ON FEATURES AND THE MLC

Saša Vukić, Ph.D.

University of Connecticut, 2003

This thesis investigates the proper characterization of triggers and locality conditions governing the displacement property of human language. Following the introductory statement of issues and background assumptions in Chapter 1, Chapter 2 presents the currently predominant view of displacement on which it involves a matching relation between uninterpretable features of the target/probe and interpretable features of the goal. It is shown that such a view faces serious problems in both undergenerating grammatical constructions and allowing ungrammatical ones, and in incorporating complex and vague notions in the computational system such as the mechanism of 'equidistance', the suppressive role of inherent Case, the 'activating' role of uninterpretable features, etc.. An alternative on which only uninterpretable features are subject to the operation Attract/Move is presented in Chapter 3, and it is shown that such an analysis makes correct empirical predictions and avoids conceptual problems inherent in the opposite view. Chapter 4 proposes a principle of Late Expletive Insertion on which the 'superraising' and several other construction are ruled out without the need for the mechanism of 'defective intervention'. Chapter 5 extends the proposed analysis to the core cases of A'-movement, showing that it provides a straightforward way to capture the

effects of Pesetsky's Path Containment Condition in the Minimalist framework. I show that the proposal can be reconciled with evidence for the successive cyclic character of movement under a version of Chomsky's Form Chain condition. I also show that a view on which structural Case is but a reflex of agreement relation faces serious empirical and conceptual problems. Chapter 6 recapitulates the main results of the thesis.

ON FEATURES AND THE MLC

Saša Vukić

B.A., University of Sarajevo, 1985

M.A., West Virginia University, 1994

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

at the

University of Connecticut

2003

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

UMI Number: 3089771



UMI Microform 3089771

Copyright 2003 by ProQuest Information and Learning Company. All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

> ProQuest Information and Learning Company 300 North Zeeb Road P.O. Box 1346 Ann Arbor, MI 48106-1346

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

APPROVAL PAGE

Doctor of Philosophy Dissertation

ON FEATURES AND THE MLC

Presented by

Saša Vukić, B.A., M.A.

Major Advisor

Howard Lasnik

Associate Advisor

Željko Bošković

Wittin " Any Associate Advisor

William Snyder

University of Connecticut

2003

ii

ACKNOWLEDGEMENTS

This thesis has benefitted greatly from the expertise and guidance extended to me by the members of my Advisory Committee: Howard Lasnik, Željko Bošković, and William Snyder. I wish to thank Howard Lasnik not only for the vast knowledge of generative linguistics and the extraordinary pedagogical gift which he generously shares with his students and colleagues, but also for his kindness, patience, and helpfulness. It has been a rare privilege to sit in his classes and to have him as a main advisor. I am grateful to Željko Bošković for always being available when I needed his help in exploring the consequences of my often vaguely formulated ideas. Željko's ability to instantly see the positive as well as negative sides of my linguistic proposals has been of invaluable help to me, and his dedication to empirical matters and success in elucidating aspects of Serbian/Croatian grammar from the generative perspective have been most reassuring. I am also thankful to William Snyder for his willingness to take time while discussing my ideas and for his thoughtful comments. It was impossible to leave his office without feeling good about linguistics.

I would also like to thank other members of the UConn Linguistics faculty or visiting scholars whose classes have shaped my linguistic knowledge: Andrea Calabrese, Steven Franks, Diane Lillo-Martin, Ignatius Mattingly, and David Michaels, as well as other members of the faculty that I had the honor of meeting: Arthur Abramson, Mona Anderson, Sigrid Beck, Harry van der Hulst, Eva Bar-Shalom, and Yael Sharvit. Finally, I would like to acknowledge my indebtedness to the professors in the Department of Foreign Languages at West Virginia University who started me on the generative enterprise: Michael Ryder and Johan Seynnaeve.

I would also like to mention the fellow students in the Linguistics Department whose friendship has been a source of great joy to me: Penka Stateva, Arthur Stepanov, and Sandra Stjepanović, as well as my classmates Marcela Depiante, Deborah Chen Pichler, and Maki Yamane. It has also been a great pleasure to keep the company of Cedric Boeckx, Tien-Hsin Hsin, Masao Ochi, Heehon Park, Suba Rangaswmai, Doug Wharram, and other students from the Linguistics and other departments in Storrs, who are too numerous to mention here. I also wish to thank Judith Marcus for her friendliness and help with administrative issues.

Finally, I want to acknowledge the generous help and financial support that I received from the University of Connecticut and the Linguistics Department as a graduate student in Storrs.

CONTENTS

Chapter 1	Displacement in the Minimalist Program: Issues and Background Assunmptions	1
1.1.	The Issues	5
1.2.	Background Assumptions	10
	1.2.1 Triggers of Displacement	11
	1.2.2. Structural, Inherent, and Quirky Case	15
	1.2.3. Structural Case and Object Shift	19
	1.2.4. Object Shift and Split-VP	27
Chapter 2.	Feature Interpretability and MLC	36
2.1	Superraising	36
2.2	'Crossing Structures'	39
	2.2.1 'Seem-to' Construction	40
	2.2.2 Psych Construction	51
	2.2.3 Crossing in Ergative Languages	61
	2.2.4 To-dative Ditransitive Construction	70
2.3.	'Freezing' Environments	87
Chapter 3.	Movement and Interpretability	94
Chapter 4.	Expletive Insertion	103
Chapter 5.	Further Issues	129
5.1	A'-Movement	129
5.2	Form Chain and Successive Cyclicity	136
5.3.	Agreement vs. Case	150
Chapter 6.	Conclusion	154
References		156

Chapter 1

Displacement in the Minimalist Program: Issues and Background Assumptions

The phenomenon of displacement, i.e. the fact that phrases in natural language may occur in positions other than those in which they are interpreted, has been the main research topic in the generative approach to language from the beginning. In particular, the research in the field has been trying to answer two important questions: first, why is there displacement in language (sometimes formulated as "what drives movement in natural language"), and second, what locality constraints govern this phenomenon. Studies of locality effects in the Principles and Parameters framework include Chomsky (1986a), Rizzi (1990), Cinque (1990) Lasnik and Saito (1992), Manzini (1992) and many others.

In the Minimalist Program (Chomsky 1993 and subsequent work), displacement (i.e. movement/attraction) is taken to be a last resort operation driven by the need to eliminate uninterpretable features from lexical items. In Chomsky (1995), uninterpretable features are eliminated through the mechanism of feature checking, which obtains when one feature is attracted by a head containing its matching feature. The elimination of uninterpretable features is taken to be the requirement imposed on the computational system by 'legibility conditions' which hold at the interface levels of PF and LF. A crucial locality constraint holding of movement/attraction is assumed to be the Minimal Link Condition (MLC), which is formulated in Chomsky (1995) as the Closeness part of the definition of the operation *Attract/Move* given in (1).

1

(1) Attract/Move

K *attracts* F if F is the closest feature that can enter into a checking relation with a sublabel of K. (p.297)

Closeness-Version I

If β c-commands α and τ is the target of raising, then β is closer to τ than α unless β is in the same minimal domain as (a) τ or (b) α . (p.356)

Closeness – Version II

If β c-commands α and τ is the target of raising, then β is closer to τ than α . (p.358)

Minimal domain of a feature/category α is defined in (2)¹:

- (2) Suppose α is a feature of an X⁰ category. Then
 - a. Max (α) is the smallest maximal projection including α .
 - b. The *domain* $\delta(\alpha)$ of α is the set of categories included in Max(α) that are distinct from and do not contain α .
 - c. The *minimal domain* Min(δ(α)) of α is the smallest subset K of δ(α) such that for any γ ∈ δ(α), some β ∈ K reflexively dominates γ. (p.299)

Version (I) of the definition of *Closeness* in (1) defines locality in terms of c-command and 'equidistance.' It was adopted in Chomsky (1993) (under a slightly different formulation) to allow for the covert raising of the object DP across the trace of the subject in the

specifier position of a VP. In Chomsky (1995), a version of this movement (i.e. movement/merger of an XP in the outer specifier position of a head) is allowed by case (a). Case (b) allows arguments in the same minimal domain of a predicate or specifiers of the same head to be equidistant from a potential target. Both cases are represented in what Chomsky takes to be a proper analysis of object shift structures in Icelandic in (3). Case (a) is represented by the attraction/movement of the object DP *bókinai* ('the book') to the outer Spec of vP, whereas case (b) is illustrated by the subsequent movement of the subject DP *nemandinn* ('many students') from the inner specifier position of the vP to the specifier position of the matrix T.

(3) [TP Nemandinn_j las $[_{vP}$ bókinai_i t_j t_{v-V} $[_{vP}$ ekki t_v t_i]]] students-the(NOM) read book-the(ACC) not 'The students didn't read the book'

The same pattern is claimed to be illustrated by object movement in English questions.

(4) $[_{CP} What_i did [_{TP} John_j T [_{vP} t_i t_j v [say t_i]]]]$

¹ The definition is modified to accommodate for Chomsky's (1995) suggestion to refer only to domains of trivial chains (p.356).

Here the movement of the object DP to the outer SPEC-vP crosses over the subject DP, which then raises across the object DP to SEC-TP. Subsequent movement of the wh-DP object to SPEC-CP yields the surface order.²

The simpler and more restrictive Version II eliminates equidistance from the computation of locality, and refers only to the structural c-command relation between the elements whose closeness to the target is being determined.

In Chomsky (2000, 2001), uninterpretable features are assumed to delete through the operation *Agree* which obtains between an uninterpretable feature of a lexical item and the matching feature of another lexical item in the former's domain. The former set of features is referred to by Chomsky as the 'probe', and the latter, the 'goal'. One difference between this formulation and the one given in (1) is that *Agree* does not involve movement/attraction of the matching feature to the local (i.e. 'checking') domain of the attracting head, but takes place 'at a distance'. The locality must still be satisfied, in the sense that the closest matching goal G in the domain D of the probe P enters the operation *Agree*. The relevant assumptions are given in (5):

(5) (i) matching is feature identity

(ii) D(P) is the sister of P

(iii) locality reduces to closest c-command (Chomsky 2000, p.38, ex.(40))

 $^{^2}$ In Chomsky (2001), the analysis is modified in a way that appears to allow countercyclic operations but is consistent with a phase based approach he espouses. Also noteworthy is the fact that such an analysis appears to allow Chomsky (2001) to dispense with equidistance, although little discussion is devoted to the case that it was originally introduced for, the raising of object across the subject DP.

(5iii) is qualified to accommodate for 'equidistance', as in (6) and (7).

- (6) Terms of the same minimal domain are 'equidistant to probes' (p.38 ex..41)
- (7) The minimal domain of a head H is the set of terms immediately contained in projections of H (op. cit. p.38)

1.1 The Issues

The above definitions of the operation *Move/Attract* and the locality constraint it incorporates raise several issues whose resolution is of crucial importance for the proper characterization of the displacement property of human language and which are therefore the central topic of this thesis. One issue that arises is whether the complex mechanism of 'equidistance' is necessary in the system. It will be argued in this study that this mechanism can and must be dispensed with. This result is conceptually desirable given that version II provides for a much simpler and more constrained computation of the locality of movement.³

The central concern of this study, however, is the question of what features are subject to *Attract*. A distinction introduced in Chomsky (1995) which plays a crucial role in this thesis and in much of the current research is the distinction between interpretable and uninterpretable features of lexical and functional categories. Interpretable features are taken to include categorial and ϕ -features of nominal categories, while uninterpretable features are assumed by Chomsky to be structural Case features on nominals and

³ Koizumi (1995) proposes a split-VP hypothesis which he claims eliminates the need for equidistance in the system. See the next chapter for some discussion of differences between Koizumi's analysis and mine.

 ϕ - and structural Case features on functional heads T(nominative) and, as I assume in this thesis, Asp (accusative), as well as wh-features on DPs and on C. Chomsky (1995 and subsequent work) assumes that it is interpretable features of a goal that are being attracted/sought by and that eliminate the matching uninterpretable features of the attractor/probe. Thus, in Chomsky (1995) the feature attracted in A-movement is the categorial feature D, whereas in Chomsky (2000, 2001) A-movement involves *Agree* of interpretable ϕ -features of a DP with the matching uninterpretable ϕ -features of T(ense). The elimination of an uninterpretable structural Case feature from the host DP is taken to be parasitic on this basic relation. In fact, Chomsky (2000 and later work) takes structural Case to be just a reflex of the agreement relation between the probe and the goal.

The above assumptions require a set of additional stipulations in order to explain the core movement paradigm. First, in order to account for the grammaticality of the sentence in (8), specifically the fact that the DP *John* is allowed to cross the closer DP *Mary*, one has to assume that inherent (i.e. semantic) Case 'suppresses' the ϕ -features of nominals, which renders them invisible to the computational system.⁴

(8) John_i seems to Mary t_i to be intelligent.

 $^{^4}$ I follow Chomsky in assuming that the experiencer DP c-commands into the embedded clause. The evidence for this assumption comes from structures such as (i) in which the impossibility of coreference between *him* and *John* is plausibly attributed to Condition C of the Binding Theory:

⁽i) *They seem to him, to have insulted John,

Other evidence comes from Condition A and NPI-licensing effects. The Condition A example is given in section 2.2 below. NPI effects can be observed in (ii) where the NPI *any good argument* is presumably c-commanded and licensed by *no linguist* at some point in the derivation:

⁽ii) John seemed to no linguist to have made any good argument.

This assumption requires a whole theory of relations among features from which such a suppressive role would naturally follow, a theory that is as yet unavailable.⁵

Second, one has to assume that uninterpretable features play a role of making their host 'active' in the system, in order to account for the fact that a DP whose structural Case has been checked off is frozen for A-movement. This assumption rules out ungrammatical structures such as the one in (9), in which the inactive DP *it* has been attracted to check the EPP feature of the matrix T. Again, a theory in which such active status of DPs would follow without stipulation is clearly preferable.⁶

(9) *It_i seems [$_{TP}$ t_i was told John that...] ...

Finally, Chomsky assumes that there is a 'defective intervention' mechanism which prevents the matching features of a DP from being attracted in case there is a closer inactive DP (i.e. DP whose uninterpretable features have been eliminated) which blocks such attraction despite the fact that the DP's interpretable features themselves cannot be attracted due to its inactive status. The main motivation for this assumption is the

⁵ Among other things, such a theory needs to explain why inherent Case suppresses some features and not others, like features relevant for wh-movement and A-scrambling (see examples from Serbo-Croatian and English in the next Chapter). The problem is worse if these are also taken to be ϕ -features.

⁶ But see the proposal in Vukić (1997, 1998, 1999a), where the inertness for A-movement of a DP whose structural Case has been checked off was taken to be a consequence of a general condition which requires interpretable features to be licensed by uninterpretable features for the purposes of the checking theory, which I termed the *Condition on Feature Licensing* (CFL). The main difference between Chomsky's proposal and mine in the relevant respects is that I had assumed that unlicensed features do not count in locality considerations since they cannot enter a checking relation and consequently cannot be attracted. That is, I did not assume a mechanism parallel to Chomsky's 'defective intervention'.

ungrammaticality of the 'superraising' in (10) which the assumption correctly rules out, as explained in the next section.

(10) *John_i seems that it was told t_i that ...

My main proposal in the dissertation is that only uninterpretable features play a role in the operation *Attract*. That is, I assume that the feature attracted in A-movement is uninterpretable structural Case, and not interpretable D- or ϕ -features. Similarly, A'-movement involves attraction of the uninterpretable wh-feature in the wh-phrase, and not some interpretable feature.

It follows from this assumption that inherent Case is not a candidate for attraction since, being associated with Θ -role assignment (Chomsky 1986b), it has semantic content and is therefore not an uninterpretable feature. However, an inherently Case-marked DP can freely be attracted in my system if some other uninterpretable feature, say wh-feature, of the relevant DP is being attracted.

It also follows from my proposal that only DPs with uninterpretable features are active, given that uninterpretable features are the only candidates for attraction. Finally, since the computational system only sees uninterpretable features, there is no defective intervention, only attraction of the closest matching uninterpretable feature.

The organization of the thesis is as follows. In the remainder of this chapter, I spell out the general background of assumptions relevant to the topic. In Chapter 2, I show that Chomsky's (1995 and later work) analysis, particularly the assumption that

interpretable features are subject to the operation Attract, runs into several empirical problems, both undergenerating grammatical sentences and allowing ungrammatical ones. In Chapter 3, I propose a modification of the definition in (1) to the effect that only uninterpretable features are subject to the operation Attract/Move. I then show that my proposal correctly accounts for data discussed in Section 2 and for other empirical data. In Chapter 4, I offer an analysis of (10) in terms of the timing of expletive insertion in the structure. I propose that the timing is determined by the principle of Late Expletive Insertion (LEI), which is shown to follow from independently needed principles in Chomsky's (1995) system, specifically from the ban on deletion of syntactic terms which, in the case of expletives, necessitates the presence of unerased (though deleted) uninterpretable features at the interface. Chapter 5 extends the analysis presented in Chapter 3 to the core cases of wh-movement and discusses some further issues related to the main proposal of the thesis. It is shown that my proposal offers a way to derive Pesetsky's (1982) Path Containment Condition in the minimalist framework and to account for the well known argument/adjunct asymmetries in extraction. I also show that the analysis presented in the foregoing chapters can be reconciled with evidence for the successive cyclic character of movement by adopting one of the current versions of Chomsky's (1993) Form Chain conception of movement. The chapter also discusses the relationship between Case and agreement and presents arguments against Chomsky's assumption that structural Case is but a reflex of the agreement relation. Chapter 6 recapitulates the main conclusions reached in the thesis.

1.2 Background Assumptions

The framework adopted in this study is the Minimalist Program (Chomsky 1993 and subsequent work), a version of the Principles and Parameters approach to natural language whose general goal is to characterize a native speaker's knowledge of his language in terms of invariant (universal) principles of language structure and parameters which are conceived of as choices that particular languages make from a highly restricted set of universally available options (see Chomsky 1981, 1986b for a foundational statement of the P&P framework).

I assume with Chomsky that human language is a generative procedure which maps lexical items to expressions that are interpreted at the two interface levels of PF and LF. I also follow Chomsky in assuming that particular derivations are based on arrays (or, more precisely, 'numerations') selected from the lexicon, and that syntactic objects are constructed from these through employment of the two basic operations: Merge and Move. The former operation takes two syntactic objects and combines them into a new object. The latter rearranges lexical items in a constructed object by displacing its constituent(s) from one position in the structure to another. *Merge* yields two basic relations among constituents: *sister* and *immediately contain*. Other basic relations such as *contain, identity*, and *c-command* follow from these in the obvious manner. Particularly important to bear in mind is the definition of c-command in terms of sister(contain), because this relation plays a crucial role in constraining admissible configurations for displacement, as in (1) and (5) above.

1.2.1 Triggers of Displacement

A central distinction in Chomsky (1995 and thereafter) is that between interpretable and uninterpretable features of lexical items. Chomsky assumes that the human language faculty makes available a universal pool of features from which particular languages make specific selections to build lexical items. The features come in three types: phonological, semantic and formal. Phonological features are those that are accessed at the PF interface, while semantic features are interpreted at LF. Formal features are those which are accessed by the computation, and which may or may not be interpretable (see above for Chomsky's and our assumptions about which features are interpretable and which are not). Objects that contain uninterpretable features at one of the interface levels are 'illegible' and cause a derivation to 'crash'. Thus, elimination of uninterpretable features is a necessary condition for convergence.⁷

We can be quite precise in isolating the very uninterpretable feature that drives movement in Chomsky's system. This is the EPP feature, which, according to Chomsky determines whether a probe α "offers a position for movement and if so, what kind of category can move to that position" (Chomsky 2001, p.3). What exactly the content of EPP is and how it determines what kind of category can move to that position is a notoriously difficult problem and is left unspecified, but it is explicitly assumed to be the feature that drives movement by requiring that its host have an XP merged in its specifier

⁷ This is not to say that elimination of uninterpretable features necessarily yields a coherent interpretation at the interface. Since derivations are not driven by the 'search for intelligibility' (Chomsky 1993), elimination of uninterpretable features can result in objects that are 'legible' but not semantically coherent. To use Chomsky's (1995), phrase, derivations can 'converge as gibberish'. In other words, elimination of uninterpretable features is a necessary, but not sufficient, condition for intelligibility.

position. We know that for Chomsky other uninterpretable features allow movement (by making the XPs that contain them 'active') but do not seem to require it since they can be checked off *in situ* as, for example, in the case of 'long distance agreement' between the ϕ -features of T and a nominative object, or under Chomsky's analysis of ECM constructions with expletives such as (11)

(11) [C [we $[v^*P v^*-expect [there to arrive a man]]]]$

According to Chomsky, the derivation of (11) involves the checking of ϕ -features between the v*, the expletive, and the associate (whose Case features are also checked off in the process). Thus, uninterpretable ϕ - and Case-features are taken to be checked off without movement, which is therefore only required by EPP.

I adopt a view in this thesis according to which overt movement is related to a property of a feature, but this feature is not EPP.⁸ We take the features responsible for

⁸ In this study, I will stay rather silent about the EPP. In general, the analysis adopted in this study is compatible with two approaches to EPP. On the first approach, the EPP is a real condition of the grammar, but one which cannot be responsible for driving movement. This is consistent with positions such as the one adopted in Lasnik (2001) on which the EPP is not a feature: Since I take displacement to be a feature-driven phenomenon, it follows that it cannot be implemented by EPP. On this view, EPP resembles other principles such as, for instance, those of the Binding Theory, which are operative in the grammar but depend on other mechanisms in order to be implemented. In other words, movement cannot be driven by the need to satisfy EPP any more than it can be driven by the need to provide binders for anaphors. Similarly, if independent mechanisms fail to implement the EPP condition, the resulting sentence is degraded just as it is degraded if independent mechanisms fail to provide antecedents for the purposes of the BT. I would like to leave open the exact module in which the EPP operates, although I am aware that at first blush such a view would tend to suggest that EPP is an interpretive condition.

overt movement to be structural Case and wh-features, perhaps uninterpretable features generally. There are various ways of implementing this idea, and not much in this thesis will hinge on a particular implementation, I believe. For concreteness, I propose an implementation that follows closely in the spirit of Chomsky (1995). I suggest that uninterpretable features on functional heads can act as *attractors*, in the sense that they may force the matching feature with which they enter into a checking relation, or a host XP containing the attractee, to merge in the vicinity of their head. I assume that the matching and attraction takes place once the functional head is merged into the structure, as in Chomsky (1995). The amount of the material affected by attraction depends on morphophonological constraints operative in a particular language, and is subject to parametric variation. Alternatively, languages may differ in whether they allow feature separation at all.⁹ Thus English does not allow feature separation and, consequently, requires category movement, while Japanese is a good candidate for a language that allows feature separation. On the latter view, we do not need to assume the 'twomovements' proposal by Chomsky (1995) in which feature movement and pied-piping are separable processes resulting in two distinct chains. Movement always results in a single chain, i.e. feature chain or category chain, depending on how the language in question

On the second approach, EPP does not exist as an independent condition in the grammar, and the effects attributed to this principle are derivable from other independent conditions. This is the view adopted in Bošković (2001) and references therein (see also below).

I believe that my analysis is for the most part neutral between the two approaches indicated in the preceding paragraphs, although there will be cases in which adopting a particular solution to a problem would favor one approach over the other. See end of chapter 4 for one such example.

⁹ For arguments for the former view, see Lasnik (to appear) and section (2.2.3) of this thesis.

relates to the above proposed parameter.¹⁰ In any event, I assume in the spirit of Chomsky (1995) that a strictly cyclic derivation is accomplished under the requirement that an uninterpretable feature on a functional head must be eliminated before the phrase which the element is the head of is embedded via subsequent structure building operations.

If a feature is not an attractor, or if it has satisfied its attracting property,¹¹ then it can enter a checking relation with a matching feature 'at a distance', i.e. without any feature or category movement, provided the above mentioned cyclic derivation is satisfied. I suspect that these options suffice to yield a good deal of language variation regarding displacement.

A departure from Chomsky's proposal is that we locate the attracting property in uninterpretable features of functional heads generally, rather than in a single (categorial)

¹⁰ This is not too different from what appears to me to be the desirable conception of Move even in Chomsky's recent proposals. That is, the operation Move appears to be not just complex in the sense that it comprises various subcomponents, but also to an extent *syncretic*, in the sense that these subcomponents appear to be blended into a single operation. To see this, consider a sentence such as (i) at the point when T with uninterpretable ϕ -features is merged in the structure.

⁽i) [t was indicted the president]]

Assume that *Move* combines *Agree*, *Select*, and *Pied-pipe a category*. If these subcomponents are taken to operate independently, then one would have to explain why, once *Agree* has taken place between the ϕ -features of T and those of the object DP with the resulting elimination of the uninterpretable Case feature on the latter, the DP is not immediately deactivated, barring *Select* and *Pied-pipe*. The simplest account, i.e. the one involving no look-ahead or extrinsic ordering, is to assume that Move is syncretic in the suggested sense.

¹¹ The latter option may be relevant for cases where one feature enters a checking relation with more than one 'goal' as, for instance, in multiple wh-constructions, or in dative-nominative constructions in languages with quirky Case. For these cases, we need either to assume with Chomsky (1995) that a feature can enter multiple checking relations, or that a functional head can contain multiple instances of a same feature type, perhaps as a strictly Last Resort option. Another option would be to incorporate *Attract All* mechanism in the grammar proposed in Bošković (1999), on which multiple features can be attracted in a single step. I leave the exploration of this issue and possible reductions for later research.

feature of a functional head. There are other ways in which our proposal may converge with or differ from Chomsky's, depending eventually on what turns out to be the proper resolution of some current controversies. For example, whether the attracting property of uninterpretable Case features needs to be relativized or not to a particular value of the feature, or to the head hosting the attractor, in the manner of Chomsky (1995), will depend on such questions as whether object movement in ECM constructions is nonexistent (as Chomsky assumes), optional (Lasnik 1999b), or obligatory (Lasnik 1995c, Bošković 1997, 2001). The same holds for object shift in simple transitive constructions. Thus, if Chomsky's proposal is correct, we would need to assume that structural Case is an attractor on T but not on v. If Lasnik (1999b) is correct, the conclusion would be that Case is allowed but not required to act as attractor on v. Bošković's arguments would require the Case feature to be attractor on both heads. I return to Lasnik's and Bošković's arguments directly. But firstly, I need to spell out my assumptions about the Case system.

1.2.2 Structural, Inherent, and Quirky Case

I will assume that natural languages exhibit a tripartite distinction between structural, inherent, and 'quirky' Case.¹² The distinction between the structural and inherent Case is based on Chomsky (1981, 1986b, and subsequent work). In the earlier of the cited works, structural Case is taken to be assigned by heads of certain lexical

¹² For a recent overview and an attempt to collapse various categories of Case by using some of the diagnostics I mention below, see Svenonius (2002). For a particularly well worked-out configurational approach to Case cast in the Government and Binding framework, see Bittner and Hale (1996a).

categories (namely those characterized by feature [-N], that is V, P and finite Infl/Agr) in certain structural configurations involving such notions as government, adjacency, or specifier-head configuration.¹³ In Chomsky (1989, 1993), those Case-assigning categories are taken to be various members of the Agr system, and the requisite 'licensing' configuration is taken to be only Spec-head. In Chomsky (1995), which is essentially the system that I assume, structural Case is taken to be a semantically uninterpretable feature of nominals and functional heads such as v or T. I also assume with Chomsky (2000) that structural Case features on nominals are *unvalued*, given that they are fully dependent on the value of the specific structural Case assigner, such as T or v (or aspectual head Asp, which we will assume to be in fact the locus of accusative Case assignment). I find this to be a natural and elegant way to capture the traditional distinction between the Case assigners, whose Case features are invariant, and Case assignees, whose Case features depend on the particular structural configuration., i.e. the vicinity of a particular Case assigning head. Note, however, that whether features are valued or not as they enter the structure is not correlated to their interpretability. That is, Case features are uninterpretable on both the assigners and the assignees, but are only fully specified on the former. So the distinction is not dependent on interpretability, but rather on predictability of feature values.

Inherent Case, on the other hand, is assigned by lexical items, such as verbs or nouns, in the local configuration of sisterhood, and is associated with θ -role assignment.

¹³ Chomsky (1986b) takes Ps to be *oblique*, i.e. inherent Case assigners, in contrast to the majority of transitive verbs which assign *objective* structural Case.

This association explains the inability of a noun to assign an inherent genitive Case to the subject of its infinitival complement with which it is not θ -related, as argued in Chomsky (1986b):

(12) *Mary's belief [TP (of) John to be innocent]

Crucially for our purposes, this association with θ -role assignment makes inherent Case *interpretable*, which, under my reformulation of the operation *Attract* given in Chapter 3, has the consequence of rendering the host DP ineligible for A-movement. This is because I claim that only uninterpretable features play a role in the checking theory, in contrast to Chomsky (1995, 2001, 2001), who posits that it is interpretable features that are sought in the operation *Agree*.

Apart from this difference in interpretability, structural Case is traditionally taken to be distinguishable from inherent Case by such diagnostics as the predictability of their assignment and the ability to alternate with other structural Cases. That is, apart from being 'licensed' in predictable structural configurations such as SPEC-head, structural Case changes its value when the requisite syntactic configuration is altered in some way, typically as a consequence of addition of various morphology onto the verb. Thus, the accusative Case borne by direct object in Serbian alternates with the structural nominative in, for example, passives constructions (13a-b), which is not possible for an inherent Case such as instrumental (13c-d). This is a well known and widely attested phenomenon. (13) a. Petar je obavijestio Mariju_{ACC}

Peter has informed Mary

- Marija_{-NOM} je obavještena.
 Mary has been informed
- c. Petar je zavladao masama-INSTR.

Petar has gotten to control the masses

d.*Mase-NOM su zavladane.

The masses are controlled.

'The masses have been subjected to control'

Other distinguishing characteristics of structural cases are taken to be their *unmarked* value (e.g. nominative DP is sometimes taken to be *Caseless*, as in Bittner and Hale 1996a, or lacking any positive feature specification, as in Jakobson 1936), appearance in 'neutral contexts' (cf. Franks 1995 on Case in Slavic, for example), etc.

Finally, 'quirky' Cases fall between the two discussed categories in that, although they appear to be related to a particular thematic role, they behave syntactically as if they were structural (see Zaenen, Maling, and Thráinsson 1985 for a battery of tests showing that quirky DPs can and do move to subject positions). A well studied case is the 'quirky' dative in Icelandic. Following rather general practice, I take a quirky Case to be a combination of an inherent and structural Case (see Bošković 2001, Belletti 1988, Chomsky 2000:127, Cowper 1988, Frampton and Gutmann 1999, Freidin and Sprouse 1991, Holmberg and Platzack 1995). For the purposes of this thesis, quirky Case is treated on a par with 'regular' structural Cases in that it undergoes a checking relation with the matching Case feature on a functional head which results in its elimination. We now turn to the discussion of whether structural Case features are universally attractors or whether their attracting property needs to be relativized to properties of their hosts. This turns on the question on whether raising is obligatory or optional in ECM and transitive constructions, as discussed above.

1.2.3 Structural Case and Object Shift

Lasnik discusses overt object shift in several papers, perhaps most extensively in Lasnik (1999b). The first argument that object shift in ECM constructions is possible comes from binding facts illustrated in (14).¹⁴

(14) The DA proved [two men to have been at the scene of the crime] during each other's trial.

Lasnik observes that *two men* in (14) can license the reciprocal *each other* in the matrix clause. Under the standard assumption that anaphor binding requires c-command, it must be the case that the DP has moved to a position in the matrix clause from which binding is possible. As Lasnik observes, the argument can be reproduced to cover other

¹⁴ This argument follows in essence the 'height' arguments presented in Postal's (1974) and reiterated in Lasnik and Saito (1991).

cases standardly assumed to fall under the purview of c-command, such as weak crossover related phenomena and NPI-licensing.

Another argument comes from ECM constructions with pseudogapping, such as (15a) which at first glance appear to affect a noncontiguous string (as in the representation in (15b). Taking the position that ellipsis must involve syntactic constituents and assuming the 'Split-VP' phrase structure proposed in Koizumi (1995), Lasnik argues that the derivation of (15a) involves VP deletion which follows the movement of the 'remnant' to a position higher in the structure (which he assumes to be AgrO), as in (15c). Specifically, the derivation of (15a) involves the raising of the infinitival subject *every Fibonacci number* to the specifier position of the matrix AgrOP, followed by deletion of the matrix VP which includes both the matrix verb and the embedded infinitival.

(15) a. Mary proved every Mersenne number not to be prime, and John will every Fibonacci number.

b. Mary proved every Mersenne number not to be prime, and John will *[prove* every Fibonacci number -not to be prime].

c. Mary proved every Mersenne number not to be prime, and John will [every
Fibonacci number]_i *[prove t_i not to be prime*].

Thus, in addition to explaining the binding facts noted above, the assumption that the subject of the embedded infinitive raises to a position in the matrix clause provides a

simple explanation of the pseudogapping facts, allowing one at the same time to maintain a rather standard assumption that deletion operations affect only syntactic constituents.^{15,16}

Another interesting aspect of (15) is the scope of the universal quantifier with respect to negation, which brings us to the issue of whether the movement in ECM constructions is optional or obligatory. Lasnik claims that it is optional. His argument is based on the interplay of the scopal relations and particle positioning in ECM constructions such as (17). He observes that, in contrast to the sentence in (16) and (17b), in which the universal quantifier can have either narrow or wide scope with respect to negation, the sentence in (17a) is unambiguous: the universal quantifier must have wide scope in (17a).

(16) I believe everyone not to have arrived yet.

(17) a. The mathematician made every even number out not to be the sum of two primes.

b. The mathematician made out every even number not to be the sum of two primes.

Assuming that A-movement does not reconstruct,¹⁷ and taking the particle position to be fixed and indicative of the border between the matrix verb and the

¹⁵ See Kennedy (to appear) for arguments that ellipsis operations target syntactic configurations.

¹⁶ Other accounts have been proposed that have the same effect of preserving the constituency requirement on deletion, such as, for example, analyses under which the 'remnant' of ellipsis undergoes extraposition or 'heavy NP shift' prior to the deletion operation. See Jayaseelan (1990) for such an analysis, and Lasnik (1995c) for some problems.

¹⁷ Evidence for this assumption is provided by examples such as *Everyone is believed not to have arrived yet.*, which, in contrast to (16), is scopally unambiguous in that it only allows for the wide

embedded infinitival, Lasnik claims that the above scope facts follow from overt raising of the subject in (17a) and lack thereof in (17b), a state of affairs which is only consistent with the assumption that raising in ECM constructions is an optional, not an obligatory, process. Lasnik then goes on to show that the analysis makes correct predictions for other data that we have discussed, such as anaphor binding.

(18) a. The DA made the defendants out to be guilty during each other's trials.

b. *The DA made out the defendants to be guilty during each other's trials.

Lasnik's arguments point strongly to the conclusion that raising in ECM constructions is optional, a conclusion which, as I have stated above, would appear to require a modification of the assumption that structural Case features generally are attractors by relativizing it to the properties of a head hosting the feature, an unwanted complication. So one is led to seek an alternative interpretation of the data presented by Lasnik.

Bošković (2001) points to a way towards a reinterpretation which keeps object shift obligatory in ECM constructions. First, with respect to the data in (16), he assumes with Boeckx (2001a) that the wide scope of negation can be obtained by a QR-like raising of *not*, which places the latter in a position above the universal quantifier. To account for the lack of ambiguity in (17a), he proposes that, instead of assuming with Lasnik that the particle position is fixed in these structure, one could take the position of the NP to be fixed and the position of the particle to vary. Then, if one assumes that the scope of

scope reading of the universal. See Boeckx (2001a) for some relevant discussion on this issue, i.e. for evidence that at least some DPs such as indefinites must be allowed to reconstruct.

negation is delimited by the phrase headed by the particle (which, for expository convenience, Bošković takes to be OutP), the scopal facts would follow. The relevant structures resulting from Bošković's proposal are given below.

(19) a. The mathematician made [Agrop every even number [OutP out [TP not to be the sum of two primes.]]]
b. The mathematician made [OutP out [Agrop every even number [TP not to be the sum of two primes.]]]

As for the binding facts in (18), Bošković suggests that they may follow from specifics of adverb placement in verb-particle constructions, namely from the assumption that adverbs are adjoined to OutP. Since, under this proposal, the anaphor in the adverbial phrase would be c-commanded in (18a) but not in (18b), the binding facts are accounted for.¹⁸

A similar situation holds with respect to object shift in simple transitive constructions. Here, too, there appears to exist strong evidence that movement of the object DP is optional. Consider the pair in (20), taken from Bošković (2001):

(20) a. Who did Bill select [a painting of t]?

b. ?*Who was [a painting of t] selected?

¹⁸ As noted by William Snyder (p.c.), however, this account must assume that adverbs are *necessarily* adjoined to OutP, which raises the questions about where such adverbs are placed in

(20b) is a standard example of Subject Condition violation or, in terms of Takahashi's (1994) theory, a violation of the ban on extraction out of heads of nontrivial chains. Bošković's point is that, if objects shift also takes place in (20a), then one would expect that it would be as degraded as (20b) since it too would involve extraction from a head of a nontrivial chain. If, on the other hand, object shift is optional, then the contrast in (20) follows.

A related argument is given in Lasnik (1999). Lasnik assumes that adjuncts are base-generated higher in the structure than complements of a verb. Thus, the only way for an object to bind an anaphor in an adjunct in (21a) is if the object DP *two aides of a senator* has raised to a position from which the requisite binding can take place.

- (21) a. The special prosecutor questioned two aides of a senator during each other's trials.
 - b. ??Which senatori did the special prosecutor question [two aides of ti]during each other's trials.

Under this account, the degraded status (21b) then follows from the interaction of independent constraints of the grammar. On the one hand, Condition A of the Binding Theory requires that the reciprocal be c-commanded by its antecedent, which in (21b) can only be accomplished through movement of the object. On the other hand, such movement turns the raised DP into a head of a nontrivial chain from which extraction is disallowed. Thus, (21b) must involve a violation of one of the two requirements. This contrasts with

structures without particles and, consequently, what prevents such placement in the structures under consideration.

(20a) above in which the extraction out of the object DP has taken place and yet the sentence is grammatical, which therefore must mean that the object DP has not raised, supporting the argument for optionality of object shift.

Parallel reasoning holds of the pair in (22). In (22a), the object DP *a photograph of Mary* must have raised from the subsequently elided VP, on Lasnik's assumptions. This is supported by (22b) whose ungrammaticality then involves an extraction from a head of a nontrivial chain. Again, the contrast between (22b) and (20a) suggests that object shift, though allowed (and in some cases required in order to derive a grammatical sentence), is not obligatory.

(22) a. Bill selected a painting of John, and Susan should [a photograph of Mary] $_{i}$

 $[v_P select t_i]$

?*Who will Bill select [a painting of t], and who j will Susan [a photograph of t j]i

[vp select t i]

Does the data above force us to give up the position that structural Case features are generally attractors? Bošković (2001) suggests that the data can still be made consistent with the position that object shift driven by structural Case is obligatory, if one assumes that the object DP can bear either structural or inherent Case. In the latter Case, it does not need to raise since the inherent case is 'checked' in situ (see the following section). If it bears the structural accusative Case, the DP has to raise in order for the feature checking to take place.¹⁹ This means that, in the grammatical (20a) the extraction has taken place out of the inherently Case marked DP. In the ungrammatical examples, the extraction has taken place out of the structurally Case marked DP which must have raised in order for the Case checking to take place. Bošković concludes that object shift is obligatory in ECM and optional in simple transitive constructions, but that this optionality does not provide a counterexample to the assumption that structural Case obligatorily drives overt movement in English.

In the light of the above discussion, and in the absence of decisive arguments to the contrary, I will assume that object shift is obligatory in both ECM and 'regular' transitive constructions when driven by structural Case, along the lines of Bošković (2001). This allows me to maintain the strongest assumption that structural Case, and uninterpretable features generally, are attractors on all functional heads, at least in English.²⁰ Note that Lasnik's and Bošković's data provide us with another diagnostic for determining whether a DP has structural Case, under our assumptions. That is, since, on both Chomsky's analysis and mine only structurally Case-marked DPs can undergo raising (with 'quirky' DPs being a special case), it follows that, if we can show that a DP has

¹⁹ Bošković assumes the analysis of Epstein and Seely (1999) and Boeckx (1999) under which the elimination of Case features requires the probe and the goal to enter mutual c-command, which can only take effect via overt movement.

 $^{^{20}}$ This is not to suggest that all uninterpretable features are always present on their potential hosts, or even that the structural Case feature is always present on the relevant functional element such as the aspectual head. The only conclusion to be drawn from the above discussion is that, when the structural Case feature *is* present, it acts as an attractor.

raised, its Case must be structural.²¹ To determine whether a DP has raised, we can use Lasnik's and Bošković's diagnostics: DPs base-generated in the VP that are pseudogapping remnants must have undergone movement, according to Lasnik. Similarly, DPs from which extraction is barred are likely to be heads of nontrivial chains, which by definition means that they have raised.

1.2.4 Object Shift and Split-VP

Given that we assume that structural Case features generally trigger movement in English, we need to specify exactly where the attracting Case feature is located. In this regard, there have been two main proposals in the literature. As noted above, Chomsky (1993 and subsequent work) assumes that accusative Case is checked in SPEC-AgroP or, in more recent work, in SPEC-vP. Since, by assumption, the relevant head on the latter version already contains the subject DP in its specifier position, such movement raises locality issues given that the subject appears to be closer than the object to the specifier position c-commanding the two DPs. This in turn appears to require a mechanism which would allow for such a locality violation and for the violation involved in the subsequent movement of the subject to the Spec-TP position. This is why Chomsky introduced the notion of equidistance, as noted above.

²¹ I'm focusing here on A-movement, since that is the movement that requires an uninterpretable Case feature. As noted, nothing in my analysis prevents A'-movement of inherently Case-marked DPs as long as these DPs contain the uninterpretable feature relevant for such movement, which I take to be the wh-feature. I am also abstracting here from scrambling, which may affect inherently Case-marked DPs but which may require a different uninterpretable feature shared by structurally and inherently Case-marked DPs.

Another approach, such as Koizumi's (1995 and references therein) and Lasnik's (1995 and subsequent work), takes object shift to target a position below the position in which the subject DP is generated. Thus, Koizumi (1995) proposes a 'split-VP' account according to which an agreement projection intervenes between the vP hosting the subject DP and the VP in which objects are generated. This solution, on which the object raises to the Specifier of this intervening Agr projection, allows Koizumi to reconcile the evidence that object DPs undergo overt raising in simple transitive constructions in English, for which he offers several arguments, with empirical evidence that the landing site of this object shift is below the position in which the subject DP is base-generated. For example, Koizumi follows Sportiche (1988) in assuming that floating quantifiers signal the positions of the DPs they are associated with. Thus, the quantifier *all* in (23) is taken to signal the position occupied by the object DP is AGRoP:

(23) I gave [
$$_{AGRoP}$$
 the books_i [$_{VP}$ all t_i t_v to John]]] (p.106, ex. (15a))

Koizumi shows that, if the subject DP were base-generated in a the position lower than the one to which the object moves, the sentences in (24d, e) below would be incorrectly predicted to be grammatical. For example, the ungrammatical (24e) would have the structure given in (25).

(24) a. The men all will have given a book to John.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
- b. The men will all have given a book to John.
- c. The men will have all given a book to John.
- d. *The men will have given all a book to John.
- e. *The men will have given a book all to John.

(25) The men_i will have given $[_{AGRoP} a \text{ book}_j [_{VP} \text{ all } t_i t_j \text{ to John.}]]$

The introduction of the split-VP structure allows Koizumi to propose that the Domain Extension Condition of Chomsky (1993) (and, in effect, equidistance, though see his footnote 11 on p.113) be eliminated from the system.

A small digression is in order at this point. In this study I arrive at the same conclusion about the dispensability of equidistance. However, it is important to note that Koizumi's motivation is crucially different from mine. One of Koizumi's main objections to the Domain Extension Condition is that it allows violations of Rizzi's (1990) Relativized Minimality, which he takes to be an inviolable constraint. In fact, strict adherence to Relativized Minimality in its strongest form leads Koizumi to posit additional structures (such as additional AGR projections) which in effect prevent any crossing pattern of A-movements in the grammar.²² As will be clear from the discussion in the following chapters, his proposal makes crucially different predictions from mine, since my aim is precisely to allow the kinds of derivations that Koizumi wants to prevent, namely

²² See, for example, Koizumi's treatment of the double object construction on pp. 134-136.

the raising of one DP over another DP in an A-position under well defined structural conditions. Consequently, all the constructions that I present as problematic for Chomsky (1995, 2000, 2001) are also problematic for Koizumi (1995). For example, Koizumi (1995) does not discuss the raising 'seem-to' construction in English or the Albanian data from Snyder (1992) which show clearly that Rizzi's RM in its strongest form cannot be maintained. In fact, it was precisely the conclusion reached in Snyder (1992) that "Relativized Minimality, as formulated in Rizzi (1990), does not apply to A-movement (p.1)". Thus, although I reach the same conclusion about dispensability of equidistance as Koizumi, I reach it on different grounds and by proposing an analysis that differs in important respects from Koizumi's. Koizumi dispenses with equidistance because in his analysis no two paths of the same type ever cross or overlap. I dispense with equidistance because, although strong evidence exists that movements of the same type do overlap, they do that via a mechanism that makes equidistance unnecessary. In other words, Koizumi analyzes equidistance away by positing structures in which this mechanism can't possibly play any role. I do away with equidistance by showing that, in structures in which it could potentially be relevant, it makes wrong predictions by being overrestrictive in some cases and too permissive in others.

I do, however, share with Koizumi an analysis in which the paths of the subject DP and object DP in transitive and ECM constructions do not cross, although my reasons are independent of equidistance or Rizzi's version of Relativized Minimality. That is, I find empirical arguments for a position between the 'base generated' positions of subjects and the Case-checking position of objects convincing, such as quantifier float distribution discussed above. What I also take to be highly significant in deciding on the position of accusative Case checking is the absence of connectivity effects in simple transitive constructions, which is unexpected if there is a point in a derivation of these sentences at which the direct object c-commands the subject. For example, take the principle A of the Binding Theory which, as argued by Belletti and Rizzi (1998) appears to be an 'anywhere' condition in the sense that it can be satisfied at any point in the derivation of a sentence. This accounts for the connectivity effects in sentences such as (26) in which there exists a point in the derivation at which the subject of the embedded infinitival is c-commanded by the experiencer DP.

(26) [Each other's photos]_i seem to John and Mary t_i to be too explicit.

(27) *Each other's friends telephoned John and Mary

However, the ungrammaticality of (27) is mysterious under the assumption that objects raise across subjects in simple transitives since on this approach there exists a point in the derivation of (27) at which the anaphor is c-commanded by the object DP. The same argument can be brought with respect to other connectivity effects discussed in Chapter 3 below which are standardly taken to involve c-command, such as (alleviation or lack of) WCO effects, interaction of wh-movement and quantifier scope, negative polarity licensing, etc.

Another reason for assuming that the structural accusative Case is checked below the vP is that it provides an immediate and principled account for why the subject DP does

31

not get Case-marked (or, more precisely, have its Case feature valued and checked off) in the SPEC-head relation with the v which, on the alternative proposal, bears an accusative Case feature. In Chomsky (1995), it is simply stipulated that only expletives can check their Case features via pure Merge. It is totally unclear why this should be the case, particularly for the expletive *it* which in other respects patterns with other DPs and is, in fact, taken to have interpretable nominal features.²³ One might try to relate this situation with the notion of the *domain* and the assumption that expletives are X^0 elements. That is, since the domain of a head α is defined in Chomsky (2000, 2001) as the set of terms contained in the sister of α . XPs merged in the specifier position of α would be outside its domain, hence feature checking would not be possible in such a configuration. However, the problem with this approach is raised by PRO and other elements such as lexical pronouns that can plausibly be claimed to be X^0 s. For example, Bošković (1997) shows that different behavior between full XPs and pronouns with allege-class verbs can be accounted for if one assumes that the latter can act as X⁰s, for which he provides some evidence. Furthermore, Chomsky (2001) explicitly takes PRO to be X^0 (p.19), so this line of reasoning would not be in concert with such an assumption. The assumption that accusative Case is checked below the head with which the subject XP is merged resolves this problem in a simple and principled way.

 $^{^{23}}$ This accounts for its presumed defective intervention effect. That is, if features of the expletive were uninterpretable, they would delete upon checking, and the defective intervention effect of superraising would be lost.

To conclude, I assume that raising in ECM and transitive constructions targets a position below the vP which hosts the subject DP. I will follow Travis (1991) in assuming that the position in which accusative Case is checked is aspectual in nature, although I do not rule out the possibility that it is some head of the category Agr.²⁴ I take the aspectual head to intervene between the vP containing the subject and the VP in which internal arguments of lexical verbs are generated. According to this proposal, the representation of a simple transitive construction and of ECM-constructions would be as in (28) and (29), respectively.²⁵



²⁴ See also Svenonius (2002) for arguments associating Case with aspectual domain.

²⁵ Here and throughout I refrain from committing to a particular theory of verb movement, which may well be a PF phenomenon, as suggested in Chomsky's latest works (see also Boeckx and Stjepanović 2001).

I also assume that agent DPs are 'base-generated' higher in the structure than experiencer/goal/benefactor DPs, which in turn are first merged in a position higher than theme DPs.²⁶ More precisely, the interpretive system reads off thematic roles from merge positions of syntactic constituents whose relative prominence in the constituent structure is consistent with such thematic hierarchies. In contrast to goal/experiencer/benefactor and theme arguments, agents are first merged in a position outside a maximal projection of a lexical verb, i.e. in the specifier of vP.

Finally, I will adopt a version of the proposals on which internal arguments are generated in distinct lexical VP shells (see Marantz 1984, Larson 1988, Koizumi 1995, Lasnik 1995 and subsequent work, McGinnis 1998), although not much in this thesis will depend on the additional VP shell.²⁷ In contrast to Koizumi and Lasnik, however, I do not assume that any Agreement or other functional projections intervene between lexical verbal projections. As noted, Koizumi's proposal is designed to prevent any crossing or overlapping pattern of A-movement in natural language, which is in contrast to the evidence for such movements that has been established and that we present in

 $^{^{26}}$ This view, while configurational in the sense indicated immediately below, is consistent with the thematic hierarchies proposed by numerous authors. For example, Kiparsky (1987) proposes the hierarchy in (i).

agent > benefactive/goal > theme > location

See also Bresnan and Kanerva (1989) Jackendoff (1972), and many others.

²⁷ The one structure whose analysis in this thesis is potentially sensitive to the presence of the additional VP shell is the 'to-dative' ditransitive construction. This construction is discussed in section 2.3.3, where I also provide evidence that the proposed structure is correct. Evidence for this structure of 'standard' triadic constructions is also frequently taken to involve various verbal morphology that surfaces in these structures, such as applicative morphemes in ergative and other languages (see below). Since such evidence is less forthcoming in 'psych' and 'seem-to' constructions (with experiencer arguments) which also involve multiple internal arguments, I will for the most part ignore the possibility of an extra verbal shell in these structures, but it will be clear that not much will depend on this decision.

the following sections. The general ditransitive monoclausal structure with arguments in their 'base-generated' positions is therefore as in (30).²⁸



We now turn to the main focus on this study, In the next chapter, I present Chomsky's theory of feature checking and locality and point to several problems that implementations of such a theory which take interpretable features to be subject to operation Agree/Attract are faced with.

²⁸ This, of course, abstracts from a possibly much more articulated structure argued for by Cinque (1999) and Rizzi (2000). Furthermore, since I follow Chomsky in assuming that vacuous projections do not exist, the positional distinction between the verb's internal arguments is either neutralized when one of the arguments is missing, or involves an unpronounced implied argument, as Hale and Keyser (1993) argue, for example, to be the case with unergatives. Nothing in the discussion below will depend on which of these two possibilities is actually instantiated.

Chapter 2

Feature Interpretability and MLC

As discussed in the introductory chapter, Chomsky assumes that uninterpretable features in a merged lexical item α cause the latter to act as a probe that seeks the closest matching interpretable feature in its domain, where the domain of α is its *sister* (see (5)). Once a matching feature is found in an active goal, *Agree* takes place between the probe and the goal, whereby uninterpretable features of the two are eliminated. Defective intervention obtains when an 'inactive' DP containing features matching those of the probe intervenes between the probe and a lower DP. We now turn to some problems with this account, in particular with the assumption that it is interpretable features that are subject to the operation *Attract* (Chomsky 1995) or *Agree* (Chomsky 2000, 2001).

2.1 Superraising

Let us first illustrate the above set of assumptions by considering how they interact to rule out the ungrammatical superraising structure in (10), repeated here as (31):

(31) *John_i seems it was told t_i that ...

To derive the ungrammatical (31), the target, which in this case is the matrix T, would have to attract the DP *John* across a closer DP, the expletive *it*, which is correctly prevented by the MLC, under the assumption that the categorical D-feature is targeted by the operation. Specifically, at the relevant point in the derivation given in (32), the D-

feature of *John* cannot be attracted to check the EPP feature of the matrix T since there is a closer feature, the D-feature of *it*, which can be attracted for the same purpose.

(32) $[_{TP} T \text{ seems } [_{IP} \text{ it was told John } [_{CP} \text{ that } IP]]] (p.295)$

Chomsky's (1995) D-feature account correctly rules out the derivation of (31). However, as observed in Vukić (1998) (see also Uriagereka and Raposo 1995), this account which relies on the interpretable D-feature is not without problems as it allows a derivation of an ungrammatical construction from the same point represented in (32). Specifically, since the D-feature of *it*, which, perhaps somewhat problematically, is taken to be interpretable by Chomsky, is still present at this point, it is not obvious how to prevent a derivation in which *it* is attracted by the matrix T, with a subsequent attraction of the formal features of *John* by T, whereby the structural Case feature of the latter would be eliminated. In other words, unless additional assumptions are made, the structure in (33) would be predicted to be grammatical:²⁹

(33) *It_i seems t_i was told John that ...

Such additional assumptions are adopted in Chomsky (2000), where the operation through which uninterpretable features are checked is taken to be *Agree* (see the next section for some discussion), and where the relevant features are taken to be ϕ -features. Here the

²⁹ Note that for Chomsky traces or copies of movement do not count as interveners. Thus, the copy of *it* in (33) cannot block the attraction of the formal features of *John*.

blocking effect is attributed to 'defective intervention' of the closer matching ϕ -features in the expletive which prevent *Agree* with the ϕ -features of the DP *John* regardless of the fact that the former cannot enter *Agree* themselves. This inability is taken to be the consequence of the fact that, having already checked off its uninterpretable Case feature, the expletive does not contain an uninterpretable feature and is therefore 'inactive' for the operations of the computational system. (31) is then ruled out as a locality violation, while (33) violates the ban on *Agree* with ϕ -features of 'inactive' DPs.³⁰

The example in (33) illustrates two basic assumptions in Chomsky (1995, 2000) that are relevant for my analysis: 1) it is the closest *interpretable* features (here D or ϕ -features) of lexical items that are being attracted, and b) interpretable features prevent by defective intervention the movement/attraction of lower features in cases when the former cannot be attracted themselves.

As I explained in Vukić (1997, 1998, 1999a,b), these assumptions run into several problems when employed in the analysis of the core movement paradigms. These problems involve inability to account for several grammatical structures without extra stipulations as well as 'overgeneration' of several ungrammatical constructions. The former include 'crossing' structures such as the raising experiencer construction, 'psych'-constructions,

³⁰ Chomsky assumes that, in contrast to the pure expletive *there*, the expletive *it* contains interpretable ϕ -features, a problematic assumption in my view.

to-dative ditransitives (34c), etc.³¹ The latter include ungrammatical structures that violate (a part of) Chomsky's (1986a) Chain Condition or ban on A'-movement from an operator position (Epstein 1993), as well as sentences whose ungrammaticality has standardly been taken to involve 'improper movement.' We consider these cases in turn.

2.2 'Crossing Structures'

Consider the structures in (34).

- (34) a. $[_{TP} [Each other's photos]_i \text{ seem to John and Mary} [_{TP} t_i \text{ to be scandalous}]]$
 - b. [These nasty rumors about each other]_i annoyed John and Mary t_{i.}
 - c. $[_{TP} I \text{ showed [each other's photos]}_i \text{ to John and Mary [VP } t_i]]$

It is argued in Vukić (1997, 1999a, 1999b), and, in view of the connectivity effects illustrated in (34) more-or-less standardly assumed,³² that in each of the structures above one DP crosses another DP (given in boldface) in an A-position. A question then arises as to how these structures can derived in a grammatical system which incorporates locality constraints such as MLC.

³¹ 'Crossing' here only means that one DP raises across another in the derivation of a sentence. It is not means to imply the pattern of movements in the derivation of a sentence exhibits a crossing, in the sense of 'partially overlapping' or 'intersecting', pattern. In fact, I will show below that a legitimate pattern of overlapping A and A'-dependencies is 'nesting' rather than 'crossing'.

³² For a raising analysis of 'seem-to' construction, see Chomsky (1995, 2000, 2001), as well as his earlier works, and a vast body of generative transformational literature which assumes that the subject position in this construction is non-thematic. For a crossing analysis of 'to-dative' ditransitive construction, see, for example, Kitagawa 1994 Vukić (1998, 1999) and, more recently,

2.2.1 'Seem-to' Construction

To begin with, it is clear that in at least some of these structures the mechanism of equidistance is of no avail. Consider the raising experiencer construction in (34a) whose partial structure at the relevant point in the derivation is given in (35).³³



Boeckx (2001). For a representative set of analyses of 'psych' construction, see Belletti and Rizzi (1988), Bouchard (1995), Pesetsky (1995), Grimshaw (1991).

³³ The structure in (35) is based on Chomsky (1995). Note that Chomsky assumes that the experiencer DP is base-generated within the maximal projection of the lexical verb, which is also what we assume in this thesis. Where Chomsky's structure potentially differs from ours is in the absence of an extra projection between the vP and VP. Not much depends on this difference as far as (35) is concerned, I believe. The simplest assumption in this regard is that the aspectual head, which we have posited in ECM and simple transitive constructions above, is not present in this structure. If the aspectual head is present, which is not obviously required, we can assume that it does not bear a Case feature in this construction (which in that case must be optional on Asp) since if it did, the derivation would inevitably crash given that either accusative Case on Asp or the nominative Case on T would be unchecked. Some degree of such optionality must be allowed in any system, I believe. Note, for example, that Chomsky also has to assume that accusative Case feature is optionally present on v, otherwise the structure in (35) would contain an unchecked uninterpretable feature and the derivation of (34a) would crash. Note that, even under this second option, we need not give up the assumption on the obligatory nature of object shift. The Case feature, which is always an attractor, need be present only when there is a DP with whose structural Case feature it needs to enter a checking relation. If there is no such DP, the option of not assigning the Case feature to Asp yields a convergent derivation. If there is, a Case feature must be present on Asp, triggering obligatory object shift.

Before raising, the DP John and Mary is neither in the same minimal domain with the target SPEC-TP nor with the DP *each other's photos*. Thus, neither of the two conditions for equidistance is satisfied at this point. Therefore, the D-feature (and ϕ features) of John and Mary is closer than the corresponding feature(s) of *each other's photos* and the movement of the latter across the former should be disallowed (hence the asterisk in the figure above). This makes the grammaticality of (34a) unexpected.

Chomsky (1995) observes the problem and concludes that "the status of the English construction still remains unexplained, along with other related questions" (p.306). In Chomsky (2000), the grammaticality of (34) is accounted for by assuming that a) Case of the DP *John and Mary* is inherent, and b) inherent Case suppresses the ϕ -features of the relevant DP which makes them invisible for the computational system. This makes possible the attraction of the lower DP *each other's photos*. Note, however, that resorting to a suppressive role of inherent Case is a complex move that implies some theory of feature interdependencies and relations which is not currently available. Clearly, an analysis that can account for the derivation in (34a) without such mechanisms is preferable on conceptual grounds. Finally, the assumption of a suppressive role of inherent Case raises empirical problems. For example, Boeckx (1999) observes that agreement between the verb and the subject of the embedded TP appears to be blocked by the presence of the experiencer DP:

(36) a. There seems to Mary to be a man in the room.

b. ?There seem to Mary to be men in the room.

As Boeckx observes, if ϕ -features of the experiencer were suppressed by the inherent Case, this blocking effect would be quite surprising. The same blocking effect of agreement between the verb and the embedded subject in the presence of the experiencer DP is reported for Icelandic in Holmberg and Hróarsdóttir (2002).

(37) Það virðist/*virðast einhverjum manni [hestarnir vera seinir]
there seems/seem some man_{DAT} the-horses_{NOM} be slow
It seems to some man that the horses are slow.TM

A highly significant piece of data reported in Holmberg and Hróarsdóttir (2002) is that raising of the embedded subject to the matrix T(ense) is possible *even when agreement between the two is blocked* (emphasis mine), as in Cases where the experiencer DP undergoes wh-movement. The sentence in (38) shows that the trace of the wh-moved experiencer block agreement between the verb and the embedded subject.

(38) Hvaða manni veist þú a virðist/*virðast t_{wh} [hestarnir vera seinir]
which man_{DAT} know you that seems/seem the horses be slow
To which man do you know that the horses seem to be slow

42

the sentence in (39) shows that raising of the embedded subject is allowed in such contexts.

(39) Hverjum hafa hestarnir virst t_{wh} [t NP vera seinir]?who_{DAT} have the-horses_{NOM} seemed be slow

I find these and Boeckx's data two be highly significant in two respects. First, they show that inherent (English) or quirky (Icelandic) Case does not suppress ϕ -features. More importantly, they show that raising of a DP is allowed even when agreement between the attractor and that DP is blocked, which strongly suggest that that the two phenomena are independent. The analysis of English and Icelandic raising constructions consistent with these data is offered in Chapter 3 of this thesis.³⁴ For now, we conclude that the raising construction in (35) poses a serious problem for Chomsky's conception of movement and for the notion of equidistance it incorporates.

We have just seen that equidistance seems to be of no help in explaining the raising in (34a). There are several lines that one can pursue at this point. One line, taken in Boeckx (1999), is to redefine or reinterpret the notion of equidistance in a way that brings

³⁴ Since the focus of my study is displacement which, as we have just seen, is independent of agreement phenomena, I will not be discussing the agreement facts presented above, except to point out that an obvious way to approach the distinction in the framework suggested in this study is in terms of different features targeted by the two processes, namely uninterpretable Case and wh-features on the one hand, and interpretable ϕ -features on the other.

the experiencer DP and the subject DP into the same minimal domain. I do not pursue this line here, since it appears to further complicate the already complex computation of locality that this notion implies. Besides, Boeckx makes several assumptions that I do not share. For example, Boeckx assumes that the embedded subject is a part of the same minimal domain as the TP that it is the subject of. Also, the preposition dominating the experiencer is assumed not to count when it comes to determining whether the experiencer is in the minimal domain of the verb and, by the mentioned reinterpretation, equidistant with the embedded subject from the target, at the same time that it has to be assumed to be actively present in order to block, by another assumption, the feature relation between the experiencer and the matrix T.³⁵ Boeckx's account also relies to an extent on Chomsky's (1995) speculation that at LF formal features of objects also adjoin to T and that all X⁰s cluster together into a conglomerate head, assumptions which many, including Chomsky and Boeckx, have meanwhile abandoned (see, for example, Boeckx and Stjepanović 2001, where, in concert with Chomsky's latest writings, it is argued that head movement is, in fact, a PF process). For these reasons, and particularly since we have seen and will see other problems with equidistance, I conclude that an account without this notion is preferable. Nevertheless, Boeckx (1999) and (1998) offer valuable discussion and insights into the experiencer constructions which, even though they will not

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

³⁵ Perhaps because of these complexities, Boeckx concludes the relevant section by deciding to put off the precise characterization of equidistance for future research, suggesting at the same time that one might try to resort to Chomsky's notion of phase in defining equidistance, whereby two targets or two elements that are part of the same phase would be equidistant for movement. While this proposal avoids many complexities of the above solution it would appear to incorrectly rule in the superraising in (31) given that, under Chomsky's current conception of phase, on which passive sentences do not have a vP phase, the expletive *it* and the DP *John* would be phasemates and equidistant from the matrix T.

play a role in locality considerations under my analysis, might turn out to be helpful in explaining how the experiencer DP gets to c-command into the embedded clause. In particular, Boeckx's (1998) proposal that reanalysis obtains between the verb and the preposition might be relevant in this regard, if properly construed. I offer a speculation below.

Another approach to the problem with (34a) is taken in Kitahara (1997) where it is suggested that the experiencer DP does not c-command into the embedded clause at the point of the derivation when the experiencer undergoes raising. This is because the experiencer DP is embedded in the containing PP and, under the standard interpretation of c-command (e.g. Reinhart 1976, under which α c-commands β if the first branching node dominating α dominates β and neither of α and β dominates the other), does not qualify as a closer DP as construed in (1). However, we have seen in the discussion in footnote 4 that there is good evidence that the experiencer DP does, in fact c-command into the embedded clause. One of the relevant examples, namely the one that illustrates Condition C violation, is repeated here.

(40) *They seem to him_i to have insulted John_i.

To account for cases such as (40), Kitahara claims that the requisite c-command relation is established *after* the experiencer raises to SPEC-TP, via a covert adjunction of the formal features of *him* to the head of the prepositional phrase. The latter movement is

taken to be motivated by Case-checking. This would result in an LF structure like the one below.



Kitahara assumes that, since the preposition determines the label of the maximal projection, the complex head can count as the binder once the formal features of *him* have adjoined to the preposition. Thus the requisite binding configuration is established at that point, yielding the Condition C effect.

There are a couple of problems with Kitahara's account. First, the movement from a complement to a head-adjoined (or adjoined) position seems rather vacuous, given that the features of the complement are in a sense already present in the 'label' by the nature of Merge (see below). If, indeed, Case features of *him* need to be checked, it would be more natural to assume that they are checked *in situ* if the Case is inherent, or by adjoining to some higher functional head, if it is structural.³⁶

³⁶ It is another matter that structural Case checking in English results in a Spec-head configuration, Crucially, structural Case assigners are functional heads F, which, in contrast to prepositions, do

Furthermore, Kitahara's account might require extra mechanisms to capture the anaphor licensing in (34a), repeated here as (42).

(42) a. $[_{TP}$ [Each other's photos]_i seem to John and Mary $[_{TP}$ t_i to be scandalous]]

If the experiencer DP c-commands into the embedded clause only after the embedded subject has raised to SPEC-TP, then how is the anaphor in the subject DP licensed? One has to posit a subsequent lowering operation which would bring the anaphor back into the c-command domain of the experiencer. Whether such a lowering exists is a matter of some controversy. In fact, Lasnik (1999b) provides some evidence against A-movement reconstruction, as we have indicated in the previous chapter.

Finally, Kitahara seems to assume Chomsky's (1993 and previous work) model of grammar in which overt and covert movement belong to different components with their own independent cycles. Such a model has been largely abandoned since Chomsky (1995) in favor of one on which all movement operations take place in the same component under a single cyclic derivation. Under such a model, either the movement in question is just the movement of the set of formal features of the DP *him*, or the whole DP moves *overtly*. In either case, the movement must happen by cycle before the raising of *they*, which brings back into the foreground the 'locality violation' which Kitahara seeks to eliminate. In the former case, i.e. if the movement is just the covert movement of formal features, it runs afoul of what seems to be a rather robust generalization pointed out in

not take DP complements. The latter therefore must raise to SPEC-F for reasons we have discussed.

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

b. The DA proved [two men to have been at the scene]during each other's trials.

Under the assumption that formal features of the 'associate' raise covertly to the matrix T (which, presumably, accounts for 'long distance agreement' facts), the contrast between the grammatical sentences in (43b) and its ungrammatical counterparts in (43a) is unexplained if covert raising can create new binding configurations. On the other hand, if, as Lasnik suggests, covert feature movement does not affect interpretive relations such as those relating to the Binding theory, the contrast between (43a) and (43b) follows.³⁸ Thus, Kitahara's account faces serious problems on either the overt or the covert movement approach.

³⁷ See also den Dikken (1995).

³⁸ Note that, under the *Agree* conception of feature checking or, more precisely, under a version in which checking of ϕ -features does not involve feature movement, Lasnik's argument above becomes vacuous, as no movement of formal features takes place in (43a). This is also true under an alternative mentioned above under which English disallows feature separation (but see section 2.2.3 below). This leaves open the possibility that, in a language which allows feature separation, the 'covert' feature movement might affect binding relations. This has been claimed to be the case in Miyagawa (1993), for example. The reader is referred to this article for data and argumentation.

Before I turn to other structures in (34), I would like to offer some speculation on how the experiencer DP gets to c-command into the embedded clause. It is noteworthy, however, that on either of the speculations that I offer below, the problem for the locality of movement as construed in Chomsky (1995, 2000, 2001) remains. I offer the solution for this in the next chapter, after I spell out the main proposal of this study.

One possible approach to the fact that the experiencer DP c-commands into the embedded clause despite the presence of the apparently dominating prepositional phrase would be to agree with Chomsky that the case on the experiencer is inherent and then to generalize Chomsky's (1986a) approach to inherent Case under which a preposition such as *of* in (44) is just a morphophonological reflex of a genitive Case that the head noun 'assigns' to its complement.

(44) the destruction (of) the city

Recall that this assumption, along with the claim that inherent case is associated with θ -assignment, accounts for the ungrammaticality of sentences such as (12), repeated here as (45).³⁹

In any event, this does not affect our conclusions about Kitahara's analysis of English experiencer constructions.

³⁹ More precisely, Chomsky (1986b) takes *of*-insertion to be one of the alternative morphological realizations of inherent Case in NPs, the other being the 'POSSessive' marker 's. Both realizations are taken to be subject to a Case-marking mechanism associated with θ -role assignment (thus preventing expletives as subjects of nominals).

(45) *Mary's belief [TP (of) John to be innocent]

If we generalize this account to the experiencer construction, it would entail that the preposition *to*, being just a morphophonological reflex of the inherent dative case assigned to the DP by the verb *seem*, is not present in the structure in the syntax. This in turn would mean that the experiencer DP c-commands into the embedded clause.

A potential problem with this approach and, it seems, for Chomsky's proposal in general, would be to account for cases of preposition stranding, which suggest a fair degree of independence of the preposition from the DP it is associated with.⁴⁰

I leave the discussion in this somewhat speculative state. The bottom line of the above discussion is that there is good evidence that the experiencer DP c-commands into the embedded clause in overt syntax and that such a configuration is established prior to the raising of the embedded subject to the matrix T. A proper formulation of the Minimal Link Condition must allow for such raising to take place and at the same time disallow unwanted derivations. I offer such a formulation in Chapter 3. We now turn to the other crossing structures in (34) which show that the crossing pattern is not limited to the 'seem-to' construction.

⁴⁰ This is, in fact, more of a problem for the case that Chomsky based his proposal upon, namely the case of 'of-insertion'. As can be seen in (i), *of* can be freely stranded in such constructions, which is hard to reconcile with its being a morphophonological reflex.

 ⁽i) What would you like to read an analysis of t? In contrast, preposition stranding in the 'seem-to' construction is degraded, suggesting that the proposal might not be too far off track
 (ii) ??Who does it seem to t that John is nervous.

Furthermore, as pointed to me by Željko Bošković (p.c.), it is not obvious that preposition stranding necessarily implies its 'independence', particularly under the copy theory of movement.

2.2.2 Psych Construction

In this section, we show that the 'psych-construction' in (34b) poses a similar problem for an analysis which takes interpretable features to be subject to the operation *Agree* or *Attract*. The sentence is repeated here as (46).

(46) [These nasty rumors about each other]_i annoyed John and Mary t_i.

Belletti and Rizzi (1988) provide a battery of arguments in support of their claim that derivations of structures such as (34b) involve the raising of the theme DP *these nasty rumors about each other* across the experiencer DP *John and Mary*.⁴¹ The binding facts illustrated in (34b) present one of these arguments. Assuming with Belletti and Rizzi that Principle A of the Binding Theory can be satisfied at any point in the derivation, the possibility of anaphoric binding in (34b) follows if the experiencer DP c-commands the theme DP before the latter raises to SPEC-TP.⁴² This explains the contrast with simple

⁴¹ Pesetsky (1995) offers a critique of many of Belletti and Rizzi's arguments. However, he also ends up adopting an analysis in which one copy of the subject DP is base generated below the experiencer DP and another copy of the same DP is base generated higher than the experiencer DP. However, the only way of interpreting the claim that two identical copies of the same DP occupy two different positions in a structure under current assumptions and without significantly complicating the theory is to assume that the two copies are two links in a chain. Pesetsky's claim that the subject position is thematic, if correct, can be accommodated by allowing movement into θ -position, a possibility argued for in Bošković (1994) and Hornstein (1999), and in fact also assumed in Pesetsky (1995). For strong recent arguments in favor of the 'ergative' analysis of psych-constructions based on Dutch data, see Bennis (2002), whose main points relevant to the construction are summarized below.

⁴² We will see other arguments for this view in the next section, where it is shown that the psychconstruction follows the same pattern of connectivity effects as 'seem-to' and 'to-dative' ditransitives.

transitive structures which disallow such binding. The data and judgments in (47) are taken from McGinnis (2001).⁴³

(47) a. These rumors about himself worry John more than anything else.

- b. *These rumors about himself describe John better than anything else.
- c. Each other's supporters worried Freud and Jung.
- d. * Each other's supporters telephoned Freud and Jung.

Furthermore, there is evidence that Case on the experiencer DP is structural accusative rather than inherent. Recalling our assumptions and Lasnik's and Bošković's diagnostics from the previous chapter, if we can show that the experiencer DP has undergone raising, this will necessarily mean that it is structurally Case-marked. This conclusion is also true of Chomsky's (2000, 2001) analysis, given that raising necessarily involves the operation Agree which cannot obtain with inherently Case-marked DPs, as we have seen.

Recall that, under Lasnik's analysis, pseudogapping involves the deletion of VP. This means that the remnant object DP raises from the VP to a higher A-position before the VP is elided. With this in mind, consider the structure in (48)

(48) Those nasty rumors didn't annoy Mary, but they_i will $Tom_j [v_P - t_j - annoy t_i]$

⁴³ See section 2.2.4 for other evidence of the crossing pattern in psych-constructions.

If pseudogapping involves VP ellipsis, then the experiencer DP in (48) must have moved overtly to a higher A-position, we assume for Case-checking.^{44,45} Since inherently Case-marked DPs are unable to undergo raising, this means that the DP must be structurally Case-marked. As noted, the assumption on the inability of inherently Case-marked DP to raise is also crucial for Chomsky (2000, 2001), where it is attributed to the 'inactive' status of the host DP.

The raising of the experiencer DP removes it from the minimal domain of the lexical verb which also includes the theme DP. Thus, the subsequent movement of the theme DP *they* crosses over a DP which is closer to the target T and which should therefore defectively intervene with such movement. The derivation of the relevant part of (48) is represented in (49). It is clear that, before the raising of *They*, the DP *Tom* is neither in the same minimal domain with the target SPEC-TP (which belongs to the minimal domain (min)T={DP, vP}), nor with the DP *they*, which belongs to (min)V= {t_i,

⁴⁴ See Baltin (2000) for evidence that the landing site of the ellipsis remnant is an A-position.

⁴⁵ Lasnik (2000) argues that the movement is triggered by EPP, which is taken to be located in AgrO. One argument for this view would be that the movement in question can involve PPs, which don't need Case. However, many recent studies have shown that EPP, which is an ill-understood concept, is to a large extent eliminable from the grammar. Bošković (2001) in fact shows that eliminating the EPP has considerable empirical advantages (see discussion and references in Chapter 6). Furthermore, Bresnan (1991) and Conway (1996) have argued that PPs that have traditionally been claimed to satisfy the EPP are best analyzed as NPs. An analysis which assumes that PPs can raise to satisfy the EPP would need to account for why, for example, the PP in the experiencer construction in (i) cannot do so despite the fact that it is closer to the target T than the embedded subject, as we have seen.

⁽i) *To John seems t Peter to be innocent.

Howard Lasnik (p.c.) suggests that (i) can be ruled out as a failure of the matrix T to check its nominative Case. But this seems questionable since the nominative on the matrix T should be able to check against the nominative on *Peter*, by *Agree* or feature movement.

Perhaps the two conflicting views can be reconciled by analyzing the PP movement as targeting some A' or 'mixed' position such as Focus Phrase, while NP remnants would target a Case-checking position such as AspP, perhaps subsequently moving to the FocusP.

They}. Thus, the raising of *they* should be impossible and (48) (as well as (46)) should be ungrammatical, contrary to fact.⁴⁶



Evidence for the proposed analysis is also offered in Bennis's (2002) insightful analysis of psych-predicates in Dutch. Bennis shows convincingly that the relevant construction must be taken to involve an 'ergative' derivation in which the theme DP crosses the experiencer DP on its way to SPEC-TP, as in our analysis. What is also important is that Bennis provides arguments that in Dutch the crossed experiencer DP

⁴⁶ Note that we do not represent here a possible additional VP shell between the experiencer and the theme arguments. In any event, the only effect such a shell would have would be to make the movement of the theme DP even longer and therefore more problematic for the 'equidistance' accounts.

must be analyzed as having a structural 'objective' Case. Let us review some of Bennis's arguments for both claims.

Evidence that psych-constructions in Dutch involve the 'ergative' pattern comes from several constructions which show that the subject DP behaves as an underlying object. First, Bennis shows that, just like in passive constructions, the internal argument of the unergative construction (50a) shows up as the subject of the corresponding ergative construction (50b), while the external argument can be spelled out in an optional PP (*van* phrase) (50c.).

(50) a. Jan amuseert / ontroert / verbaast / intereseert/...mij met dat gedrag.

John amuses / moves / astonishes / interests /... me with that behavior. (Bennis

2002, ex.42)

- b. dat gedrag amuseert / ontroert / verbaast / intereseert /... mij. (ibid. ex (41)
 That behavior amuses / moves / astonishes / interests / ... me
- c. Dat verbaast / irriteert / ...mij van hem. (ibid., ex (46)
 that astonishes / irritates /...me of him.

Second, Bennis shows that the two arguments in the psych-construction can participate in structural inversion (52a,b), an option which, as he demonstrates, is strictly limited to cases in which the surface subject is an underlying direct object, as in passives and ergatives (51b,c). The sentences in (51a) show that the order IO-SUBJ is impossible in 'unergative' ditransitive constructions. In contrast, the pair in (51b,c) shows that the order IO-SUBJ is grammatical with passives and ergatives.

- (51) a. dat DIE JONGEN HEM /*HEM DIE JONGEN beloofde naar huis te gaan that that boy_{DO} him_{IO} / him_{IO} that boy_{DO} promised to go home (ibid.ex(5a))
 - b. dat DIE JONGEN HEM / HEM DIE JONGEN werd voorgestelf
 that that boy_{DO} him_{IO} / him_{IO} that boy_{DO} was introduced (ibid. ex(5b))
 c. dat DIE JONGEN HEM /HEM DIE JONGEN opviel
 - that that $boy_{DO} him_{IO} / him_{IO}$ that $boyD_{DO}$ struck (ibid. ex(5c))

The fact that the inversion is also possible in psych-construction, as illustrated in (52a,b), strongly suggests that this construction requires a 'derived-subject' approach.

- (52) a. dat die voorstelling mij amuseert / behaagt / irriteert / ...(ibid. (ex. 47a.)that that performance me amuses / pleases/ irritates / ...
 - b. dat mij die voorstelling amuseert / behaagt / irriteert / ... that me that performance amuses / pleases / irritates...

Finally, Bennis observes that, as argued in Stowell (1987), and Cinque (1989) for German and Italian, the gap in *as-clauses* in Dutch always corresponds to an underlying sentential object. This can be seen in the sentences in (53), where the gap corresponds to (53) a. Zoals ik zei e, houdt Jan van slakken. (ibid., ex (8a))

as I said, John likes snails.

- b. Zoals *e* door iedereen beweerd wordt, houdt Jan van slakken. (ibid., ex (8b))
 as by everybody is said, John likes snails.
- c. Zoals *e* mij opgevallen is, houdt Jan van slakken. (ibid., ex (8c). as me struck is, John likes snails.

The sentences in (54) illustrate that the gap cannot correspond to an underlying subject.

- (54) a. *Zoals *e* zijn onschuld bewijst, had Jan geen slakken gegeten. (ibid., ex (9a)) as proves his innocence, John had not eaten snails.
 - b. *Zoals e mij van zijn gelijk overtuigt, lust Jan geen slakken. (ibid., ex. (9b))
 as me of his right convinces, John doesn't like snails.

As can be seen in the sentences in (55), the psych-constructions pattern with the sentences in (53), strongly supporting the derived subject analysis.

(55) a. Zoals mij telkens weer verbaast, houdt Jan van slakken. (ibid., ex. (48a)) as me again-and-again surprises, loves John snails.

b. Zoals mij altijd irriteert, wast Jan zijn handen niet voor het eten. (ibid.,ex (48b)) as me always irritates, washes John his hands not before dinner

Thus, there is very strong evidence that the subject of the psych-constructions in Dutch is derived by raising from the object position of the verb to SPEC-TP. What is more, Bennis provides evidence that the Case on the experiencer DP is structural. His argument involves the pattern of occurrence of the experiencer argument in the present participle constructions which are based on psych predicates. Bennis shows that, although their 'external' distribution is restricted to environments where adjectives typically occur, internally these constructions can be either adjectival or verbal when headed by stative and psych predicates, in contrast to dynamic predicates with which the constructions are necessarily verbal. This is clear from the pair in (56), for example, where the distribution of stress in verb-particle combinations shows that the experiencer *predicate op-winden* 'excite' can be either verbal (56a), or adjectival (56b). This is because Dutch particle verbs have stress on the particle, whereas stress on the verbal stem in participles is characteristic of adjectives. The sentences in (56c) and (56d) show that the present participles of psych predicates can be modified by the intensifier *heel* 'very', or prefixed with the negative *on*, both of which target exclusively adjectives in Dutch

(56) a. de mij ópwindende gebeurtenis

the me exciting happening

- b. de (*mij) opwindende gebeurtenis.
- c. de (*mij) heel opwindende gebeurtenisthe (me) very exciting happening
- d. de (*mij) onopwindende gebeurtenis
 - the me unexciting happening. (ibid., ex. (40a,c,d,e)

The interesting fact about the examples in (56) is that they show that the experiencer argument can only occur with participles of psych-verbs if the latter are verbal, as in (56a). If the participle is adjectival, (as in (56b, c, d), the experiencer object must be omitted. Given that other adjectives in Dutch can freely assign inherent Case, as can be seen in (57), Bennis concludes that the restriction of the experiencer argument to verbal predicates follows naturally if the presence of the object experiencer argument is dependent on structural Case.

(57) a. Deze mensen zijn mij bekend.

These people are me known

b. Deze mensen zijn mij trouw.

These people are me loyal. (ibid., ex. (4a,b))

Bennis then proposes a derivation in which the experiencer argument, which is basegenerated higher than the theme object, raises to SPEC-vP to check the objective Case. Subsequently, the theme DP raises to SPEC-TP to check nominative. It is clear that Bennis's proposal is relevant respects identical to the one I proposed above and in Vukić (1999a, b). However, Bennis's suggestion on how the proposed derivation relates to locality constraints in the system differs sharply from mine.

Bennis claims that the verb raising to v, and then to T, makes the specifier positions of Spec-TP and SPEC-vP, as well as SPEC-VP from which the experiencer argument has raised, equidistant for the raising of the theme DP. That is, Bennis not only crucially relies on Chomsky's (1993) proposal that verb movement extends the locality domain, but in fact seems to be reinterpreting it in allowing for such movement to have a cumulative extension effect, which Chomsky explicitly argued against in his original proposal. In effect, all the specifier positions in the sentence become equidistant from the deeply embedded theme DP, which can therefore freely raise to SPEC-TP.

It is clear that such a conception of equidistance cannot be maintained. First, given that most Germanic languages exhibit a verb-second phenomenon, which is standardly taken to locate the verb in C, Bennis's proposal would predict that there are generally no superiority effects in such languages, a controversial result.⁴⁷

Second, given that Case on both internal arguments is structural, Bennis's proposal would allow a derivation in which the theme argument can raise across the experiencer DP to get/check accusative, whereupon the experiencer DP can raise to SPEC-TP to get/check nominative, thus erroneously predicting that the surface string *Peter annoyed rumors about his resignation* can be interpreted as rumors annoying Peter.

⁴⁷ See, for example, Wiltschko (1997), Grewendorf (2001), and Featherston (2003), for evidence that superiority does apply in German.

Third, given that English does not have verb movement to anywhere higher than Asp, or AgrO, Bennis's analysis cannot be generalized to cover the identical situation in English, as we have seen.

Finally, domain extension due to verb movement even in its more restrictive original formulation has been largely abandoned, in part because of the growing body of research suggesting that verb movement may be a PF phenomenon.

In view of these problems, we conclude that Bennis's equidistance-based account cannot be maintained, in contrast to his strong arguments about the derived nature of the subject DP and structural Case on the experiencer DP. We will offer in the next chapter a simple analysis which maintains Bennis's arguments and at the same time allows the relevant movement pattern in a principled manner.

We conclude this section by observing that the psych-construction poses the same problem for the conception of movement which assumes that interpretable features are targeted by *Agree/Attract*, and for the mechanism of equidistance which the associated theory of locality incorporates. We now turn to even more drastic 'violations' of locality that occur in ergative languages.

2.2.3 Crossing in Ergative Languages

Our next example comes from deep ergative languages such as Dyirbal or Inuit.. As is well known, ergative/absolutive languages differ from nominative/accusative languages in that the patient argument of a transitive verb appears in the same (absolutive) Case and/or triggers the same kind of agreement as the sole argument of an

61

intransitive verb, while agent arguments of transitive verbs have a distinct ergative Case and agreement. Agreement with the absolutive is generally marked further on the verb stem than agreement with the ergative, which is another aspect in which the former patterns like nominative in most Indo-European languages, as observed in Baker (1997).

(58) a. Payi yara-Ø paninyu. DYIRBAL

there man(abs) come

'The man is coming.'

- b. Palan jukumpil-Ø pangkul yara-ngku palkan.there woman(abs) there(erg) man-erg hit'The man is hitting the woman.'
- (59) a. (Uanga) qungujup-p-u-nga. GREENLANDIC

me(abs) smile-ind-intrans-1s

'I smiled.'

b. Anguti-p (uanga) urnip-p-a-a-nga.

man-erg me(abs) approach-ind-trans-3s-1s

'The man approached me.'

Numerous other processes have been found to isolate the absolutive arguments in a language like Dyirbal. For example, only absolutive arguments are licensed in embedded purposive control structures, only absolutive arguments can be relativized or coordinated in 'clause-chaining' (Dixon 1972, Murasugi 1992, Baker 1997), and have default wide scope relative to any sentential operator (Bittner 1988). This, coupled with the fact that the unmarked word order in Dyirbal is Patient - Agent - verb (Dixon 1972) has lead many researchers to assume that the absolutive DPs, namely the theme arguments of transitive and sole arguments of intransitives, are licensed in the same syntactic environment, the Spec-TP/IP (Murasugi 1992, Bittner and Hale 1996a, 1996b, Baker 1997). Furthermore, evidence from binding and morphological processes such as incorporation and compounding suggests that the theme DP is base generated lower in the structure than the ergative agent DP. In view of such facts, Baker (1997) and Bittner and Hale (1996a,b) have argued that in this language transitive structures such as the one in (60) involve a derivation in which the agent DP "remains in the specifier of the VP and receives ergative Case (Baker 1997, p.83)" while the patient DP raises across it to the specifier of TP to receive nominative Case and trigger agreement. Note that these assumptions fit naturally into our system, as well as Chomsky's, given that A-movement is crucially dependent on Case (or, in Chomsky's analysis, agreement) properties.

(60) Dyirbal

Payi parrkan pangkul yara-ngku jurrka-nyu.CL(NOM) wallaby(NOM) CL(EGR) man-ERG spear-NFUT'The man is spearing the wallaby.' (Bittner and Hale 1996a, p.15)

The structures for basic transitive clauses in English and Dyirbal/Inuit that are consistent with Chomsky's current assumptions on phrase structure are given in (61) (based on Baker (1997), but with the added vP shell in which Chomsky (as well as the current study) assumes the agentive DP is licensed.).



It is important to observe that many authors (e.g. Marantz 1984, Levin and Massam 1985, Murasugi 1992, Bobaljik 1993, Bittner and Hale 1996a,b), assume that the structural distinction between the nominative and accusative Case in accusative languages corresponds to a structural distinction in ergative languages, although they differ in what the exact correspondence is. Thus, Levin and Massam (1983) and Bobaljik (1993) identify ergative case with the nominative and absolutive with the accusative, while Marantz (1984), Murasugi (1992), Bittner and Hale (1996a, b) identify absolutive with the nominative. As observed in Stepanov (2002), the latter is the majority view which better captures the cross linguistic symmetries between nominative and absolutive Cases, such as
the above mentioned unmarked value and appearance in single argument sentences. The position that ergative Case is also structural is certainly the null assumption given that it parallels the accusative case in nominative/accusative languages in being the standard 'other' case in transitive structures, in being predictable from syntactic configuration (see Bittner and Hale 1996a), in alternating with other Cases, etc. In other words, the arguments for treating ergative Case as structural are essentially parallel to those for treating accusative Case as structural in nominative/accusative languages. Thus, Bittner and Hale (1996a) treat ergative and Accusative as 'direct Cases' subject to universal Case-licensing conventions, in contrast to 'oblique' Cases (which may also show some characteristics of structural Cases, see below), whose realization is subject to language-specific requirements.

Given that the derivation sketched in (61b) seems well justified in its essential content, a question arises as to how the derivation of (61b) relates to the proposed locality constraints in the system. Obviously, under the assumption that the movement of the theme DP to the Spec. of TP is triggered by the attraction of its ϕ -features, such a derivation presents the same problem for Chomsky's analysis as the sentences discussed above. That is, neither the two DPs (the agent DP and the patient DP) nor the two specifier positions, i.e. SPEC-T and SPEC-v are in the same minimal domain of a predicate, hence equidistance is of no avail. Therefore, the ϕ -features of the agent DP should defectively intervene and prevent the attraction of the ϕ -features of the lower theme DP, and the sentence in (60) should be ungrammatical, contrary to fact.⁴⁸

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

⁴⁸ In fact, the problem is compounded in ditransitive structures, where the raising of the direct object to Spec-TP crosses two DPs in A-positions which Bittner and Hale claim to be structurally

It is worth mentioning that Chomsky (1995, Chapter 3) treats absolutive Case as licensed in SPEC-AgrOP, i.e. in the same configuration where Chomsky at that time assumed the patient DPs to be 'covertly' Case-marked with accusative at LF. The difference in overt movement pattern between English and ergative transitive constructions is then taken to be the consequence of AgrSP being 'active' in English and AgrOP being active in ergative languages. This is essentially the line taken in Bobaljik (1993) and Levin and Massam (1993), as noted. However, in addition to the above conceptual problems with identifying absolutive with accusative, this proposal does not relate comfortably to Chomsky's current view of phrase structure in which Agr is taken

Case-marked. Thus, Bittner and Hale give the derivation of the 'applicative' triadic construction in (ia) as roughly represented in (ib). (The authors assume that the surface order in the case of (ia) results from a subsequent PF fronting of the ergative subject.)

- (i)a. Juuna-p atuakka-t Kaali-mut nassi-up-p-a-I Juuna-_{ERG} book-_{NOM} Kaali-_{DAT} send-_{APPL-IND}-[+tr]-_{3SG-3PL} 'Juna sent the books to Kaali'
 - b. [TP atuakka-t_i [VP Juuna-p [VP Kaali-mut [VP t_i t_v-appl-] v.nassi-up-p-a-i]]]]

Bittner and Hale show that not only is the ergative Case structural, which we have said is rather generally assumed, but also that the dative Case in Inuit should be taken to be structural since a) it is completely predictable from the syntactic configuration and b) it participates in alternations with other structural cases, as is clear from the comparison of the 'applicative' triadic (ia) with the 'standard' triadic construction in (ii) which shows that the dative alternates with the nominative This means that the patient DP in (ib) has raised across two structurally marked DPs that should defectively intervene with such movement.

(ii) Juuna-p Kaali atuakka-nik nassip-p-a-a J.-_{ERG i} K._j book-_{PL.INSk} send-IND-[+tr]-3SG i .3SG j 'Juuna sent Kaali the books.'

I will not discuss these case further, since the oblique dative and instrumental are language-specific realizations and are less uncontroversially structural. If they are indeed inherent, the pattern illustrated in (i) would fit perfectly into the analysis of ditransitive structures we propose in the next section. If not, the presence of additional 'V-shells' in double-object and other ditransitive constructions (see the next section) could again make the data more easily analyzable under our assumptions than under alternatives.

not to exist, and with the cyclic or phase-based conception of derivation on which uninterpretable features are eliminated as soon as they are introduced in the structure or at least moved to the phase periphery to save the derivation from crashing. As we have indicated above, Chomsky currently assumes that the head licensing overt object movement is 'active' in the just mentioned sense in such languages as Icelandic, therefore 'activeness' of vP cannot possibly be the distinction between accusative and ergative languages.

A somewhat similar proposal is put forward in Murasugi (1992). Murasugi takes the nominative and absolutive case to be checked by T, and the accusative and ergative to be checked by v (or Tr(ansitive) head), in accordance with the above, we believe strong, arguments for treating these Cases in a parallel manner. Murasugi then proposes that differences in overt movement between accusative and ergative languages follows from different Case features being 'strong' on the respective heads in the two classes of languages. She proposes the Ergative Parameter in (62).

(62) Ergative Parameter (Murasugi (1992, p 24)

In an accusative language, the Case features of T are strong.

In an ergative language, the Case features of Tr are strong.

The different settings of the parameter in accusative and ergative languages yield the crossing and nested movement pattern, respectively, with the movement satisfying the strong feature taking place overtly, and the other movement satisfying weak Case feature

being forced to wait until LF, by Procrastinate. She claims that language design is such that the two movements cannot take place overtly in a language (p.26-27, but see below). The resulting patterns in the two types of languages are shown in (63), where the solid line represents overt movement and the dashed line covert movement.



Murasugi's approach shares some weaknesses with the previous proposals, though her study has a major advantage in capturing the inter-linguistic and cross-linguistic symmetries and generalizations on marked (accusative and ergative) and unmarked (nominative and absolutive) Cases, as we have noted above. One weakness of her account is that she assumes a model of the grammar now generally abandoned, on which there are

68

two distinct components, overt and covert, with independent cycles.⁴⁹ Also, we have seen that strong arguments can be offered for overt object shift and subject movement in ECM constructions in English. In addition, since she claims that apparent counterexamples to the ban against two-movements in the same language (e.g. subject movement and overt object shift in Icelandic and, on the other hand, overt agent and patient DP movement to in Dyirbal), follow from EPP, i.e. independently of Case, it is not clear why the different movement pattern in the two types of languages should obtain in such cases. That is, one would expect that Icelandic should behave like Dyirbal, since the strong EPP on Tr in this language would presumably attract the closer agent DP. The only principled way out of this contradiction would be to introduce another cycle in the grammar, clearly an undesirable result.

We have seen that the pattern of movement in ergative languages poses a serious problem for locality effects in a theory which assumes interpretable features to be subject to Agree/Attract. This is because such features, since they are semantically charged and therefore undeletable, should prevent by defective intervention the movement of a more remote DP. We have also seen that, as in the 'seem-to' construction, the mechanism of equidistance is not sufficient to allow for the crossing pattern. In the next section, we show another case where equidistance does not allow a derivation of a grammatical sentence. We will also see a case which provides a different kind of argument against this mechanism, this time showing that it is too permissive in allowing unwanted derivations.

⁴⁹ Koizumi points to the strong evidence that movement to Spec-TP occurs overtly in the ergative language Dyirbal, in apparent violation of the above parameter. She claims that such movement is driven independently of Case, by EPP. We discuss this below.

2.2.4 To-dative Ditransitive Construction

I argued in Vukić (1998, 1999a) that the to-dative ditransitive construction (henceforth TDC) in (34b) also involves a derivation in which the direct object raises across the c-commanding indirect object. This accounts for numerous asymmetries that exist between this construction and the Double Object Construction (DOC) with respect to anaphor licensing, WCO effects, quantifier scope ambiguities, and similar effects that are standardly associated with c-command. I argued that the contrasts to be discussed follow if, in contrast to TDC, the derivation of DOC does not involve a crossing pattern of A-movement. The structures I propose for the derivations of TDC and DOC that accord with our general schema in (30) and with our assumptions on accusative Case 'assignment' by Asp are represented in (64a) and (64b), respectively. ⁵⁰



⁵⁰ The structures is (64) are fully consistent with the thematic hierarchy adopted in the previous section, on the standard assumption that such hierarchies refer to base-generated positions of arguments.

As can be seen from the structural representations in (64), I claim that in TDC the theme DP raises across the goal DP to check accusative Case. In DOC, the accusative Case is assigned to the goal which therefore moves to SPEC-AspP, whereas the theme bears an inherent Case and stays *in situ*.

One argument for the proposed structure for TDC comes from the 'backward binding' facts illustrated in (34c) in which the anaphor *each other's photos* is licensed by the antecedent DP *John and Mary*. The same backward binding is impossible in the double object construction.

(65) *I showed [each other's_i friend] John and Mary_i

Under the previously discussed conditions on anaphor licensing, the contrast between (34c) and (65) follows if in the former, but not in the latter, the DP containing the anaphor crosses over a c-commanding DP which provides the antecedent.

The same pattern obtains with respect to WCO effects and quantifier scope interaction. As noted by Kitagawa (1994), WCO effects in TDC (66a) are much weaker than in the corresponding DOC (66b) which, as shown in (64) and contra Larson (1988), does not exhibit the crossing pattern. The pair in (66) is based on Barss and Lasnik (1986) where the asymmetrical pattern of c-command in the DOC was first extensively documented.

(66) a. ?I gave his, paycheck to [every worker]_i.

b. *I gave its_i owner [every paycheck]_i.

Assuming that, as in the case of anaphors, the conditions on bound variable readings of pronouns can be satisfied anywhere in the derivation,⁵¹ the above asymmetry is explained if in the TDC the theme DP starts out lower than the goal DP and then raises across it at some point in the derivation, as it does in (64a), but not (64b).

Since we assume the same crossing pattern in the 'seem-to' and psychconstruction, we predict that these constructions will also exhibit 'reconstruction' effects.⁵² This is well known to be the case:

(67) a. His_i mother seems to every child_i to be the best woman in the world.

b. His_i poor performance on the exam annoyed every student_i.

Other tests also support the 'crossing' analysis of TDC. Quantifier scope interactions show the same contrast between TDC and DOC:

⁵¹ See Lebeaux (1991, 1998), among others, for evidence for this assumption.

⁵² Here and throughout this study, I use the term 'reconstruction' to refer generally to a situation in which a particular interpretive effect is plausibly accomplished in a position lower than the position which the relevant XP occupies 'at surface structure.' I take this effect to be accomplished derivationally, as the XP in question is displaced from one position in the structure to another. In other words, I do not commit myself to an analysis which literally lowers the XP back to a position in which the interpretive effect obtains.

- (68) a. I gave a check to every worker. $(\exists > \forall, \forall > \exists)$
 - b. I gave a worker every paycheck. $(\exists > \forall, *\forall > \exists)$

Aoun and Li (1989) propose that, for two QNPs, α and β , α has scope over β if α ccommands β or its trace. On our analysis the universally quantified DP c-commands the trace of the existential DP in (68a) but not in (68b), and the contrast between the two sentences is predicted. Again, the same scopal interactions obtain in the psych and seem-to constructions.

Aoun and Li's formulation also accounts for the cases of wh/quantifier interactions such as (69a,b), also discussed in May (1985) and others:

(69) a. What did every guest buy for the party?

(for which pairs x y, x a guest and y a thing, x bought y for the party)

b. Who bought every dish for the party?

(#for which pairs x y, x a person a y a dish, x bought y for the party)

Example (69a), but not (69b) allows for a pair-list reading of the question, i.e. the reading in which the set of dishes is paired with the set of guests. (69b) allows only for a 'singular' reading of the wh-phrase *who*, on which there is a unique person that bought all the dishes for the party. Many authors, including those mentioned above (see also Chierchia 1993) have related this to the fact that only in (69a) does the wh-phrase cross over the quantified DP on its way to SPEC-CP.⁵³ The pattern is also expected under Aoun and Li's formulation since only in the former sentence does the universal quantifier c-command the trace of the extracted wh-phrase. Thus, the possibility of the pair-list (or 'functional') can be taken as a test for whether a crossing pattern obtains between a wh-phrase and a quantified noun phrase in a sentence. We therefore predict that the pair-list reading will be available in TDC, and not DOC, since we have argued that only in TDC does the crossing pattern obtain. The data are given in (70) and, as indicated, are based on judgments from David Braze.⁵⁴

- (70) English (David Braze, p.c.)
 - a. Which book did you show to every woman.

(?for which pairs x, y, x a woman and y a book, you showed y to x)

b. Which woman did you show every book?

(*for which pairs x, y, x a woman and y a book, you showed x y)

There is a clear contrast between (70a) and (70b) in that only the former allows for a pairlist reading of the question. If in the TDC, in contrast to the DOC, the universal quantifier c-commands the trace of the wh-phrase, as it does in the structure we have proposed, the

⁵³ Chierchia (1993) takes pair-list readings to be a special case of 'functional readings' of whphrases. These readings, on his analysis, require an anaphoric element left by the extraction of the wh-phrase to be bound by a c-commanding antecedent, which further supports our crossing analysis.

⁵⁴ Howard Lasnik (p.c.) informs me that his judgments on (67), (68), (70), (69) are different, so these data are not uncontroversial. It is also well known that many native speakers of English reject

75

contrast between (70a) and (70b) is accounted for. The questions in (71a) and (71b) show that the pair list readings are also available in 'seem to' and psych-constructions, as expected.

- (71) a. Which man seems to every woman to be the most intelligent?(for which pairs x, y, x a man and y a woman, x seems to y to be the most intelligent)
 - b. What kind of criticism annoyed every woman?

(for which pairs x, y, x a kind of criticism y a woman, x annoyed y)

Thus, there is overwhelming evidence for the 'crossing' derivation of the TDC given in (64) as well as for the other structures in (34). But if that is the case, then TDC presents the same problem for Chomsky's analysis as the structures in (34a,b), since the crossed DP is not in the same domain with either the target or the theme DP, and the structure is wrongly predicted to be ungrammatical. As is clear from the structural representation in (72), the target τ is in the min(Asp), which also includes the top VP node. The theme DP is in the domain of the innermost V which includes no other term. Hence, the goal DP, which is in the domain of the preposition P or the higher V (if the preposition does not count given that it does not prevent c-command) is closer to the target and the movement of the theme DP is wrongly disallowed.

wh-extraction of indirect object in the DOC. Obviously, for those speakers the contrast reported in (70) does not apply.



Note that this is the one structure we have discussed where the presence of an additional VP shell might have an effect on the computation of equidistance (provided that the preposition is also ignored, as observed above). In other words, if the VP shell between the two internal arguments were not present and if the preposition did not count, the goal DP would be in the same minimal domain as the theme DP, namely the minimal domain of V. While this account would 'work' for the case at hand, I will show in the next section that dispensing with multiple VP shells in ditransitive structures runs into problems with cross-linguistic data on DOC. That is, on such an analysis equidistance will turn out to be too permissive in that it would allow unwanted derivations. Given that we have already presented evidence that it is too restrictive in other cases, this would further compromise its status in the theory and would constitute an additional problem for the analyses which rely on it. The evidence will come from the interaction of passivization and focus data in Icelandic double object constructions. Before we discuss Icelandic data, though, I wish to point out that there appears to be good evidence for the presence of the

76

additional VP shell in these constructions, as well as for the crossing analysis that we have proposed. The evidence comes from pseudogapping TDC constructions with the PP remnant, as discussed in Vukić (1999).

Let us again follow Lasnik (1995) and assume that pseudogapping is a PF process that elides a VP under identity with its correspondent in the antecedent clause. Then, consider the sentence in $(73)^{55}$.

(73) Although John wouldn't give money to Bill, he would to Susan.

The problem raised by (73) is that, assuming a Larson-type structure for the TDC, it seems difficult to maintain the standard assumption that ellipsis affects syntactic constituents. To avoid this problem, one might follow in the spirit of Jayaseelan (1990) and assume that the PP object *to Susan* in the second clause in (73) undergoes extraposition before deletion takes place. Alternatively, one may assume that the object PP raises across the theme object to a higher verbal projection before deletion. However, Lasnik (1999b) points to some problems with the former, and Vukić (1999a) with the latter approach.

⁵⁵ Lasnik (1999) takes PP remnants in TDC-based pseudogapping structures to be degraded. However, Baltin (2000) shows that they are quite acceptable if one controls for the 'like-subject restriction on pseudogapping' (p.15), as the sentence in (0) illustrates. The judgments in Baltin (2000) are consistent with those that were given in Vukić (1999a), and in fact turn out to support Lasnik's crucial assumptions about pseudogapping, as will be clear below.

An interesting fact observed in Vukić (1999a)⁵⁶ about pseudogapping in the TDC is that bound variable readings of pronouns, as well as NPI and anaphor licensing, which are otherwise perfectly possible in non-ellipsis structures, are unavailable in pseudogapping construction. The data is given in (74).

(74) English (Kazuko Hiramatsu, p.c.)

a. Jane will send each boy to his mother and Susan will send each boy to his father.

b. *Jane will send each boy to his mother and Susan will to his father.

c. Jane will send each son to his caretaker and Susan will each grandson.

d. Jane will give each soldier his paycheck and Susan will each officer.

(75) a. Jane will give no money to any student and Peter will give no money to any man.b. *Jane will give no money to any woman and Peter will to any man.c. Jane would show no gun to any small child and Susan will no rifle.d. Jane will give no man any money, and Peter will no man.

(76) a. Peter will show Jane to himself, and Bon will show Jane to herself.

b. *Peter will show Jane to himself, and Bob will to herself.

Under the extraposition or PP-raising analyses which do not assume a crossing pattern in the TDC, it is not entirely clear why the sentences in (b) are deviant, given that the necessary licensing configuration is established prior to deletion. The well-formedness of the examples (a) shows that there is nothing semantically odd with these sentences while

⁵⁶ The same kind of data have been subsequently reported in Baltin (2001).

the (c) and (d) examples in (74) and (75) show that NPIs and pronouns as bound variables are possible in ellipsis sites, provided the structural condition of c-command by the appropriate antecedent is satisfied. To account for the degraded status of the sentences in (b), a proponent of the extraposition-type analysis would have to assume either that NPI, pronoun-as-bound-variable, and anaphor licensors must not occur in ellipsis sites, or that extraposing a structure containing an NPI item or a pronoun as bound variable/anaphor is disallowed, which is clearly shown to be wrong by the data in (77). Thus, I reported in Vukić (1999) that my informants fully accepted (77a-b), which contain elided licensors of NPIs and anaphors. Examples in (77c-d) show that structures in which these elements are extraposed are perfectly grammatical.

- (77) a. John denied any allegation that he was a thief, and Mary any claim that she was his accessory.
 - b. Rumors about each other annoy Tom and Bill, but jokes about each other do not.
 - c. The professors returned no report yesterday to any of the students who had taken the exam.
 - d. The DA introduced Peter and Mary yesterday to each other's legal representatives.

I then proposed that these facts would follow naturally on my analysis if deletion takes place before the verb and the theme object raise across the PP. I than showed that such an analysis is fully consistent with Lasnik's observation that the raising of formal features does not affect properties relevant to binding, NPI-licensing, etc., and with his analysis under which a strong feature of a target can be checked by just raising the formal features of an item which is subsequently deleted at PF.

If we now assume with Lasnik that pseudogapping elides VPs, it is clear that such an assumption requires the structure in (64a), and not the one in which the goal and the theme DPs are contained in the same VP. The relevant part of (73) is given in (78).



We have seen that we have a good reason to assume a VP shell between the goal DP and the theme DP which means that the theme DP in (34c) has raised across a goal DP which is closer to the target, in violation of equidistance.

We now turn to the double object constructions in Albanian and Icelandic. As will be shown, the pattern of c-command relations in these and the associated passive constructions indicates further instances of A-movement of one DP across another DP in an A-position, and present further problems for the concept of equidistance and for analyses which assume that movement targets closest interpretable features.

The Albanian example is based on Snyder (1992). In this language, the standard tests show that in the DOC the goal DP is higher in the structure than the theme DP. This is shown by the sentences in (79).

- (79) a. Agimi ia tregoi secilit djale_i baben e tij_{i/j}.
 Agim clcl show each boy-_{DAT} father his-_{ACC}
 'Agim showed each boy his father'
 b. Agimi ia tregoi babait te tij_{i/*j} secilin djale_j.
 - Agim clcl show father- $_{DAT}$ his each boy- $_{ACC}$ 'Agim showed his father each boy'

Of the two sentences in (79), only (79a) allows for the bound variable reading of the pronoun in "his father". This fact obtains even if the order of the two relevant DPs is switched, as in (80) which is ungrammatical on the bound variable reading of the pronoun.

(80) Agimi ia tregoi secilin djale_i babait te tij_{*i/j}.

Agim clcl show each boy-ACC father his-DAT

"Agim showed his father each boy'

These facts indicate that there is no derivation in which the theme DP occupies a higher Aposition than the goal DP in the active DOC in Albanian. This in turn means that in the passive construction in (81) the theme DP must have raised to the subject position across the c-commanding goal DP.

(81) Secili djale iu tregua babes te tij
each boy-_{NOM} cl show father his-_{DAT}
'Each boy was shown to his father'

On the basis of the above paradigm, Snyder (1992) concludes that "Relativized Minimality, as formulated in Rizzi (1990), does not apply to A-movement (p.1)". We conclude that the data also pose a problem for a conception of *Move/Attract* which is based on the attraction of interpretable (D- or ϕ -) features.

To account for the Albanian data under the assumption that *Attract* targets interpretable features, one again needs to resort to the stipulation that the relevant features which should but do not defectively intervene are suppressed by inherent Case. This assumption is not only conceptually problematic (see fn.5) but appears to run afoul of empirical facts, as we have seen on the example of 'seem-to' constructions. Potentially problematic for this assumption is also the fact that other types of movement can freely target inherently Case-marked expressions. For example, inherently Case-marked DPs can freely undergo wh-movement or scrambling in SC, so the relevant attracted feature is not suppressed. (82) a. Kome si poslao pismo.

Whom-_{DAT} have-_{2sg} sent letter

'To whom did you send a letter'

b. Nekome je Petar poslao pismo

Someone-DAT has Peter-NOM sent a letter

'Peter sent someone a letter'

Furthermore, it is clear that even in English inherently Case marked DPs undergo whmovement. Recall that Chomsky (1986a) assumes that prepositions generally assign 'oblique' Case which he takes to be inherent. Recall also that he takes the genitive Case on complements of nouns to be inherent. Finally, we have seen that it is plausible to assume that the theme DP in the DOC is inherently Case marked, since it doesn't appear to raise and does not passivize. The sentences in (83) show that inherently Case-marked DPs can clearly undergo wh-movement.

(83) a. What_i did you talk about t_i

b. What_i would you like to read an analysis of t_i

c. What did you give John t?

We now turn to Icelandic data and further discuss the notion of equidistance and its validity in the grammar. We have already seen that this mechanism is not sufficient to

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

explain the grammaticality of (34a-c), and that it is not needed to explain the grammaticality of simple transitive constructions in English and elsewhere (Koizumi's data). Skepticism about its validity is further strengthened by data from the DOC in Icelandic. As noted by Holmberg and Platzack (1995), Collins and Thráinsson 1996), the dative (indirect) object in Icelandic can be shown to c-command the accusative DP (direct object) in the DOC construction (84a,b), but not vice versa. What is particularly interesting is that in this language either object may passivize (85a,b).

(84) a. Ég hafðI gefið konunginuum_i ambáttina sina_i

I-NOM had given the king-DAT the maidservant-ACC his-RFL

'I had given the king his maidservant'

b. *Ég hafðI gefið konungi sinum ambáttina.

I-NOM had given king-DAT her-RFL the maidservant-ACC

'I had given her king the maidservant' (Collins and Thráinsson 1996:398)

(85) a. Jóni var gefin bókin.

Jon-DAT was given the book-NOM

b. Bókin var gefin Jóni.

The book-NOM was given John_{DAT} (Holmberg and Platzack 1995:189)

A question that seems to be more problematic for the current analysis is how the passive structure in (85b) is derived. That is, since the dative DP c-commands the

accusative DP in the DOC, the former should block the attraction of the latter in the corresponding passive structure, incorrectly predicting (85b) to be ungrammatical.

The solution to this problem is provided by the observation of Holmberg and Platzack that, in a restricted set of circumstances, the accusative DP can c-command the dative DP in the Icelandic DOC. The restriction is that, if the accusative DP c-commands the dative DP, then the latter must be focused (indefinite, stressed, or complex):

(86) a. Ég ætla að gefa bókina einhverju bókasafni.

I will give the book(A) some library(D).

b. *Ég ætla að gefa bókina bókasafninu.

I will give the book(A) the library(D). (H&P, p.206)

In view of the focus effects, I assume that the configuration in (86a) results from the scrambling of the theme DP across the goal DP.⁵⁷ Since it will be shown that this movement feeds movement to SPEC-TP in the related passive constructions, I assume that it must take place in overt syntax as well, which in turn means that it must be triggered by some feature that is optionally assigned to the head in whose specifier position the moved phrase lands.

The pair of sentences in (87) shows that when the theme DP is passivized, as in (85b), than the dative goal DP has to be focused.

⁵⁷ See Reinhart (1996), among others, for the association between scrambling and focus.

(87) a. Hún var gefin einhverjum börnum.

it(N) was given some children(D)

*b. Hún var gefin þeim.

it(N) was given them(D). (ibid. p. 217)

As will be clear from the proposal which we put forward in the next chapter, the correlation between the focus restriction and the passive sentence in (85b) is expected under an approach which eliminates equidistance and thus requires the theme DP to c-command the dative goal DP in order to be attracted. In approaches which incorporate equidistance and defective intervention this correlation does not follow without extra assumptions. That is, it is not clear why the theme DP could not be attracted from its base generated position, which by assumption is in the same minimal domain as the dative DP. Since the focus restriction does not obtain in 'uninverted structures', the passivization of the theme DP which by equidistance could take place from such structures would fail to predict the focus restriction since it would not depend on the option of selecting the feature which induces the focus effect.⁵⁸

Thus, the mechanism of equidistance seems to be overrestrictive in some cases (those in (34), ergative constructions, etc.), unnecessary in others (transitive and ECM constructions), and too permissive in others, which strongly argues for its elimination from the system.

⁵⁸ Note that this account does not imply any look-ahead property of the grammar, or at least no more than is the case with any other optional feature. The account simply relies on the option of the grammar freely applying its resources, with the convergence or crash, depending on choices taken.

We have seen that there are a host of constructions in which one DP undergoes Amovement across another DP in an A-position. We have also seen that an analysis which assumes that Attract/Agree targets interpretable features makes incorrect predictions by disallowing such structures. Finally, we have seen that this fact remains even if the complex mechanism of equidistance is employed to extend the reach of the relevant movements. In the next section, we turn to constructions that an analysis based on the attraction of interpretable features wrongly predicts to be grammatical.

2.3 'Freezing' Environments

It is a well known fact that A-movement ends in a Case-marking position, and that wh-phrases 'freeze' in place in operator positions. The former restriction was captured in Chomsky (1986b) via a condition on argument chains given in (88)

(88) if $C = (a_1, ..., a_n)$ is a maximal CHAIN, then a_n occupies its unique θ -position and a_1 its unique Case-marked position. (p. 137, ex. (171))

(88) rules out structures such as (89) in which the DP Jane has moved from a position in which it receives/checks nominative Case, to another A-position.

(89) *Jane seems t was believed.

(89) is successfully ruled out in the Chomsky (2000, 2001) system since the movement of *Jane* to the matrix SPEC-TP would affect an 'inactive' phrase, which is not allowed. However, there are several constructions where the ungrammatical movement is still allowed in Chomsky's system. Consider the sentences in (90)

- (90) a. *the belief [$_{TP1}$ John_i to seem [$_{TP2}$ t_i is sick]]
 - b. Who dispelled the belief [CP that which man seems to be sick]]
 - c. $*[_{TP}$ Who dispelled the belief [which man_i to seem t_i is sick]]
 - d. *I wonder who_i t_i seems to t_i was told John that ...
 - e. *I wonder who_i seems t_i was told t_i John to be innocent.

Recall that, to get the defective intervention effect in the superraising construction in (31), Chomsky needs to couple the assumption that *Attract* targets interpretable features with the requirement that a DP whose ϕ -features are targeted contain an uninterpretable feature which makes it 'active'. The same considerations apply to whmovement. To get the defective intervention in 'wh-island' configurations, Chomsky has to assume that this movement also involves attraction of some interpretable feature, while uninterpretable features are only there to make such a DP active. For example, the ungrammatical sentence in (91) would be ruled out under the assumption that some interpretable feature INT is targeted in *who* by the interrogative probe head. The movement of *how* over *who* would then be prevented by the latter's INT feature(s) defectively intervening between the matrix interrogative head and the INT features of *how*, while the movement of *who* to the matrix interrogative C would be barred since *who* does not contain uninterpretable features and is therefore 'inactive'.

(91) *How_i did you wonder who can fix the problem t_i.

At this point, it is important to note that, in this system, the only principled interpretation of this 'activeness' requirement is that any uninterpretable feature makes a DP that contains it active for the computational system. In other words, that a DP whose structural Case has been checked off is 'frozen' for A-movement follows here not from the fact that its Case has been checked off, but rather from the fact that such a DP has no uninterpretable features left which would make it eligible for further movement. That structural Case in and of itself is not necessary for A-movement is obvious from Chomsky's treatment of the raising of the expletive *there* which, in his view, contains only uninterpretable ϕ -feature(s) (uninterpretable because *there* is a true expletive), which suffices for raising. For example, the successive cyclic movement of *there* in (92) is assumed to be allowed by the latter's single uninterpretable ϕ -feature.

(92) there_i seems t_i to be likely t_i to be a riot in the city.

Thus, all that matters for making a DP active is the presence of some uninterpretable feature in that DP, whatever that feature may be. Consider now the structures in (90). (90a) is readily ruled out on the assumption that the DP *John*, whose (only) uninterpretable Case feature has been checked off by the embedded T, is inactive and cannot raise to check the EPP feature of the higher T. However, the structure in (90c) is wrongly predicted to be grammatical because the illicit movement is allowed by the presence of another uninterpretable feature which makes the relevant DP active. In other words, assuming that it is the D-feature that needs to be attracted to check the EPPfeature of the embedded T in (90c) (the same holds if the relevant features are ϕ -features), there is nothing to prevent the attraction of the DP *who* for that purpose since this DP is still active, containing the uninterpretable wh-feature (or whatever feature makes the whphrases active and hence eligible for movement).⁵⁹ Thus, (90c) converges and the contrast between this structure and the much better (90b) is unaccounted for.

Similarly, the sentence in (90d) is wrongly predicted to be grammatical. Since *who* contains an uninterpretable wh-feature even after its Case has been eliminated via a relation with the preposition *to* (in 90d), it is still active and there is nothing to prevent its attraction by T(ense), whose nominative Case feature can subsequently be checked against the matching Case feature of *John*.⁶⁰ The account carries over straightforwardly to (90e),

⁵⁹ The above argument holds, I believe, even if, contrary to what we assume in this study, one adopts an analysis on which not every wh-phrase must have an uninterpretable wh-feature. That is, since the uninterpretable feature must be allowed to be present in wh-phrases in order to make, say, the wh-phrase *what* in (ii) eligible for movement to SPEC-CP, one would have to explain why such an option would not be utilized to make the wh-phrase *who* in the embedded sentence in (90c) eligible as well.

⁽i) What will you bring for the party.

The uninterpretable wh-feature of *which man* in (90c) would presumably be checked against those of C under the assumption, adopted in Chomsky (1995, 2000, and, in particular, 2001) that deleted but unerased features can be accessed by the computation.

⁶⁰ Note that (90d) cannot plausibly be ruled by, for example, pointing to the ungrammaticality of (i) as an example of the inability of wh-movement from the complement position of to and then claiming that whatever rules out (i) would also rule out (90d).

⁽i) ?Who did it seem to t that John is sick.?

which has an added property of ruling out any account that would resort to 'suppressive role' of inherent Case.

Parallel reasoning applies to structures that violate the ban on wh-movement from an operator position (Epstein 1992, and references therein). (93a) is correctly ruled out since the wh-phrase *who* has had all its uninterpretable features checked in the embedded CP and is inactive regardless of which interpretable feature is targeted by *Attract*. But (93b) is incorrectly ruled in since the uninterpretable Case feature on *who* in the most deeply embedded operator position makes the DP active and the relevant interpretable feature available for attraction by the higher C.⁶¹

(93) a. *Who_i did you wonder [$_{CP}$ t_i [$_{IP}$ t_i would buy what]]

(for which person x and a thing y, you wondered whether x would buy y)

b. *I know [CP who you wondered [CP t_i [P it was told t_i that John would buy what]

(ii) ??Who did you give a book?

However, even those speakers who do not accept (ii) accept (iii), in which the underlying indirect object has wh-moved from the derived subject position.

(iii) Who was given a book?

⁶¹ Note that (93b) cannot be ruled out by requiring that *wonder* have a wh-phrase in the specifier of its complement (or the 'Wh-Criterion' of Rizzi 1996). First, such a requirement is a stipulation that should follow from independent principles of grammar. Second, on the copy theory of movement, there *is* a wh-phrase in the specifier of the complement of wonder in (93b). Furthermore, the relation between a head and the specifier of its complement is not a natural relation in the Minimalist program, which dispenses with the notion of *government* from the inventory of basic relations.

There are couple of problems with such a argument. First, (90d) is far worse than (i), and the 'account' would not explain the contrast. Second, given that, obviously, *who* in (90d) is wh-moved from the subject position, the argument would have to rely on an untenable position that, if an XP cannot undergo wh-movement from its base-generated position, it cannot do that from the derived subject position either. That this is an untenable claim is obvious from, say, the pair in (ii-iii). (ii) is an example of a wh-movement of an indirect object in the double-object construction, which most speakers of English (though not all, see below) consider ungrammatical.

To rule out (90a,c,d,e) and (93b) in a uniform manner, one would have to assume a) that uninterpretable Case or ϕ -features, but not an uninterpretable wh-feature, makes a DP active for A-movement and, conversely, that an uninterpretable wh-feature, but not uninterpretable Case or ϕ -features, makes a DP active and eligible for wh-movement. These assumptions, coupled with the already fuzzy and complex notion of activation, would further complicate the system.⁶²

Note also that, on the accounts discussed in the previous chapter which assume that ECM verbs drive overt movement to their specifier position, (93b) also violates the ban against 'improper movement', i.e. the movement from one A-position to another via an A'-position. (94a) is another example of improper movement for which I don't see an obvious account in Chomsky's system. This is because, whatever feature of the noninterrogative Cs is responsible for the attraction of DP's, it is certainly not a structural Case feature. Thus, the structural Case on the DP *John* in (94a) is available at the point

Finally, (93b) cannot be ruled out as a violation of the Case filter, given that *know* can ECM an XP in the similar structural configuration:

⁽i) I've known him for a long time now to be unreliable.

Additional evidence for the ECM properties of *know* comes from the contrast between (ii) and (iii), reported in Bošković (1997) (note that Bošković argues that the verb *conjecture* does not assign structural Case):

⁽ii) *Who did John conjecture t to know French?

⁽iii) ?* I know who John conjectured t to know French.

Note also that, since Chomsky does not assume that ECM verbs have the EPP property, (93b) does not involve an 'improper movement', so this 'account' of (93b) is also not available in his system.

⁶² The problem gets worse if the interpretable features attracted in wh-movement are again agreement (ϕ -) features, as is sometime suggested.

when it has moved to the embedded Spec. of CP^{63} and its ϕ -features should be available for attraction by the matrix T. (94b) adds to the problem by introducing a DP which bears an additional uninterpretable feature.

- (94) a. *John_i seems [$_{CP}$ t_i [$_{IP}$ it was told t_i that ...]]
 - b. *Who_i seems [$_{CP}$ t_i [$_{IP}$ it was told t_i that]]

To sum up, we have seen that the assumption that *Agree/Attract* targets interpretable features makes incorrect predictions by allowing ungrammatical structures to be generated as well as disallowing grammatical ones. In addition, it uses computationally complex or vague notions such as equidistance, 'activation' of XPs by uninterpretable features, suppression of ϕ -features by inherent Case, and the like. In the next Chapter, I will show that a reformulation of (1) to the effect that only uninterpretable features are targeted by *Agree/Attract* not only provides a simple account for all the structures that we have discussed, but it also allows us to dispense with equidistance and all the problematic concepts mentioned above. We now turn to our main proposal.

⁶³ Such a movement would be triggered in Chomsky's system by utilizing the option of assigning an EPP feature to the CP phase (see Chomsky 2000) for details.

Chapter 3

Movement and Interpretability

Let's restate Chomsky's definitions of the operation *Attract/Move/Agree* in terms of the definition in (95):

(95) Feature Checking

K enters a checking relation with F if F is the closest matching uninterpretable feature in the domain of K.

Closeness

If β c-commands α and τ is the target of raising, then β is closer to τ than α .

Under this reformulation of the MLC, the grammaticality of the structures in (34) is straightforward since, at the relevant point in the derivation, the higher DPs have either checked the relevant uninterpretable feature, i.e. structural Case, or, in the case of inherently Case-marked DPs, do not have uninterpretable features to begin with. As there is no defective intervention in my system,⁶⁴ the uninterpretable Case feature of the lower DP can be attracted and the raising of the lower DPs in (34) is predicted to be OK. Let's illustrate with the derivation of (34b), whose 'S-structure' is given in (96).⁶⁵

⁶⁴ This is because the probe searches for the closest matching uninterpretable feature. If such a feature has been eliminated from an otherwise closer DP, the probe is free to search further down the tree for the next closest matching feature.

 $^{^{65}}$ I assume that in ditransitive structures, of which the 'psych' construction under consideration is a special case, the light verb v selects another functional category, Asp(ect), which 'checks' accusative Case feature of the goal. I also abstract here from verb movement, which may well be a PF process, as Chomsky (2001) argues.

(96) $[_{TP} [each other's friends]_j T [_{\nu P} annoy [_{ASP} [John and Mary]_i [_{\nu P} t_i [_{\nu'} t_j]]]]]$

When the aspectual head is merged in the structure, its structural Case feature seeks the closest matching feature in the tree, and finds it in the DP *John and Mary*, which we have assumed to be generated in the Spec. of VP. This results in the movement of the containing DP *John and Mary* to the Spec. of ASP. Subsequently, when T is merged, its structural Case feature seeks the closest structural Case feature in the constructed object. Since the Case feature of the DP *John and Mary* has been eliminated in the previous step, it cannot be attracted nor can it interfere with the attraction of the next closest structural Case feature, the nominative Case feature of the DP *each other friends*. The movement of this DP into the Spec. of TP yields the surface order.

In the 'seem-to' construction in (34a), repeated here as (97), the DP *John* can be attracted and raised to Spec. of TP after the Case on *Mary* has been checked off by the preposition to.⁶⁶

(97) [TP [Each other's photos]_i seem to John and Mary [TP t_i to be scandalous]]

When the matrix T is merged, it seeks the closest uninterpretable Case feature in its domain. Since at this point the Case feature of the DP Mary has been eliminated, the

⁶⁶ If the Case on *Mary* is not structural but inherent and semantically contentful, then it cannot block the attraction of the lower DP since the probe seeks only uninterpretable Case features, on my account.

closest structural Case feature is the one on the DP John. Thus, the latter can be attracted to SPEC-TP, as required.

We predict that, in a language in which the structural Case on the experiencer DP is available at the point that matrix T is merged in the structure, the attraction of the lower DP will be blocked. This prediction is borne out by data from Icelandic (Sigurðsson 1989) in which the dative experiencer is assumed to also contain a structural (nominative) Case (see Belletti 1988, Chomsky 2000:127, Cowper 1988, Frampton and Gutmann 1999, Holmberg and Platzack 1995, and Freidin and Sprouse 1991, among others).⁶⁷ In Icelandic, when the experiencer DP is present in the structure, it blocks the raising of the subject of the embedded sentence:

(98) a. Hafði Ólafur, virst [t, vera gáfaður]

Had Olaf seem intelligent

'Did Olaf seem intelligent?'

(i) a. Ég vil að pú hjálpir Jóni

I want that you_{NOM} help Jon_{DAT}

b. Ég vil að Jóni verdi hjálpad

I want that Jon_DAT be_SUBJ helped

- c. *Ég vil að Jóni vera hjálpad
 - I want that Jon_{DAT} be_{INF} helped (H&P, p.29)

In (ic), the derived subject of the embedded infinitival clause is assigned inherent Case by the verb, yet the sentence is ungrammatical. To capture the fact that dative DPs in Icelandic seem to be subject to the distributional restrictions parallel to those of DPs with structural Cases, Holmberg and Platzack suggest that dative DPs be specified as [+/- nom, dat], and that each such DP passes through both lexical checking with the verb assigning the inherent Case and structural Case with the appropriate functional head. Example (ic) is then ruled out since the non-finite Tense (their Infl) cannot check the nominative Case. In a similar vein, Frampton and Gutmann (1999) propose that the relevant structural Case is Generic Case (GC) which, as in Holmberg and Platzack, must be checked by finite Infl.

⁶⁷ For example, Holmberg and Platzack show that the morphological Case assigned by the verb does not suffice to license the dative DP. The relevant data are given in (i).

b. Hafði peimi virst ti [Olafur vera gáfaður]

had them seemed Olaf intelligent

'Did it seem to them that Olaf was intelligent.

c. * Hafði Olafur_i virst peim [t_i ivera gáfaður] (Sigurðsson 1996, p. 29)

As pointed to me by Żeljko Bošković (p.c.), and as we have seen in the previous Chapter, some languages in which raising across the full lexical experiencer DP is barred allow the movement if the experiencer is a clitic or a wh-phrase which moves to SPEC-CP, as in French or Icelandic. While data are often controversial (e.g. French speakers vary greatly in judgments about these sentences, as noted in Bošković 2001), the allowed crossing pattern can be easily accommodated under our proposal if we follow Baker (1988) in assuming that Case checking of nominals can alternatively take place under incorporation. As for wh-phrases, they would presumably have to check Case before moving on to Spec-CP, so their Case feature cannot interfere with the attraction of the embedded subject. The problem that remains to be accounted for is the blocking effect of the full DPs for those speakers of French that disallow raising, given that there appears to be no evidence in French that the experiencer DP bears structural Case. My speculation is that there may be a change under way in French towards a structural reinterpretation of Case on the experiencer DP among those speakers, but, again, independent evidence is required.

The derivation of the TDC (34c) closely parallels the derivation for the 'seem-to' construction that we have just sketched. After the goal DP has had its case checked in

relation with the preposition *to*, there is nothing to interfere with raising of the theme DP to SPEC-AspP.

Finally, the crossing pattern exhibited in ergative languages follows directly if ergative Case on the agent DP is checked off before the raising of the theme DP. There are at least two approaches to ergative Case checking that one can pursue. On one approach, ergative languages would differ from nominative-accusative languages in that the functional head hosting the relevant (ergative/accusative) Case feature would be v rather than Asp. The agent DP would then have its Case feature checked off immediately upon the merger in SPEC-vP. On the other approach, ergative languages would differ from nominative/accusative languages in allowing a different ordering between the functional heads v and Asp, on which the former would be dominated by the latter. The agent DP would then have the Case feature checked off by raising to SPEC-AspP, which would be followed by the raising of the theme argument to SPEC-TP. I leave the choice between these and other alternatives for further research. The point is that that our analysis offers a simple way to allow for the raising of the theme DP to take place, in contrast to the alternative on which the movement would be disallowed by defective intervention.

As for the ungrammatical sentences in (90), repeated below, they are also straightforwardly ruled out since in none of them does the target T share an uninterpretable feature with the goal DP. In (90c), for example, the embedded T does not have a wh-feature which would attract the corresponding uninterpretable wh-feature of *who*. Similarly for (90d,e): neither structure contains a feature that the target T can attract in the relevant DP. And since *Move* is dependent upon *Attract*, the movement of the DP to the target T is also impossible.

(90) a. *the belief [$_{TP1}$ John_i to seem [$_{TP2}$ t_i is sick]]

- c. $*[_{TP}$ Who dispelled the belief [which man_i to seem t_i is sick]]
- d. *I wonder who_i t_i seems to t_i was told John that ...
- e. *I wonder who_i seems t_i was told t_i John to be innocent.

The account carries over straightforwardly to the ungrammatical (93b).

(93b) *I know [CP who you wondered [CP t_i [P it was told t_i that John would buy what]

The illicit movement from the most deeply embedded operator position to the next higher Spec, CP is barred since it does not involve the matching of any uninterpretable features. That is, we know from structures such as (99) that successive cyclic A'-movement is not driven by structural Case, since the structural accusative Case in (99) is checked off in the relation between Asp and the DP *what* before the wh-phrase undergoes A'-movement.

(99) What_i did you say [$_{CP}$ t_i that John bought t_i]

Since the only plausible candidate, the wh-feature of *who*, has already been checked off in the relation with the matching feature of the embedded interrogative C in (93b), there is

no uninterpretable feature shared by the higher C and the wh-phrase *who* and the movement from the operator position is correctly barred. This differs from the situation in (99) in which the embedded C does not have a wh-feature that would erase the wh-feature of *what*, which therefore can be attracted by the matrix interrogative C.

A note is, perhaps, in order here about the character of the grammar that the above proposal, as well as Chomsky's (1995, 2000, 2001) set of assumptions, imply. In particular, under Chomsky's analysis, a question can be raised whether syntax should or can be 'knowledgeable' of the distinction between interpretable and uninterpretable features so that, for example, it would 'know' when an XP is 'active' in order to establish whether *Agree* can apply, or where it knows not to delete interpretable features once *Agree* has applied. In other words, while Chomsky has consistently claimed in his works that syntax is a 'dummy' component of the grammar which is not driven by interpretive conditions (see footnotes 7 and 8 above), assuming that feature checking (or the operation *Agree*) affects only 'active' XPs in the above sense, and that it deletes or erases features only if they are uninterpretable, comes pretty close to implying that syntactic operations are knowledgeable of semantic distinctions. Similarly, the conception of feature checking in (95) appears to imply such knowledge.

There are a number of ways in which one can approach this issue. One can claim, for example, that syntax has a free option of deleting or not deleting the checked features, where derivations in which interpretable features are deleted and uninterpretable features are preserved are discarded by some overarching and independently needed principles such
as Full Interpretation. However, this proposal appears to undermine the Last Resort conception of syntactic operations, given that one would have to allow for cases where *Attract/Agree* does not result in feature deletion. In other words, to maintain the Last Resort conception of movement on this analysis, it appears that one would need to incorporate some global principle which would rule out derivations involving vacuous applications of syntactic operations, with the attendant complexity issues which global economy conditions introduce in the grammar. In addition, while this might resolve the question of feature deletion consequent upon *Agree*, it does not appear to resolve the issue of how the system knows if an XP is 'active', so that *Agree* can apply in the first place, or 'inactive' so that the XP can induce defective intervention effects.

A conception of feature checking that avoids these problems is one on which feature checking necessarily involves feature deletion/erasure, or on which feature checking *is* feature deletion/erasure. Under this conception, the design of the grammar has to be such that interpretable features are invisible to *Attract/Agree*, in order to avoid radical violations of Full Interpretation. On this proposal, we can maintain the view that syntax is insensitive to semantic distinction since feature checking blindly and generally affects all features involved in the operation.⁶⁸ No global economy is needed under this

Case (i) Wh Phon. Sem. Feat. Feat.

⁶⁸ Technically, one way to implement this would be to have bundles of uninterpretable features constitute the roots in the representation of functional and lexical categories, in something like the representation in (i), and to assume that only roots are affected by syntactic operations.

proposal and the Last Resort nature of syntactic operations can be maintained in the strictest form. Crucially, though, this proposal is only consistent with the conception of the nature of the operation *Attract* underlying (95), and not with the alternative that I argued against in the previous chapter on which feature checking can involve interpretable features.

We still have to account for the case of superraising in (31), and for the cases of 'improper movement' in (94). It will turn out that our explanation for the superraising structure will translate directly to the cases in (94). The explanation relies on rethinking the timing of expletive insertion in the structure.

This would also provide a natural mechanism to capture Chomsky's (2001) proposal that "the phonological component spells out elements that undergo no further displacement" (p.10), which otherwise appears to be raising the issue discussed above in a different form: in order to know whether an XP undergoes 'no further movement' in Chomsky's implementation, the phonological component, i.e. the Spell-Out mechanism, would need to 'know' whether the XP is 'active' or not, since inactive DPs cannot move. But to know if an XP is inactive, the mechanism would need to know whether it contains uninterpretable features, which implies semantic knowledge. Under the implementation illustrated in (i), we can just assume that phonological Spell-Out takes place once the root of an XP has lost all its features. Elaborating further on this proposal, one can assume that deletion of all features from the root splits the XP into its immediate constituents, which at that point must be accessed by PF and LF interfaces. Other technical implementations are, of course, possible.

Chapter 4

Expletive Insertion

Recall that the main reason that Chomsky (2000) assumes the mechanism of defective intervention is to block 'superraising' cases in (10), repeated here as (100):

(100) *[$_{TP}$ John_i seems [it was told t_i that ...]]

This assumption, however, leads to undesirable blocking effects with respect to the structures in (34), as we have seen. But now we have to make sure that, by allowing the structures in (34) we do not incorrectly rule in the superraising in (100). Before taking up this example, however, let us review Chomsky's relevant assumptions on expletive constructions.

An important economy principle in Chomsky (1995) is that the operation *Merge* is preferred to the operation *Move/Attract*. This is taken to follow from the conceptual necessity and simplicity of the operation *Merge*, and is also claimed to be a consequence of another economy principle, *Procrastinate*, which prefers covert to overt movement. In Chomsky (2000, 2001), the preference of Merge over Move is taken to be a matter of less vs. more, since the operation Move combines the operations *Agree and Merge*. In fact *Move* is more complex not only than each of the operations *Agree* and *Merge* but also their combination, since *Move* involves an extra step of selecting the category that is to be pied-piped in the operation.

To see how this works, let us briefly consider the derivation of the sentences in (101-103). Both of the derivations represented in (101) and (102) involve the same

number of overt movements. However, the derivation in (102) is selected since, at the point of the derivation shown in (103), the merger of the expletive *there* in the SPEC-TP is a more economical way of checking the EPP feature of T than the raising of the associate *a man*.

- (101) *There seems a man_i to be t_i in the room.
- (102) There, seems t_i to be a man in the room
- (103) $[_{TP} T \text{ to be a man in the room}]$

In Chomsky (1995), these assumptions lead to a problem with constructions involving the expletive *it*, as it was not clear how to derive the structure in (104), given that at the stage of derivation shown in (105) the merger of *it* should prevent the raising of *John* to the embedded SPEC-TP. In other words, not only was (104) incorrectly predicted to be underivable, but, as observed above, the grammar allowed the derivation of unacceptable (106).⁶⁹

- (104) It seems [John was told that ...]
- (105) $[_{TP}$ was told John that ...]
- (106) It_i seems [t_i was told John that ...]

⁶⁹ For problems with these structures in Chomsky's (1995) system, see Uriagereka and Raposo (1995), Vukić (1997, 1999).

This problem is avoided in Chomsky (2000) under the assumption that a) the computational system cyclically accesses the subarrays of the numerations that are called phases, b) phases are propositional categories CPs and vPs, and c) the preference of *Merge* over *Move* involves the operations affecting the same phase. Thus, since *it* and *John* in the above structures belong to different phases, *John* is allowed to raise to the embedded SPEC-TP.

Since I do not assume the existence of phases in this thesis,⁷⁰ I would like to suggest another approach to expletive insertion, along the lines advocated in the work of Murasugi (1992), Frampton (1997), and Vukić (1999a) and closely related to the recent proposal in Bošković (2001). Each of the works cited suggests treating expletive insertion as a last resort operation to be employed only if other options are unavailable. Let me briefly review their arguments.

Murasugi (1992) assumes that the expletive *it* is a non-thematic element inserted to satisfy feature requirements of [+tense] T. She proposes that the insertion of *it* is similar in the relevant respects to English *do*-support, an operation which she claims inserts *do* in the modal position in the overt syntax, in that "both these processes involve a last-resort strategy, occurring after all other s-structure movement." (p53). She refers to Chomsky's (1991) observation that, although *do*-support in English may appear to be a more economical operation than verb raising, the latter 'wins' as it involves the application of the principles of Universal Grammar, unlike the former which is a language-specific option. Since UG principles are 'wired in' and therefore less costly than language

 $^{^{70}}$ For problems with the concept of phase, see Boeckx (2001) as well as Bošković (2001), where some alternatives are also sketched.

particular operations,⁷¹ the latter are used only to 'save' a representation which otherwise would be uninterpretable. She then suggests that "both *do*-support and *it*-insertion, which are language specific requirements, are applied only when universal movement options are not available." (p. 53) Similar observations apply to the insertion of the other English expletive, *there*, which Murasugi takes to be a last resort strategy to satisfy the Case requirements of [+tense] T. She also adopts Belletti's (1988) and Lasnik's (1992) proposals that, in existential constructions, partitive Case is assigned inherently by the verb.

Frampton (1997) also argues that expletives are inserted as a last resort operation (under his principle of Last Resort Insertion (LRI)) when movement operations do not suffice to yield a convergent derivation. He claims that expletives are excluded from the initial representation (i.e. Numeration) and are inserted directly in the course of derivation via the operation *Insert*, which he takes to be more costly⁷² than the operations *Merge* and *Move*. Frampton also assumes that nominals can be either NPs or DPs, and that only the latter can check the EPP feature of the category T. To see how his proposal works, consider the pair in (107)

(107) a. A man was in the room.

b. There was a man in the room.

⁷¹ I guess that 'cost' here is estimated with reference to language acquisition where, in contrast to language specific mechanisms, those that are wired in come 'for free'.

 $^{^{72}}$ I don't think that the assessed cost here follows from any deeper principle. See the analysis in Vukić, sketched below, where late expletive insertion is shown to be practically forced by other, I think plausible, assumptions in Chomsky's (1995) system.

In (107b), Frampton claims that the category of the associate is NP, hence the merger of the expletive is required to check the EPP feature of T. In (107a), the category of *a man* is taken to be DP, thus the EPP feature is checked by the raising of the associate rather than by the more costly insertion of the expletive.⁷³

Frampton claims that his proposal offers a simple account of several structures whose nonexistence is otherwise rather mysterious. To begin with, it accounts for the well known definiteness effect in these constructions. Since definite nominals are DPs, they are able to check the EPP feature of T and their raising will always preempt the insertion of the expletive as a consequence of LRI. This explains the ungrammaticality of (108).

(108) *There was the man in the room.

In addition, Frampton shows that his proposal accounts well for some missing structures in Icelandic. First, he observes that the distribution of the Icelandic expletive $pa\partial$ is wider than that of its English counterpart, *there*. For example, in contrast to *there*, $pa\partial$ does not require an associate:

(109) a. það rignt

there rained.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

⁷³Frampton must be assuming that D is the locus of information on specificity, which can be positively or negatively specified.

b. það var hlegið að okkur.

There was laughed at us (p.50, ex.33)

Given this wider distribution, one needs to account for the ungrammaticality of (110a,b).

(110) a. *Það var maðurin i garðinum

there was man-the in the garden

b. *Það var hjálpað okkur.

There was helped us (p.51, ex.36)

Frampton claims that the LRI accounts for both of these structures. That is, the examples in (110) are blocked by the following, in which the raising of the definite DP in (111a) and the pronoun in (111b) prevents the assumed more costly insertion of the expletive.

(111) a. Maðurin var í garðinum

man-the was in garden-the

- b. Okkur var hjálpað
 - us was helped

Furthermore, one does not find (112) in Icelandic, although one does find structures such as (113a,b):

(112) *Það virðist [bækurnar hafa verið lesnar]

there seem books-the(N) have been read. (p.52, ex.39)

(113) a. Bækurnar virðast [t hafa verið lesnar]

books-the(N) seem have been read

'books seem to have been read'

b. Mér virðast [bækurnar hafa verið lesnar]me(D) seem books-the have been read

'it seems to me that the books have been read' (p.51, ex.38)

Again, the paradigm receives a natural account under LRI. In (112), the expletive það is inserted at the point in the derivation where a less costly raising alternative given in (113a) is available.

Vukić (1999a) offers an economy account of late expletive insertion which is based on Chomsky's (1995) set of assumptions. As we have seen Chomsky assumes that deletion of uninterpretable features drives movement. If an uninterpretable feature survives at LF (or PF for features uninterpretable at that level), the derivation will crash since the LF representation will contain an uninterpretable object. Once the feature is checked, it is deleted and, if possible erased. This is explicitly stated as the Economy Condition:

(114) a. A checked feature is deleted when possible.

b. Deleted α is erased when possible. (p.280, ex (52)

From the formulation in (114) it is obvious that the deletion plus erasure of a checked feature is favored over deletion alone. Erasing checked feature is crucial for the workings of the whole system, given Chomsky's assumption that deleted features are accessible to the computational system though invisible at the interface. As Chomsky observes, if the Case feature of *John* were just deleted but not erased upon checking against the corresponding feature of the embedded T in (115), then this feature could again be accessed to check the Case feature of the matrix T, and the sentence would be predicted to be grammatical:

(115) *John_i seems [t_i is intelligent. (p.284, ex (17)

Thus, within the overall system of assumptions in Chomsky (1995, 2001), deletion without erasure is a marked operation to be allowed only if some overriding principle forces it.

One such principle in Chomsky (1995) stipulates that erasure "cannot create an illegitimate object, so that no derivation is generated"(p.281). In particular, the erasure of a checked uninterpretable feature cannot lead to the erasure of the entire term from a syntactic object. This principle is given in (116):

(116) A term of Σ cannot erase (Chomsky 1995, p.281)

One situation where the violation of (116) arises is with structures involving expletives, which are pure bundles of uninterpretable features (apart from phonological features accessed at the PF interface, of course). That is, even the categorial D-feature and ϕ -features, which are generally interpretable on nominals, are uninterpretable on expletives which lack semantic content. However, these features can never be erased, since their erasure would result in the erasure of the entire syntactic object, in violation of (116). Therefore, the structures involving expletives necessarily lead to the presence of deleted ,but unerased uninterpretable features, an option that is otherwise strictly disfavored, as we have seen.⁷⁴

Since the presence of expletives forces the computational system to choose a highly marked option of non-erasure, it seems natural to propose that they are inserted in the structure only if the derivation would not otherwise converge, i.e. as a last resort

⁷⁴ Recently, Chomsky has suggested that a term of Σ can erase provided it does not determine the label. Lasnik (p.c.) also observes that the preference of erasure over deletion is in contradiction with Chomsky's rationale for Merge-over-Move preference, which is based on the principle that 'fewer is more economical.' Perhaps these issues can be resolved in the following way. Assume, contra Chomsky (1995, 2001), that feature checking necessarily results in their erasure. This seems plausible, given that feature erasure is the ultimate end of the whole operation and is inevitable in any way to avoid crash. Then, one could still argue that deletion of a term of Σ would necessarily entail subsequent adjustment of the syntactic object. For example, projections would become vacuous, or categories that are [-minimal, -maximal] would become [-minimal, +maximal], which could impact the interpretive mechanisms, given that intermediate projections are often taken to be invisible to them. In any case, if such an adjustment would be necessary, it would involve 'doing more' and late expletive insertion would still fall under an economy account.

operation employed to save the derivation from crashing. In Vukić (1999a), I proposed the principle of Late Expletive Insertion (LEI):

(117) Expletives are inserted as late in the derivation as possible.

However, as I observed in Vukić (1999a), the principle in (117) is somewhat odd in a system where the derivation of a sentence is based upon a pre-chosen fixed set of lexical items in the numeration. Once an expletive element is chosen for the numeration, there seems to be no choice but to use it, if the derivation of a sentence is to converge. It is therefore much more natural to conceive of (117) as a principle governing the operation *Select*, which chooses lexical items for the numeration from the lexicon. We may, then, essentially follow Frampton (1997) and assume that expletives are excluded from the numeration⁷⁵ (his 'initial representation'), and are accessed in the course of the derivation from the lexicon when the set of items selected in the numeration is insufficient to yield a convergent derivation.^{76,77}

⁷⁵ Frampton assumes this only for the expletive *there*, and not for *it*.

⁷⁶ In fall '97 class lectures and in Chomsky (2000, 2001), Chomsky argues against analyses which do not assume that derivation of a sentence is based on a pre-selected set of lexical items. The argument rests on the assumption that accessing the whole lexicon throughout the derivation poses a tremendous load on the working memory. (Howard Lasnik (p.c.) notes that this is not obvious, given that the lexicon is always there in the background, while the numeration has to be put in the working memory.) Note that this objection does not arise in my analysis, since the only elements that can be accessed directly from the lexicon are expletives, which in every language represent a highly limited set. Technically, one can imagine that expletive elements occupy a separate part of the lexicon, which is marked as accessible at all times to the computational system.

⁷⁷ Bošković (1997) proposes that the initial numeration also excludes functional categories, on somewhat similar grounds.

(102) There_i seems t_i to be a man in the room

(118) [TP T to [be a man in the room]]

If the verb is chosen with partitive Case, then the same Case must be assigned to the associate if the derivation is to converge. And since partitive is an inherent Case, which we have assumed to be checked off through relation with the verb, the associate will not be able to raise and check the EPP feature of T in (118). Furthermore, since at this point the numeration contains only the verb *seem*, the expletive *there* must be accessed from the lexicon to check the EPP feature. The subsequent merger of the matrix verb and the raising of the expletive yield (102), as desired. Note that for the raising of the expletive to be possible, we must assume that the expletive bears structural Case, which has been independently argued to be the case in Belletti (1988), Bošković (1997) and Lasnik.(1992, 1995a, and subsequent work).⁷⁸

If the verb in a structure like (118) does not bear partitive Case, the DP *a man* will necessarily bear nominative Case, in which case it can raise to the embedded T once it

⁷⁸ I would like leave open the question of apparent long distance agreement between T and the associate in the expletive constructions. For some possibilities, see Lasnik (1992, 1995), Frampton (1997), Bošković (1995, 1997).

enters the derivation. At this point, the DP will still have its nominative Case unchecked and will therefore be eligible for further raising. When the matrix verb and T enter the structure, the DP raises to check the latter's Case and EPP features and its own Case, yielding the sentence in (119).

(119) A man seems to be in the room.

Recently, Bošković (2001) has argued that expletives never undergo any movement. The claim is embedded in Bošković's analysis which argues for elimination of the EPP from the grammar. Bošković shows that this observation on expletives' immobility provides a straightforward explanation for the otherwise puzzling facts related to several constructions. Here I present one of his arguments, based on French experiencer constructions.

As discussed in the previous chapter, Bošković observes that some speakers of French disallow raising of the embedded subject in experiencer constructions parallel to English (120), if the experiencer DP is a full lexical DP.

(120) Two soldiers seem to the general to be smart.

(121) a. *Deux soldats semblent au général manquer (être manquants) à la caserne two soldiers seem to-the general to-miss to-be missing at the barracks 'Two soldiers seem to the general to be missing from the barracks.'

b. *Deux soldats semblent au général être arrivés en ville.

two soldiers seem to-the general to-be arrived in town

Bošković attributes the ungrammaticality of (121) to a locality violation, as suggested in Chomsky (1995) and elsewhere. He then observes that the same blocking effect does not show up in the parallel constructions with expletives.

(122) a. Il semble au général y avoir deux soldats manquants à la caserne.

there seems to-the general to-have two soldiers missing at the barracks 'There seem to the general to be two soldiers missing from the barracks.'

b. Il semble au général être arrivé deux soldats en ville.

there seems to-the general to-be arrived two soldiers in town

'There seem to the general to have arrived two soldiers in town.'

Bošković claims that the contrast follows if a) expletives don't move, and b) EPP does not operate in the grammar. This means that no movement of expletives takes place in the above sentences, which are therefore inserted directly in the SPEC-TP position, hence no locality violation takes place.

While Bošković's proposal clearly differs from mine,⁷⁹ the two proposals converge on dispensing with the Merge-over-Move preference, as I discuss below. In other words, the ungrammaticality of (101) is accounted for in both approaches independently of the Move-over-Move preference, though on different grounds. It should also be noted that, if EPP indeed does not operate in the grammar (or at least not as a force that can drive movement/merger, see footnote 8), as I will assume with Bošković in the chapter on successive cyclicity, then the immobility of expletives would practically follow in both Bošković's analysis and mine under the plausible assumptions that a) expletives are just bundles of uninterpretable features, and b) feature deletion without erasure is not an option (alternatively, deleted features cannot be attracted). Given that, on the no EPP hypothesis, the heads (finite T and Asp, or AgrO) with which expletives enter a checking relation always contain Case and ϕ -features, the last-resort merger of expletives and the resulting feature checking would necessarily delete all their features, rendering them invisible for any further syntactic operation.⁸⁰

Before we take up structures with the expletive *it*, I wish to point to some other potential problems in Chomsky's (2000, 2001) system which involve the expletive *there*, one conceptual and a few empirical.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

⁷⁹ For example, in contrast to the present proposal, Bošković's analysis does not relate directly to the superraising and other constructions which we discuss below.

⁸⁰ Note, however, that on this proposal, the superraising construction in (31) would be more problematic for analyses that do not assume late expletive insertion. That is, since on such analyses *it* can be inserted in the embedded SPEC-TP, the deletion of all of its features would render it invisible for defective intervention. This problem does not arise if LEI holds.

The conceptual argument involves a curious asymmetry that exists in Chomsky's (2000, 2001) system and which involves the expletive *there*. Chomsky (2001) assumes (123) as the main condition governing the applicability of the operation *Agree* under which uninterpretable features are deleted:

(123) i. Probe and goal must both be active for Agree to apply.

ii. α must have a complete set of ϕ -features (it must be ϕ -<u>complete</u>) to delete uninterpretable features of the paired matching element β .

It is important to recall that Chomsky assumes both the expletive *there* and the raising T to have only the feature [person]. Thus, neither of the two is ϕ -complete. So given (123), neither should be able to delete the other's uninterpretable ϕ -feature and the raising of *there* to the SPEC of the 'defective' T should be unavailable, i.e. should not result in feature deletion and should therefore violate the last resort condition on movement. However, Chomsky assumes, contrary to what (123) should lead us to expect, that *Agree* between the matching uninterpretable [person] features of *there* and T delete the latter but not the former. Thus, an unexplained asymmetry exists in the system.⁸¹

⁸¹ As pointed out in Bošković (2001, fn 50), Chomsky's assumptions about expletives, in particular the position that expletives, being X^0 s can act as probes, would predict that *there* would be 'frozen' in the embedded SPEC-TP position in (i) since it would lose its uninterpretable feature upon probing the embedded clause and finding the matching feature in *someone*.

⁽i) There seems to be someone in the garden.

Empirically, it seems to me that Chomsky's assumption that *there* has no Case may allow unwanted structures in the grammar. To see this, recall that Chomsky assumes that T and v have no Case feature and that A-movement can therefore never involve Case, since matching between features involves feature identity. Thus, even in control structures such as (124), one has to assume that *Agree* obtains between the matching ϕ -features of T and PRO, and that the latter's null Case is deleted as a reflex of this agreement.

(124) It's good [PRO_i to be invited t_i]

However, if there has no Case, then it's not clear how (125) is ruled out.

(125) *it's good [there to be invited PRO]⁸²

(125) should converge unless one stipulates that control T has no ϕ -features, or that its ϕ -features cannot delete the ϕ -feature of the expletive, which I think cannot be successfully argued in Chomsky's system given that he has to assume that control T is ϕ -complete (since Agree between the T and PRO deletes the latter's Case feature). On the other hand, if the expletive *there* has Case, then (125) is simply ruled out by the Case filter.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

⁸² As is well known, structures such as (125) used to be ruled out by the PRO theorem in 'standard' Principles and Parameters approach. However, since Minimalism dispenses with the notion of government on which the PRO theorem relies, and since it seeks to account for the distribution of PRO in terms of null Case, the structure in (125) clearly raises a problem.

Another empirical problem is the one noted by Lasnik (1995a), which refers to the simple intransitive constructions. Namely, if *there* has no Case, the structure in (126) poses a problem:

(126) * There someone laughed.

(126) should converge since it satisfies all the conditions: EPP is checked by the expletive, while *Agree* deletes the matching features between the T and the 'associate' DP *a man*. If, on the other hand, the expletive has a Case feature, than the structure is correctly ruled out as a Case filter violation.⁸³

Consider next the structures in (127), the likes of which were also mentioned in Frampton (1996) but were left in a somewhat inconclusive state (as Frampton himself observes).⁸⁴

(127) a. a man_i was laughed at t_i

b. *There was laughed at a man.

⁸³ For additional arguments that the associate in expletive constructions has partitive Case, see, in particular, Lasnik (1995b) and Bošković (1997, 2001).

⁸⁴ Frampton used examples with pronouns as 'associates'. One problem that a structure like (127b) would appear to raise under Frampton's analysis is that *a man* being ambiguous between an NP and a DP, would seem to allow the derivation of (127) under the NP 'reading'. This problem does not arise on our analysis, as we show directly.

Given that, based on (127a), one has to assume that case-assigning properties of prepositions are neutralized in the pseudo-passive constructions in English, the ungrammaticality of (127b) is mysterious under the account which favors early merger of expletives. That is, the structure would in essential respects be completely parallel to (128), where the merger of the expletive results in the checking of the latter's uninterpretable ϕ -feature, while the uninterpretable features of T and the associate are checked off by the subsequent *Agree* between the two.

(128) There was a man in the room.

On our analysis, the ungrammaticality of (127b) is straightforward. Assuming that the presence of the preposition prevents the relation between the associate and the potentially partitive bearing verb, the only option is for the 'associate' to be assigned nominative Case. But in that case, the principle of Late Expletive Insertion would force the raising of the associate, barring (127b).

We can now turn to structures with the expletive it, in particular the case of superraising in (100), repeated here as (129):

(129) *[$_{TP}$ John_i seems [it was told t_i that ...]]

On our assumptions, (129) is ruled out because, at the relevant point of the derivation given in (130), there exists a mechanism to check the Case feature of embedded T by an element selected for the numeration and already incorporated in the structure, i.e. by raising of the DP *John*.⁸⁵

(130) $[_{TP}$ was told John that ...]

This simple explanation carries out directly to the 'improper movement' structures in (94), repeated here:

(131) a. *John_i seems [$_{CP}$ t_i [$_{IP}$ it was told t_i that ...]]

b. *Who_i seems [$_{CP}$ t_i [$_{IP}$ it was told t_i that]]

In both cases, the expletive *it* is inserted from the lexicon at the point when raising of an element incorporated in the structure, i.e. *John* in (94a) and *who* in (94b) is possible, in violation of the principle of late expletive insertion. ⁸⁶

⁸⁵ In fact, (100) was also ruled out by Chomsky's (2000) 'Phase Impenetrability Condition' which states that in a phase with the head H, the domain of H is not accessible to operations outside that phase. So in this case, the domain of the head of the embedded CP should not be accessible for attraction to elements outside this CP. However, Chomsky (2001) reformulates PIC in a way that no longer rules out (129) without adopting defective intervention.

⁸⁶ A word is in order here about pairs such as the one in (i)

⁽i) a. That John likes Mary is likely.

b. It is likely that John likes Mary.

Obviously, the proper analysis would depend on whether clauses can have Case, whether they are indeed clausal elements when in subject positions, etc. The simplest approach most consistent with our account and with Bošković's (1995b) analysis of Case properties of clauses would be to assume that the clause bears Case in (ia), but not in (ib). When it does not have Case, the only way

Thus, the assumption that expletives are inserted as a last resort allows us to account both for the superraising in (129) and the improper movement constructions in (94). Before we turn to the next section, I want to point to another issue with the Merge over Move principle which seems to dissolve under our assumptions.

The issue is that, since the preference for Merge over Move is taken to be determined by Principles of UG, one has to assume that such principles are violable (or ranked lower that other principles of UG). This is because of the sentences such as (132a,b)

(132) *a. I_i expected [α t_i to be a proof discovered]

b. I expected [α a proof_i to be discovered t_i] (Chomsky 2000, p.18)

In (132a) the resort to the preference of Merge over Move leads to an ungrammatical structure. The ungrammaticality is explained in terms of a violation of another principle of UG given in (133).

(133) Pure merge in theta position is required of (and restricted to) arguments.

(Chomsky 2000 p.16)

to check the Case feature of the matrix T is by insertion of *it*. When the clause does have Case, the merger of the expletive is prevented by LEI. Obviously, this cannot be the whole story, as one has to specify what determines a particular choice of expletive, among other things. I would like to leave this issue at this admittedly tentative state. For some proposals on how the association between the expletives and their 'associates' is determined, see, in particular, Lasnik (1995b) and Bošković (1997).

Thus, in order to explain the grammaticality of (132b) which satisfies (133), one has to abandon the assumption that violation of the principles of UG necessarily leads to ungrammaticality, and is led to adopt something like a ranked constraints model in which (133) is ranked higher than the Merge over Move principle. On the other hand, it seems that (133), along with the principle of late expletive insertion, suffices to account for the ordering of Merge and Move operations in a way that preserves inviolability of the UG principles.⁸⁷

Before I close this section, I wish to mention two other constructions, both pointed out to me by Željko Bošković (p.c.). First, my approach to expletive insertion provides a straightforward account for the sentence in (134).

(134) *I know [CP what_i [TP it was remarked t_i]]

The problem that (134) raises for the alternative theory is how to rule out the derivation on which *what* gets exceptionally Case-marked by *know*, which can otherwise be shown to Case-mark the specifiers of its complements, as in (135)

(135) I've known John for a long time now to be a liar.

⁸⁷ See Castillo, Drury and Grohmann (1999) and Bošković (2001) for additional problems with merge-over-Move principle, and Bošković for problems with (133)

One way of ruling out (134) would be to assume that such Case-checking would have to attract *who* to SPEC-AspP (or Spec-AgrOP, on alternative accounts), and then to resort to claim that such a pattern would represent 'improper movement'. Alternatively, one could claim that any movement of a wh-phrase from an operator position is impossible. On my account, the status of (134) falls out straightforwardly as an LEI violation, since the expletive *it* is inserted in the structure at the point when raising of *what* to SPEC-TP is possible.

A potential problem for my analysis brought to my attention by Željko Bošković (p.c.) is that the analysis appears to allow the derivation of (136), based on the ditransitive construction in which one of the complements is a 'raising' infinitival.



Since the Case of Mary is checked off in SPEC-TP of the embedded clause, nothing under my analysis would appear to block the raising of the subject of the embedded infinitival to the matrix SPEC-TP.

There are at least two ways in which one can approach this problem under my analysis. First, observe that our discussion of A-movement so far has been restricted to inter-clausal structures or movements out of the raising TP. In other words, none of the cases that we have discussed involves A-movement out of a CP or A-movement across a IP-CP pair. On the other hand, the movement of *John* in (136) crosses the IP-CP pair, given that the embedded finite sentence is plausibly a CP.⁸⁸ Now Bošković (1997) provides numerous examples of the impossibility of such movement. Thus, the difference in grammaticality between the sentences in (137) from Romanian is accounted for if, in contrast to the complementizerless subjunctive (137a), (137b) involves the crossing of the IP/CP boundary.

- (137) a. Studenti trebuiau sa plece
 Students-the must_{-3PL} SUBJ.PART. leave
 'The students must have left'
 b. *Studenti trebuiau ca sa plece
 - Students-the must-3PL that SUBJ.PART. leave

Similarly, in scrambling languages such as Serbian or Japanese, (Bošković 1997, Saito 1992), scrambling of a DP across the clausal boundary necessarily causes such movement to have A'-properties, in contrast to clause-internal scrambling which can create new binders, suppress WCO effects and exhibit other properties characteristic of A-movement. The contrast between (138a) and (138b) from Serbian illustrates this:

⁸⁸ The CP status of the embedded clause would follow under Chomsky's (2000) assumption that a T(ense) with a full set of features must be selected by a functional category. Since the embedded T(ense) is finite and has a full set of features, it must be that it is selected by C rather than a verb, which selects 'defective' T(ense).

- (138) a. Nekoga_i će njegov_{i/j} otac kazniti
 Someone will his father punish
 'Someone, his father will punish'
 - b. Nekoga njegov otac vjeruje da će Marija pozvati
 Someone, his_{j/*?i} father believes that Mary will invite
 'Someone, his father believes that Mary will invite'

The loss of 'A-properties' in long-distance scrambling follows directly from the ban on Amovement across the IP/CP pair.

Finally, consider the well known contrast between (139a) and (139b) in English which has standardly been attributed to the assumption that the embedded infinitival in (b) is of the category CP.

- (139) a. John is certain $[_{TP} t to win]$
 - b. *John is probable [CP t to win]

The point of all these examples is to illustrate that A-movement across the CP boundary seems to be impossible even when 'defective intervention' account does not appear to be available or plausible. Given that the sentence in (136) involves a movement across the CP boundary, its ungrammaticality will also follow irrespectively of defective intervention, from whatever turns out to be a proper account of this constraint.

Another possibility is to rule out (136) as a violation of distributional requirements on raising infinitivals. The relevant restriction may be that infinitivals are generally excluded from inherently Case-marked positions, and we have seen that the lower object in the English DOC is plausibly taken to bear inherent Case, since it does not raise or alternate with other Cases. This account is admittedly stipulative, but appears to be right in excluding movement from an explanation. In fact, there is independent evidence that movement has nothing to do with the ungrammaticality of (136), based on Bošković's treatment of expletives.

Recall that Bošković argues that expletives never undergo any movement, a condition which we tentatively attributed to their having only Case and, potentially, ϕ -features which are necessarily checked off in relation to the same head, given that there is no T(ense) with which expletives are required to Merge that does not contain both sets of features. With that in mind, consider the structure in (140).

(140)*There seems [TP someone was shown t [TP to be a man outside]]

Given hat expletives never move, and given that the raising T does not have a Case feature that would force their merger, as Bošković argues, one has to conclude that the expletive in (140) is merged in the matrix SPEC-TP. However, (140) is as ungrammatical as (136). This suggests that their ungrammaticality should fall out from the same condition which therefore has nothing to do with movement. We conclude that (136) does not pose a problem for the analysis of displacement proposed in this study, although, of course, the data in (136) and (140) remain to be explained in a principled manner.

We have seen that our analysis on which only uninterpretable features are subject to the operation *Attract/Move* provides a simple and natural explanation of many constructions that pose a problem for the alternative account. In the next Chapter, we will show that the proposed analysis extends straightforwardly to cover the core cases of A'movement. It will be shown that the analysis offers a natural way to incorporate the effects of Pesetsky's (1982) Path Containment Condition in the Minimalist framework, and that a uniform pattern of overlapping dependencies holds of both A- and A'movement types. We will also discuss some further issues relevant to our proposal, such as successive cyclicity, and the relationship between agreement and structural Case. It will be shown that our analysis meshes well with a current version of Chomsky's Form Chain conception of movement proposed in Bošković (2001). We will also show that the view on which structural Case is but a reflex of an agreement relation between the probe and the goal faces serious empirical and conceptual problems.

Chapter 5

Further Issues

In this Chapter, we extend the analysis presented above to the core instances of A'-movement. It will be shown that our proposal provides a natural way to capture the effects of Pesetsky's (1982) Path Containment Condition in the Minimalist framework, and that the same pattern of nested dependencies obtains in both A- and A'-movement. We then address the issue of successive cyclicity and suggest that the proposed analysis of displacement be reconciled with evidence for the successive cyclic character of movement under a current implementation of Chomsky's Form Chain condition proposed in Bošković (2001). Finally, we address the issue of relationship between agreement and structural Case and point to several problems with the assumption that structural Case checking is but a reflex of the agreement relation.

5.1 A'-Movement

We have already seen that my analysis explains the ungrammaticality of structures involving 'improper movement' across expletives. We have seen that the analysis also offers a simple account for the ungrammaticality of the structures such as those given in (141), i.e. for the well known fact that a wh-phrase in an operator position is frozen for further wh-movement (cf. Epstein 1992).

(141) a. *Who_i did you wonder [_{CP} t_i [_{IP} t_i would buy what]]
b. *Who_i was believed [_{CP} t_i you wondered [_{CP} t_i [_{IP} it was told t_i that ...]
(for which person x was it believed that you wondered whether x was told that ...)

(141a) is underivable since there are no matching uninterpretable features between the matrix C and the wh-phrase. This is because the relevant feature has been deleted from the wh-phrase in the embedded Spec CP in a checking relation between the wh-phrase and the interrogative C.

Similarly, in (141b), the relevant step, i.e. the raising of the wh-phrase *who* from the embedded interrogative SPEC-CP to the next higher SPEC-CP, is ruled out since it does not involve matching uninterpretable features. The fact that the wh-phrase still contains a structural Case feature is irrelevant under our analysis since this is not the feature shared by the higher SPEC-CP

In this chapter, I will extend my analysis to some other instances of wh-movement In particular, I will concentrate on cases where one wh-phrase undergoes movement across another phrase in an A'-position. It will be shown that my proposal makes it possible to derive the core effects of Pesetsky's (1982) Path Containment Condition (PCC) in a straightforward and conceptually simple Minimalist manner. I will then also show that, under plausible and well motivated assumptions, my analysis extends to cover the well known cases of argument-nonargument asymmetries in wh-movement. As is well known, the grammar of English allows 'long-distance extraction' of some wh-phrases and bars extraction of others, as shown in (142): the sentence in (142a) is judged by many speakers to be perfectly grammatical, in contrast to (142b,c), which are completely unacceptable.⁸⁹

(142) a.. Which car_i did you wonder who_j t_j can fix t_i ?

- b. *Which man_i did you wonder what_i t_i can fix t_i ?
- c. *How_i did you wonder who can fix you car t_i ?

The sentences in (142) represent a well known argument-nonargument asymmetry in whextraction that has received a tremendous amount of attention in the generative literature (see references on p.1 and, in particular, Huang 1982, representative studies). The question we want to concentrate is how the paradigm relates to the analysis of locality and movement proposed in the previous chapter. We will begin with the contrast between (142a) and (142b), returning to (142c) at a later point.

⁸⁹ Note that the clear contrast between (142a) and (142b) indicates that D-linking is not sufficient for well-formedness of extraction from wh-islands. Nor is it necessary for the crossed wh-phrase to be subject. For example, none of these two conditions is satisfied in (i), which is none the lessperfectly grammatical in a situation where one is seen 'madly searching through the dictionary', as observed by Szabolcsi and Zwarts (1993)

⁽i) What the hell do you still not know how to spell?

We will see other examples below in which the extracted phrase is D-linked, yet the extraction is impossible (see, for example, Pesetsky's data in (149b).

This is not to deny, of course, that semantics can facilitate some readings or filter out others, as will be clear from my discussion of adjunct extraction in the next chapter.

To begin with, we can observe that the path of the subject movement in (142a) is, in a sense to be made precise below, included in a path which represents the movement of the object: the two paths exhibit a 'nesting' pattern, as indicated by arrows.

(142) a.. Which car_i did you wonder who_j t_j can fix t_i ?

In contrast, the paths in the ungrammatical (142b) exhibit a 'crossing' pattern in that one path partially overlaps with the other.

(142) b. *Who_i did you wonder which car_j t_i can fix t_j

This distinction in the movement pattern was taken to distinguish grammatical from ungrammatical sentences already in Kuno and Robinson (1972), where it was first proposed in linear terms that overlapping dependencies must be nested and must not intersect.

Pesetsky (1982) further develops a theory of movement paths in order to derive ECP effects from their interdependencies or interactions. Pesetsky proposes a structural reinterpretation of the Kuno and Robinson's observation and shows that the structural account is superior to Kuno and Robinson's linear constraint. Pesetsky proposes that the constraint be reformulated in terms of the Path Containment Condition given in (143). (143) Path Containment Condition (PCC)

If two paths overlap, one must contain the other.

where Path is, roughly, the set of nodes in the tree that connect the maximal projections dominating the foot and the head of the chain. The precise formulation is given in (144)

(144) Definition of Paths (p.289)

Suppose t is an empty category locally A'-bound by b. Then

(i) for α the first maximal projection dominating <u>t</u>

- (ii) for β the first maximal projection dominating <u>b</u>
- (iii) the path between t and b is the set of nodes P such that

 $P = \{x \mid (x=\alpha) \lor (x=\beta) \lor (x \text{ dom. } \alpha \& \neg x \text{ dom. } \beta)\}$

Under PCC, the grammaticality contrasts in the structures in (145)–(149) (based on Pesetsky's (1982) examples (34)-(41)) receive a simple and natural account, since it is only in the (a) examples that one extraction path is contained in the other.

(145) Questions

a. [CP what books_i [TP do you know [CP who_j [IP PRO to persuade t_j [PRO to read t_i]]]]]

b. *[_{CP} who_i [_{TP} do you know [_{CP} what books_j [_{TP}PRO to persuade t_i [PRO to read t_j]]]]



If one inspects the general movement pattern allowed under my analysis, one notices that in structures where two movements of the same kind (i.e. case-driven or whdriven) take place, their paths are either disjoint or, if they intersect, one must be nested within the other. In other words, legitimate patterns are either (150) or (151).

This is the consequence of 1) MLC and 2) lack of defective intervention. In other words, at the point when P_1 is merged in the structure in (151), it has to attract G_1 and not G_2 , by MLC. When P_2 is merged, it can attract G_2 because the relevant feature has been eliminated from G_1 . In A-movement cases, we have already seen that this pattern is illustrated in the structures in (34) such as the psych-construction repeated here in (152):



Here the movement of the experiencer DP to the Spec. of aspectual head (whereby its accusative Case is checked) is followed by the movement of the lower theme DP to the Spec. of T for the purpose of nominative Case checking.

In A'-movement, the crossing pattern is illustrated in the (a) examples in (145)-(149) and in structures such as (142a) in which the object DP has moved across the subject DP in the embedded Spec. of TP and which, as noted, many native speakers judge to be perfectly acceptable. In contrast, the ungrammatical (142b), which exhibits a crossing pattern, is ruled out straightforwardly as a Minimal Link Condition violation. This account carries over to the other structures. In all of the grammatical cases above, the closer eligible XP has been attracted by a c-commanding target, which was then followed by attraction of a lower XP by a higher c-commanding head. In other words, the grammatical sentences follow the pattern shown in (151). On the other hand, the ungrammatical examples involve the attraction of an XP across a closer eligible XP, in violation of the MLC.

To conclude this section, we have seen that my analysis offers a way not only to derive in part Pesetsky's (1982) Path Containment Condition (the requirement that intersecting A'-paths exhibit a nesting pattern) in the Minimalist framework, but in fact shows that the uniform pattern of nesting paths obtains for both kinds of movement.

5.2 Form Chain and Successive Cyclicity

As things stand at this point, it is not clear how my analysis makes a distinction between the grammatical argument extraction and the ungrammatical adjunct extraction across a wh-island. In particular, one ungrammatical structure that appears to be allowed is the extraction of adjunct wh-phrase across the embedded subject wh-phrase, as in (142c), repeated here as (153).

(153) *How_i did you wonder who_i t_i can fix your car t_i .
This is because, at the point that the wh-feature from the subject DP has been eliminated, there is nothing to prevent the attraction of the lower wh-phrase under my analysis. The problem is to find a solution that will allow (142a) and the grammatical sentences in (145)-(149)., and at the same time rule out (153).

Note that so far we have assumed that feature attraction and the movement of the host XP are part of a single process which moves the XP in one fell swoop to the target position. However, there seem to be rather strong arguments in favor of the (traditional) view that phrasal movement occurs in a more local, 'successive cyclic' fashion. Let us briefly review some of the prevalent arguments in the literature.⁹⁰

For A-movement, the main arguments for successive cyclicity, i.e. for the presence of intermediate landing sites, come from quantifier float phenomena and the binding theory. Following Sportiche (1988), a floating quantifier such as *all* in (154) is generally taken to mark the position of extraction of the associated XP such as *the students* which is taken to be base generated as forming a constituent with the quantifier.⁹¹

(154) The students_i seem (all) to be tired.⁹²

⁹⁰ The following sections rely heavily on the reviews given in Bošković (2001), Boeckx (2001b), and McCloskey (2001).

⁹¹ For a different view on floating quantifiers, see Bobaljik (1995). For recent arguments in support of Sportiche's approach, see McCloskey (2000), Bošković (2002).

 $^{^{92}}$ Of course, other options for the placement of *all* are also available for (154) (see example (162)). At this point, I am only recounting the arguments for the standardly assumed intermediate positions. The other positions at which *all* can occur should then follow from a particular formulation of locality of chain links, as I propose later in this section.

Under this assumption, the possibility of having *all* in the embedded SpecTP position in (154) provides evidence that the XP *students* passes through this position on its way from the VP internal subject position to the matrix SpecTP.

An argument for the presence of the intermediate SPEC-TP position is based on the binding facts reported in Castillo, Druri, and Grohmann (1999, p.19):⁹³

(155) a. Mary seems to John [TP to appear to herself/*himself to be in the room]

In (155), the DP *Mary* can bind a reflexive in the embedded TP, which is not possible for the DP *John*. How can this difference be accounted for? The data fall in place if the embedded TP contains a copy of the DP *Mary* in its specifier position, which it does if the movement of this DP proceeds in a successive cyclic fashion. In that case the copy of the DP *Mary* in the embedded SpecTP qualifies as an antecedent for *herself* in (155), and preventing the alternative on which John would be an antecedent to *himself*. If, on the other hand, the movement of the DP *Mary* occurred in one fell swoop, the inability of the *John* to bind the reflexive would be mysterious, given that the experiencer DP can otherwise be shown to c-command out of its containing PP and participate in the binding relations, as we have assumed to be the case in the discussion of the crossing structures such as (34a) above.

⁹³ It should be noted, though, that the authors, who attribute the data to David Pesetsky and Danny Fox, question the assumptions on which the above argument is based and, in fact, argue against the successive cyclic character of A-movement.

Our final argument comes from reconstruction facts originally observed by Lebeaux (1991). Consider the contrast between the pair of sentences in (156).

- (156) a. *[His_i mother's_j bread] seems to her_j _____ to be known by every man_i ______ to be the best there is.
 - b. [His_i mother's_j bread] seems to every man_i _____ to be known by her_j _____ to be the best there is.

The ungrammaticality of (156a) can be plausibly attributed to a condition C violation, on the coreferential reading of *her* and *his mother*. This is because, in order for the bound variable reading to be licensed, the subject DP has to be 'reconstructed'⁹⁴ into the most deeply embedded clause, where it is c-commanded by the universal quantifier. However, in this position it is also c-commanded by the DP *her*, which gives rise to a Condition C violation. On the other hand, the grammaticality of (156b) can be readily accounted for if the subject DP can be interpreted in the specifier position of the higher

⁹⁴ I am using the term 'reconstruction' here only to refer to the 'derivational history' of movement, without committing myself to the view under which the actual lowering of XPs take place. Whether interpretive access of arguments in intermediate positions takes place 'on-line' or under actual lowering of XPs is a matter of some controversy. For example, Lasnik (1999) argues that A-movement does not reconstruct, for which he provides some evidence (see pp. 21-21). See also Boeckx (2001) for arguments that some XPs, namely indefinites, do undergo lowering. I have nothing much to contribute to this interesting debate, except to point out that the data in (156) suggest rather strongly that arguments must be allowed to be interpreted in intermediate positions, which suggests that they have at some point occupied that position, and that is what the above given argument seeks to establish.

As for the contrast between the ungrammatical (155b) and the grammatical (156b), I take the former not to be the consequence of a forced reconstruction, but of fact that there is no point in the derivation of (155b) at which a proper binding configuration is established, while there is a point in (156b) at which both the Condition C and the configuration necessary for bound pronoun licensing obtains.

infinitival TP, a position where the bound variable reading can be licensed without inducing Condition C violation. But this is only possible if the subject DP passes through that position on its way to the matrix SpecTP, in other words if A-movement proceeds successive cyclically.

Similar arguments can be constructed for A'-movement. As pointed out by Barss $(1986)^{95}$, structures such as (157a) provide an empirical argument for intermediate landing sites, such as the one indicated as t_i ' in (85), since it is only in this position that the anaphor *himself* can be licensed by the antecedent *John*, on standard assumptions about Condition A.

(157) a. [Which picture of himself]_i does John think t_i that Mary likes t_i

b. *John thinks that Mary likes some pictures of himself.

That the anaphor *himself* is not licitly bound in its base-generated position is evident from the ungrammaticality of (157b), in which the relevant locality condition is not satisfied. But note that the intermediate landing site in (157a) will only be available if the movement of the wh-phrase is successive cyclic.

Another argument for the successive cyclicity of A'-movement has been standardly taken to be supported by 'agreeing complementizers' in languages in which the form of

140

⁹⁵ See also Huang (1995), Takahashi (1994).

(158) a. Creidim gu-r inis sé bréag.

I-believe go-Past tell he lie

'I believe that he told a lie.'

b. an t-ainm a hinnseadh dúinn a bhi – ar an áit
the name aL was-told to-us aL was on the place
'the name that we were told was on the place'

Finite complement clauses are normally introduced in Irish by various allomorphs of the particle *go*, as shown in (158a). However, finite clauses out of which A'-movement takes place are introduced by a different particle which is conventionally written aL, as in (158b). The interesting point here is that this morphological change affects not only the target, topmost complementizer whose specifier position represents the final landing site of the movement, but also all of the embedded complementizer positions intervening between the base generated position of the wh-phrase (or a null operator) and the topmost C. This fact has been standardly taken to reflect a local relation between the complementizers and the XP undergoing A'-movement, under the assumption that such morphological change requires a local relationship.⁹⁶

⁹⁶ McCloskey takes this paradigm to develop an analysis of feature-driven successive cyclicity along the lines of Chomsky (2000, 2001). However, Boeckx (2001b) shows that the paradigm is also compatible with Takahashi's (1994) approach to successive cyclic movement. See the discussion below.

Similar pieces of evidence come from other morphophonological reflexes of successive-cyclic wh-movement such at deletion of the verbal prefix *men*- in Malay/Bahasa Indonesia (Cole & Hermon 2000, Saddy 1991), tonal downstep in Kikuyu (Clements et al. 1983, Sabel 2000) or pronoun selection in Ewe (Collins 1993, 1997). Additional standard arguments in support of successive-cyclic *wh*-movement include the distribution of floating quantifiers in Irish English dialects (McCloskey 2000b,c), partial *wh*-movement in languages such as Quechua (Cole 1982) and Malay/Bahasa Indonesia (Saddy 1991, Cole & Hermon 2000), children's use of 'medial' wh-movement in long-distance questions such as *Who do you think who's in the box?* (De Villiers et al. 1990, McDaniel et al. 1995, Thornton 1990), and *wh*-copying in languages like German and Romani (Höhle 2000, McDaniel 1986, Reis 2000).

Given that the successive cyclic character of both types of movement seems to be well established, a question that arises is how to capture this fact in the current framework. Two approaches have been proposed in the recent literature: feature-driven approaches such as Chomsky's (2000, 20011) or McCloskey's (2000a), and 'shortest move' approaches such as those proposed by Takahashi (1994), Bošković (2001), or Boeckx (2001b). The latter seem to adopt in one way or another a version of Chomsky's (1995) proposal that a tension between two Economy principles governing displacement fewest steps in the derivation and shortest move - can be eliminated if one takes the movement to be implemented through the operation Form Chain under which the local chain links are formed in one step, though the movement itself is triggered by some feature requirement that must be satisfied at the head of the chain and is thus compliant with the Last Resort requirement. Thus, following Takahashi (1994) (who in turn develops the ideas of Chomsky and Lasnik's (1993) Minimize Chain Link Principle (MCLP)), Bošković (2001) proposes an analysis of successive cyclic movement as a process that results from the requirement that links in a chain be minimal. According to this proposal, an XP undergoing movement is required to pass through specifier position of each appropriate dominating head on its way to the ultimate landing site because such movement reflects an inherent property of chain formation.⁹⁷ In other words, that movement occurs in a successive cyclic fashion is not the result of EPP or some (other) feature checking operation taking place in the intermediate landing sites. Instead, feature checking is restricted only to the ultimate landing site, i.e. to the head whose uninterpretable feature triggers the movement/attraction.

Bošković (2001) offers several arguments in favor of his proposal over an EPP based account of successive cyclicity. In fact, following many recent studies (Boeckx 2000, Castillo, Drury, and Grohman 1999, Epstein and Seely 1999, Grohmann, Drury and Castillo 2000, Martin 2000), he seeks to eliminate EPP completely from the grammar. His arguments rely on the observation that, if EPP does not exist, then, the intermediate landing sites in A and A'-movement will only be projected if they are required by the MCLP, i.e. if they constitute intermediate steps in a movement to the final landing site

⁹⁷ Bošković follows Takahashi (1994) in relativizing the appropriate landing sites to the type of movement. However, as noted by Boeckx (2001b), the traditional A/A' division has no clear status in the current framework. Furthermore, since we have seen that the core cases of improper movement can be handled independently of the A/A' distinction, perhaps we can generalize Bošković's proposal and allow the movement to pass through the specifier position of every dominating head, as has been suggested in the literature, perhaps most 'closely' in Manzini (1994). This is what I assume below.

which is the only site in which feature checking takes place. Furthermore, no spec-head relationship is expected to obtain in the intermediate landing sites, since no feature checking takes place at these sites between the relevant head and the moved phrase. On the other hand, an EPP approach requires that positions such as Spec-TP always be projected since EPP is taken to be an inherent property of the functional head T. Bošković then goes to show that some structures receive a more principled explanation on his approach.⁹⁸

Adopting the ellipsis analyses by Lobeck (1990) and Saito and Muasugi (1990) under which functional heads can license ellipsis of their complements only when they undergo specifier-head agreement, Bošković points out that his approach offers a simple analysis of the contrast between control and ECM infinitivals with respect to ellipsis observed, among others, by Martin (1996): only the former allow the ellipsis of the complement of the relevant T head. This contrast is illustrated in (159), taken from Bošković (2001).

(159) a. John was not sure he could leave, but he tried PRO to.

b. *John believed Mary to know French, but Peter believed Jane to.

Bošković shows that the contrast is easy to explain under his approach, which denies any spec-head relationship between *Jane* and the embedded infinitival head *to* in (159b), since the movement of *Jane* through the embedded spec-TP is only required by MCLP and not

144

⁹⁸ For arguments in favor of Takahashi's over an EPP approach, see also Boeckx (2001b).

by any feature checking.⁹⁹ On the other hand, the control structure in (159a) involves a spec-head relation between PRO and the embedded T which is induced by the requirements of Case checking and which therefore licenses the elision of the complement.

The argument carries over directly to A'-movement. Bošković (2001) presents the sentence in (160)

(160) *John met someone but I don't know who i Peter said [CP ti [C C John met ti]].

The ungrammatical status of (160) follows if the movement through the embedded SPEC-CP does not involve feature checking. Under the alternative approach, the sentence in (160) is incorrectly predicted to pattern with (159a) instead of (159b).

Another argument for Bošković's approach is based on the structures in (161), which we have discussed in the section on expletive insertion.

(161) *There seems a man to be in the room.

Pointing to several weaknesses of Chomsky's Move-over-Merge principle¹⁰⁰, Bošković argues that the ungrammaticality of (161) falls out on his approach under which nothing requires the embedded spec-TP to be projected. That is, since there is no EPP, nothing requires that the expletive be merged in this position, which can then be merged

⁹⁹ Bošković claims that ECM constructions involve obligatory object shift to Spec-AgroP.

¹⁰⁰ On this, see also Vukić (1998, 1999), as well as Chapter 2 of this thesis.

directly in the matrix Spec-TP to enter a Case-checking relation. Furthermore, since there is no feature of the embedded T that would trigger the movement of *a man* to its specifier position, such a movement would violate the Last Resort condition.

I follow Bošković's (2001) and Boeckx's (2001) approach to successive cyclicity in this thesis. Specifically, I adopt the idea that, whereas feature attraction may cross an arbitrary number of nodes in a tree (all the way down to the closest matching feature), in the manner that I have suggested in the previous chapters, the pied piping of the relevant DP occurs in a strictly local fashion, by passing through a specifier position of every dominating head on its way to the ultimate landing site. If, indeed, quantifier float indicates the position which a subject DP has occupied at some point in the derivation of a sentence, then this approach receives support from constructions such as (162), from Sportiche (1988) and Bošković (2001).¹⁰¹

(162) The carpets (all) will (all) have (all) been (all) being (all) dusted for two hours.

We are now in the position to return to the ungrammatical sentence in (142c/153), repeated here as (163), which we saw creates a potential problem for our account.

(163) *How_i did you wonder who_i t_i can fix your car t_i .

¹⁰¹ See also Koizumi's data in (24).

Recall that, once the uninterpretable wh-feature on *who* is checked off against the matching feature of the embedded interrogative C, there seems to be nothing that would prevent the attraction of the lower wh-phrase *how*, and (163) is wrongly predicted to be grammatical.

Under the approach to successive cyclicity adopted above, the ungrammaticality of (163) would follow from a semantic filter such as the one proposed in Cresti (1995). Cresti examines an interplay between available interpretations and extraction configurations in sentences such as (164), originally discussed in Longobardi (1987).

(164) a. How many people should I talk to?

- b. How many people do you think I should talk to?
- c. How many people do you wonder whether I should talk to?
- d. I know which book you think that no student read
- e. I know which book you wonder whether no student read.

Following Longobardi (1987), Cresti claims that the sentences in (164a,b) are ambiguous between what she takes to be a wide scope and a narrow scope reading of *n many people* with respect to *should*, unlike the sentence in (164c) which only allows for the wide scope. Cresti takes the ambiguity to be a function of a relative position which the variable \mathcal{P} whose semantic type corresponds to the QNP *many people* occupies with respect to *should* in the tree. Thus, the lack of the narrow scope reading of *n many people* is taken to be the consequence of an impossibility to have \mathcal{P} in the embedded context in (164c). The analysis carries over to the pair in (164d-e), where the lack of the 'functional reading' in (164e) is attributed to the same condition.

Cresti then suggests that the exact point where the semantic 'reconstruction' (which in her system is accomplished without lowering, i.e. by λ -conversion and the possibility of shifting semantic types of 'traces' left by wh-phrase in the course of derivation) is blocked is the position of adjunction to CP, which in her system means that \mathcal{P} , or any variable except those of type e, cannot be available in that position. The existence of the adjoined position is taken to be supported by sentences such as (165), based on Barss (1986), in which the binding facts indicate that there must be a position above the filled SPEC-CP in which the anaphor is licensed.

(165) ?Which picture of himself is John wondering why his mother wants to enlarge?

Cresti then proposes the filter in (166)

(166) *[$_{CP}$ [$_{\delta}$ X] [$_{CP}$...]] where X is not of type *e*.

The filter bars adjunction of an XP to an embedded CP which constitutes the wh-island in cases where X is not interpreted as an expression of type e.

Going back to our example in (153), recall that the adjunction of *how* would be forced by MCLP, given that the specifier position of the CP is filled and given that the

movement must be via every maximal projection dominating the extraction site. Under the assumption that, in contrast to the trace/copy of the DP *which car* in structures such as (142a), the trace (or copy) of *how* cannot be interpreted as an expression of type e,¹⁰² the relevant part of the structure of (153) would exhibit exactly the barred configuration, as shown in (167), and the sentence would be filtered out by (166).

(167) *how ... [CP t_{how} [CP who [TP ... t_{who} t_{how} ...]]]

It should also be noted at this point that there are other semantics accounts of the ungrammaticality of (163) which do not depend on the successive cyclic character of movement. Thus, one could resort to proposals such as the one in from Kiss (1993) given in (168), or its development in terms of algebraic semantics by Szabolcsi and Zwarts (1993).¹⁰³

(168) If Op_i has scope over Op_j and binds a variable in the scope of Op_j, Op_i must be specific.

Given that wh-phases *how* and *who* are scope-taking operators, and given that, as can be seen in the abstract representation in (167), *why* has scope over *who* and binds a variable in the latter's domain, the filter in (168) requires *how* to be specific, which it clearly can

¹⁰² See Szalbocsi and Zwarts (1993) on the interpretation of *how* and similar adverbials.

¹⁰³ For an overview of various semantic and syntactic approaches to islands, see Szalbocsi and den Dikken (1999).

not be. Thus (142c) is also ruled out by (168).¹⁰⁴ We conclude that there appears to be substantial semantic grounds for the ill-formedness of (163), and that this structure therefore does not present a problem for the extension of our analysis to A'-movement.

To conclude this section, we have seen in this that, in addition to accounting for the previously discussed 'freezing' and improper movement, our proposal offers a straightforward and simple way to derive the core effects of Pesetsky's (1982) Path Containment Condition and to account for the basic cases of argument-adjunct asymmetries in extraction.

5.3 Agreement vs. Case

In this section, I address the relationship between agreement and Case. As noted, Chomsky (2000, 2001) considers structural Case to be a reflex of agreement relation between a probe and a goal. This proposal has some technical as well as conceptual problems. First, the assignment and disappearance of structural Case from narrow syntax is rather mysterious since the mechanism under which the uninterpretable Case features are deleted is unclear (in contrast to ϕ -features, uninterpretable Case-features do not disappear under Match with a corresponding valued feature). And second, if structural Case is but a reflex of agreement relation, then one would not expect to find the former without the latter in a language which exhibits overt agreement patterns. Yet this is what

¹⁰⁴ See also Szabolcsi and Zwarts (1993), for an algebraic approach in which structures (142c) are ruled out in terms of inability to perform the Boolean operation required by the lower wh-phrase in the domain determined by manner expressions such as *how*.

happens in Italian Aux-to-Comp constructions in which a nominative DP appears with a nonfinite form:¹⁰⁵

(169) Avendo io partita all'alba...

Having I left at dawn ... (from Giorgi and Pianesi (1997, p.64)

Based on examples such (169), Giorgi and Pianesi conclude that "relation between nominative and agreement holds only in one direction – that is, agreement implies nominative, but not vice versa" (p.64). In other words, nominative is possible without, hence independent from, agreement, though not vice versa.

Parallel evidence comes from Inuit finite clauses, exemplified in (170), where in the embedded clause agreement obtains with the ergative subject in Spec-vP and the nominative object in Spec-IP.

(170) Juuna-p miqqat taku-ga-mi-git

nuannaar-p-u-q.

[Juunai-ERGi children-NOMj see-DPST-3SG.PROXi-3PLj] pro-NOMi happy-IND-[-tr]-3Sgi

'Because Juuna_i has seen the children, he_i's happy.' (Bittner and Hale 1996a, ex.31b)

¹⁰⁵ I am not assuming, of course, that the presence of agreement must always be overtly manifested in agreement morphology. That is, it seems reasonable to assume that agreement with the verb obtains in sentences such as *They walk.*, where it is not overtly manifested. I take it that even an impoverished inflectional paradigm in English suffices to establish that finite forms exhibit agreement. The point about the Italian example is that, given that the language has rich overt morphological agreement which is consistently lacking in non-finite forms, it is plausible to assume that the nonfinite forms are non-agreeing.

However, in infinitival clauses, the ergative Case on the subject is still present, regardless of the fact that the corresponding agreement with the verb is lost. This is shown in (171):

(171) Juuna-p miqqat taku-llu-git __ nuannaar-p-u-q.

Juuna-ERG_i children_j see-INF-3PL_j] pro happy-IND-[-tr]-3SG_i 'Seeing the children, Juuna was happy'

Bittner and Hale (1996a) conclude that "...agreement is independent of structural Case."(p.17) and that "...structural Case may be preserved while agreement is lost."(p.17). Furthermore, from their theory, in which predictable oblique cases such as dative or instrumental are also treated as structural, they conclude that "structural Case can be assigned even to arguments that never control agreement."(p18).

Finally, something needs to be said about the lack of agreement in Icelandic structures such as (98b), repeated here as (172), in which the dative experiencer has moved to the matrix SpecTP position.

(172) Hafði peim_i virst t_i [Olafur vera gáfaður]

had them seemed Olaf intelligent 'Did it seem to them that Olaf was intelligent. Under Chomsky's set of assumptions, it has to be assumed that the movement of the experiencer DP follows after the operation *Agree* has taken place between the ϕ -features of T and those of the DP. But then one needs to explain why such agreement is not manifested on the verb which either agrees with the (embedded) nominative, as in (172), or takes a default form if a nominative DP is not available, as in (173)

(173) Stepunum var hjálpad

The girls_{DAT-PL-FEM} was_{3SG} helped_{NEUT,SG}

'The girls were helped"

To account for the lack of manifest agreement with the experiencer DP in (172), Chomsky would have to assume that dative Case on the DP suppresses its ϕ -features. But if ϕ -features are suppressed, then how is *Agree* established between the DP and T? To find a way out of this paradox, one would have to divorce the operation *Agree* from overt agreement facts, a highly undesirable result in Chomsky's system. On the other hand, if Amovement is Case-based and independent from agreement, than the data above do not pose a problem. Since on such an account, the movement of the experiencer DP in (172) is triggered by the latter's Case feature that is not tied to agreement, it is not surprising that agreement can obtain with another DP under an independent relation, or as a 'default' 3rd person in (173).

Chapter 6

Conclusion

In this study, we have investigated the phenomenon of displacement in an attempt to properly identify its triggers and to characterize the locality constraints governing this crucial property of human language. We have shown that the core movement paradigm falls in place if

- only uninterpretable features play a role in the checking theory, and
- expletives are inserted directly from the lexicon as a last resort.

These assumptions make it possible to account for the main cases and at the same time avoid problems that analyses which assume that it is the interpretable features that are attracted are faced with. The problems involved undergeneration of a number of grammatical constructions whose derivation must be taken to involve raising of one DP across another c-commanding DP to a higher position in the structure. we have shown that such crossing receives a natural account under our proposal. Similarly, the ungrammatical construction involving the 'freezing' structural environments receive a straightforward explanation, in contrast to the alternative proposal on which they are wrongly predicted to be grammatical. My proposal offers a number of conceptual advantages. On the proposed analysis, there is no need to stipulate the suppressive role of inherent Case, an assumption that otherwise requires a whole theory of interrelations between features from which such a suppressive role would naturally follow. In addition, I have shown that resorting to the suppressive role of inherent Case, which is unavoidable in the accounts of locality violation which assume that Attract targets interpretable features, has a number of empirical problems that do not arise on my account

Similarly, my proposal eliminates the need for other complex notions, such as the one involving the 'activating' role of uninterpretable features proposed in Chomsky (2000, 2001) or the mechanism of feature licensing proposed in Vukić (1997, 1998, 1999b). On the analysis argued for in this study, the activating role of uninterpretable features follows directly, given that these are the only features that play a role in the operation *Attract*. 'Defective intervention' has also been eliminated from the system without extra stipulations, i.e. as a consequence of isolating uninterpretable features as the only possible attractees in the movement theory.

As a direct consequence of my proposal, the determination of *Closeness* is simplified in that it involves the simple computation of c-command. The complex mechanism of equidistance, which has been shown to be inadequate on both conceptual and empirical grounds, is eliminated in favor of the simpler and more restrictive version.

All these positive results indicate that the proposed analysis is on the right track. And to the extent that this is the Case, the study supports Minimalism as a productive framework for the study of human language.

155

REFERENCES

- Aoun J. and A. Li. 1989. Constituency and Scope. Linguistic Inquiry 20:141-72
- Baker, M. Incorporation: A Theory of Grammatical Functions Changing.. Chicago and London: University of Chicago Press
- Baker, M. 1997. Thematic Roles and Syntactic Structure. In L. Haegeman (ed.) *Elements* of Grammar, Kluwer
- Baltin, Mark (2000) "Implications of Inner VP-Ellipsis for Binding and The Representation of Information Structure", unpublished ms., New York University
- Barss, A. 1986. Chains and Anaphoric Dependence. Doctooral dissertation, MIT.
- Barss, A and H. Lasnik. 1986 A note on anaphora and double objects. *Linguistic Inquiry* 17:347-354.
- Belletti, A. 1988. The Case of Unaccusatives. Linguistic Inquiry 19:1-34.
- Belletti, A. and L. Rizzi (1988) Psych Verbs and Theta-Theory. *Natural Language and Linguistic Theory* 6:291-352.
- Bennis, H. 2002. Unergative adjective and psych verbs. to appear in A. Alexiadou, E. Anagnostopoulou & M. Everaert (eds.) Unaccusativity. Oxford University Press.
- Bittner, M. and K. Hale. 1996. The Structural Determination of Case and Agreement. Linguistic Inquiry 27:1-68
- Bittner, M. and K. Hale. 1996. Ergativity: Toward a Theory of Heterogeneous Class. Linguistic Inquiry 27:531-604.
- Bobaljik. J.D. 1995. *Morphosyntax: The Syntax of Verbal Inflection*. Doctoral dissertaion, Cambridge: MIT.
- Boeckx, C. 1998. Quirky Agreement in Icelandic and Elsewhere. Working papers in Scandinavian Syntax 61, pp.1-35
- Boeckx, C. 1999a. EPP eliminated., Ms. University of Connecticut, Storrs.
- Boeckx, C. 1999b. Conflicting C-command Requirements, *Studia Linguistica* (53(3): 227-250.

- Boeckx, C. 2001a. Arguments reconstructed. *Natural Language and Linguistic Theory* 19:503-548.
- Boeckx, C. 2001b. *Mechanisms of Chain Formation*. Doctoral dissertation, University of Connecticut, Storrs.
- Boeckx, C. 2001. Scope Reconstruction and A-Movement. *Natural Language & Linguistic Theory* 19: 503–548
- Boeckx, C and S. Stjepanović. 2001. Head-ing toward PF. *Linguistic Inquiry* 32: 345-355.
- Bošković, Ž. 1994. D-structure, theta-criterion, and movement into theta-positions. *Linguistic Analysis* 24, 247-286.
- Bošković, Ž. 1995a. Principles of Economy in Nonfinite Complementation. Doctoral Dissertation, University of Connecticut.
- Bošković, Ž. 1995b. Case properties of clauses and the Greed Principle. *Studia* Linguistica 49: 32-53.
- Bošković, Ž. 1997. The Syntax of Nonfinite Complementation. MIT Press.
- Bošković, Ž. 2001. A-Movement and the EPP, Ms. University of Connecticut, Storrs.
- Bošković, Ž. 2002. Be careful where you float your quantifiers, Ms. University of Connecticut, Storrs.
- Bobaljik, Jonathan D. (1993) "On Ergativity and Ergative Unergatives." in Colin Phillips, ed., *MIT Working Papers in Linguistics 19: Papers on Case and Agreement II.* 45–88.
- Bouchard, D. 1995. The Semantics of Syntax. Chicago: University of Chicago Press
- Bresnan, J. and J. Kanerva. 1989 Locative Inversion of Chichewa: A Case Study of Factorization in Grammar. *Linguistic Inquiry* 15:355-380.
- Bresnan, J. 1991. Locative case vs. locative gender. In *Proceedings of the Seventeenth* Annual Meeting of BLS, 53-68. Berkeley, Calif.: Berkeley Linguistics Society.
- Castillo, J.C., J.E. Drury, and K. Grohmann. 1999. The status of Merge over Move preference. Ms., University of Maryland, College Park.

Chomsky, N. 1957. Syntactic Structures. Mouton & Co.

158

Chomsky, N. 1981. Lectures on Government and Binding. Dordrecht: Foris Publications.

- Chomsky, N. 1986a. Barriers. MIT Press.
- Chomsky, N. 1986b. Knowledge of Language: Its Nature, Origin and Use. New York: Praeger.
- Chomsky, N. 1993. A Minimalist Program for Linguistic Theory. In *The View from Building 20: Essays in Honor of Sylvain Bromberger*, ed. Kenneth Hale and Samuel Jay Keyser, 1-52. MIT Press.
- Chomsky, N. 1995 The Minimalist Program. MIT Press.
- Chomsky, N. 2000. Minimalist Inquires: The Framework. Ms. MIT.
- Chomsky, N. 2001. Derivation by Phase. In *MIT Occasional Papers in Linguistics* 18. MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, Mass.
- Chomsky, N and H. Lasnik. 1993. The theory of principles and parameters. In J. Jacobs,
 A. voon Stechow, W. Sternfeld, and T. Vennemann (eds.). Syntax: An international handbook of contemporary research, Vol. 1, 506-569, Berlin: Walter de Gruyter.
- Cinque, G. Types of A'-Dependencies. MIT Press
- Clements, G., J. McCloskey, J. Maling and A. Zaenen. 1983. String-vacuous rule application. *Linguistic Inquiry* 14: 1-17.
- Cole, P. 1982. Imbabura Quechua. The Hague: North Holland Publishing Company.
- Cole, P. and G. Hermon. 2000. Partial wh-movement: Evidence from Malay. in U. Lutz, G. Müller, and A. von Stechow (eds.). *Wh-Scope Marking*, 101-130, Amsterdam & Philadelphia: John Benjamins
- Collins, C. 1993. Topics in Ewe Syntax. Doctoral Dissertation, MIT
- Collins, C. 1997. Local Economy, MIT Press
- Collins, C. and H. Thráinsson. 1996. VP-Internal Structure and Object Shift in Icelandic. *Linguistic Inquiry* 27:391-444.

Conway, L. 1996. NPs in disguise. Ms., University of Connecticut, Storrs.

- Cowper, E. 1988. What is a subject? Non-nominative subjects in Icelandic. In *Proceedings* of the North East Linguistic Society 18, 94-108. Amherst, Mass.: GLSA.
- Cresti, D. 1995. Extraction and Reconstruction. *Natural Language Semantics* 3:79-122.
- De Villiers, J., T. Roeper, and A. Vainikka. 1990. The acquisition of long-distance rules. in L. Frazier and J. de Villiers (eds.). *Language Processing and Language Acquisition*, 257-297. Dordrecht: Kluwer,
- den Dikken, M. 1995. Binding, Expletives and Levels. Linguistic Inquiry 26:347-353.
- Epstein, S. 1992. Derivational Constraints on A'-Chain Formation. *Linguistic Inquiry* 23:235-260.
- Epstein, S. and T.D. Seely. 1999. SPEC-ifying the GF "subject"; eliminating A-chains and the EPP within a derivational model. Ms., University of Michigan and Eastern Michigan State University.
- Featherston, S. (2003) Universals and grammaticality: Wh-constraints in German and English, Ms. Eberhard-Karls-Universität, Tübingen
- Frampton, J. 1997. Expletive Insertion. In *The role of Economy Principles in Linguistic Theory*, eds. Chris Wilder, Hans-Martin Gärtner and Martin Bierwisch. Academie Verlag.
- Frampton, J. and S, Gutmann. 1997. Eliminating Non-Local Computation in Minimalist Syntax. Ms. North-Eastern University, Boston.
- Franks, S. 1995. Parameters of Slavic Morphosyntax. Oxford University Press.
- Friedin, R. & R. Sprouse. 1991. Lexical case phenomena. In *Principles and parameters in comparative grammar*, ed. R. Freidin, 392-416. Cambridge, Mass.: MIT Press.
- Giorgi, A. and F. Pianesi. 1997. Tense and Aspect. Oxford University Press.
- Grewendorf, G. 2001. Multiple Wh-Fronting. Linguistic Inquiry 32: 87-122.

Grimshaw, J. 1991. Argument Structure. MIT Press.

Grohmann, K., J. Drury, and J.C. Castillo. 2000. No more EPP. In *The Proceedingss oof the Nineteenth West Coast Conference on Formal Linguistics*, 139-152. Somerville, Mass.: Casscadilla Press.

- Höhle, T. 2000. 'The *w...w* construction: Appositive or scope indicating?', in U. Lutz, G. Müller and A. von Stechow (eds.), 249-270.
- Holmberg C. and A. Platzack. 1995. *The Role of Inflection in Scandinavian Syntax*. Oxford University Press, Oxford.
- Holmberg, A. and T. Hróarsdóttir. 2002. Agreement and Movement in Icelandic raising Constructions., Ms. University of Durham and University of Tromsø.
- Hornstein, N. 1999 Movement and Control. Linguistic Inquiry 30:69-96.
- Huang, J. 1982. Logical relations in Chinese and the theory of grammar. Doctoral dissertation, MIT.
- Huang, J. 1993. Reconstruction and the structure of VP: Some theoretical consequences. *Linguistic Inquiry* 24:103-138.
- Jackendoff, R. 1972. Semantic Interpretation in Generative Grammar. MIT Press, Cambridge, Massachusetts.
- Jakobson, R. 1936/1962. Beitrag zur allgemeinen Kasuslehre: Gesamtbedeutungen der russischen Kasus. Travaux de Cercle Linguistique de Prague VI 240-299. Reprinted as 'Contribution to the general theory of case: General meanings of the Russian cases' in Waugh and Halle 1984: 50-103
- Jayaseelan, K.A. 1990 Incomplete VP-deletion and Gapping. *Linguistic Analysis* 20:64-81.
- Johnson, K. 1991. Object Positions. Natural Language and Linguistic Theory 9:577-636.
- Kennedy, C. (to appera) Ellipsis and Syntactic Representation. To appear in Schwabe, K. and S. Winkler, (eds.), *The Syntax-Semantics Interface: Interpreting (Omitted) Structure*, John Benjamins, Amsterdam.
- Kiparsky, P. 1987. Morphology and Grammatical Relations. Ms., Stanford University.
- Kiss, É. Katalin. 1992. Wh-movement and specificity. *Natural Language and Linguistic Theory* 11: 85-120.
- Kitagawa, Y. 1994 Shells, Yolks and Scrambled Eggs. in *Proceedings of NELS 24*. University of Massachusetts, Amherst.

Kitahara H. 1997. Elementary Operations and Optimal derivations, MIT Press.

Koizumi, M. 1995. Phrase Structure in Minimalist Syntax. Doctoral Dissertation, MIT.

Larson, R. 1988. On the Double Object Construction. *Linguistic Inquiry* 19:335-391

- Lasnik, H. 1992. Case and Expletives: Notes Towards a Parametric Account. *Linguistic Inquiry* 23:381-405.
- Lasnik H. 1995a. Case and expletives revisited: On Greed and other human failings. Linguistic Inquiry 26:615-633
- Lasnik, H. 1995b. Last Resort and Attract F. In Papers from the 6th Annual Meeting of the Formal Linguistics Society of Mid-America, Vol. 1, pp. 62-81.
- Lasnik, H. 1995c. A note on pseudogapping. In *Papers on Minimalist Syntax, MIT Working Papers in Linguistics 27*, ed. Rob Pensalfini and Hiroyuki Ura, 143-163. MITWPL, Department of Linguistics and Philosophy, MIT.
- Lasnik, H. 1999a Pseudogapping puzzles. In E. Benmamoun & H. Lappin (eds) Fragments: Studies in Ellipsis and Gapping, 141-174. Oxford: Oxford University Press.
- Lasnik, H. 1999b Chains of Arguments. in *Working Minimalism.* ed. Samuel Epstein and Norbert Hornstein, 185-215, MIT Press.
- Lasnik, H. 2001a. Subjects, Objects, and the EPP. In W. Davies and S. Dubinski (eds.). *Objects and Other Subjects: Grammatical Functions, Functional Categories, and Configurationality*, Kluwer Academic Publishers, Dordrecht

Lasnik, H. 2001b. A note on the EPP. *Linguistic Inquiry* 3: 356-362.

- Lasnik, H. to appear . Feature movement or agreement at a distance? In *Remnant movement*, *F-movement and the T-model*, A. Alexiadou, E. Anagnostopoulou, S. Barbiers and H.-M. Gaertner (eds). Amsterdam: John Benjamins.
- Lasnik, H. and M. Saito. 1991. On the subject of infinitives. In *CLS 27*. Part 1, *The General Session*, ed. By L. M. Dobrin, L. Nichols, and R. M. Rodriguez, 324-343, Chicago Linguistic Society, University of Chicago, Chicago, Ill.
- Lasnik, H. and M. Saito. 1992. Move a, MIT Press.
- Lebeaux, D. 1991. Relative clauses, licensing and the nature of the derivation. In *Syntax* and Semantics 25: Perspectives on phrase structure, ed. S. Rothstein 209-239. New York: Academic Press.

- Lebeaux, D. 1998. Where Does the Binding Theory Apply. Technical Report 98-044, NEC Research Institute.
- Lobeck, A. 1990. Functional heads as proper governors. In *Proceedings of NELS* 20:348-362. GLSA, University of Massachusetts, Amherst.
- Lutz, U., G. Müller, and A. von Stechow (eds.). 2000. *Wh-Scope Marking*. Amsterdam & Philadelphia: John Benjamins.
- Manzini, R. 1994 Locality. MIT Press
- Manzini, R. 1994. Locality, minimalism, and parasitic gaps. *Linguistic Inquiry* 25:481-508.
- Marantz, A. 1984. On the Nature of Grammatical Relations. MIT Press
- Marantz, A. 1992. Implications of Asymmetries in Double Object Constructions. In *Theoretical Aspects of Bantu Grammar.*, ed. by Sam A. Mchombo. Stanford, CA:CSLI Publications
- Martin, R. 1996. *A Minimalist Theory of PRO and Control*. Doctoral dissertation, University of Connecticut, Storrs
- Martin, R. 1999. Case, the Extended Projection Principle, and minimalism. In *Working Minimalism*, ed. S. Epstein and N. Hornstein, 1-25. Cambridge: MIT press
- Massey, V. 1990. Experiencers, Themes and C-Command in Albanian. in *MIT Working Papers in Linguistics.* ed. by T. Green and A. Uziel. Vol 12
- May, R. 1985. Logical Form: Its Structure and Derivation. MIT Press.
- McDaniel, D. 1986. *Conditions on wh-chains*. PhD dissertation, City University of New York.
- McDaniel, D. 1989. Partial and multiple *wh*-movement. *Natural Language and Linguistic Theory* 7: 565-604.
- McDaniel, D., B. Chiu, and T. Maxfield. 1995. Parameters for wh-movement types: Evidence from child Englis. *Natural Language and Linguistic Theory* 13, 709-753.
- McCloskey, J. 2000a. Resumption, successive cyclicity, and the locality of operations. Ms., to appear in S. Epstein and D. Seeley (eds.), *Prospects for Derivational Explanation*. Blackwell.

- McCloskey, J. 2000b. Quantifier float and *wh*-movement in an Irish English, *Linguistic Inquiry* 31, 57-84.
- McCloskey, J. 2000c. The prosody of quantifier stranding under *WH*-movement in West Ulster English. Ms. University of California, Santa Cruz.
- Murasugi, K.G. 1992. Crossing and Nested Paths: NP Movement in Accusative and Ergative Languages. Doctoral Dissertation, MIT
- McGinnis, M. J. 1996. Equidistance and Binding. *Proceedings of the 1996 CLA Annual Conference. Calgary Working Papers in Linguistics*, 257-268. Department of Linguistics, University of Calgary.
- Miyagawa, S. 1993. LF Case Checking and Minimal Link Condition. in In *Case and Agreement II, MIT Working Papers in Linguistics 19*, 213-254. Department of Linguistics and Philosophy, MIT, Cambridge, Mass.
- Ochi, M. 1999. Constraints on Feature Checking, Doctoral Dissertation, University of Connecticut
- Perlmutter, D. and P. Postal. 1984. The 1-Advancement Exclusiveness Law. in *Studies in Relational Grammar.* ed. by D. Perlmutter and C. Rosen. Chicago: University of Chicago Press.
- Pesetsky, D. 1982 Paths and Categories. Doctoral Dissertation. MIT
- Pesetsky, D. 1995. Zero Syntax. MIT Press.
- Postal, P. On Raising: One Rule of English Grammar and Its Theoretical Implications. MIT press.
- Richards, N. III 1997. What Moves Where When in Which Language. Doctoral Dissertation, MIT
- Reinhart, T. 1976. The Syntactic Domain of Anaphora. MIT dissertation
- Reinhart, T. 1996. Interface Strategies. *OTS Working Paper*. Utrecht: Utrecht Institute of Linguistics
- Reis, M. 2000. On the parenthetical features of German was...w constructions and how to account for them. in U. Lutz, G. Müller and A. von Stechow (eds.), 359-407.

Rizzi, L. 1990. Relativized Minimality. MIT Press.

- Rizzi, L. 1996. Residual Verb Second and the Wh-Criterion. In A. Beletti and L. Rizzi, eds., *Parameters and Functional Heads*. Oxford University Press, New York.
- Saddy, D. 1991. Wh scope mechanisms in Bahasa Indonesia. in L. Cheng and H. Demirdache (eds.) *MIT Working Papers in Linguistics* 15, 183-218. Cambridge, MA: MITWPL
- Saito, M. 1992. Long-Distance Scrambling in Japanese. *Journal of East-Asian Linguistics* 1:69-118.
- Saito, M. and K. Murasugi. 1990. N'-deletion in Japanese. In University of Connecticut Working Papers in Linguistics 3. Ed. J Ormazabal and C. Tenny, 87-108. University of Connecticut, Storrs.
- Schütze, C. and K. Wexler. 1996. Subject Case Licensing and English Root Infinitives. In *Proceedings of BUCLD20*, v2: 670-681.
- Sigurðsson, H. 1988. NP-Movement with Special reference to Icelandic. In Groninger Arbeiten zur Germanistischen Linguistik. 29:1-36
- Sigurðsson, H. 1992. The Case of Quirky Subjects. Working Papers in Scandinavian Syntax.49:1-26
- Sigurðsson, H. 1996. Icelandic Finite Verb Agreement. Working Papers in Scandinavian Syntax.57:1-46
- Snyder, W. 1992 Chain Formation and Crossover. Ms., MIT
- Sportiche, D. 1988. A theory of floating quantifiers and its corrolaries for constituent structure. *Linguistic Inquiry* 19:425-449
- Stepanov, A. to appear. Ergativity, Case and the Minimal Link Condition. *In Minimality effects in syntax*, Ed. Arthur Stepanov, Gisbert Fanselow and Ralf Vogel. Berlin: Mouton de Gruyter.
- Svenonius, P. (2002). Case is Uninterpretable Aspect. To appear in *Proceedings of Perspectives on Aspect*, Utrecht 2002
- Szabolcsi, A. and Frans Zwarts. 1993. Weak islands and an algebraic semantics of scope taking. *Natural Language Semantics* 1: 235-284.

- Szabolcsi, A. and M. den Dikken. 1999. Islands. *GLOT International* 4. To be reprinted in *The Second Glot International State-of-the-Article Book*, ed. by Cheng and Sybesma. Mouton--de Gruyter
- Takahashi, D. 1994. Minimality of Movement. Doctoral Dissertation, University of Connecticut.
- Thornton, R. 1990. Adventures in long-distance moving: The acquisition of complex whquestions. PhD dissertation, University of Connecticut, Storrs.
- Travis, L. 1991. Derived Objects, Inner Aspect and the Structure of VP. Paper presented in NELS 22
- Uriagereka, U. and E. Raposo. 1995. Indefinite Se. Natural Language and Linguistic Theory 14:749-810.
- Ura, H. 1994. Varieties of Raising and the Feature-Based Phrase Structure Theory. *MIT* Occasional Papers in Linguistics. 7. MTWPL, Department of Linguistics and Philosophy, MIT.
- Ura. H. 1996. *Multiple Feature Checking: The Theory of Grammatical Function Splitting*, doctoral dissertation, MIT.
- Vukić, S. 1997 Backward Binding and Minimal Link Condition. Ms. University of Connecticut
- Vukić, S. 1998 feature Dependence and MLC. Syntax Generals Examination paper, University of Connecticut
- Vukić, S. 1999a On Attract F and the Minimal Link Condition. *Linguistic Analysis* 28: 185-226.
- Vukić, S., 1999b On Case and the Minimal Link Condition, *Proceedings of NELS* 29 v2:177-192
- Waugh, L. and M. Halle. 1984. Roman Jakobson: Russian and Slavic Grammar. Number 106 in Janua Linguarum. Berlin: Mouton
- Wiltschko M. 1997. D-linking, scrambling and superiority in German. Groninger Arbeiten zur Germanischen Linguistik 41, 107-142
- Zaenen, A., J. Maling, and H. Thráinsson. 1985. Case and grammatical functions: the Icelandic passive. *Natural Language and Linguistic Theory* 3, 441-483.