varieties of Hungarian allow both Russian option and the English option. Recall that the *wh*-scope marker *mit* in Hungarian is not an affix. Under the proposal we were developing, the existence of these dialects is accounted for by positing a null counterpart of the *wh*-scope marker, viz. MIT. In contrast, in the dialects that do not allow (35b) MIT is not available. Consequently, the long-distance *wh*-extraction is not allowed.

The dialectal split seems to hold with regard to the counterpart of (34a) in (36): it is reported grammatical in those sources which also report (35b) as grammatical. In contrat, those Hungarian speakers, which do not accept long-distance *wh*-extraction in general, (36) is degraded:

(36) #Mit   gondolsz hogy Mari kit   mondott hogy látta Jánost?

what-acc you-think that Mary who-acc said   that saw John

'What do you think, who did Mary say saw John?'

Thus the relevant parametric difference between languages that do and do not have long-distance *wh*-extraction, is reduced to a difference in the lexical inventory, always a welcome move under the Minimalist perspective.[13]

Note that in positing the silent counterpart of the *wh*-scope marker in languages that do allow long-distance *wh*-extraction it was crucial for us that it have an affixal character, which, by hypothesis, justifies its 'incorporation' into the matrix V. One

---

[13] We found no speakers which accept long-distance *wh*-extraction, but not *wh*-scope marking, in German. The *wh*-scope marker *was* seems to be available in all dialects.

may wonder whether the consistently affixal character of the silent *wh*-scope marker cross-linguistically is simply a coincidence, or a reflection of a principle. In this connection, Ormazabal (1995) argues independently that all null heads are affixes (see also Pesetsky (1992) for related discussion; and Bošković (1997c) for criticism of this view). If Ormazabal's view is correct, it suggests a correlation between the affixal character of a *wh*-scope marker and its phonologically null status.

A number of dialects of German allow the following counterpart of the long-distance *wh*-questions, in addition to *wh*-scope marking (cf. McDaniel (1989), Fanselow and Mahajan (1996), examples from the latter):[14]

(37)   a. Wen denkst Du wen sie meint   wen Harald liebt

who think you who she believes who Harald loves

'Who do you think that she believes that Harald loves?'

b. Wovon glaubst du wovon sie träumt?

of-what think you of-what she dreams

'What do you thing that she dreams of?'

In (37), the contentful *wh*-phrase *wen* occurs in the matrix Spec-CP, but also in intermediate Spec-CPs. It is generally believed (see the above works) that questions such as (37) which we, following the standard practice, will refer to as copy

---

[14] The copy construction with these properties is attested, in addition to German, in Frisian (Hiemstra (1986)), Afrikaans (Plessis (1977)), Romani (McDaniel (1989)), and child English (Thornton (1990)).

construction, should receive an analysis similar to that of *wh*-scope marking questions (see the above works for discussion).

(37) presents no problem for our approach. We analyze (37) as an instance of long-distance movement of *wen*, and incorporation of the silent *wh*-scope marker WAS, along the lines above. Thus (37a) has the structure in (38):

(38)   Wen WAS$_2$+denkst Du [t$_{WAS2}$[wen sie WAS$_1$+meint [t$_{WAS1}$[wen Harald t$_{wen}$

        who think you   who she   believes   who Harald

        liebt]]]]

        loves

Under our analysis (and under most traditional ones), it must be the case that *wen* in the matrix and the embedded clauses are copies of the same item; otherwise, there is no place to generate *wen* in the matrix and the first embedded clause. The fact that the (copies of the) *wh*-phrases are located in Spec-CP is in line with the general pattern of overt *wh*-movement in this language. This is expected in our analysis, given that all intermediate CPs are in fact questions, that is, intermediate Cs have a [+Q] feature. See Fanselow and Mahajan (1996) and Nunes (1999) for proposals concerning the issue of why the non-highest copies are not deleted in (38).[15]

---

[15] Fanselow and Mahajan (1996) and Nunes (1999) attribute non-deletion of intermediate copies in (37) to their 'cliticization' to each intermediate the complementizer C^0, by head adjunction, which, arguably, makes the copy 'invisible' for the deletion procedure. This cliticization is seen overtly in Frisian (Hiemstra (1986):

(i)   Wa tinke jo   wa't   il sjoen haw?
        who think you who-that I seen have
        'Who do you think that I have seen?'

6. Conclusion

In this chapter we advanced an analysis of long-distance *wh*-questions in languages like English, and related constructions. This analysis is advantageous in several respects. First, it provides an explanation of the successive cyclic property of long-distance *wh*-constructions, as a residual effect of the syntactic structure of *wh*-scope marking understood in the sense of Indirect Dependency (cf. Chapter 5). It also places apparently unrelated construction types - long-distance and *wh*-scope marking questions - under one conceptual umbrella, giving an account to their complementarity in certain languages, and coexistence in others. The analysis also reduces parametric variation to differences in morpholexical inventories of languages, which is the most natural locus of variation from the point of view of the Minimalist program.

A number of issues remain to be addressed, in light of the proposed analysis. One question is whether and how the analysis of *wh*-movement proposed in this chapter extends to other types of A'-movement, usually claimed to be successive cyclic, in particular, topicalization (cf. Chomsky (1977), Lasnik and Saito (1992)). Another task is to spell out a compositional analysis of long-distance *wh*-constructions, taking into account the adjusted syntactic structure and the semantic effect of adjunction of the *wh*-scope marker to V. Yet another question concerns a more precise

---

Note that these proposals correctly explain the fact that the copy construction is impossible if the *wh*-phrase is a nonminimal maximal projection, viz. a *which*-phrase; a *which*-phrase cannot cliticize to C:
(ii) *Welchen Mann glaubst du   welchen Mann sie liebt?
       which   man   believe you which   man   she loves
       'Which man do you believe that she loves?'

characterization of properties of interrogative Cs in various languages, with respect to the kind of element (X^0 or XP) that can satisfy the EPP property of C in each language.

To the extent an analysis along the lines of this chapter, and of this dissertation, is on the right track, it further strengthens the fundamental idea that the local character of A'-movement, in particular, *wh*-movement, does not need to be forced by imposing external constraints. Every step of movement is independently motivated, in accord with principles such as Last Resort. Locality arises as a by-product of the derivational procedure, or, alternatively, certain interface conditions (cf. Chapter 2), in accord with the Minimalist design of grammar.

References

Alexiadou, Artemis, and Elena Anagnostopoulou. 1998. Parameterizing AGR: Word Order, V-Movement and EPP-Checking. *Natural Language and Linguistic Theory* 16: 491-539.

Alexiadou, Artemis, and Elena Anagnostopoulou. 1999. EPP without Spec, IP. In *Specifiers: Minimalist approaches*, ed. David Adger et al., 93-109. Oxford: Oxford University Press.

Avrutin, Sergey, and Maria Babyonyshev. 1994. Obviation in subjunctive clauses and Agr: Evidence from Russian. In *Proceedings of the II Annual Workshop on Formal Approaches to Slavic Linguistics*, ed. Sergey Avrutin et al., 1-24. Michigan Slavic Publications, Ann Arbor.

Bach, Emmon, and George M. Horn. 1976. Remarks on 'Conditions on Transformations'. *Linguistic-Inquiry* 7: 265-299.

Bailyn, John F. 1995. *A configurational approach to Russian "free" word order*. Doctoral dissertation, Cornell University, Ithaca, NY.

Baker, Mark. 1988. *Incorporation: A theory of grammatical function*. Chicago: University of Chicago press.

Barss, Andrew, Ken Hale, Ellavina Tsosie Perkins, and Margaret Speas. 1991. Logical Form and barriers in Navajo. In *Logical structure and linguistic structure*, ed. C.-T. James Huang and Robert May, 25-47. Dordrecht: Kluwer.

Barss, Andrew, and Howard Lasnik. 1986. A note on anaphora and double objects. *Linguistic Inquiry* 17: 347-354. [Reprinted in Essays on anaphora, Lasnik, Howard 1989, 143-148 Dorrecht: Kluwer]

Barwise, J., and R. Cooper. 1981. Generalized quantifiers and natural language. *Linguistics and philosophy* 4: 159-220.

Bayer, Josef. 1996. *Directionality and Logical Form*. Dordrecht: Kluwer.

Beck, Sigrid. 1996. Quantified structures as barriers for LF movement. *Natural Language Semantics* 4: 1-56.

Beck, Sigrid, and Stephen Berman. 1996. Wh-scope marking: Direct vs. indirect dependency. In *Papers on wh-scope marking*, ed. Uli Lutz and Gereon Müller,

315

59-83. Stuttgart/Tübingen: Universität Stuttgart/Universität Tübingen/IBM Deutschland.

Belletti, Adriana, and Luigi Rizzi. 1988. Psych-verbs and theta theory. *Natural Language and Linguistic Theory* 6: 291-352.

Benedicto, Elena. 1993. AGR, phi-features and V-movement: identifying pro. In *University of Massachussetts Occasional papers*.

Bobaljik, Jonathan. 1995a. *Morphosyntax: The syntax of verbal inflection*. Doctoral dissertation, MIT, Cambridge, Mass.

Bobaljik, Jonathan D. 1995b. In terms of Merge: Copy and head movement. In *Papers on minimalist syntax, MIT working papers in linguistics* 27, ed. Rob Pensalfini and Hiroyuki Ura, 41-64. MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, Mass.

Bobaljik, Jonathan D., and Samuel Brown. 1997. Interarboreal operations: Head movement and the Extension requirement. *Linguistic Inquiry* 28: 345-356.

Boeckx, Cedric. 1998. Agreement constraints in Icelandic and elsewhere. In *Working papers in Scandinavian syntax* 62, 1-35, Lund, Sweden.

Boeckx, Cedric, and Sandra Stjepanovic. 2001. Heading toward PF. *Linguistic Inquiry* 32: 345-355.

Borer, Hagit. 1994. The projection of arguments. In *University of Massachusetts Occasional Papers 17: Functional projections*, ed. Elena E. Benedicto and Jeffrey T. Runner, 19-47. GLSA, University of Massachusetts, Amherst.

Bošković, Željko. 1994. D-structure, theta-criterion, and movement into theta-positions. *Linguistic Analysis* 24: 247-286.

Bošković, Željko. 1995. Case properties of clauses and the Greed principle. *Studia Linguistica* 49: 32-53.

Bošković, Željko. 1997a. Fronting *wh*-phrases in Serbo-Croatian. In *Proceedings of the V Annual Workshop on Formal Approaches to Slavic Linguistics*, ed. Martina Lindseth and Steven Franks, 86-107. Michigan Slavic Publications, Ann Arbor.

Bošković, Željko. 1997b. On certain violations of the superiority, AgrO, and economy of derivation. *Journal of Linguistics* 33: 227-254.

Bošković, Željko. 1997c. *The syntax of nonfinite complementation: An economy approach*. Cambridge, Mass.: MIT Press.

Bošković, Željko. 1998a. LF movement and the Minimalist Program. In *Proceedings of the North Eastern Linguistic Society* 28, 43-57. GLSA, University of Massachusetts, Amherst.

Bošković, Željko. 1998b. *On the interpretation of multiple questions*. Ms., University of Connecticut, Storrs.

Bošković, Željko. 1999. On multiple feature checking: Multiple *wh*-fronting and multiple head movement. In *Working minimalism*, ed. Samuel D. Epstein and Norbert Hornstein, 159-187. Cambridge, MA: MIT Press.

Bošković, Željko. 2001. *On the nature of syntax-phonology interface: Cliticization in Slavic*. Elsevier.

Bošković, Željko, and Howard Lasnik. 1999. How strict is the cycle? *Linguistic Inquiry* 30: 691-703.

Bošković, Željko, and Daiko Takahashi. 1998. Scrambling and last resort. *Linguistic Inquiry* 29: 347-366.

Bowers, John. 1988. Extended X-bar theory, the ECP, and the Left branch condition. In *Proceedings of the West Coast Conference on Formal Linguistics* 6, 47-62. Stanford Linguistic Association, Stanford University, Stanford, Calif.

Bowers, John. 1993. The syntax of predication. *Linguistic Inquiry* 24: 591-656.

Brecht, Richard D. 1972. *Problems of deixis and hypotaxis in Russian: Towards a theory of complementation*. Doctoral dissertation, Harvard University, Cambridge.

Bresnan, Joan. 1994. Locative inversion and the architecture of Universal Grammar. *Language* 70: 72-131.

Bresnan, Joan W. 1976. Evidence for a theory of unbounded transformations. *Linguistic Analysis* 2: 353-393.

Browning, Maggy. 1991. Bounding conditions on representation. *Linguistic Inquiry* 22: 541-562.

Burrow, J. A., and Thorlac Turville-Petre. 1992. *A book of Middle English.* Cambridge, MA: Blackwell.

Cattell, Ray. 1976. Constraints on movement rules. *Language* 52: 18-50.

Čavar, Damir, and Gisbert Fanselow. 1997. Split constituents in Germanic and Slavic. *Paper presented at the International Conference on Pied-Piping*, University of Jena.

Chametzky, Robert. 1996. *A theory of phrase markers and the extended base.* Albany: State University of New York Press.

Chametzky, Robert. 2000. *Phrase structure: from GB to Minimalism.* Malden, MA: Blackwell.

Cheng, Lisa Lai-Shen. 1997. "Partial" *wh*-movement. In *UCI Working papers in Linguistics* 3, ed. Luther Chen-Sheng Liu and Kazue Takeda, 27-50. University of California, Irvine, CA.

Chomsky, Noam. 1955. *The logical structure of linguistic theory.* Ms., Harvard University, Cambridge, Mass. [Revised 1956 version published in part by Plenum, New York, 1975; University of Chicago press, Chicago, 1985].

Chomsky, Noam. 1965. *Aspects of the theory of syntax.* Cambridge, Mass.: MIT press.

Chomsky, Noam. 1970. Remarks on nominalization. In *Readings in English transformational grammar*, ed. Roderick A. Jacobs and Peter S. Rosenbaum, 184-221. Waltham, Mass.: Ginn and Co. [Reprinted in Studies on semantics in generative grammar, 11-61. The Hague: Mouton, 1972]

Chomsky, Noam. 1973. Conditions on transformations. In *A festschrift for Morris Halle*, ed. Stephen Anderson and Paul Kiparsky, 232-286. New York: Holt, Rinehart and Winston.

Chomsky, Noam. 1976. Conditions on rules of grammar. *Linguistic Analysis* 2: 303-351.

Chomsky, Noam. 1977. On *wh*-movement. In *Formal syntax*, ed. Peter Culicover et al., 71-132. New York: Academic Press.

Chomsky, Noam. 1981. *Lectures on government and binding.* Dordrecht: Foris.

Chomsky, Noam. 1982. *Some concepts and consequences of the theory of government and binding*. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 1986a. *Barriers*. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 1986b. *Knowledge of language*. New York: Praeger.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, ed. Kenneth Hale and Samuel J. Keyser, 1-52. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 1995a. Bare phrase structure. In *Government and binding theory and the minimalist program*, ed. Gert Webelhuth, 383-439. Cambridge, MA: Basil Blackwell. [Reprinted in Evolution and revolution in linguistic theory: Essays in honor of Carlos Otero, eds. Héctor Campos and Paula Kempchinsky, 51-109. Washington D. C.: Georgetown University Press]

Chomsky, Noam. 1995b. Categories and transformations. In *The minimalist program*, 219-394. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 1995c. *The minimalist program*. Cambridge, Mass.: MIT press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by step: Essays in minimalist syntax in honor of Howard Lasnik*, ed. Roger Martin et al., 89-155. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, ed. Michael Kenstowicz, 1-50. Cambridge, MA: MIT press.

Chomsky, Noam, and Howard Lasnik. 1993. The theory of principles and parameters. In *Syntax: An international handbook of contemporary research*, Vol. 1, ed. Joachim Jacobs et al., 506-569. Berlin: Walter de Gruyter.

Chung, Sandra, William Ladusaw, and James McCloskey. 1995. Sluicing and Logical Form. *Natural Language Semantics* 3: 1-44.

Cinque, Guglielmo. 1978. Towards a unified treatment of island constraints. In *Proceedings of the Twelfth International Congress of Linguists. Innsbrucker Beiträge zur Sprachwissenschaft*, ed. W. U. Dressler and W. Meid, 344-348.

Cinque, Guglielmo. 1990. *Types of A'-dependencies*. Cambridge, Mass.: MIT Press.

Collins, Chris. 1994. Economy of derivation and the generalized proper binding condition. *Linguistic Inquiry* 25: 45-61.

Collins, Chris. 1997. *Local economy*. Cambridge, Mass.: MIT press.

Collins, Chris, and Höskuldur Thráinsson. 1996. VP-internal structure and object shift in Icelandic. *Linguistic Inquiry* 27: 391-444.

Comrie, Bernard. 1972. *Aspects of sentence complementation in Russian*. Doctoral dissertation, Cambridge, Cambridge University.

Davidson, Donald. 1969. On saying that. In *Words and objections*, ed. Donald Davidson and Jaakko Hintikka, 158-174. Dordrecht: Reidel.

Davies, William, and Stanley Dubinsky. 2001. Bypassing subjacency effects: How event structure amnesties extraction out of object NPs. In *Proceedings of the North Eastern Linguistic Society* 31. GLSA, University of Massachusetts, Amherst, MA.

Dayal, Veneeta. 1994. Scope marking as indirect wh-dependency. *Natural Language Semantics* 2: 137-170.

Dayal, Veneeta. 1996a. *Locality in WH quantification: Questions and relative clauses in Hindi*. Dordrecht: Kluwer Academic Publishers.

Dayal, Veneeta. 1996b. Scope marking: In defence of indirect dependency. In *Papers on wh-scope marking*, ed. Uli Lutz and Gereon Müller, 107-130. Stuttgart/Tübingen: Universität Stuttgart/Universität Tübingen/IBM Deutschland.

De Haan, G. 1979. *Conditions on rules: The proper balance between syntax and semantics*. Dordrecht: Foris.

De Hoop, Helen. 1996. *Case configuration and noun phrase interpretation*. New York: Garland.

De Rijk, R. 1972. Relative clause in Basque: A guided tour. In *Chicago Which Hunt*. Chicago: University of Chicago.

Diesing, Molly. 1992. *Indefinites*. Cambridge, Mass.: MIT Press.

Enc, Murvet. 1991. The semantics of specificity. *Linguistic Inquiry* 22: 1-25.

Epstein, Samuel D. 1999. Un-principled syntax: the derivation of syntactic relations. In *Working minimalism*, ed. Samuel D. Epstein and Norbert Hornstein, 317-345. Cambridge, MA: MIT Press.

Epstein, Samuel D., Erich M. Groat, Ruriko Kawashima, and Hisatsugu Kitahara. 1998. *A derivational approach to syntactic relations*. Oxford: Oxford University Press.

Erteschik, Naomi. 1973. *On the nature of island constraints*. Doctoral dissertation, MIT, Cambridge, MA.

Fanselow, Gisbert. 1997. *Partial movement*. Ms., University of Potsdam.

Fanselow, Gisbert, and Damir Cavar. 2000. Remarks on the economy of pronunciation. In *Competition in syntax*, ed. Gereon Müller and Wolfgang Sternefeld, 107-150. Berlin: Mouton de Gruyter.

Fanselow, Gisbert, and Anoop Mahajan. 1996. Partial movement and successive cyclicity. In *Papers on wh-scope marking*, ed. Uli Lutz and Gereon Müller, 131-161. Stuttgart/Tübingen: Universität Stuttgart/Universität Tübingen/IBM Deutschland.

Fiengo, Robert, and James Higginbotham. 1981. Opacity in NP. *Linguistic Analysis* 7: 395-421.

Fillmore, Charles J. 1963. The position of embedding transformations in a grammar. *Word* 19: 208-231.

Fodor, Janet D., and Ivan Sag. 1982. Referential and quantificational indefinites. *Linguistics and Philosophy* 5: 355-398.

Fox, Danny, and Jon Nissenbaum. 1999. Extraposition and scope: a case for overt QR. In *Proceedings of West Coast Conference on Formal Linguistics* 18, ed. Sonya Bird et al., 132-144. Cascadilla, Somerville, MA.

Frampton, John, and Sam Gutmann. 1999. Cyclic computation, a computationally efficient minimalist syntax. *Syntax* 2: 1-27.

Frank, Robert, Paul Hagstrom, and K. Vijay-Shanker. 1999. Roots, constituents and c-command. *Paper presented at the GLOW 22*, Berlin.

Frank, Robert, and K. Vijay-Shanker. 1998. *Primitive c-command.* Ms., John Hopkins University and University of Delaware, Baltimore, MD and Newark, DE.

Franks, Steven. 1994. Parametric properties of numeral phrases in Slavic. *Natural Language and Linguistic Theory* 12: 570-649.

Franks, Steven. 1995. *Parameters of Slavic morphosyntax.* New York, New York: Oxford Universsity press.

Freidin, Robert. 1978. Cyclicity and the theory of grammar. *Linguistic Inquiry* 9: 519-549.

Freidin, Robert. 1986. Fundamental issues in the theory of binding. In *Studies in the acquisition of anaphora,* Vol. 1, ed. Barbara Lust, 151-188. Dordrecht: Reidel.

Freidin, Robert, and Howard Lasnik. 1981. Disjoint reference and *wh*-trace. *Linguistic Inquiry* 12: 39-53. [Reprinted in Essays on anaphora, Howard Lasnik, 110-124. Dordrecht: Kluwer, 1989]

Freidin, Robert, and Rex A. Sprouse. 1991. Lexical case phenomena. In *Principles and parameters in comparative grammar,* ed. Robert Freidin, 392-416. Cambridge, MA: MIT Press.

Fukui, Naoki. 1986. *A theory of category projection and its applications.* Doctoral dissertation, MIT, Cambridge, Mass.

Gärtner, Hans-Martin. 1999. Phrase-linking meets minimalist syntax. In *Proceedings of West Coast Conference on Formal Linguistics* 18, ed. Sonya Bird et al., 159-169. Cascadilla, Somerville, MA.

Georgopoulos, Carol. 1991. *Syntactic variables: Resumptive pronouns and binding in Palauan.* Dordrecht: Kluwer.

Georgopoulos, Carol, and Roberta Ishihara, eds. 1991. *Interdisciplinary approaches to language: Essays in honor of S.-Y. Kuroda.* Dordrecht: Foris.

Giejgo, Janina Aurelia. 1981. *Movement rules in Polish syntax.* Doctoral dissertation, University College, London.

Godard, Danièle. 1985. French relative clauses with DONT: A'-chains and binding principles. In *Grammatical representation,* ed. J. Guéron et al. Dordrecht: Foris.

Goodall, Grant. 1987. *Parallel structures in syntax*. Cambridge: Cambridge University Press.

Greenberg, Gerald R. 1988. Relativization and Subjacency in Russian. In *Proceedings of ESCOL '88*, 187-197. Cornell University, Ithaca.

Grewendorf, Gunther. 1989. *Ergativity in German* 22). Dordrecht: Foris.

Hagstrom, Paul. 1998. *Decomposing questions*. Doctoral dissertation, MIT, Cambridge, Mass.

Haider, Hubert. 1986. Affect α: A reply to Lasnik and Saito, On the nature of proper government. *Linguistic Inquiry* 17: 113-126.

Haider, Hubert. 1997. Extraposition. In *Rightward movement*, ed. Dorothee Beerman et al., 115-151. Amsterdam/Philadelphia: John Benjamins.

Haider, Hubert. 2000. OV is more basic than VO. In *The derivation of VO and OV*, ed. Peter Svenonius, 45-67. Amsterdam/Philadelphia: John Benjamins.

Hale, Kenneth, and Samuel J. Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, ed. Kenneth Hale and Samuel J. Keyser, 53-110. Cambridge, Mass.: MIT Press.

Halle, Morris, and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, ed. Kenneth Hale and Samuel Jay Keyser, 111-176. Cambridge, Mass.: MIT Press.

Hamblin, C. L. 1973. Questions in Montague English. *Foundations of Language* 10: 41-53.

Hand, Michael. 1991. On saying that again. *Linguistics and Philosophy* 14: 349-365.

Hankamer, Jorge, and Laura Knecht. 1976. The role of the subject/non-subject distinction in determining the choice of relative clause participle in Turkish. In *Proceedings of the North Eastern Linguistic Society* 6, 123-135. McGill University/Université de Montréal, Montreal.

Harris, Alice C. 1981. *Georgian syntax: A study in relational grammar*. New York: Cambridge University Press.

Hegarty, Michael. 1991. Adjunct extraction without traces. In *Proceedings of West Coast Conference on Formal Linguistics* 10, 209-222. Stanford Linguistic Association, Stanford University, Stanford, Calif.

Heim, Irene, and Angelika Kratzer. 1998. *Semantics in generative grammar*. Malden, MA.: Blackwell Publishers.

Herburger, Elena. 1994. A semantic difference between full and partial *wh*-movement in German. *Paper presented at the 68th Annual meeting of the Linguistic Society of America*, Boston, Mass.

Hiemstra, Inge. 1986. Some aspects of wh-questions in Frisian. *NOWELE: North Western European Language Evolution* 8: 97-110.

Higginbotham, James. 1983. A note on phrase-markers. *Revue Québécoise de Linguistique* 13: 147-166.

Hiramatsu, Kazuko. 1999. Subject and Adjunct island asymmetries: Evidence from syntactic satiation. In *Proceedings of West Coast Conference on Formal Linguistics* 18, ed. Sonya Bird et al., 183-192. Cascadilla, Somerville, MA.

Höhle, Tilman N. 1996. The W-...W-construction: Appositive or scope indicating. In *Papers on wh-scope marking*, ed. Uli Lutz and Gereon Müller, 37-58. Stuttgart/Tübingen: Universität Stuttgart/Universität Tübingen/IBM Deutschland.

Holland, Gary B. 1984. Subordination and relativization in Early Indo-European. In *Proceedings of the Tenth Annual Meeting of the Berkeley Linguistic Society*, ed. Claudia Brugman and Monica Macaulay, 609-622. University of California, Berkeley, Berkeley.

Honcoop, Martin. 1998. *Dynamic excursions on weak islands*. Doctoral dissertation, Leiden University, Leiden.

Hornstein, Norbert. 1999. Movement and control. *Linguistic Inquiry* 30: 69-96.

Horvath, Julia. 1986. *Focus in the theory of grammar and the syntax of Hungarian*. Dordrecht: Foris.

Horvath, Julia. 1997. The status of 'wh-expletives' and the partial *wh*-movement construction of Hungarian. *Natural Language and Linguistic Theory* 15: 509-572.

Horvath, Julia. 1998. Multiple *wh*-phrases and the *wh*-scope marker strategy in Humgarian interrogatives. *Acta Linguistica Hungarica* 45: 31-60.

Huang, C. -T. James. 1982. *Logical relations in Chinese and the theory of grammar*. Doctoral dissertation, MIT, Cambridge, Mass.

Ishii, Toru. 1997. *An asymmetry in the composition of phrase structure and its consequences*. Doctoral dissertation, University of California, Irvine.

Ishii, Toru. 1998. Derivational selectional restriction and reconstruction. *English Linguistics* 15: 28-49.

Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar*. Cambridge, Mass.: MIT Press.

Johnson, Kyle. 1987. Against the notion "SUBJECT". *Linguistic Inquiry* 18: 354-361.

Johnson, Kyle. 1991. Object positions. *Natural Language and Linguistic Theory* 9: 577-636.

Kawashima, Ruriko, and Hisatsugu Kitahara. 1996. Strict cyclicity, linear ordering and derivational c-command. In *Proceedings of the West Coast Conference on Formal Linguistics* 14, 255-269. CSLI, Stanford, Cali.

Kayne, Richard. 1981. ECP extensions. *Linguistic Inquiry* 12: 93-133. [Reprinted in Kayne 1984]

Kayne, Richard. 1983. Connectedness. *Linguistic Inquiry* 14: 223-249. [Reprinted in Kayne 1984]

Kayne, Richard. 1984. *Connectedness and binary branching*. Dordrecht: Foris.

Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, Mass.: MIT Press.

Kayne, Richard. 1998. Overt vs. covert movement. *Syntax* 1: 128-191.

Kayne, Richard, and Jean-Yves Pollock. 1978. Stylistic Inversion, successive cyclicity and Move NP in French. *Linguistic Inquiry* 9: 595-621.

Kikuchi, Akira. 1987. *Comparative deletion in Japanese*. Ms., Yamagata University.

King, Tracy Holloway. 1993. *Configuring topic and focus in Russian*. Doctoral dissertation, Stanford University, Stanford.

Kiparsky, Paul, and Carol Kiparsky. 1970. Fact. In *Progress in linguistics*, ed. Manfred Bierwisch and Karl E. Heidolph, 143-173. The Hague: Mouton.

Kiss, Katalin É. 1987. *Configurationality in Hungarian*. Dordrecht: D. Reidel.

Kiss, Katalin É. 1991. An argument for movement. In *Representation and derivation in the theory of grammar*, ed. Hubert Haider and Klaus Netter, 199-215. Dordrecht: Kluwer.

Kiss, Katalin É. 1994. Sentence structure and word order. In *Syntactic structure of Hungarian (=Syntax and Semantics Volume 27)*, ed. Ferenc Kiefer and Katalin É. Kiss, 1-90. New York: Academic Press.

Kitagawa, Yoshihisa. 1986. *Subjects in Japanese and English*. Doctoral dissertation, University of Massachusetts, Amherst.

Kitahara, Hisatsugu. 1994. A Minimalist analysis of cross-linguistically variant CED phenomena. In *Proceedings of the North Eastern Linguistic Society* 24, 241-253. GLSA, University of Massachussets, Amherst.

Kitahara, Hisatsugu. 1995. Target α: deducing strict cyclicity from derivational economy. *Linguistic Inquiry* 26: 47-77.

Kitahara, Hisatsugu. 1997. *Elementary operations and optimal derivations*. Cambridge, MA: MIT Press.

Koizumi, Masatoshi. 1995. *Phrase structure in minimalist syntax*. Doctoral dissertation, MIT, Cambridge, Mass.

Koopman, Hilda, and Anna Szabolcsi. 2000. *Verbal complexes*. Cambridge, MA: MIT Press.

Koster, Jan. 1978. *Locality principles in syntax*. Dordrecht: Foris.

Koster, Jan. 1987. *Domains and dynasties*. Dordrecht: Foris.

Kuno, Susumu. 1973. *The structure of the Japanese language*. Cambridge, Mass.: MIT Press.

Kuno, Susumo. 1997. *Binding theory in the minimalist program.* Ms., Harvard University, Cambridge, MA.

Kuno, Susumu, and Ken-ichi Takami. 1997. Remarks on negative islands. *Linguistic Inquiry* 28: 553-576.

Kural, Murat. 1993. V-to (I-to) -C in Turkish. In *UCLA Occasional Papers in Linguistics* 11, ed. Filippo Beghelli and Murat Kural, 17-54. UCLA, Los Angeles.

Kural, Murat. 1997. Postverbal constituents in Turkish and the Linear Correspondence Axiom. *Linguistic Inquiry* 28: 498-519.

Kuroda, S. -Y. 1988. Whether we agree or not: A comparative syntax of English and Japanese. In *Papers on the Second International Workshop on Japanese Syntax,* ed. William J. Poser, 103-143. CSLI, Stanford University, Stanford, Calif.

Larson, Richard K. 1988. On the double object construction. *Linguistic Inquiry* 19: 335-391.

Lasnik, Howard. 1995a. Case and expletives revisited: On Greed and other human failings. *Linguistic Inquiry* 26: 615-633.

Lasnik, Howard. 1995b. Verbal morphology: *Syntactic Structures* meets the Minimalist Program. In *Evolution and revolution in linguistic theory: Essays in honor of Carlos Otero,* ed. Héctor Campos and Paula Kempchinsky, 251-275. Washington, D. C.: Georgetown University Press.

Lasnik, Howard. 1997. A gap in an ellipsis paradigm: Some theoretical implications. *Linguistic Analysis* 27: 166-185.

Lasnik, Howard. 1998. Some reconstruction riddles. In *University of Pennsylvania working papers in linguistics* 5.1, ed. Alexis Dimitriadis et al., 83-98. Penn Linguistics Club, University of Pennsylvania, Philadelphia, PA.

Lasnik, Howard. 1999a. Chains of arguments. In *Working minimalism,* ed. Samuel D. Epstein and Norbert Hornstein, 189-215. Cambridge, Mass.: MIT Press.

Lasnik, Howard. 1999b. *Minimalist analysis.* Malden, MA.: Blackwell.

Lasnik, Howard. 1999c. On feature strength: Three minimalist approaches to overt movement. *Linguistic Inquiry* 30: 197-217.

Lasnik, Howard. to appear. When can you save the structure by destroying it? In *Proceedings of the North Eastern Linguistic Society* 31. GLSA, University of Massachusetts, Amherst.

Lasnik, Howard, and Joseph J. Kupin. 1977. A restrictive theory of transformational grammar. *Theoretical Linguistics* 4: 173-196. [Reprinted in Essays on restrictiveness and learnablity, Howard Lasnik, 17-41. Dordrecht: Kluwer, 1990]

Lasnik, Howard, and Mamoru Saito. 1984. On the nature of proper government. *Linguistic Inquiry* 15: 235-289. [Reprinted in Essays on restrictiveness and learnablity, Howard Lasnik, 198-255. Dordrecht: Kluwer, 1990.]

Lasnik, Howard, and Mamoru Saito. 1991. On the subject of infinitives. In *Papers from the 27th regional Meeting of the Chicago Linguistic Society Part One: The General Session*, ed. Lise M. Dobrin et al., 324-343. Chicago Linguistic Society, University of Chicago, Chicago, Ill.

Lasnik, Howard, and Mamoru Saito. 1992. *Move α: Conditions on its application and output*. Cambridge, Mass.: MIT Press.

Law, Paul. 1993. On the base position of wh-adjuncts and extraction. *Paper presented at the 67th Annual Meeting of the Linguistic Society of America*, Los Angeles, Calif.

Lebeaux, David. 1988. *Language acquisition and the form of the grammar*. Doctoral dissertation, University of Massachusetts, Amherst.

Lebeaux, David. 1991. Relative clauses, licensing and the nature of the derivation. In *Syntax and Semantics 25: Perspectives on phrase structure*, ed. Susan Rothstein, 209-239. New York: Academic Press.

LePore, Ernest, and Barry Loewer. 1989. You can say that again. *Midwest Studies in Philosophy* XIV.

Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice theoretical approach. In *Meaning, use, and interpretation of language*, ed. Rainer Bauerle et al., 302-323. Berlin: Walter de Gruyter.

Longobardi, Giuseppe. 1985. Connectedness and island constraints. In *Grammatical representation*, ed. J. Gueron et al., 169-185. Dordrecht: Foris.

Lutz, Uli, and Gereon Müller, eds. 1996. *Papers on wh-scope marking*. Stuttgart/Tübingen: Universität Stuttgart/Universität Tübingen/IBM Deutschland.

Mahajan, Anoop. 1990. *The A/A-bar distinction and movement theory*. Doctoral dissertation, MIT, Cambridge, Mass.

Mahajan, Anoop. 1992. The Specificity Condition and the CED. *Linguistic Inquiry* 23: 510-516.

Mahajan, Anoop. 1996. *Wh*-expletives and the syntax of partial *wh*-movement. In *Papers on wh-scope marking*, ed. Uli Lutz and Gereon Müller, 163-177. Stuttgart/Tübingen: Universität Stuttgart/Universität Tübingen/IBM Deutschland.

Manzini, Rita. 1983. *Restructuring and reanalysis*. Doctoral dissertation, MIT, Cambridge, MA.

Manzini, Rita. 1998. A minimalist theory of weak islands. In *The limits of syntax. Syntax and semantics vol. 29*, ed. Peter Culicover and Luise McNally, 185-210. New York, NY: Academic Press.

Marácz, László. 1989. *Asymmetries in Hungarian.* , University of Groningen, Groningen.

Marantz, Alec. 1984. *On the nature of grammatical relations*. Cambridge, Mass.: MIT Press.

May, Robert. 1985. *Logical Form: Its structure and derivation*. Cambridge, Mass.: MIT Press.

McCawley, James D. 1982. Parentheticals and discontinuous constituent structure. *Linguistic Inquiry* 13: 91-106.

McCloskey, James. 1991. Resumptive pronouns, A'-binding, and levels of representation in Irish. In *Syntax and Semantics 23: The Syntax of the Modern Celtic Languages*, ed. Randall Hendrick, 199-248. New York: Academic Press.

McDaniel, Dana. 1989. Partial and multiple *wh*-movement. *Natural Language and Linguistic Theory* 7: 565-604.

Merchant, Jason. 1999. *The syntax of silence: Sluicing, islands, and identity in ellipsis*. Doctoral dissertation, University of California, Santa Cruz.

Milsark, Gary. 1974. *Existential sentences in English*. PhD, MIT, Cambridge.

Mitchell, Bruce, and Fred C. Robinson. 1992. *A guide to Old English*. Cambridge, MA: Blackwell.

Miyagawa, Shigeru. 1997. Against optional scrambling. *Linguistic Inquiry* 28: 1-25.

Molotkov, A. I., ed. 1967. *Frazeologicheskij slovar' russkogo jazyka (=Phraseological dictionary of the Russian language)*. Moscow: Sovetskaja Enciklopedija.

Moltmann, Frederike. 1992. *Coordination and comparatives*. Doctoral dissertation, MIT, Cambridge.

Moro, Andrea. 2000. *Dynamic antisymmetry*. Cambridge, MA: MIT Press.

Muadz, H. 1991. *Coordinate structures: A planar representation*. Doctoral dissertation, University of Arizona, Tucson.

Müller, Gereon. 1995. *A-bar syntax: A study of movement types* (=Studies in Generative Grammar 42). New York: Mouton de Gruyter.

Müller, Gereon. 1996. Partial *wh*-movement and Optimality Theory. In *Papers on wh-scope marking*, ed. Uli Lutz and Gereon Müller, 179-230. Stuttgart/Tübingen: Universität Stuttgart/Universität Tübingen/IBM Deutschland.

Müller, Gereon. 1998. *Incomplete category fronting* 42). Dordrecht: Kluwer.

Müller, Gereon. 1999. Shape conservation and remnant movement. In *Proceedings of the North Eastern Linguistic Society* 30. GLSA, Amherst, MA.

Müller, Gereon, and Wolfgang Sternefeld. 1995. Extraction, lexical variation and the theory of barriers. In *Lexical knowledge in the organization of language*, ed. Urs Egli et al., 35-80. Amsterdam/Philadelphia: John Benjamins.

Müller, Gereon, and Wolfgang Sternefeld. 1996. A'-chain formation and economy of derivation. *Linguistic Inquiry* 27: 480-511.

Munn, Alan. 1993. *Topics in the syntax and semantics of coordinate structures*. Doctoral dissertation, University of Maryland, College Park, MD.

Murasugi, Keiko. 1992. Locative/temporal vs. manner/reason phrases. In *Papers in English linguistics and literature*, 153-170. Department of English Literature, Kinjo Gakuin University.

Nakamura, Masanori. 1998. Global issues. In *Proceedings of the North Eastern Linguistic Society* 28, ed. Pius Tamanji and Kiyomi Kusumoto, 301-318. University of Massachusetts, Amherst.

Neeleman, Ad. 1994. Scrambling as a D-structure phenomenon. In *Studies on scrambling: Movement and non-movement approaches to free word-order phenomena*, ed. Norbert Corver and Henk van Riemsdijk, 387-429. Berlin: Mouton de Gruyter.

Neeleman, Ad, and Tanya Reinhart. 1998. Scrambling and PF interface. In *Projecting from the lexicon*, ed. W. Gueder and Miriam Butt. Stanford: CSLI.

Nishigauchi, Taisuke. 1986. *Quantification in syntax*. Doctoral dissertation, University of Massachusetts, Amherst.

Nishigauchi, Taisuke. 1990. *Quantification in the theory of grammar*. Dordrecht: Kluwer.

Nissenbaum, Jon. 1998. Movement and derived predicates: evidence from parasitic gaps. In *MIT Working Papers in Linguistics* 25, ed. Uli Sauerland and Orin Percus, 247-295. MITWPL, Cambridge, Mass.

Nissenbaum, Jon. 2000. Covert movement and parasitic gaps. In *Proceedings of NELS 30*. GLSA, University of Massachusetts, Amherst, MA.

Nunes, Jairo. 1995. *The copy theory of movement and linearization of chains in the Minimalist Program*. Doctoral dissertation, University of Maryland, College Park, MD.

Nunes, Jairo. 1998. Bare X-bar theory and structures formed by movement. *Linguistic Inquiry* 29: 160-168.

Nunes, Jairo. 1999. Linearization of chains and phonetic realization of chain links. In *Working minimalism*, ed. Samuel D. Epstein and Norbert Hornstein, 217-249. Cambridge, Mass.: MIT Press.

Nunes, Jairo. 2001. Sideward movement. *Linguistic Inquiry* 32: 303-344.

Nunes, Jairo, and Juan Uriagereka. 2000. Cyclicity and extraction domains. *Syntax* 3: 20-43.

Ochi, Masao. 1999a. Multiple Spell-Out and PF adjacency. In *Proceedings of the North Eastern Linguistic Society* 29. University of Massachusetts, Amherst.

Ochi, Masao. 1999b. Some consequences of Attract F. *Lingua* 109: 81-107.

Oka, Toshifusa. 1993. *Minimalism in syntactic description*. Doctoral dissertation, MIT.

Ormazabal, Javier. 1995. *The syntax of complementation: On the connection between syntactic structure and selection*. Doctoral dissertation, University of Connecticut, Storrs.

Pafel, Jürgen. 1996. Absolute and relative: On scope in German *wh*-sentences including *w...w*-constructions. In *Papers on wh-scope marking*, ed. Uli Lutz and Gereon Müller, 231-256. Stuttgart/Tübingen: Universität Stuttgart/Universität Tübingen/IBM Deutschland.

Partee, Barbara, Alice ter Meulen, and Robert E. Wall. 1993. *Mathematical methods in linguistics*. Dordrecht: Kluwer.

Pesetsky, David. 1982. *Paths and categories*. Doctoral dissertation, MIT, Cambridge, Mass.

Pesetsky, David. 1987. *Wh*-in situ: Movement and unselective binding. In *The representation of (in)definiteness*, ed. Eric J. Reuland and Alice G. B. ter Meulen, 98-129. Cambridge, Mass.: MIT Press.

Pesetsky, David. 1992. *Zero Syntax Vol. 2*. Ms., MIT, Cambridge, Mass.

Pesetsky, David, and Esther Torrego. 2000. T-to-C movement: Causes and consequences. In *Ken Hale: A life in language*, ed. Michael Kenstowicz. Cambridge: MIT Press.

Platero, P. 1974. The Navajo relative clause. *International Journal of American Linguistics* 40: 202-246.

Plessis, Hans du. 1977. *Wh*-movement in Afrikaans. *Linguistic Inquiry* 8: 723-726.

Postal, Paul. 1997. Strong crossover violations and binding principles. *Paper presented at the ESCOL '97*, Yale University.

Postal, Paul. 1999. *Islands*. Ms., New York University, New York, NY.

Postal, Paul M. 1974. *On raising: One rule of English grammar and its theoretical implications*. Cambridge, Mass.: MIT Press.

Postal, Paul M. 1998. *Three investigations of extraction*. Cambridge, Massachusetts: The MIT Press.

Progovac, Ljiljana. 1993. Locality and subjunctive-like complements in Serbo-Croatian. *Journal of Slavic Linguistics* 1: 116-144.

Raposo, Eduardo, and Juan Uriagereka. 1990. Long-distance Case assignment. *Linguistic Inquiry* 21: 505-537.

Reinhart, Tanya. 1976. *The syntactic domain of anaphora*. Doctoral dissertation, MIT, Cambridge, Mass.

Reinhart, Tanya. 1995. *Interface strategies*. Ms., Research Institute for Language and Speech, Faculty of Letters, University of Utrecht. hardcopy

Reinhart, Tanya. 1998. Wh-in-situ in the framework of the Minimalist Program. *Natural Language Semantics* 6: 29-56.

Reis, Marga. 1996. On *was*-parentheticals and *was...w*-constructions in German. In *Papers on wh-scope marking*, ed. Uli Lutz and Gereon Müller, 257-288. Stuttgart/Tübingen: Universität Stuttgart/Universität Tübingen/IBM Deutschland.

Richards, Norvin. 2000. An island effect in Japanese. *Journal of East Asian Linguistics* 9: 187-205.

Riemsdijk, Henk van. 1981. The Case of German adjectives. In *Markedness and Learnability - University of Massachusetts Occasional Papers in Linguistics* 6, ed. James Pustejovsky and Victoria A. Burke, 148-173. University of Massachusetts, Amherst.

Riemsdijk, Henk van. 1982. Correspondence effects and the Empty Category Principle. In *Tilburg papers in Language and Literature* 12. University of Tilburg.

Rizzi, Luizi. 1990. *Relativized minimality*. Cambridge, Mass.: MIT Press.

Rizzi, Luigi. 1992. Argument/adjunct (a)symmetries. In *Proceedings of the North Eastern Lingusitic Society*, 365-381. University of Massachusetts, Amherst, Mass.

Ross, John Robert. 1967. *Constraints on variables in syntax*. Doctoral dissertation, MIT, Cambridge, Mass.

Ross, John Robert. 1969. Guess who? In *Papers from the Fifth Regional Meeting of the Chicago Linguistic Society*, ed. Robert I. Binnick et al., 252-286. Chicago Linguistic Society, University of Chicago, Chicago, Ill.

Rothstein, Susan D. 1995. Pleonastics and the interpretation of pronoun. *Linguistic Inquiry* 26: 499-529.

Rullman, H. 1995. *Maximality in the semantics of wh-constructions*. Doctoral dissertation, University of Massachusetts, Amherst.

Rullmann, Hotze, and Sigrid Beck. 1998. Presupposition projection and the interpretation of *which*-questions. In *Proceedings of SALT VIII*, 215-232. Cornell University, Ithaca.

Rumfitt, Ian. 1993. Content and context: The paratactic theory revisited and revised. *Mind* 102 (407): 429-454.

Sabel, Joachim. 2001. Wh-questions and extraction asymmetries in Malagasy. *Paper presented at the AFLA VIII*, MIT, Cambridge, MA.

Safir, Ken. 1986. Relative clauses in a theory of binding and levels. *Linguistic Inquiry* 17: 663-689.

Saito, Mamoru. 1985. *Some asymmetries in Japanese and their theoretical implications*. Doctoral dissertation, MIT, Cambridge, Mass.

Segal, Gabriel, and Margaret Speas. 1986. On saying [that]. *Mind and Language* 1: 124-132.

Sigurðsson, Halldór Ármann. 1992. The Case of quirky subjects. In *Working papers in Scandinavian syntax* 49, 1-26, Lund, Sweden.

Snyder, William. 2000. An experimental investigation of syntactic satiation effects. *Linguistic Inquiry* 31: 575-582.

Speas, Margaret. 1990. *Phrase structure in natural language*. Dordrecht: Kluwer.

Sportiche, Dominique. 1981. Bounding nodes in French. *The Linguistic Review* 1: 219-246.

Srivastav, Veneeta. 1991. *Wh-dependencies in Hindi and the theory of grammar*. Doctoral dissertation, Cornell University, Ithaca, NY.

Stateva, Penka. to appear. Possessive clitics and the structure of nominal expressions. *Lingua*.

Stechow, Arnim von, and Wolfgang Sternefeld. 1988. *Bausteine syntaktischen Wissens*. Opladen: Westdeutscher Verlag.

Stepanov, Arthur. 1997. Case properties of nominalization DPs in Classical Arabic. In *Perspectives on Arabic linguistics X*, ed. Mushira Eid and Robert R. Ratcliffe, 69-90. (=*Current issues in linguistic theory, 153*.) Amsterdam and Philadelphia: John Benjamins.

Stepanov, Arthur. 2000a. The timing of adjunction. In *Proceedings of the North Eastern Linguistic Society* 30. GLSA, University of Massachusetts, Amherst.

Stepanov, Arthur. 2000b. *Wh*-scope marking in Slavic. *Studia Linguistica* 54: 1-40.

Stepanov, Arthur. 2001. Late adjunction and minimalist phrase structure. *Syntax* 4: 94-125.

Stepanov, Arthur, and Carol Georgopoulos. 1997. Structure building and the conceptual interface: An analysis of Russian long-distance *wh*-questions. In *Proceedings of the V Annual Workshop on Formal Approaches to Slavic Linguistics*, ed. Martina Lindseth and Steven Franks, 275-294. Michigan Slavic Publications, Ann Arbor, MI.

Stjepanović, Sandra. 1997a. Is inherent Case structural? In *Proceedings of the V Annual Workshop on Formal Approaches to Slavic Linguistics*, ed. Martina Lindseth and Steven Franks, 295-311. Michigan Slavic Publications, Ann Arbor, MI.

Stjepanović, Sandra. 1997b. VP ellipsis in a verb raising language: Some implication for verbal morphology. In *Proceedings of Eastern States Conference on Linguistics* 14, 192-203. Cornell University, Ithaca, NY.

Stjepanović, Sandra. 1999. *What is common between multiple wh-fronting, VP-ellipsis and clitic movement*. Doctoral dissertation, University of Connecticut, Storrs.

Stowell, Timothy. 1978. What was there before there was there. In *Papers from the Fourteenth Regional Meeting, Chicago Linguistic Society*, 458-471. Chicago Linguistic Society, University of Chicago, Chicago, Ill.

Stowell, Timothy. 1981. *Origins of phrase structure*. Doctoral dissertation, MIT, Cambridge, Mass.

Svenonius, Peter. 1994. C-selection as feature-checking. *Studia Linguistica* 48: 133-155.

Szabolcsi, Anna, and Frans Zwarts. 1992-1993. Weak islands and an algebraic semantics for scope taking. *Natural Language Semantics* 1: 235-284.

Takahashi, Daiko. 1994. *Minimality of movement*. Doctoral dissertation, University of Connecticut, Storrs.

Tancredi, Chris. 1992. *Deletion, deaccenting, and presupposition*. Doctoral dissertation, MIT, Cambridge, Mass.

Tellier, Christine. 1990. Subjacency and Subject Condition violations in French. *Linguistic Inquiry* 21: 306-311.

Thornton, Rosalind. 1990. *Adventures in long-distance moving: The acquisition of complex wh-questions*. Doctoral dissertation, University of Connecticut, Storrs.

Torrego, Esther. 1983. More effects of successive cyclic movement. *Linguistic Inquiry* 14: 561-565.

Torrego, Ester. 1984. On inversion in Spanish and some of its effects. *Linguistic Inquiry* 15: 103-129.

Torrego, Esther. 1996. Experiencers and raising verbs. In *Current issues in comparative grammar*, ed. Robert Freidin, 101-120. Dordrecht: Kluwer.

Torrego, Ester, and Juan Uriagereka. 1993. *Indicative dependents*. Ms., University of Massachusetts and University of Maryland, Amherst, MA and College Park, MD.

Toyoshima, Takashi. 1997. Derivational CED: A consequence of the bottom-up parallel-process of Merge and Attract. In *Proceedings of the West Coast Conference on Formal Linguistics* 15, ed. Brian Agbayani and Sze-Wing Tang, 505-519. CSLI, Stanford.

Tsai, W.-T. Dylan. 1994. *On economizing the theory of A'-dependencies*. Doctoral dissertation, MIT, Cambridge, Mass.

Turano, Giuseppina. 1998. Overt and covert dependencies in Albanian. *Studia Linguistica* 52: 149-183.

Uriagereka, Juan. 1988. *On government*. Doctoral dissertation, University of Connecticut, Storrs.

Uriagereka, Juan. 1999. Multiple Spell-Out. In *Working minimalism*, ed. Samuel D. Epstein and Norbert Hornstein, 251-282. Cambridge, MA.: MIT Press.

von Stechow, Arnim. 1996. Against LF pied-piping. *Natural Language Semantics* 4: 57-110.

Wahba, Wafaa Abdel-Faheem Batran. 1991. LF movement in Iraqi Arabic. In *Logical structure and linguistic structure*, ed. C.-T. James Huang and Robert May, 253-276. Dordrecht: Kluwer.

Watanabe, Akira. 1992a. Subjacency and S-structure movement of *wh*-in situ. *Journal of East Asian Linguistics* 1: 255-291.

Watanabe, Akira. 1992b. *Wh*-in-situ, Subjacency, and chain formation. In *MIT Occasional Papers in Linguistics* 2. MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, Mass.

Watanabe, Akira. 1995. Conceptual basis of cyclicity. In *Papers on minimalist syntax, MIT working papers in linguistics* 27, ed. Rob Pensalfini and Hiroyuki Ura, 269-291. MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, Mass.

Webelhuth, Gert. 1989. *Syntactic saturation phenomena and the modern Germanic languages*. Doctoral dissertation, Universite de Geneve, Geneva.

Wexler, Kenneth, and Peter Culicover. 1981. *Formal principles of language acquisition*. Cambridge, Mass.: MIT Press.

Whitman, John. 1986. Configurationality parameters. In *Issues in Japanese syntax*, ed. Takashi Imai and Mamoru Saito, 352-374. Dordrecht: Foris.

Yatsushiro, Kazuko. 1999. *Case licensing and VP structure*. Doctoral dissertation, University of Connecticut, Storrs.

Zaenen, Annie. 1983. On syntactic binding. *Linguistic Inquiry* 14: 469-504.

Zaenen, Annie. 1985. *Extraction rules in Icelandic*. New York: Garland.

Zubizarretta, Maria-Luisa. 1998. *Prosody, focus and word order*. Cambridge, MA: MIT press.