Movement and Optionality in Syntax

by

Anna Margaret Pettiward

Thesis submitted in partial fulfilment
of the requirements for the degree of Doctor of Philosophy

School of Oriental and African Studies
University of London
20th December 1996

Abstract

This thesis concerns itself with the core syntactic phenomenon traditionally thought of within Principles and Parameters approaches in terms of *movement*. The point of departure is the observation that in two important respects, the characterization of this phenomenon in the recent Minimalist model of grammar (Chomsky 1993, 1995) (in terms of the operation *Move*) seems to fall short of that in the earlier GB model (Chomsky 1981, 1986) (in terms of the rule *Move-α*): first, the notion that movement operations apply freely seems impossible to maintain - a theoretical inadequacy; second, there is no obvious way of dealing with "optionality" phenomena - an empirical inadequacy. This thesis argues, however, that these apparent serious inadequacies of the Minimalist framework are in fact principledly soluble, and crucially without reverting to a GB-type model.The thesis falls into two parts, corresponding to the theoretical and empirical problems noted above.

The central proposal of Part I is the Copy Hypothesis (Chapter 2): "all copies in a chain are active in the computational system". The relevance of this proposal is that, due ultimately to very fundamental properties of the standard Minimalist model, it actually appears impossible to maintain the notion that movement operations apply to any element - contrary to the Copy Hypothesis. However, I show how general conditions on movement are in fact sufficient to properly regulate the activity of traces, and give detailed arguments against Chomsky's (1995) proposal that "trace is immobile". Further to this, I show that the Copy Hypothesis has empirical applications involving the behaviour of *wh*-objects and associates of *there* in English, as well as computational complexity implications (Chapters 3 and 4).

The Copy Hypothesis of Chapter 2 goes on to play an important role in Part II, in which I takes up the topic of optionality. The economy principles assumed to constrain derivations (in particular *Last Resort*) seem to exclude the possibility of optionality within the computational system. Since there is a certain amount of data which do appear to involve such optionality, the Minimalist framework evidently faces a major empirical problem, again seeming to lose out earlier models in which optionality data could be characterized simply in terms of optional application of Move-α. In Chapters 5 and 6, I show that there is in fact scope for some syntactic optionality within the derivational economy system. A system is developed whereby economy conditions in conjunction with feature properties of lexical items can derive variation in the timing of movement relative to Spell-Out. In this way, I account for optionality data (plus associated non-optionality effects) from French (optionality of participle agreement), English and Swedish (optional partial associate-movement with non-Case/agreement-checking expletives *there* and *det* 'it'), Icelandic, German and Dutch (optional overt Object Shift).

Contents

Abstract 2
Table of Contents 3
Acknowledgements 7

Chapter 1 Introduction 8

1 Introduction 8
2 The Minimalist hypothesis 9
3 The Minimalist model 11
3.1 Levels of representation 12
3.2 Phrase structure 14
3.3 Formatives and clause structure 17
4 Principles of economy 19
4.1 Economy of Representation 19
4.1.1 Movement, features and checking 21
4.1.2 Overt movement 24
4.2 Derivational economy and other conditions on movement 26
4.2.1 Last Resort and Greed 26
4.2.2 Procrastinate 29
4.2.3 Minimal Link Condition 31
5 Summary and thesis preview 37

PART I 40

Chapter 2 Copies and derivations 41

1 Introduction 41
2 The copy theory of movement and its consequences 43
2.1 Full versus empty categories 43
2.2 Copies at LF 45
2.3 Copies in C_{HL} 50
3 Arguments for the Copy Hypothesis 52
3.1 Chomsky's (1995) theory of movement 52
3.2 The independent constrainedness of trace activity 57
4 Potential arguments against the Copy Hypothesis 63
4.1 Chomsky's (1995) proposal 63
4.2 Problems with Chomsky's (1995) proposal 64
4.3 Section summary 70
5 Movement to copies 71
6 Copies, derivations, and representations 75
7 Summary and conclusion 77

Chapter 3 When does the English object shift? 78

1 Introduction 78
2 Current views on objects in English 80
2.1 Preamble 80
2.2 The standard object phenomena 82
2.3 Chomsky 1995 86
2.4 Koizumi 1993/1995 88
3 Arguments and counterarguments for (c)overt Object Shift 91
3.1 Adverb construal arguments 93
3.1.1 For overt Object Shift 93
3.1.2 Counterargument 94
3.1.3 For covert Object Shift 96
3.1.4 Counterargument 99
3.2 Particle construction argument 100
3.2.1 For overt Object Shift 100
3.2.2 Counterargument 101
3.3 Deletion argument 104
3.3.1 For overt Object Shift 104
3.3.2 Counterargument 105
3.4 Extraction argument 107
3.4.1 For covert Object Shift 107
3.4.2 Counterargument 107
3.5 Floating Quantifier arguments 108
3.5.1 For overt Object Shift 108
3.5.2 Counterargument 109
3.5.3 For covert Object Shift 110
3.5.4 Counterargument 111
3.6 Section summary 113
4 On the relevance of "Case Adjacency" effects 114
4.1 Case, Adjacency, and a Misconception 115
4.1.1 Not only XPs needing Case are subject to Adjacency effects 116
4.1.2 Not all XPs needing Case are subject to Adjacency effects . . 118
5 Summary and conclusion 121

Chapter 4 The Case of *wh*- objects and associates 122

1 Introduction 122
2 *Wh*- objects and associates of *there* 124
2.1 Theoretical questions 124
2.2 Proposal 126
2.3 *Wh*- phrase associates of *there* 131
2.4 Comparison with Chomsky1995 133
3 Floating Quantifiers and *wh*- objects 135
3.1 Sportiche's (1988) movement analysis 135
3.2 *Wh*- objects and FQs in English (and other languages) 142

3.3 Potential counterarguments addressed 142
3.3.1 Are FQs *not* stranded by movement? 142
3.3.2 Are FQs *never* licensed by A'-movement? 145
3.3.3 FQs and A'-movement in French 148
4 Approaching extraction asymmetries 150
4.1 A Minimalist dilemma 150
4.2 Kitahara 1994 152
4.3 Extracting associates of *there* 153
4.4 Extraction asymmetries among objects: German, Dutch and French . . 157
5 The Copy Hypothesis, economy and complexity 163
5.1 Preamble 163
5.2 Procrastinate, necessary and unnecessary degrees of violability 164
6 Summary and conclusion 167

Part II 169

Chapter 5 Syntactic optionality in theory and practice 170

1 Introduction 170
2 Optionality is a Minimalist issue 172
2.1 An informal discussion of syntactic optionality 173
2.2 Scope for syntactic optionality in the Minimalist system 178
3 Optional agreement in French as optional overt Object Shift 179
3.1 The classical analysis: Kayne 1989 180
3.2 Arguments for a retaining a syntactic approach 181
3.2.1 Syntactic, lexical and dialectal approaches 181
3.2.2 The syntactic roots of phonologically overt agreement 183
3.3 Characterizing the options 191
3.3.1 Preliminary assumptions 191
3.3.2 Derivations 193
3.3.2.1 *Wh*-movement 193
3.3.2.2 Accusative clitics 197
3.4 Shortest Derivations, Form Chain and anti-Procrastinate effects 200
4 When agreement is not optional in French 210
4.1 Two situations where agreement is impossible 210
4.2 Obligatory and impossible agreement in *être* environments 211
4.3 Kayne, expletives and agreement: a short comparison 221
5 Expletives and optional associate movement in English and Swedish 222
6 Summary and conclusion 229

Chapter 6 Extensions 230

1 Introduction 230
2 Object Shift (non)optionalities in SVO and SOV Germanic 231
2.1 Icelandic 231
2.1.1 Data and background assumptions 231

2.1.2 Derivations and anti-Procrastinate effects 234
2.1.3 "Modular economy" and other issues 240
2.1.4 When Overt OS is not an option in Icelandic: verbs *in situ* . . 245
2.2 Optional overt Object Shift in German and Dutch 247
2.2.1 Data 247
2.2.2 An old-fashioned solution: "covert" overt V-raising 249
2.2.3 Why optionality persists with non-finite verbs 256
2.2.4 The Germanic-French difference 258
2.3 Section summary 260
3 Three comparisons 262
3.1 Kitahara 1994 262
3.2 Branigan 1992 267
3.3 Adger 1994 269
4 On the status of "interpretive effects" 274
5 Summary and conclusion 278

Chapter 7 Concluding remarks 279

References 282

Acknowledgements

Before finally letting go of this piece of work, I would like to acknowledge and thank some of the people who have been involved in one way or another in its development.

First and foremost, thanks to Wynn Chao, my supervisor for the last three years, for her encouragement, support and belief in me throughout this time. I also extend my sincere thanks to Ruth Kempson, my first supervisor and the person who first got me interested in syntax. Another person to whom I am extremely indebted is Jonathan Kaye, whose Spring 1993 Research Training Seminars were not only a formative experience intellectually, but also enabled me to develop some much-needed practical skills. Without those classes, unnerving though they seemed at the time, I doubt if I would have been able to upgrade my studies to Ph.D status, let alone write a thesis, present conference papers, or make the most of the teaching opportunities which have come my way. Finally, many thanks to Elabbas Benmamoun, who has been an important source of feedback, encouragement and inspiration since he arrived in the department last year.

I would next like to pay my respects to five fellow-students whose friendship, support and linguistic companionship has contributed significantly to my work (usual disclaimers apply) and helped me to survive some stressful times: my first SOAS friends Julia Capritsa, Aslı Göksel and Stavroula Tsiplakou, as well as Margaret Cobb and Richard Breheny. Having written this thesis in the Phonetics Lab, I'd like to thank all the people who make up its discordant but calming atmosphere, but in particular Jiyoung Shin, Taeko Maeda, Justin Watkins (for numerous discussions on chenille), Jeanne Cornillon - and, making everything possible, Bernard Howard. Other people who I'd like to acknowledge for their intellectual inspiration and/or help and/or friendship at various times throughout my life as a postgraduate are: Dave Bennett, Misi Brody, Monik Charette, Annabel Cormack, Ann Denwood, Diana Gierling, Melanie Green, Howard Gregory, Nilou Hawthorne, Bruce Ingham, Sean Jensen, Rodger Kibble, Shalom Lappin, Ana Maria Madeira, Lutz Marten, Julia Martinez, Wilfried Meyer Viol, Jamal Ouhalla, Joe Phillips, Manuela Pinto, Stefan Ploch, Ian Roberts, Anne Rowlands, Andrew Simpson, David Swinburne, Hans van de Koot, and Geoff Williams.

I have been fortunate to be funded by a three year state studentship from the British Academy, to whom I am duly grateful.

The final word belongs to my closest friends and family for putting up with me for the last year: Pav (and Adam and Mrs Modelski), Adrian; William, Andrew & Manjit; Carol; Marcela; Sarah & Richard; Jerry; my parents and Dave; lastly my sister and brother, Jo and Jim, to whom, if anyone, this work should be dedicated.

Chapter 1
Introduction

1 Introduction

This research is conducted within the *Minimalist framework*, a version of the Principles and Parameters model of grammar developed in Chomsky 1993 and subsequent works. The purpose of this introductory chapter is threefold. Firstly, I give an overview of the Minimalist framework and introduce some of its basic theoretical notions, concentrating on those which are most relevant to the thesis, that is, those which relate to movement,[1] and drawing comparisons with the previous Principles and Parameters model (Government-Binding theory - Chomsky 1981, 1982, 1986) where applicable. Secondly, I introduce the topic of the thesis and explain its relevance in the context of the Minimalist framework. Thirdly, I give a short preview of the organization of the thesis. It should be noted here that this introduction is based mainly on the 1993/94 version of the Minimalist model, some specific parts of which are now obsolete. However, I have chosen not to remove the description of the older model, since many fundamental aspects remain

[1] See for instance Lasnik 1993, Marantz 1995 for more detailed expositions/interpretations of the Minimalist Program.

unchanged; it also forms a useful context within which to interpret more recent developments.

This chapter is structured as follows. Section 2 gives a general overview of the Minimalist framework and the rationale underlying its development. In Section 3, I outline the model. Section 4 describes the workings of this model in some detail, focusing on the role of movement and constraints pertaining to it. In section 5, I give a preview of the substantive part of the thesis.

2 The Minimalist hypothesis

The Minimalist hypothesis is essentially captured in the following extract from Chomsky 1994: "...language is something like a 'perfect system', meeting external constraints as well as can be done" (p.386). The idea is that the language faculty (often known as Universal Grammar or UG[2]) is a nonredundant, or "economical", system in the context of cognitive systems as a whole, with the corresponding theoretical methodology that "the basic principles of language are formulated in terms of notions drawn from the domain of virtual conceptual necessity" (Chomsky 1993:5). Of course, the concept of *economy* is closely associated with the Minimalist program. It has a number of relevant senses, which will be illustrated throughout this chapter. It is the particular sense of economy just mentioned which uniquely characterizes Minimalism, setting it apart from previous incarnations of the Principles and Parameters (P&P) approach. It should be noted that an empirical hypothesis is involved here; i.e. from a logical point of view, UG might or might not be "economical" in sense of Minimalism. In earlier P&P models, this was explicitly taken to be false, and Chomsky indeed comments that "[I]t is...far from obvious that language should have anything like the character postulated in the minimalist program" (1995:221).

Economy in the particular sense of the Minimalist hypothesis just described

[2] On the notion of Universal Grammar, see e.g. Chomsky 1980, 1986a.

contrasts with a different type of economy consideration which we may call *conceptual economy*. This consideration is applicable to theories generally, irrespective of their specific details or subject-matter. In contrast to naturally-occuring phenomena such as human language, there is actually some *a priori* reason why theories themselves should be constructed economically. If a theory by definition which explains or makes predictions with respect to some fact/s, this more or less entails that it should not contain redundancy (although it is noted by Chomsky (1994:386) that notions relating to conceptual economy are "not precise, but not without content").

Although, as mentioned above, there is no *a priori* reason why language itself should be an economical, nonredundant, "perfect" system in the sense of the Minimalist hypothesis, there is one well-known empirical consideration, namely certain well-known facts concerning language acquisition,[3] which could be taken to suggest that UG itself is likely to have these properties, to some extent, and as Chomsky (1993:2) mentions, "a working hypothesis in generative grammar has been that...the language faculty is nonredundant, in that particular phenomena are not "overdetermined" by principles of language...". Nevertheless, although the facts about language acquisition might constitute a good reason to believe that the language faculty must be a non-redundant system in and of itself, they do not necessarily suggest that it is a perfect system in the particular sense of the Minimalist hypothesis, i.e. with respect to other cognitive systems within which it is embedded.[4] In summary, then, the special characteristic of Minimalism from a general perspective is its claim that "language is something like a 'perfect system', meeting external constraints as well as can be done", an idea which is independent of ordinary considerations of conceptual economy and of explanatory adequacy, both of which have

---

[3] The relevant facts about acquisition are that language is acquired quickly and on the basis of extremely limited and imperfect evidence. In the words of Chomsky 1981: "...it is a certainty that fundamental properties of ... attained grammars are radically underdetermined by evidence available to the language learner and must therefore be attributed to UG itself". See also Chomsky 1980 for general discussion.

[4] See Chomsky 1981, Chapter 1 for detailed discussion of the relevance of considerations such as explanatory adequacy, conceptual economy and so forth.

always been relevant in the P&P approach.

It will be clear even from the brief outline just given that the Minimalist model of grammar must differ substantially from its predecessor, the GB model. Generally, it could be assumed that such "paradigm shifts" are justifiable only if the incoming system can account for as much as its predecessor did, and in equally as principled a way, in as much as these sorts of properties are quantifiable. The wider objective of this thesis is to make some contribution towards evaluating the Minimalist framework against its predecessor, by focusing on how it handles a particular core aspect of syntax traditionally thought of in terms of *movement*. Although in many other aspects, as we shall see shortly, it turns out to be very possible to re-explain data in terms of the restricted theoretical resources which the Minimalist framework allows, it is interesting to note that with respect to the characterization of movement, the GB model could be argued to be superior both in conceptual and empirical respects. Given the central importance of movement phenomena in syntactic theory, the way in which, and the extent to which, they can be characterized is undoubtedly a crucial factor in deciding just how viable the new model is. In order to see exactly what Minimalism's movement-related problems are (see §5 below), we need firstly to take a closer look at general properties of the model. This is undertaken in the next two sections.

3 The Minimalist model

Since the Minimalist hypothesis renders syntax-internal constructs and notions unavailable, in contrast to the GB model in which these were quite freely invoked as the need arose, considerable quantities of data are left to be re-analysed in different terms. Three specific illustrations of this are given in this section, in which we build up a picture of the Minimalist model itself.

3.1 Levels of representation

In the Government-Binding model which preceded Minimalism, there were taken to be four levels of representation relevant to syntax: D-structure, S-structure, PF (Phonetic Form) and LF (Logical Form) (see Chomsky 1981:4). Out of these four, only PF and LF can be said to be conceptually necessary, since these constitute the "input" to, or interface with, grammar-external systems which Chomsky refers to as Articulatory-Perceptual (A-P) and Conceptual-Intentional (C-I) respectively.[5]

The disappearance of D- and S-Structure levels raises some important questions: in general terms, how to account for the facts which motivated the introduction of these levels in the first place? As an example, consider the principles regulating argument binding, or the distribution of Negative Polarity Items. In the GB model, these principles were thought to apply at the level of S-structure. However, this cannot be the case, if S-structure does not exist. In fact, it seems reasonable in principle that binding and similar interpretation-related principles (the Θ-criterion, for example) should be relocated to LF.[6] What appears to be more problematic for the reduced-level system is the existence of more resolutely syntactic phenomena, such as the surface distribution of NPs, which was formerly dealt with for the most part by an S-structure condition, the Case Filter (Rouveret and Vergnaud 1980[7]). The Minimalist answer to these questions has involved, among other things, an increase in conditions on movement, and an active role for the notion of PF wellformedness, both of which will be discussed later in this chapter.

As far as D-structure is concerned, there were two syntactic modules which

---

[5] Cf. Chomsky 1981: "It is reasonable to suppose that the representations PF and LF stand at the interface of grammatical competence, one mentally represented system, and other systems: the conceptual system, systems of belief, of pragmatic competence, of speech production and analysis, and so on" (p. 18).

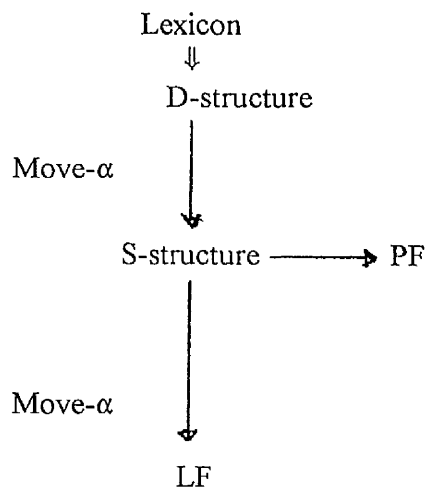
[6] See Chomsky 1993 for some discussion of Binding theory without S-structure. See also Barss 1994 who argues for an alternative Minimalism-based theory of Binding.

[7] The Case Filter applies to S-structure and states "*NP if lexical and no Case".

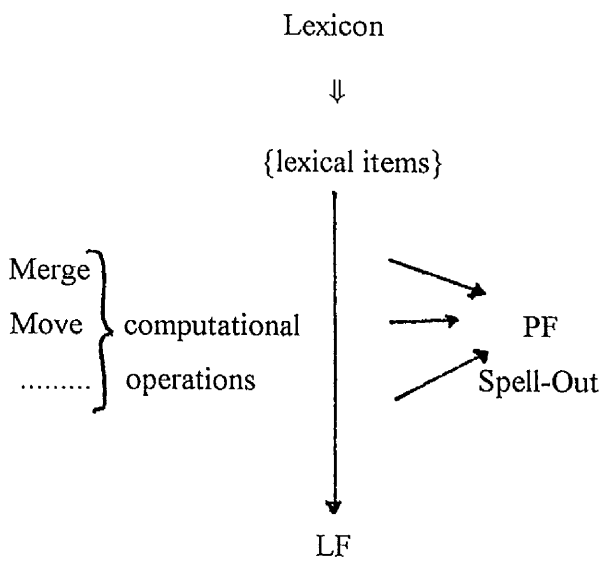
defined it: Θ-theory and X-bar theory. We will look at what replaces these shortly, starting with X-bar theory in the next subsection. The two-level Minimalist model is illustrated in (1b), with the model of Chomsky 1981 in (1a) for comparison.

(1) Principles and Parameters models of grammar

a. The GB Model (Chomsky 1981)



b. The Minimalist model (Chomsky 1993)



As depicted in (1b), a set of elements drawn from the lexicon - known as the *numeration* (N) - is constructed by computational operations (Merge, Move (and others, e.g. deletion) - see §§ 3.2, 4.2) into a syntactic representation, still known as LF. The operation *Spell-Out* may in principle apply at any point during the mapping from N to LF, to create a PF representation. By contrast, in the GB model (1a), lexical items are formed into a syntactic representation D-structure (see Chomsky 1993 for discussion). D-structure is mapped to a further level of S-structure, by the rule Move-α. S-structure is in turn mapped to LF and PF, the former mapping also involving the rule Move-α, and the latter, "the rules of morphology and phonology" (Chomsky 1986:100). An important characteristic of the Minimalist model is clearly visible in (1b): there is a single mapping relating the numeration and LF, which means that "the computational procedure C_{HL} is uniform .... any distinction pre-and post-Spell-Out is a reflex of morphology within the phonological component" (1994:411). Hence while in the GB model it is legitimate to treat "the LF rules" and "the transformational rules" as different subsystems possibly obeying different conditions - for instance, the Subjacency condition has sometimes been assumed to apply to the derivation of S-structure from D-structure but not to the derivation from S-structure to LF - this is not a theoretical possibility in the Minimalist model. Note lastly that the models shown in (1) have in common the property of non-trivial derivationality, i.e. the creation of LFs from lexical items takes place in steps, or stages.

3.2 Phrase structure

In the GB model, and also in the Minimalist framework in its earlier stages (1992, 1993), phrase structure had an independent status in the grammar, with certain properties taken to define it. The claim was, specifically, that phrase structure was regulated by a syntax-internal module, known as X-bar (X') Theory.[8] The exact details of X'-theory have been formulated in various ways; the following version comes from Chomsky 1986:

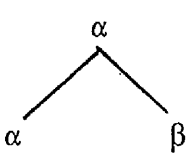
[8] See Jackendoff 1977 for early discussion.

(2) X'-Theory:

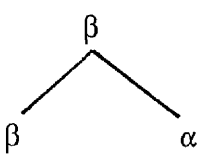
a. $X' = X\ X''^{*}$ $\qquad$ * = zero or more occurences
b. $X'' = X''^{*}\ X'$

A further restriction on phrase structure, not encoded in (2) but often assumed, is that phrase structure is at most binary branching (Kayne 1984). Although X'-theory was both empirically and conceptually superior to the system of Phrase Structure rules which it superceded (Stowell 1981), from the Minimalist point of view it is not adequate, since as Chomsky 1994 points out, the system still makes some reference to notions outside the domain of virtual conceptual necessity. What exactly is conceptually necessary in this respect? If there are assumed to exist entities larger than, and consisting of, single lexical items (for example, representations) then there must be some means by which the lexical items are combined. One possibility is that lexical items could simply be concatenated into strings; however, Chomsky notes that "bare output conditions" (on the LF side, at least) force a more complex means of combination to be assumed. He proposes an operation called *Merge*, a "computational" operation which applies to two separate items to form a single item which inherits the category of either one (but not both) of the "inputs", which may themselves be either lexical items, or complex items previously created by Merge. The "output" of an application of Merge is then said to have the *label* of one of the merged objects. For example, given two items α and ß, Merge forms either (3a), with the label α (or, if α itself is a complex object, the label of α), or (3b), with the label β (or, if β itself is a complex object, the label of β); in traditional terms, either α or β projects.

(3) a.



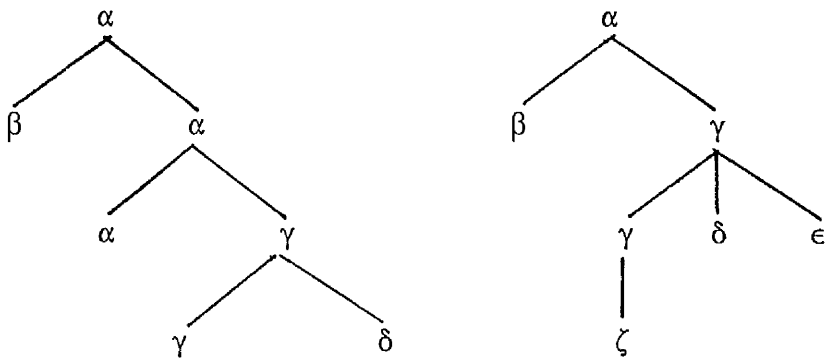
b.



The way that Chomsky sets it up, Merge creates structures like (4a) rather than (4b), thus deriving some of the results of X'-theory - for instance, the fact that phrases may not have more than one head. Unlike X'-theory, Merge also predicts that phrase structure is both minimally and maximally binary branching, with no relations like that between γ and ζ in

(4b) (single branching), or that between γ, δ and ϵ (ternary branching).

(4) a. Structure derived by Merge b. Structure not derived by Merge



As can also be seen from (4a), there are some phrase-structure properties deriving from Merge which do not derive from X'-theory. For example, the X'-schema in (2) seems to presuppose/require that two levels are invariably projected, even if no actual phrase occupies the designated position for specifier or complement. In the Merge system, on the other hand, this is not the case. There may be maximal projections with no Specifier, and maximal projections which do not project at all, i.e. which are both maximal and minimal. There is no inherent difference between Specifiers and Complements - generally, levels of projection have no independent status as such, but are derivative: "a category that does not project any further is a maximal projection XP and one that is not a projection at all is a minimal projection X^0; any other is an X', invisible at the interface and for computation" (1994:396). Thus there are no purely syntax-internal levels of projection under the Bare Phrase Structure theory (as there are no syntax-internal levels of representation).[9]

[9] For a detailed critique of many aspects of Chomsky's Bare Phrase Structure system, see Brody 1994.

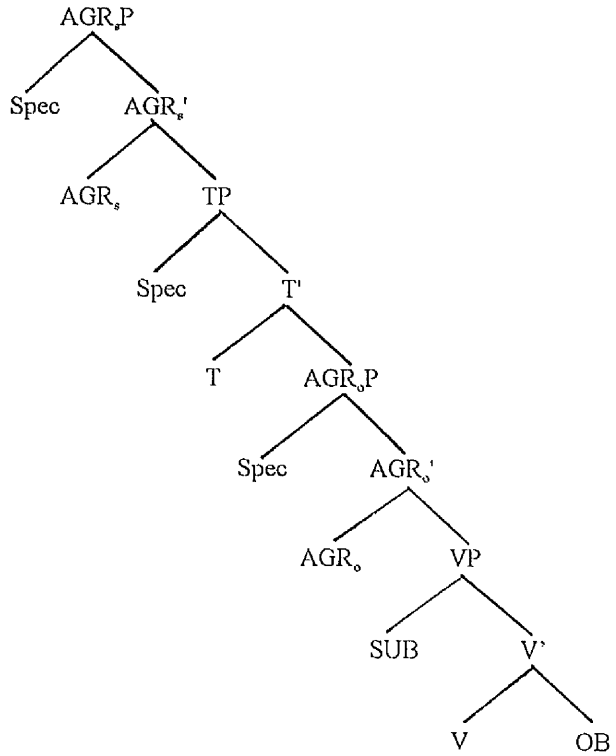
3.3 Formatives and clause structure

As a third and final example of the effects of the Minimalist hypothesis, let us consider a proposal put forward in the most recent development of the framework to date (Chomsky 1995). In this paper, Chomsky addresses the question of what lexical formatives can legitimately be assumed to exist, if "the basic principles of language are formulated in terms of notions drawn form the domain of virtual conceptual necessity". Departing from recent assumptions about clausal architecture (dating from Pollock 1989; see also Chomsky 1991), Chomsky proposes "eliminating AGR from UG entirely...keeping to functional categories with intrinsic properties that are manifested at interface levels" (1995:355).The rationale behind the relinquishing of AGR is that the category seems to have no reflex at any level of representation.[10] In this respect, AGR is taken to contrast with the element T (Tense), which has also been standardly assumed to project a phrase of its own (see again Pollock 1989).

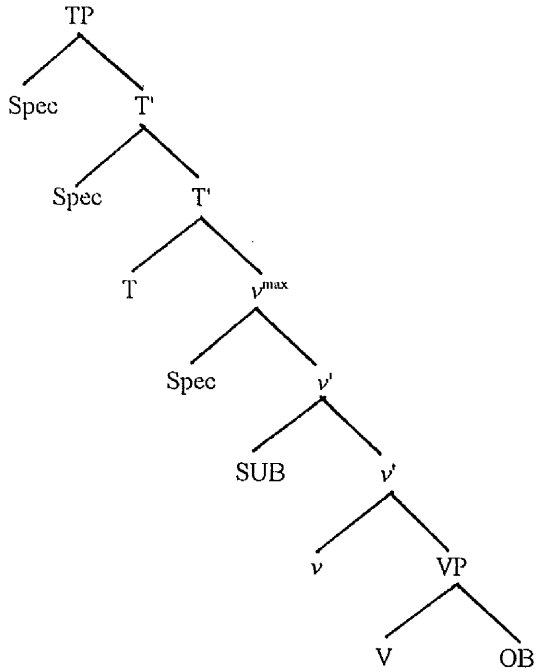
To compensate for the theoretical role formerly taken on by AGR and its projections (for example, AGR provided a position for head movement (e.g. "short movement" of infinitives in French, as discussed in Pollock 1989), while the Spec positions of AGRPs were used as landing-sites for arguments), Chomsky makes a number of new assumptions, one of which is that phrases may have multiple Specs. This provides phrasal landing-sites. An X^0 position to replace AGR_o is provided by replacing the simple VP with a more complex arrangement involving a light verb with a VP complement. (5) illustrates the structure of the (transitive) clause before and after these developments:

[10] However, it will be seen in Chapter 2 that Chomsky assumes that there are certain *features* which have no interpretation at either of the interfaces.

(5) a. Clause structure (Chomsky 1993/94)



b. Clause structure (Chomsky 1995)



Having described briefly the content of the Minimalist model, let us now see how it works Here we also find various notions of economy.

4 Principles of economy

The use of economy-type notions in syntax is not in itself unique to Minimalism. Chomsky points out that "such considerations have arisen in various forms and guises as theoretical perspectives have changed..." (1993:2). Some specific examples of these considerations are the Principle of Full Interpretation (Chomsky 1986) and various locality/minimality conditions which have been taken to characterize different kinds of syntactic relations (Rizzi's (1990) Relativized Minimality, to give just one example). In the Minimalist framework, economy takes various forms. Most basically, a division can be made between economy of representation and economy of derivation (see Chomsky 1991).

4.1 Economy of Representation

In accordance with the Minimalist hypothesis, as mentioned already (§ 3.1), there are taken to be just two levels of representation relevant to syntax: PF and LF, the interfaces with extra-grammatical systems. In the GB model, conditions of one kind or another constrained the various levels of representation. As would be expected, given that the number of these levels has been halved, constraints on the wellformedness of the remaining levels become more significant. Let us look at the nature of these constraints.

There is one basic requirement which PF and LF representations would have to obey in any theory (if it assumes these representations at all) - they must contain elements which are recognizable by the system which they are supposed to interact with - Articulatory/Perceptual in the case of PF, Conceptual-Intentional in the case of LF. On the LF side, Chomsky suggests that the legitimate objects are *chains:* "[W]e assume each legitimate object to be a chain CH = $(\alpha_1,...,\alpha_n)$: at least (perhaps at most) with CH a head,

an argument, a modifier, or an operator-variable construction" (1993:27). A similarly detailed definition of legitimate PF objects is not provided by Chomsky,[11] although, as we will see later, the notion of PF (il)legitimacy is instrumental in the theory.

While one cannot fail to assume that PF and LF must contain elements usable by the extra-grammatical systems which these levels are the "input" to, the Minimalist hypothesis leads to a tighter restriction than this, namely that representations must meet this requirement in a minimal way, i.e. they must contain no *excess* material. This requirement is known as Economy of Representation, or more commonly as the Principle of Full Interpretation (FI). The basic idea behind this is that the Articulatory/Perceptual and Conceptual/Intentional systems are not able to "overlook" extra objects in a representation, provided that the representation also contains interpretable (i.e. legitimate) objects - although this is in principle possible (the standard analogy in this connection is with vacuous quantification in artificial languages such as the predicate calculus - see Chomsky 1991:438).[12] Instead, FI requires that *every* component in a representation makes a contribution to interpretation: the necessary requirements of the grammar-external systems must be met in a minimal way.[13] Introducing some Minimalist

---

[11] For the record: "...what are the legitimate objects at PF and LF. At PF, this is the standard problem of universal phonetics" (*op. cit.*, p. 27).

[12] Cf. Chomsky 1986: "...there is a principle of full interpretation (FI) that requires every element of PF and LF...must be licensed...none can simply be disregarded" (98).

[13] There is a clear resemblance between the idea that representations should be "at least and at most" capable of interpretation by the relevant systems, on the one hand, and the two-claused Θ-criterion of the GB model, on the other, a version of which is in (i) (taken from Haegeman 1994:54):

(i) a. Each argument is assigned one and only one θ-role.
b. Each θ-role is assigned to one and only one argument.

Violations corresponding to clause (ia) and (ib) are in (iia) and (iib) respectively:

(ii) a. John saw Mary Bill
b. John likes

The exact status of violations previously attributed to the Θ-criterion , such as those in (ii), is still an open and problematic question in the Minimalist framework (and it is possible that they do not form a unitary class of violations at all), as are certain syntactic facts which used to derive from the

terminology, a derivation which results in a pair of PF and LF representations both of which obey FI is said to *converge*. If FI is violated at either level, on the other hand, the derivation *crashes*.

One last point which should be mentioned whilst on this general topic is the distinction which Chomsky draws between a representation violating FI - which amounts to being incapable of interpretation (i.e. the representation contains some object which the relevant system cannot deal with) - and a representation receiving an interpretation which is defective. All that is ensured by a representation conforming to FI is the possibility (not the fact) of receiving a non-defective interpretation. In other words, the term *interpretation* in Full Interpretation does not equate exactly with *interpretation* in its "common sense" usage.

4.1.1 Movement, features and checking

Having discussed wellformedness conditions on representations, i.e. representational economy (embodied in the principle of FI), we are now in a position to look at the status of movement, which, in the Minimalist framework, relates crucially to Full Interpretation, with respect to LF in particular. Recall that wellformed LF objects are chains of various types (see above). Thus, anything other than a chain is an illformed object at LF. Specifically, Chomsky singles out unchecked morphological features as illformed objects whose presence in an LF representation will cause the derivation to crash.[14] What is the

---

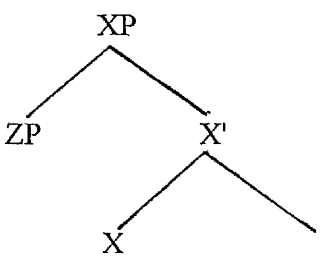
Θ-criterion. Chomsky (1993) suggests that Θ-theory violations are convergent but "defective" (p.32), while in a more recent version of the model, he maintains that they actually fail to converge, a conclusion adopted for purely theory-internal reasons. Since these questions fall well outside the scope of this thesis, I will not discuss them here. See Brody 1993 on some issues relating to Θ-theory.

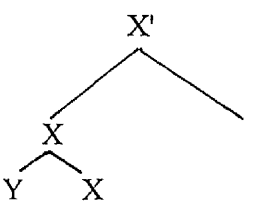
[14] Although Chomsky at one point advocates the idea that "derivations are driven by the narrow mechanical requirement of feature-checking only" (1993:33) it may be of interest to note that in general, the connection between LF legitimacy and feature-checking does not seem to be as neat and

status of these features, and how does this relate to movement?

In the Minimalist framework, continuing a trend which had already taken hold in the GB model, all syntactic licensing[15] is conceived of as a checking relation between formal features on a substantive (also known as thematic or lexical) element and features on a functional element. If the features match, they are eliminated.[16] Checking is assumed to take place under certain designated strictly local conditions, as shown in (6).

(6) a. Spec-Head relation: b. Head-Head relation (adjunction):





Given the contemporary perception of clause structure, according to which substantive elements initially all congregate at the "foot" of the tree, to the exclusion of functional elements, i.e. there is an in-built nonlocality between lexical and functional elements,[17]

---

tidy as this would imply. For instance, in dealing with reconstruction effects, Chomsky envisages a syntactic procedure "akin to QR" which applies to chains to form operator-variable constructions (see Ch. 2 of this thesis for discussion). Further, it is maintained for theory-internal reasons in Chomsky 1995 (contra Chomsky 1993) that failure of θ-role assignment, though not taken to be a feature-checking relation, prevents convergence.

[15] In the sense of Chomsky 1986a. See Rothstein 1991 for general discussion of the notion of licensing.

[16] This is an extremely simplified portrayal of the checking theory; it is discussed in more detail in Ch. 2 of this thesis, where the details become relevant.

[17] The relevant contrast is with the now largely abandoned view (though see Williams 1994) that subjects are base-generated in Spec-IP, nonlocal in some sense to the assigner of their θ-role (i.e. V

it follows that the substantive elements must invariably move, if they are to be in a sufficiently local relation to check their features with a functional element, and that the configurations for checking will be those configurations which result from movement, i.e. Spec-Head and head-adjunction. The notion of morphological features forcing movement is to a partial extent reminiscent of the GB principle known as Lasnik's Filter.[18] However, the Minimalist concept of feature-checking requirements inducing movement is of a more abstract nature, having more in common with Rizzi's (1990b) *wh*-criterion and similar criteria modelled on this.[19] Chomsky assumes a version of the Lexicalist Hypothesis, whereby items enter the syntactic computation already inflected, so that there is no question of inflections and "lexical material" needing to become physically attached to one another by syntactic operations.[20]

As an illustration of the notion of syntactic licensing as feature-checking, let us consider briefly the Minimalist theory of Case. According to this theory, all (at least structural) Cases involve movement of the argument to the Specifier of a functional projection. For Nominative Case, movement to Spec-TP, and for Accusative, movement to the Spec of a functional category whose complement is VP, and to whose head V, containing the relevant Case features, has raised - in the 1995 model, with which I

---

or V and its object). This view has been superceded by the so-called Internal Subject Hypothesis (on which see e.g. Kuroda 1988, Koopman and Sportiche 1988, Contreras 1987, 1991) according to which subjects originate within the projection of V. See Déprez 1989 for a comprehensive discussion of the theoretical implications of the Internal Subject Hypothesis. On the distinction between lexical and functional categories, see Abney 1987.

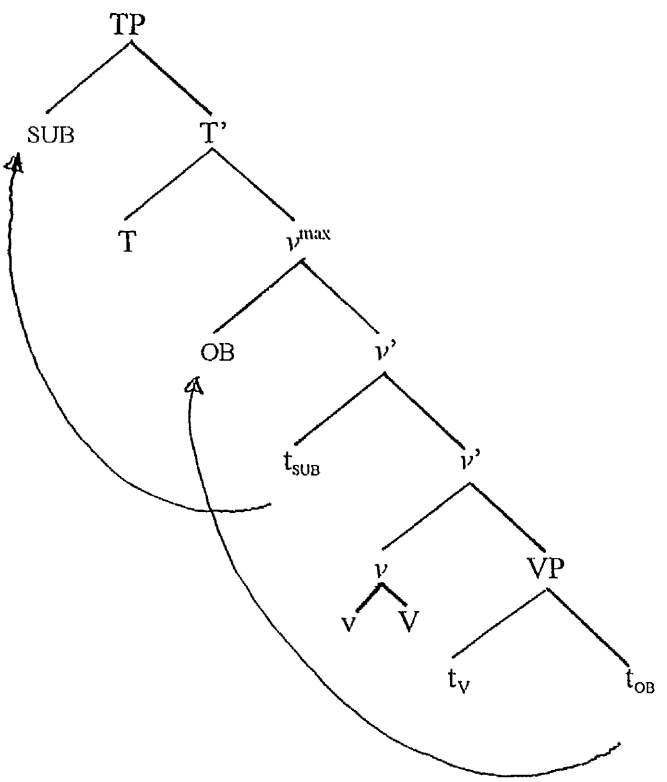
[18] Lasnik 1981. As Pesetsky describes it (1989:2), "Lasnik's Filter requires morphemes designated as affixes to be 'supported' by lexical material at PF" and is a "prime mover for the various transformations and insertions found in the English and French verbal auxiliary system". See also Chomsky 1991: "...an affix...must be 'completed' in the overt syntax..." (p.427).

[19] See Haegeman 1995 and works cited therein for extensive discussion. The *wh*-criterion originated in May 1985.

[20] For some discussion, and a different view, see Bobaljik 1995. For a theory which espouses the Lexicalist Hypothesis and lacks even head-movement for checking, see Williams 1994.

illustrate here in (7) below, the functional category is a "light verb"designated by v.[21]

(7) Spec-Head-based Case theory



4.1.2 Overt movement

Since no concept of S-structure is available, the question arises as to how to characterize the various phenomena previously explained in terms of conditions applying at this level - for example, in the GB theory, as already mentioned, the overt distribution of NPs was largely explained in terms of the Case Filter, applying at S-structure. In the Minimalist framework, such facts are dealt with in PF terms. Although it could not realistically be assumed that chains as such are relevant objects at PF, movement/chain formation can be indirectly forced by PF-related considerations in the following way. Chomsky 1993

[21] For arguments that objects in English are indeed displaced at an abstract level, see Lasnik and Saito 1991, Branigan 1992. These arguments are discussed in Chapter 2.

proposes that there are certain features which have a special property of PF illegitimacy, meaning that they are "perceived" by the articulatory/perceptual system and hence must not survive to that stage, otherwise FI will be violated. Like features generally, strong features are eliminated when checked, with the result that their presence in the numeration effectively induces movement before the operation Spell-Out applies (see illustration in (1b) above). Features with this special PF property are known as *strong features*, while the rest are *weak*.[22]

Introducing some more terminology, movement which occurs before Spell-Out, and hence is reflected in PF, is known as *overt movement*. Movement without a PF reflex is *covert*.[23] However, it must always be kept in mind that in the Minimalist model, where there is only the one syntactic level (LF), the computation of this level from the initial numeration is uniform, and so it is not the case that movement which is overt is intrinsically different to movement which is covert.

So far in this chapter, we have seen that "operations are driven by morphological necessity; certain features must be checked in the checking domain of a head" Chomsky *op. cit.*:32) (although see note 12 above); any features which fail to be checked cause the derivation to crash at LF; in addition, unchecked Strong features cause the derivation to crash at PF.

---

[22] This system has its origins in Pollock's (1989) account of various differences between the behaviour of main verbs in French and English. In Chomsky's elaboration of Pollock's theory, no relation exists between the Strength property of features and the "morphological richness" of inflection. This seems to suggest that the Strong feature system for overt movement is as it stands essentially stipulatory, as indeed is admitted by Chomsky 1995.

[23] This usage first appears in Chomsky 1991.

4.2 Derivational economy and other conditions on movement

Convergence at each level is necessary for a linguistic expression to be "grammatical", but not sufficient. Chomsky claims (following Chomsky 1991) that it is not only representations which are subject to principles of economy, but also the derivations which form them, namely, the operations of the computational system: "a [grammatical] linguistic expression is the *optimal realization* of...interface conditions" (1993:26; emphasis added - AMP). With respect to which units/dimensions economy of derivation is measured in, the exact answer to this has varied over time and is still something of a controversial issue.[24] The explicit beginnings of derivational economy are to be found in Chomsky 1991,[25] where a principle of *Least Effort* is proposed which states that "shorter derivations are always chosen over longer ones" (*op. cit.*:426). The motivation for this Least Effort effort principle came primarily from comparative data concerning the placement of verbs in English and French. I will not discuss the Least Effort principle here, since it arose within the GB framework. In what follows, I describe the system of derivational economy proposed in Chomsky 1993-1994. The description will be fairly detailed, since derivational economy conditions are of fundamental importance throughout this thesis.

4.2.1 Last Resort and Greed

In the framework of Chomsky 1993/1994, the notion of derivational economy manifests itself in more than one condition. In a similar vein to the above-mentioned Least Effort condition, Chomsky (1993:32) proposes a principle of *Last Resort*, as in (8):

(8) Last Resort: "a step in a derivation is legitimate only if it is necessary for convergence - had that step not been taken, the derivation would not have converged."

---

[24] For discussion of what counts for derivational economy, see Chapter 5 of this thesis.

[25] *Some Notes on Economy of Derivation and Representation.*

Added to this is a further requirement known as *Greed*. According to this, not only is movement forbidden unless convergence depends on it; it is only allowed if the moved element itself has some convergence-related, i.e. feature-checking, requirement:

(9) Greed: "Move-α applies to an element α only if morphological properties of α itself are not otherwise satisfied. The operation cannot apply to α to enable some different element ß to satisfy its properties". (1993:33).

Chomsky illustrates the application of Last Resort with the following example (= Chomsky's (26b), *op. cit.*).

(10) * There seems to a strange man that it is raining outside

(10) converges, on the assumption that *a strange man* is Case-checked in virtue of some (albeit unspecified) relation with the preposition *to*, and further, that the expletive *there* is itself able to check the Case feature of T.[26] The deviance which (10) nevertheless clearly suffers from is attributed to the presence of "free-standing *there*". This problem could in principle be solved, if *a strange man* were to raise covertly to adjoin to the expletive (as in the theory of expletives outlined by Chomsky 1991). However, Last Resort dictates that this is impossible in practice, since (10) is convergent as it stands and therefore no further operations can possibly be necessary for convergence. There is a relevant contrast between (10) and (11) (=Chomsky's (26a)):

(11) There is a strange man in the garden

In this case, *a strange man* is allowed to raise, since otherwise the derivation would not converge, the NP being unable to check its Case feature in its original position. The application of Greed is exemplified by (12) (=Chomsky's (27))

---

[26] Although Chomsky 1995 makes an alternative proposal about *there*. This is discussed in chapter 5, in which I examine (non)optionality of partial movement of the associate of expletives in French, English and Swedish.

(12) * Seems to a strange man that it is raining outside

In the case of (12), there are unchecked features which will cause the derivation to crash if nothing moves to check them - notably, a strong D(P)-feature on T which needs checking by overt movement. In the absence of Greed, *a strange man* would be able to raise and check this feature, resulting in (13), since obviously this movement is necessary for convergence and would be consistent with Last Resort.

(13) * A strange man$_i$ seems to t_i that it is raining outside

However, Greed disallows *a strange man* from moving, since its own requirements can clearly be otherwise satisfied, in the same way as they are in (10), in fact.[27] What differentiates Greed and Last Resort is the fact that the former is not subordinate to Full Interpretation; as the case of (12) demonstrates, it is an absolute condition which cannot be overridden, even if convergence would be otherwise prevented.[28]

---

[27] See Lasnik 1993 for discussion and modification of Chomsky's (1993) account of this range of expletive data.

[28] Note that (10) with an expletive *it* added, as in (I), is grammatical:

(i) It seems to a strange man that it is raining outside

The expletive itself evidently checks any features of T which need to be checked, allowing for convergence. Nor does (i) have a deviant interpretation, in this respect contrasting interestingly with (ii):

(ii) There seems to a strange man that it is raining outside

Chomsky claims that examples such as (ii) are convergent yet deviant: *there* has "no coherent interpretation, because [it] receives no semantic interpretation" (1993:33). Greed prevents the movement which Chomsky claims would allow for a nondeviant interpretation, i.e. adjunction of a *strange man* to *there* (which he characterizes as an LF affix). This raises the question of how a deviant interpretation is avoided in cases like (i); is *it* an LF affix, and if so, what adjoins to it? A different perspective on expletives which may have some bearing on these questions is developed in Chomsky 1995; this is discussed in Chapter 5 below.

4.2.2 Procrastinate

As we saw in § 4.1.2, strong features force overt movement, since these features violate FI at PF if still unchecked and therefore present at that level.[29] Considering features which are *not* strong, movement to check these may take place overtly *or* covertly , as far as convergence is concerned. However, Chomsky proposes a further economy condition, *Procrastinate*, which curtails this potential optionality in when movement occurs relative to Spell-Out:

(14) Procrastinate: "LF movement is 'cheaper' than overt movement...The system tries to reach PF 'as fast as possible', minimizing overt syntax."

Thus, the morphological property of feature-strength and the principle Procrastinate together account for the surface positions of elements. As a typical illustration, consider Chomsky's account of the difference between French and English with respect to the position of main verbs (Emonds 1978, Pollock 1989). Assuming that the adverbs are left-adjoined to VP, (15) illustrates the fact that in French, (finite) main verbs must move overtly, while in English, this is not possible:

(15) a. Jean mange souvent des citrons
J. eats often lemons
'Jean often eats lemons'

b. * Jean souvent mange des citrons
c. * John eats often lemons
d. John often eats lemons

Along the lines of Pollock 1989, Chomsky assumes that the features to be checked by the verb are strong in French but weak in English. The strong feature must be checked by

[29] Zubizarreta 1994, building on Cinque's (1993) theory of the relation between syntactic structure, stress-assignment and focus, develops a model in which some types of overt movement, e.g . Scrambling and Heavy NP Shift, are motivated by PF considerations, but of a prosodic rather than a morphological nature. For some discussion of this, in the context of the issue of how to deal with optionality phenomena in Minimalism, see chapter 6.

overt movement, otherwise the derivation crashes at PF, accounting for (15a,b). In English which lacks the strong feature, movement of the verb may happen covertly - and therefore is forced to do so by Procrastinate (15c,d): "In English-type languages, overt [verb] raising is not forced for convergence, therefore, it is barred by economy principles" (1993:30-31). [30, 31]

---

[30] On a conflict which arises between the principles Greed and Procrastinate in this model, see Wilder and Cavar 1992. For objections to various aspects of Procrastinate, see Brody 1994/5, who proposes a principle similar to Pesetsky's (1989) Earliness: "Satisfy filters as early as possible..." (p. 7).

[31] The question arises of why auxiliaries, e.g. *have* and *be*, *are* able to raise overtly in English, in common with lexical verbs in French rather than lexical verbs in English:

(i) John has often eaten lemons

Chomsky 1993 suggests ("adopting the intuition (but not the accompanying technology)" of Pollock's (1989) theory, which connected the possibility of raising to whether or not the verb assigns a θ-role) that auxiliaries must raise overtly because "such elements, lacking semantically relevant features, are not visible to LF rules. If they have not raised overtly, they will not be able to raise by LF rules and the derivation will crash" (p. 31). Chomsky's solution seems to be less principled than Pollock's (1989) upon which it is modelled. For example, the question arises as to what qualifies as "having semantic content". As is explicit in Pollock's account, the relevant distinction seems to be between being a theta-role assigning element or not. This explains why modals behave syntactically like *have*, *be* and *do* - see (ii) - although in contrast to the latter, they clearly have some kind of semantic, although not thematic, import of their own, as opposed to just carrying inflection. In Chomsky's system, there seems to be no reason why "LF rules" should recognize some types of semantic content but not others.

(ii) a. John must/can/should/ often eat lemons
b. Must/can/should John eat lemons?

Another question is the following: why should, in fact how can, the "LF rules" of raising which Chomsky refers to in the above quote, differ from any other rules of raising, given that the syntactic computation is uniform? In fact, it is not clear if one is forced to claim that auxiliaries and modals move overtly in any case, since they are perhaps generated in a different and higher position than lexical verbs. It may well be that these elements raise to T covertly in English, since, as Koizumi (1995) observes, adverbs are permitted to occur between them and the subject, unlike in French:

(iii) a. John probably has eaten lemons
b. * Jean probablement a mangé des citrons

4.2.3 Minimal Link Condition

In order to capture various locality effects, Chomsky introduces a further derivational economy condition similar to Relativized Minimality (Rizzi 1990). This condition is known as the *Minimal Link Condition* (MLC), or sometimes *Shortest Move*. The MLC as stated in Chomsky 1993 is as follows:

(16) MLC: "...given two convergent derivations D1 and D2, both minimal and containing the same number of steps, D1 blocks D2 if its links are shorter" (1993:34).

It is assumed that only occupied positions count for this condition (see Jonas and Bobaljik 1992).[32] We can see the MLC in action in (17) (= Chomsky's example (9)), accounting for what is traditionally known as a Superiority effect.

(17) a. Whom$_1$ did John persuade t$_1$ to visit whom$_2$
b. * Whom$_2$ did John persuade whom$_1$ to visit t$_2$

It is assumed that the head C has a Strong feature, which must be checked by overt movement of one of the *wh*-phrases. In each of (17a) and (17b), a *wh*-phrase has moved to a position (i.e. Spec-CP) where it can check this feature, so both derivations are convergent. However, the movement made by *whom*$_2$ in (17b) is longer than the step taken by *whom*$_1$ in (17a): the former involves one *wh*-phrase (*whom*$_2$) skipping the other (*whom*$_1$), hence this derivation is blocked. In order for this story to make sense, the further assumption should be noted that *wh*-phrases *in situ* do not undergo syntactic raising at all.[33]

---

[32] And in fact, in the 1994 (*Bare Phrase Structure*) framework this falls out, since positions only come into existence in virtue of movement or merge of an element.

[33] In Chomsky's system,"[t]he LF rule that associates the in-situ *wh*-phrase with the *wh*-phrase in [Spec-CP] need not be construed as an instance of Move α. We might think of it as the syntactic basis for absorption in the sense of Higginbotham and May (1981), an operation that associates two *wh*-phrases to form a generalized quantifier". (1993:26). See Reinhart 1995 for a critique and

Turning to A-movement, the MLC is also invoked to account for SuperRaising violations as discussed by Rizzi 1990 (**John seems that it is likely to win*) (see Chomsky 1993:15 for details). In addition, a relatively new issue concerning locality and A-movement has arisen from the conjunction of the VP-Internal Subject Hypothesis, whereby subject and object originate within VP, and the theory that structural Cases are exclusively checked in local relations with functional heads above this VP (see §4.1.2 above). Under such a system, it must be the case that arguments can routinely cross each others' paths, and that this possibility is allowed for by the MLC, at the same time as not compromising the condition's beneficial effects. To this end, Chomsky (1993:18) proposes to incorporate a notion of *Equidistance* into the definition of MLC:

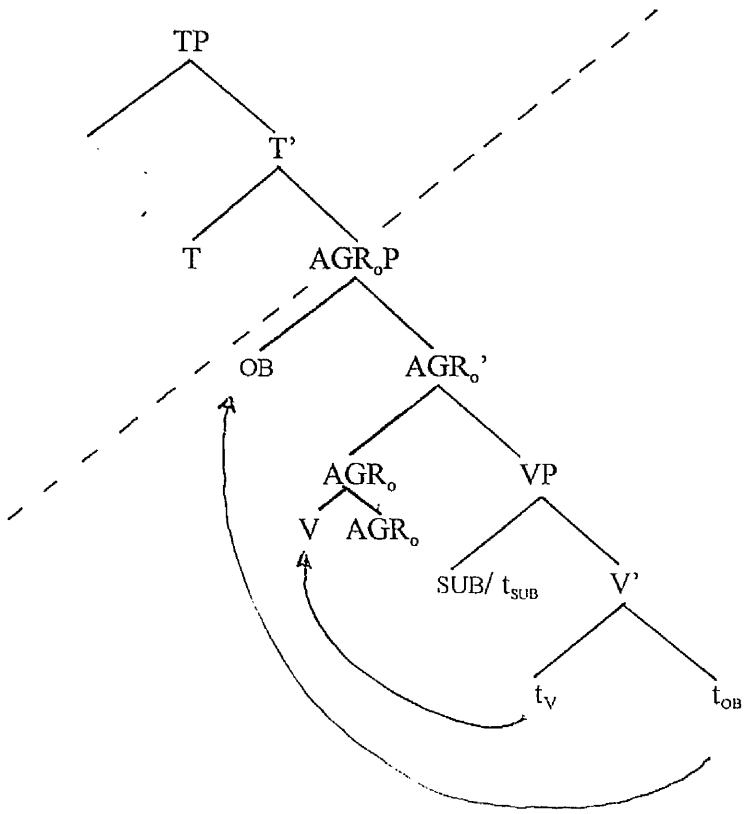
(18) Equidistance: "If α and β are in the same minimal domain, they are equidistant from γ. In particular, two targets of movement are equidistant if they are in the same minimal domain".

The precise definition of minimal domains varies according to other assumptions about the components and structure of the clause; however, some notion of Equidistance is necessary whether the clause structure in (5a) or (5b) above is adopted, since the two do not differ in the relevant respect. The main requirement is to in principle allow objects to move across Spec-VP, containing either subjects or subject traces, and subjects to move across a shifted object. In the 1993 model, in which arguments check Case in the Specifiers of AGR-phrases, Spec-VP and Spec-AGR_oP appear to be in different minimal domains, that of V and AGR_o respectively, which seems to predict, problematically, that Spec-AGR_oP should not qualify as the closest movement target for objects. To get around this, Chomsky defines minimal domains with respect to X^0-chains. When V incorporates into AGR_o, a new domain is created which includes both Spec-VP and Spec-AGR_oP which are then equidistant targets for movement; see (19):

---

extension of Chomsky's (1993) analysis of Superiority phenomena. Her account relies crucially on relinquishing the Minimalist assumption that "derivations are driven by the narrow mechanical requirement of feature checking only, not by a "search for intelligibility" or the like" (Chomsky 1993:33). For a discussion and critique of Chomsky's MLC from a diffferent perspective, see Manzini 1995.

(19) (cf. (5a) above)



As for the subject, Equidistance allows it to move past an object in Spec-AGR$_o$P and land in the Spec of the next head up, Tense, when the AGR$_o$/V complex has raised there creating a new chain and extended minimal domain which contains Spec-AGR$_o$P and Spec-TP.[34] Chomsky points out that this system derives Holmberg's Generalization (Holmberg 1986), i.e the generalization that object movement is contingent upon verb movement, which is known to hold with respect to the Scandinavian languages.[35] Concomitantly, it predicts that object A-movement cannot be overt in English, since - as it is commonly assumed - main verbs do not themselves move overtly, meaning that

[34] Whether this is actually allowed to happen seems to involve further variables - according to Jonas and Bobaljik 1992 and Branigan 1992, whether Spec-TP is available as a landing-site, which is linked to a parameterized property of T. See chapter 6.

[35] On this topic, see Chapter 6.

objects are also forced to "wait till LF".[36]

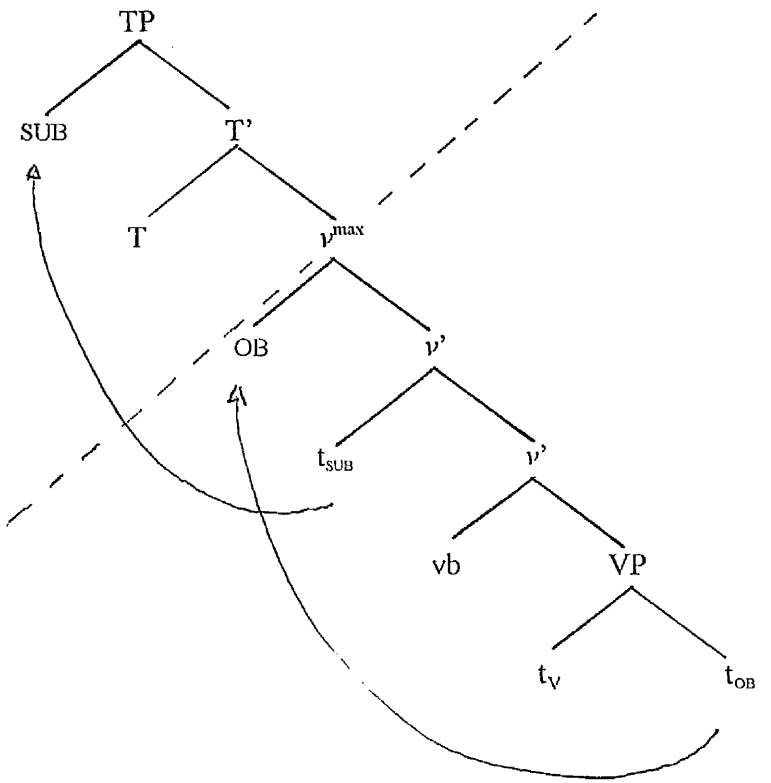
With the introduction of clause structure without AGR phrases ((5b) above), the site for Accusative Case-checking is the second Spec of v^{max} and this position is therefore in the same minimal domain as the subject (i.e the minimal domain of the light verb) - see (20) - hence in order to allow objects to skip subjects, minimal domains are no longer defined relative to head-chains but to heads themselves. For the case of subjects skipping objects, the notion of Equidistance needs to be amended so that it applies not only to targets of movement but also to the moved elements themselves: if two potentially moving elements are in the minimal domain of the same head, then they count as equidistant from any target (this much in fact is already implicit in (18) above). Fuller details of the MLC and Equidistance in the 1995 framework are given in the next chapter.[37]

---

[36] See for example Branigan 1992, Ferguson and Groat 1995 for discussion of the 1993 system. The possibility has now been acknowledged that Holmberg's Generalization as usually understood may not hold universally; there seem to exist cases in which objects can raise overtly in the absence of overt V-raising, e.g. Celtic (see Bobaljik 1995), and possibly "SOV" Germanic languages such as German and Dutch. As we will see shortly in the text, Holmberg's Generalization is not derived in the 1995 Minimalist framework. See chapter 6 of this thesis for a slightly different outlook on the link between overt V-raising and (the option of) overt object movement.

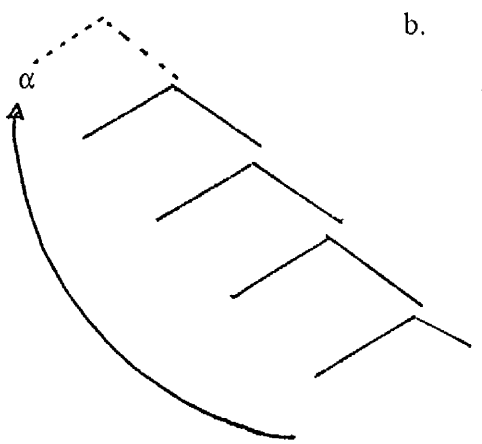
[37] Chomsky (1995) notes that the Equidistance clause could be eliminated completely if it was assumed that θ-roles can be assigned in the outer Spec of the relevant head. Were this to be the case, the movement paths of subject and object need never cross (cf. discussion of Koizumi 1993, 1995 above, and also Bobaljik 1995) and *closeness* could be defined in terms of c-command alone.

(20)

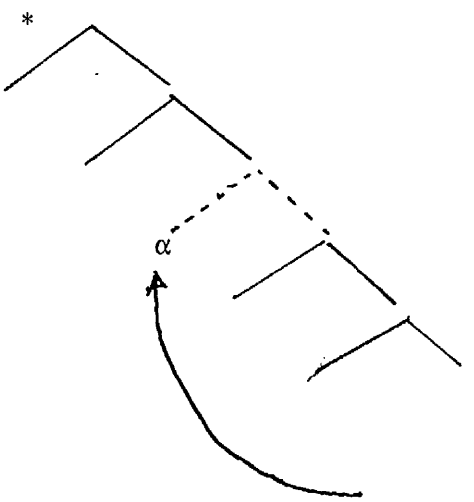


One further constraint on movement must be noted in discussing the MLC, and that is the *Extension Condition* (EC). This condition is similar to the older idea of Strict Cyclicity (see Chomsky 1977) and requires that overt computational operations must not target an embedded site in the structure (α = a Merged or Moved element):

(21) a.



b. *



Given the MLC account of Relativized Minimality effects described above, the EC is necessary as an auxiliary assumption, for in the absence of such a condition, an element could legitimately skip an appropriate empty landing-site - assuming crucially that this movement does not itself constitute an MLC violation[38] - with a second element then moving into the empty site at a later stage of the derivation. (23) illustrates how a case of SuperRaising (in the sense of Rizzi 1990:10) as in (22) could be derived without the EC:

(22) * John seems that it is likely to win

(23) a. $[_{IP}$John$_i$ seems that $[_{IP}$ is likely $[_{IP}t_i$ to $[_{VP}t_i$ win]]]]

b. $[_{IP}$John$_i$ seems that $[_{IP}$ it is likely $[_{IP}t_i$ to $[_{VP}t_i$ win]]]]

Since the countercyclic insertion of the expletive in (23b) is barred by the EC, the only possible derivation of (22) is that in (24), a derivation which the MLC excludes straightforwardly, as *John* fails to make the shortest move.

(24) a. $[_{IP}$ it is likely $[_{IP}$ John$_i$ to $[_{VP}t_i$ win]]]

b. $[_{IP}$John$_i$ seems that $[_{IP}$ it is likely $[_{IP}t_i$ to $[_{VP}t_i$ win]]]]

Some further considerations, specifically the assumption (Chomsky 1993) that movement for Accusative Case-checking in English takes place covertly, force the additional assumption that covert operations (and also head movement and adjunction generally) must be exempt from the EC (see e.g. Kitahara 1994, Branigan 1992 for discussion). Apparently, therefore, the EC cannnot strictly speaking be taken to be a condition on

[38] See note 33, and the sentence preceding it in the text. In the framework of Chomsky 1995, in which the MLC is stated in terms of the moved-to position/feature, and requires that the nearest suitable element move there, there is no question that unfilled intervening positions could be relevant.

computational operations *per se*, since it applies only to those of them which are overt, and the model admits of no intrinsic difference between overt and covert operations, as mentioned in §3.1 above. Accordingly, several attempts have been made in the literature to derive the effects of the EC from independent derivational economy conditions - see for instance Kitahara 1994, Chomsky 1994.

Note before concluding this section that, as Chomsky has pointed out (1991, 1993), derivational economy conditions, i.e. those which select "the optimal realization...of interface conditions", are necessarily global. This is so simply because the concept of *optimality* entails higher-order quantification (cf. *most* etc.). Conditions which are global in the 1993/1994 framework are Last Resort, Procrastinate and the Minimal Link Condition. These can be compared with the Principle of FI, which is a non-global, i.e. a local or absolute condition - a representation either conforms to it or it does not. Global conditions are associated with the property of computational complexity - in simple terms, it is more complicated to evaluate a candidate with respect to a global condition (e.g. Procrastinate) than with respect to an absolute condition (e.g. FI). Though computational complexity and its implications appears to be a poorly-defined area, at least among "syntacticians proper", it is nevertheless quite easy to see in general how there might be thought to be some level of complexity in excess of which a grammar entailing this level would be excluded from the plausible candidates for UG on grounds of this alone. The complexity issue has received more attention recently, including that of Chomsky 1995. The main proposal in Part I of this thesis has implications in this area; this is discussed in Chapter 4.

5 Summary and thesis preview

In this chapter, I have set out the theoretical context for the thesis with a description of the Minimalist framework, paying particular attention to the operation Move and conditions upon it. Having done this, I conclude by introducing the topic of the thesis and giving a brief preview of its organization.

We have seen that the Minimalist Program has concerned itself with the following questions: why elements move, whether they move, when they move, and how they move, with the answer taking respectively the form of the economy principles FI, Last Resort/Greed, Procrastinate, and the Minimal Link Condition (locality). But a further question one might ask is: "what moves?". As it turns out, this question is optimally relevant. In the GB model, the answer was simple; the rule Move-α stated "move anything anywhere" (Chomsky 1986), with all necessary constraints on its application provided by other principles of grammar (e.g. the ECP). It is interesting to note that this highly desirable aspect of the GB model seems to be lost in Minimalism, more or less as a direct consequence of fundamental properties of the framework. This thesis concentrates on two ways in which this change is relevant.

Since there is only one syntactic level of representation, LF (see §3.1 above), it is no longer possible to characterize "reconstruction" effects in terms of Binding conditions applying at different levels; this led Chomsky (1993) to propose that movement in fact consists of a *copying* operation. It is this copy theory of movement (described in detail in Chapter 2) which seems to have the consequence that the operation can no longer be taken to "move anything anywhere". On the contrary, it seems that those elements which were formerly thought of as a type of empty category, i.e. traces, will have to be stated to be incapable of movement, as Chomsky (1995) in fact proposes, since given the copy theory of movement, these traces will be identical in all but phonological features to the elements to which they are related, appearing to imply that they too will be moveable objects, leading potentially to a wealth of false predictions. This need to prevent movement from applying to certain types of element is the first of two important ways in which the Minimalist theory of movement seems inferior to that of the GB model.

As for the second, consider a further property of Move-α, namely that it was free to apply or not to apply, again with necessary restrictions provided by independent principles. This property apparently had a major empirical advantage, namely that it provided a straightforward means of characterizing optional movement data. There is a very long list of phenomena which were traditionally seen as involving optionality in the

application of Move α, among them Heavy NP Shift in English (Ross 1967); so-called Scrambling in many languages, for instance German (Grewendorf and Sternefeld 1989 among others) and Dutch (Vanden Wyngaerd 1989, among others), optionality of participle agreement in object *wh*-movement and certain clitic constructions in French (Kayne 1989). Here again the Minimalist theory of movement seems to be at a fairly severe disadvantage, if we compare it with GB. Since economy conditions select the "optimal realization of interface conditions", there will clearly be a strong tendency towards any given numeration yielding a unique permissible derivation. In practice, this means that truly optional movement cannot be freely assumed to exist (this problem is well-known; see e.g. Chomsky 1991, Fukui 1993), which seems to leave the optionality data with no explanation.

The thesis is divided into Parts I and II, based respectively around the conceptual and the empirical problems relating to movement in Minimalism which I have just described. In each part, I propose a solution which I believe to be Minimalist, although it departs from the standard model in various respects. In Part I, consisting of chapters 2-4, I argue for taking Move *in principle* to operate on any element, "traces" included. This proposal, which I call the Copy Hypothesis, is introduced in chapter 2 and defended against potential theoretical objections, mainly in the form of Chomsky's various (1995) arguments for stipulating that the traces of A-movement must be unavailable to Move. Chapters 3 and 4 are devoted to showing that the Copy Hypothesis has essential empirical advantages. Part II tackles the optionality problem. Showing that there is one dimension in which movement optionality of sorts is theoretically possible, viz. at what stage movement occurs relative to Spell-Out (overt/covertness), I attempt to develop an integrated theory of (non)optionality which works off derivational economy conditions (Procrastinate and the Shortest Derivation condition (see Chapter 5) and the feature-properties of elements. The theory is applied to various cases of optionality and related non-optionality phenomena: past participle agreement in French (Chapter 5), overt Object Shift in Icelandic, German and Dutch (Chapter 6), and partial movement of the associate of expletives in French, English and Swedish (Chapter 5).

Part I

Chapter 2
Copies and derivations

1 Introduction

Chomsky (1993, 1995) proposes to replace the movement operation with a process which creates a *copy* of the "moved" element at the appropriate site. This copy theory of movement is well motivated in the Minimalist framework given that it provides what appears to be the only straightforward basis for characterizing various interpretive phenomena traditionally thought of in terms of "reconstruction" (see e.g., Chomsky 1977, van Riemsdijk and Williams 1986, Lebeaux 1988, 1991, Barss 1986, 1994).

While the contribution of the copy theory to the reconstruction problem makes it virtually indispensable in the Minimalist framework, this view of movement seems to have other consequences - in the computational system (C_{HL} , in Chomsky's (1995) abbreviation) - which are potentially much less desirable. The problem is quite simple: if a chain consists of copies of an element, by definition identical to one another, then it seems to be predicted that in principle, each of these copies - not only the head of the

chain, but also those those normally known as *traces* - must be "active" in the computational system - including being subject to further applications of movement/copying. This raises an important theoretical question: is it necessary to impose special conditions upon certain copies, namely "traces", in order to control their potentially unwanted movements, or is it in fact possible to maintain some version of the notion that "Move-α applies freely" (Chomsky 1981, 1986), with all required restrictions, including those on traces, falling out from independent principles? The latter of these two alternatives is at least preferable, and ideally would be seen as the only option.

Thus, this chapter pursues the hypothesis that in principle all copies in a chain may indeed participate in operations of C_{HL}. The relevance of this project is twofold. In the first place, from a relatively general perspective, taking operations of the computational system to apply to any element is the null hypothesis (recall that the rule Move-α in the GB framework represented an improvement on the earlier system of construction-specific "transformational" rules). If it turned out that fundamental Minimalist assumptions forced a departure from this view, it could be concluded that in one serious respect, the Minimalist model does not measure up to the GB model. To lend some concreteness to this issue, note that in the most recent version of the Minimalist framework to date, Chomsky (1995b) argues that certain traces, namely those of A-movement, must be prevented from participating in movement, basically by stipulation (see § 4 below for arguments against this proposal). In the second place, from a more specific perspective, the assumption that all copies in a chain are active in the computational system is used repeatedly throughout this thesis in analyses of various data (see Chapters 4, 5 and 6). It is therefore necessary to introduce and justify the assumption independently, since it is not currently standard in the Minimalist framework.

The chapter is structured as follows. Section 2 sets out the background to the discussion. Chomsky's copy theory of movement is introduced and contrasted with the older Principles and Parameters (GB) notion that movement creates empty categories; I then describe the motivation for the copy theory within the Minimalist framework which, as already indicated, involves LF "reconstruction" phenomena; finally, I set out the

problems which the copy theory potentially gives rise to in the computational system, and introduce the main proposal of this chapter, which I call the *Copy Hypothesis*: all copies in a chain are active in the computational system. In section 3, I show how necessary constraints on trace activity are provided by independent general properties of movement itself. In section 4, the Copy Hypothesis is defended against a potential objection, in the form of the above-mentioned proposal by Chomsky (1995) that traces of A-movement are "immobile". In Section 5, I discuss the question of movement *to* traces. Finally, in Section 6, I consider the question of copies and derivations in the context of the question of whether the grammar is a derivational or a representational system. Section 7 concludes.

2 The copy theory of movement and its consequences

2.1 Full versus empty categories

Chomsky 1993, 1995 proposes that the operation Move involves the creation of a *copy* of the element it applies to:[1]

(1) **Copy theory of movement** (Chomsky 1995:251):
"A two-element chain is a pair <α, ß>, where α = ß"

Given (1), the syntactic (LF) representation of a case like (2) is assumed to include a chain something like (3); the lower element of the chain is null in the phonetic sense only:

(2) John was seen

(3) $<[_{NP}\text{John}]_i, [_{NP}\text{John}]_i>$

The copy theory as in (1) contrasts significantly with the conception of movement in earlier Principles and Parameters models. There, an application of the rule Move-α was taken to create an element known as a *trace*, described by van Riemsdijk and Williams 1986 in the following terms: "a trace is a syntactic category (such as NP) that has been

[1] This idea had already been proposed in the literature; see for instance Burzio 1986.

voided of phonological content and internal structure, retaining only an index that is identical to the index of the material that was moved out of the trace position" (p.139). The assumption that traces exist was motivated to a great extent by the Projection Principle and the Θ-criterion, and in keeping with this theoretical role, the empty elements were generally assumed to have the properties necessary to satisfy these requirements - namely the same category and referential index as the displaced element - but nothing else besides.[2] Under the trace theory of movement, then, (2) would be taken to include a chain as in (4), containing the element *John* and a coindexed empty category of the type NP:

(4) $<[_{NP}\ \text{John}]_i, [_{NP}\ e]_i>$

The empty categories which resulted from movement had independent theoretical status in the GB model, and were regulated by their own licensing conditions, embodied mainly in the Empty Category Principle (ECP).[3]

Continuing to generally contrast the trace and copy theories of movement, notice that the copy theory raises the following question: since the members of a chain are identical to one another, why is it that even in the case of overt movement, only one,

---

[2] Chomsky 1981: "One might say that trace theory in its essentials is nothing other than the minimal way of satisfying these requirements, taking the coindexed NP that satisfies the projection principle to be maximally simple, i.e. to have no unmotivated properties" (p.31). Note that similar empty categories were also allowed to be base-generated in Spec-IP at D-structure, eventually to be substituted by movement of a full category, for similar reasons, i.e. the Extended Projection Principle - see e.g. Chomsky 1986a.

[3] In the words of Chomsky (1986a:155), the ECP "imposes certain narrow 'identification conditions' on empty categories". One of the various formulations of this requirement is Rizzi's (1990) so-called conjunctive ECP, which states that "A nonpronominal empty category must be (i) properly head-governed (Formal Licensing) (ii) antecedent-governed or Theta-governed (Identification)" (p.32). Recall from the last chapter that within Minimalism, ECP-type effects are, in theory, accounted for in terms of constraints upon movement itself (Minimal Link Condition). A further condition on traces was the Proper Binding Condition (Fiengo 1977), Lasnik and Saito 1992. For an attempt to derive the latter from derivational economy considerations, see Collins 1994.

rather than e.g. all, of the copies in the chain actually "surfaces" at PF? That is, why do we have (2) instead of (2'):

(2') John was seen John

The answer suggested by Chomsky is that all but the highest copy in a nontrivial (i.e. with more than one member) chain entering Spell-Out is "deleted by a principle of the PF component" (1993:35). In the case of movement which is covert, this PF principle obviously does not delete the phonological features of the single-copy item which is input to Spell-Out. It may be the case that the PF deletion operation which applies to overtly formed chains is subject to economy considerations of some description, although these would presumably be of a different type to the derivational economy conditions described in the last chapter, given the assumption (see e.g. Chomsky 1994) that the mapping to PF involves procedures completely different to those of the computational system C_{HL} which creates LF representations.[4]

2.2 Copies at LF

The primary motivation for adopting the copy theory of movement within the Minimalist framework involves various "reconstruction" phenonmena,[5] in which, generally speaking, a moved element, or part of it, behaves for the purposes of interpretation as if it were in a different, lower, position to where it outwardly appears. One example of this phenomenon, discussed by Chomsky 1976/7 and van Riemsdijk and Williams 1986, among others, concerns *wh*-movement which pied pipes (takes with it) extra material with the *wh*-element (operator) itself. Examples of this are given in (4) and (5):

---

[4] On the other hand, see Kitahara 1994 who assumes that deletion of the phonological matrix of the lower copies of overtly formed chains does indeed count as an operation when it comes to calculating derivational economy. Kitahara's theory is critically reviewed in Chapter 6 below.

[5] Hereafter I use the term *reconstruction* (*phenomena/facts* etc.) to refer to the data in question, rather than a particular analysis of it.

(4) [which person's mother]$_i$ does Mary like t_i
(5) [in which house]$_i$ does John live t_i

In both (4) and (5), a part of the moved phrase - (*'s*) *mother* and *in* respectively - is clearly not part of the operator-variable interpretation itself; these elements belong in the "nuclear" part of the formula. Thus, a rough but adequate approximation of the interpretations associated with (4) and (5) is as follows:

(6) [which x] [x a person] [Mary likes x's mother]
(7') [which x] [x a house] [John lives in x]

Problems of a similar nature arise with respect to Binding theory (see for instance Barss 1986, 1994). Consider (8) (taken from Chomsky 1993:37):

(8) John$_i$ wondered which pictures of himself$_{i/j}$ Bill$_j$ saw

In (8), as indicated by the subscripts, *himself* may take either *John* or *Bill* as its antecedent, although, as it appears, only *John* is in the appropriate position, given Principle A of the Binding theory.[6] Consider also (9) and (10) (also from Chomsky 1993, p.40):

(9) * John wondered [which pictures of Tom$_i$] he$_i$ liked?
(10) * John wondered [which pictures of him$_i$] Bill$_i$ took

In (9), *Tom* cannot be interpreted as coreferential with *he*, suggesting a Principle C violation analagous to that in (11). Likewise in (10), *Bill* and *him* cannot be coreferential, suggesting a violation of Principle B similar to that in (12):

(11) * He$_i$ likes pictures of Tom$_i$
(12) * Bill$_i$ took pictures of him$_i$

[6] For details of which see e.g. Chomsky 1986a.

The problem is that (9) and (10) do not appear to involve configurations which should lead to violation of Principles C and B respectively. In earlier Principles and Parameters models, the unexpected interpretive properties of the examples discussed here were standardly dealt with by claiming that the overtly fronted *wh*-phrases move back, or lower, to their original position by LF, "undoing" the movement and resulting in the appropriate configurations as far as interpretation is concerned (the lowering analysis was first proposed in relation to the pied-piping problem, by Chomsky 1977). In (9), for example, the phrase *which pictures of Tom* returns to the VP-internal position it initially occupied, from where it is indeed c-commanded by *he*, with which *Tom* consequently cannot be interpreted as coreferential.[7]

But it is not clear how such problematic facts can be dealt with within the Minimalist framework. An approach making use of the notion of variation in the syntactic level at which the relevant interpretive principles apply, developed by some authors to deal with the binding cases,[8] is not really compatible with the Minimalist framework, simply because only one such level is taken to exist, i.e. LF (although see Barss 1994 for a derivational theory of the binding data within the Minimalist framework). Moreover, it would certainly be preferable to avoid invoking exceptional movement of a phrase back into a base or intermediate position in order to deal with these problematic interpretive facts - although it is another question to what extent such operations are actually excluded in principle in the Minimalist framework as it stands.[9]

---

[7] See also May 1985 on scope reconstruction effects with A-movement, illustrated in (i) (= May's (17), p.97):

(i) A hippogryph is likely to be apprehended

There is an interpretation of (i) in which the subject is inside the scope of the raising predicate. May proposes that this interpretation is obtained by what he calls quantifier lowering. On reconstruction effects involving A-movement and Binding, see Belletti and Rizzi 1988.

[8] See e.g. Belletti and Rizzi 1988 on certain Principle A reconstruction effects.

[9] Chomsky proposes a C-command Condition (1995:253) which states that "α must c-command its trace, so that there cannot be an operation that lowers α or moves it 'sideways'", although, as he presents it, the prevention of lowering or sideways movements appears to be the only

Against this background, we can now fully appreciate the usefulness of the copy theory of movement in dealing with the reconstruction phenomena discussed previously in this section. Consider firstly the pied-piping example (4), and the informal depiction of its interpretation, (6), both repeated here:

(4) [which person's mother]$_i$ does Mary like t_i
(6) [which x] [x a person] [Mary likes x's mother]

According to the copy theory of movement, (4) corresponds to a structure something like (4') below; here obviously the pied-piped material is simultaneously present in its original position, accounting for the interpretation:[10]

---

motivation for this condition, i.e. it does not seem to be the case that lowering is ruled out by any deeper principle. It is possible that lowering into previously-occupied positions - the relevant case for the reconstruction facts - might be excluded by Last Resort, particularly the version of Chomsky (1995) according to which all movements must result in feature-checking (see § 3 below). Given this, one place to which an element should be unable to move to is one that it has occupied previously, since such a position is either (i) a position in which the element is unable to check features at all (i.e. its original position) or (ii) a position where it moved in order to check features in the first place (that is, an "intermediate" position) and where, therefore, all features which could possibly be checked by the element in question, will have been checked by it. This still leaves the question of lowering to positions not previously occupied, although Minimal Link Condition and/or Extension Condition might prevent such a situation from arising (see Chapter 4, note 9 for a concrete example). See also Collins 1994 for an attempt at excluding lowering operations by economy conditions.

[10] Mirroring the situation with PF, the copy theory "overdetermines" LF. Since (4) and (5) involve an operator-variable interpretation, it must be assumed that chains like that of *which person's mother* in (4) are at some stage subject to removal of the operator part of the lower copy, and piedpiped material (*'s mother*) in the higher copy, as well as the material corresponding to the restriction on the operator in one or other of the copies. Chomsky argues that the formation of actual operator-variable constructions is effected by deletion, driven by FI: "for convergence at LF, we must have an operator-variable structure. Accordingly, in the operator position...everything but the operator must delete" (1993:36). Unlike the operation which deletes the phonological features of traces, this deletion, which Chomsky characterizes as "an operation akin to QR" (*op. cit.*:35) is

(4') [which person's mother] does Mary like [which person's mother]

Let us now look at how the copy theory of movement similarly provides an account for the binding examples (8), (9) and (10), repeated here:

(8) John$_i$ wondered [which pictures of himself$_{i/j}$]Bill$_j$ saw
(9) * John wondered [which pictures of Tom$_i$] he$_i$ liked
(10) * John wondered [which pictures of him$_i$] Bill$_i$ took

Again, according to (1), movement of the *wh*-phrases creates copies, as illustrated in (8'), (9') and (10') respectively (with some irrelevant details omitted), "providing the materials for 'reconstruction'":

(8') John wondered [$_{CP}$ [$_{wh}$ which pictures of himself] Bill saw [$_{CP}$ [$_{wh}$which pictures of himself]
(9') John wondered [$_{CP}$ [$_{wh}$which pictures of Tom] he liked [$_{wh}$which pictures of Tom]]
(10') John wondered [$_{CP}$ [$_{wh}$which pictures of him] Bill took [$_{wh}$which pictures of him]]

To account for how the necessary configuration for binding arises, the claim is that in (8'), along with obligatory deletion of the operator itself in the lower copy, the restriction -

---

obviously syntactic, i.e. an operation of the computational system C_{HL}, since it is taken to take place before LF, in order to satisfy FI at that level. Brody 1995 alternatively argues that the deletion which forms actual operator-variable constructions is not syntactic. In his system, actual operator-variable constructions are not required by FI. He proposes an alternative to FI, Partially Determined Full Interpretation (PDFI). Post-syntactic processes are responsible for creating actual operator-variable constructions. Irrespective of other differences between Chomsky's and Brody's Minimalist frameworks (see § 6 below for discussion of the latter), the post-syntactic formation of operator-variable constructions assumed by Brody (a necessary assumption in fact, since in his framework there are no derivations in the usual sense) appears to be more consistent with the restrictive notion that syntactic operations are "driven by the narrow mechanical requirement of feature checking only". In addition, as we have seen, there is a certain amount of optionality as to which parts of which copy get deleted, a fact which is perhaps more difficult to explain if deletion is an FI-driven operation of C_{HL} and as such, subject to economy considerations of some sort.

pictures of himself - is deleted either in the lower copy, which leads to the reflexive being bound by *John*, or in the higher copy, in which case the reflexive is bound by *Bill*. In (9') and (10') the requisite configurations for Principle C and Principle B violations are met if deletion of the restriction applies in the higher copy, i.e. if reconstruction takes place.[11]

2.3 Copies in C_{HL}

So far, we have seen that there is a strong case for taking "movement" to involve a copying operation; this provides a straightforward basis for explaining reconstruction-type interpretative phenomena. Moreover, as a result of fundamental properties of the Minimalist model (notably the existence of only one syntactic level of representation) there is not obviously any satisfactory alternative method of accounting for these facts; it seems then that the copy theory is indispensable within a Minimalist framework. Having seen this, let us now investigate the ramifications of the copy theory in another domain: the computational system itself. Here it seems that serious problems arise.

Consider a two-element chain $<\alpha, \beta>$. Given the copy theory of movement as it stands (see (1) above), the members of a chain are inherently identical, differing only in

---

[11] Although generally well motivated, it is well-known that the copy theory does not in and of itself explain why reconstruction is optional in the case of (8) (Principle A) but obligatory in the case of (9) and (10) - (Principles B and C), any more than the pre-Minimalist approaches do. Consider e.g. (i):

(i) ?* Which pictures of $John_i$ does Mary think that he_i likes?

The copy theory predicts that (i) there should be at least one copy of the *wh*-phrase not c-commanded by *he* and in which a coindexed *John* should therefore be able to appear without violation of Principle C - cf. (ii).

(ii) $[_{CP}\ [_{wh}$ ... John] .. does Mary think $[_{CP}\ [_{wh}$ ... John] ... that he likes $[_{wh}$ ...John]]]

But, in contrast to examples with reflexives like (8) in the text, deletion of the restriction (containing *John*) in the top copy seems to be obligatory, in other words, reconstruction to the base position is for some reason the only possibility, making violation of Principle C impossible to avoid. See Chomsky 1993 and Brody 1994/5 for different approaches to these questions.

the structural position they occupy. If β in the chain <α, β> is a category with exactly the same properties as α, as opposed to an empty category, then it is reasonable to assume that there can be no syntactic operations or principles which are allowed to apply to β, but not to α (or indeed vice versa). In other words, it seems that if α is active in the computational system, then β is likewise, in principle at least. With respect to movement (i.e. copying) in particular, α and β, since they are identical, will share the same feature-properties and so must both be potential subjects for further applications of movement; if α possesses certain checkable features, then β must also possess them.[12]

It seems that if non-head copies of chains (hereafter referred to as "traces" for convenience) are themselves eligible for further copying operations, then there will be a significant increase in the number of possible derivations generated by the grammar (cf. if lowering/sideways movements). To keep unwanted derivations at bay, it might then be felt necessary to introduce some kind of rule which prohibits computational operations from applying to traces, or a subset of them - as in Chomsky's (1995) proposal about traces of A-movement (addressed in detail in § 4 below).

However, as stressed already, it seems theoretically preferable (if not obligatory) to keep to the idea that movement can "move anything..." (cf. Chomsky 1986:74). I assume then that Move is an operation which can in principle apply to any copy in a chain, subject of course to any conditions which generally regulate this operation (i.e. Last Resort, Minimal Link Condition, and so on - see Chapter 1 above). From now on, this will be referred to as *the Copy Hypothesis*.

(13) **The Copy Hypothesis**:

All copies of a chain are active in the computational system

In due course, we shall see that the Copy Hypothesis has direct empirical applications (chapters 4,5,6). In the rest of the present chapter, I argue for the Hypothesis from a more

[12] In addition (see note 10 above), it is assumed by Chomsky that a syntactic process of deletion, driven by the principle of Full Interpretation, operates on both α and ß, potentially removing material from each copy in order to create an operator-variable construction.

theoretical perspective and address some important potential arguments against it, in the form of recent proposals by Chomsky (1995).

3 Arguments for the Copy Hypothesis

I prepare the ground for this section by firstly giving a detailed description of recent Minimalist assumptions about the nature of movement and the conditions constraining it. It will then be possible to show explicitly that the movement possibilities of traces are already severely restricted, strongly indicating that conditions specifically designed for this task may not be necessary.

3.1 Chomsky's (1995) theory of movement

The theory of movement developed in Chomsky 1995 differs in a number of respects from the 1993/1994 version of the Minimalist framework outlined in chapter 1. One change is the division of formal features into two types: interpretable and uninterpretable. This has important movement-related consequences, as we shall see shortly. Interpretable features are those which have an interpretation at LF: they include the categorial features of lexical elements (V, N, A etc.) and the ϕ- (agreement) features of argument forms, i.e. number, gender and person. Uninterpretable features are those without any import at LF:[13] they include all Case features, as well as ϕ-features associated with predicative elements (verbs and adjectives). Uninterpretable features also include strong features which (as mentioned in Chapter 1) are characterized by the property of PF uninterpretability.[14]

[13] It may be noted in passing that the property of uninterpretability, i.e. pure "formalness", is taken to be a legitimate notion in the case of *features* but not in the case of "formatives" such as AGR (cf. Chapter 1, § 3.3). See Brody 1995 for a proposal that there is no ±Interpretable difference among features; all are interpretable.

[14] Recall that weak features are invisible at PF, i.e. neither interpretable nor uninterpretable there.

Generally, the formal features which constitute checking domains will always be uninterpretable, whilst formal features *in* the checking domain (i.e. those which move there) may be interpretable or uninterpretable. What these latter facts might ultimately follow from is not clear, as Chomsky notes; in any case, the practical significance of the proposed interpretability distinction for movement is the fact that uninterpretable features must be eliminated by LF (and also by PF, in the case of Strong features), while interpretable features by comparison are not subject to this requirement: these features need not, and therefore must not, be eliminated. Nevertheless, interpretable features still have the *ability* to check the uninterpretable features of another element. But since they may not be eliminated, interpretable features may in principle enter into checking relations on as many occasions as there exist suitable uninterpretable features - subject to various constraints on movement, which we discuss next.

Recall that in the model of Chomsky 1993/94 (outlined in Chapter 1), Last Resort and the Minimal Link Condition (MLC) have the status of derivational economy conditions. The basic *modus operandi* of such conditions is to consider the convergent derivations from some numeration (initial set of lexical items) and select that (or maybe those[15]) which is (are) optimal with respect to some specified dimension - for instance, distance of movement, in the case of the MLC. The framework of Chomsky 1995 differs from this in that both Last Resort and MLC become part of the definition of Move itself, rather than being separate conditions. This innovation is motivated primarily by a perceived need to reduce computational complexity in the model - recall that it is inherent in the nature of derivational economy conditions to induce this.[16] Incorporated into the definition of Move, Last Resort and MLC become absolute rather than global conditions. If some movement operation fails to obey either of them, the derivation simply crashes.

Another difference from the 1993/94 model is that Move is reformulated in terms of the position/feature moved to, and not in terms of the element which moves, as was

---

[15] See Chomsky 1991, and Chapters 5 and 6 of this thesis.

[16] See Chapter 1, § 4.2.4; Johnson and Lappin (forthcoming) for an in-depth discussion of computational complexity and Chomsky's Minimalist framework; also Chapter 4 of this thesis.

formerly the case. Correspondingly, the operation is sometimes referred to as Attract or Attract/Move. Attract/Move is defined as follows, incorporating Last Resort:

(14) "Move raises feature F to target K only if F enters into a checking relation with a sublabel of K as a result of the operation" (Chomsky 1995:280)

Here, a *sublabel* of K refers to a feature belonging to the head (i.e. the label) of K, or to any other head/s which are adjoined to that head. Features which may enter into a *checking relation* are (i) unchecked uninterpretable features, and (ii) interpretable features (see above). In conjunction with certain further assumptions, given in (15), the definition of Attract/Move in (14) deals with the range of cases formerly attributed to the principles of Last Resort and Greed in the 1993/94 model (see Chomsky 1995:284 for details).

(15) a. A checked feature is deleted where possible
b. Deleted α is erased when possible (1995:280)

Deletion is defined as an operation which renders its subject "invisible at LF but accessible to the computation" (*op.cit.*:280). Deletion contrasts with *erasure*, which amounts to complete, irrevocable eradication of the element it applies to, rendering it invisible to the computation as well as at the interface. As with Attract/Move, these processes of Deletion and Erasure are claimed to be subject to economy, which becomes inoperative (can be overridden) should the derivation fail to converge. Deletion (and therefore erasure) are not possible for features which are interpretable;[17] nor can they apply to units larger than features, i.e. categories, since these, by hypothesis, cannot generally be lacking in interpretative import of some sort (see chapter 1, §2). As for uninterpretable features, Deletion and Erasure must apply to them when this is possible. The only occasion where it is not possible, that is, where erasure would prevent the derivation from converging, is

[17] Not surprisingly, there is also a kind of converse requirement to this, the *Inclusiveness* condition (Chomsky 1994, 1995) which requires that LFs should contain no material which was not already present in the numeration, i.e. nothing can be added in the course of the derivation, such as indexes, bar-levels etc. Inclusiveness applies to the computation from numeration to LF but not to the operation Spell-Out.

the case of so-called pure expletives exemplified by English *there*, German *es*. These items are semantically (though of course not phonetically) null, being constituted entirely of an uninterpretable categorial feature, which therefore cannot delete and erase.

Finally, let us see what becomes of the Minimal Link Condition (MLC) in Chomsky's (1995) model. This too is incorporated into the definition of Attract/Move, as follows:

(16) "K [= a target for movement] attracts F if F is the closest feature that can enter into a checking relation with a sublabel of K" (Chomsky *op.cit.*, p. 297)

Among the features which can enter into a checking relation (as determined by Last Resort - see (14) above), a feature β counts as *closer* to the target K than a feature α iff:

(17) a. β c-commands α, and
b. β is not in the same Minimal Domain as (i) τ or (ii)α

(rephrased from Chomsky 1995:357), where τ is the target of raising (adjunction to head of K, or Spec of K). (17b) is the "Equidistance" clause (see chapter 1),which it will be recalled is needed under the assumptions that subjects and objects originate within a VP to the exclusion of all functional categories, and that structural Cases are uniformly checked by movement to functional projections - both more or less standard in the Minimalist framework.[18]

There is one further change in the 1995 model, not particularly relevant to the concerns of this chapter, but which I include in this description since it will be relevant at later stages. Chomsky proposes that it is features themselves rather than categories are the

[18] Although see Koizumi 1993, 1995 for a proposal that subjects and objects, and in fact, even the internal arguments of ditransitive verbs, are each generated in a separate VP dominated by an Agreement Phrase of its own; as Koizumi notes, this apparently eliminates the need for Equidistance. Cf. note 26 below. Koizumi's theory is discussed in a different context in Chapter 3 of this thesis.

basic input to movement operations: "we should restrict α in the operation Move α to lexical features" (262). Thus, features, or more precisely those of them which participate in checking, referred to by Chomsky as *formal features*, are in some sense autonomous elements syntactically. As for how to account for cases where it is obvious that whole categories, rather than just features, are moved (that is, what we think of as overt movement) Chomsky maintains that movement of a whole category is caused indirectly by PF considerations, as is overt movement in general. The idea is that, at least in the cases where full categories are seen to move, formal features alone do not qualify as legitimate objects at PF, and therefore the category to which the features belong must accompany them on an overt movement, otherwise the derivation will crash. Furthermore, movement of a full category is not permitted unless convergence depends on this: "The operation Move F carries along 'excess baggage' only when it is 'heard' in the phonetic output" (265).[19,20]

---

[19] This "natural economy condition" is not the whole story about overt pied piping, as Chomsky himself notes. The question arises, for instance, as to why there is sometimes optionality in how much material moves overtly - for instance, in English, prepositions may pied pipe or be stranded: *which house did John live in? vs. in which house did John live?*, and the various sites in which "floating quantifiers" can appear (assuming the stranding analysis of Sportiche 1988): *all the children will have left vs. the children all will have left* vs. *the children will all have left* vs. *the children will have all left.* Of course, if floating quantifiers were analysed as adverbial elements (as in e.g. Bobaljik 1995), the optionality problem would not arise in that case (see chapter 6, § 5, note 39).

[20] The introduction of the concept of pure feature-movement makes it possible to see overtness of a movement in terms of whether the visible parts of an element move, instead of when the movement happens relative to Spell-Out , which in turn leads to the possibility of Spell-Out applying directly to LF, which may be seen as a simplification of the grammar (see Brody 1994). Chomsky seems to find it necessary to employ both the notion of pre-Spell-Out movement and movement of a full category, although it looks as if there should be some redundancy here. An interesting alternative framework is proposed by O'Neil and Groat 1995, using an idea extensionally similar to Chomsky's Move F. In their system, "the difference between "overt" and "covert" movement of a given category is not expressed in the timing of movement with respect to Spell-Out, but rather with respect to where a category is pronounced in its chain... chain formation, invariably

Further to this, Chomsky proposes that when a particular feature of some element moves (is copied), the rest of the formal features associated with that element are "automatically" moved with it. "Applied to the feature F, the operation Move...creates at least one and perhaps two "derivative chains" alongside the chain $CH_F = (F, {}_F t)$ constructed by the operation itself." (265).Chomsky calls these features *free riders.* What *automatically* seems to mean in this context is that the situation could not be otherwise (though it is unclear exactly why), and accordingly, economy considerations are irrelevant.

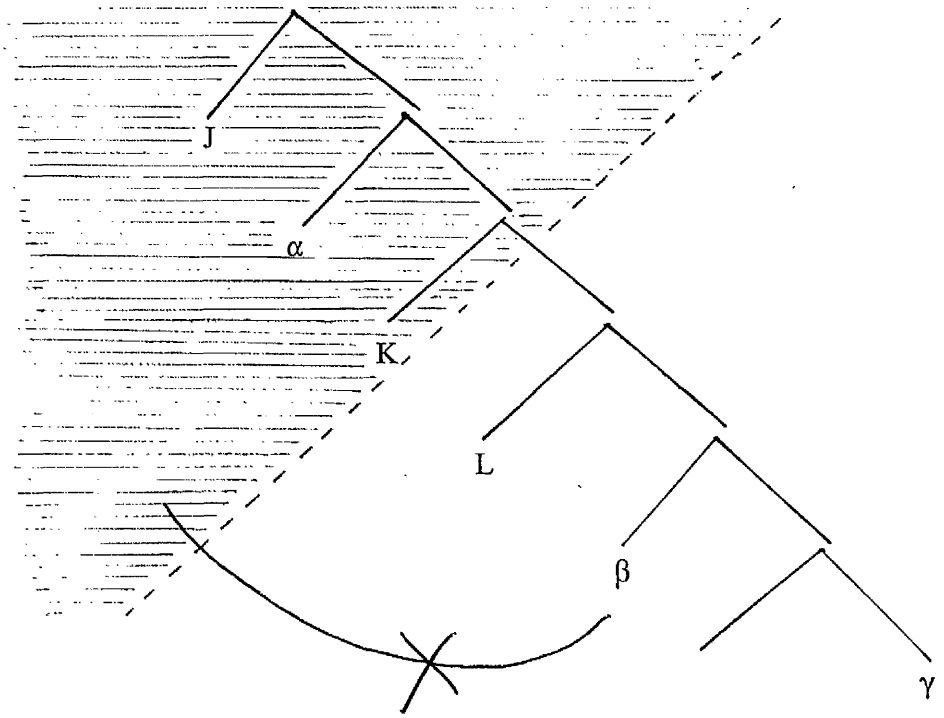
3.2 The independent constrainedness of trace activity

We are now in a position to demonstrate that the non-head copies in a chain (traces) have very limited opportunities for further movement, as a direct result of how movement itself is defined. In (18) is an abstract illustration of the movement possibilites of the trace β in a chain $<\alpha, \beta>$, created by attraction of an element by a feature associated with a head K:

---

a reflex of feature checking, copies phonological material to the head of the chain only when the features to be checked are strong". A theory similar in spirit is proposed by Brody 1994.

(18) Movement possiblities of β in the chain <α, β>



Firstly, notice that Last Resort (see (14) above) ensures that if β is to be a candidate for Attract/Move by/to any target (K or otherwise), then β must possess features which are able to enter into a checking relation. In other words, β must possess either (i) interpretable features (e.g. categorial), or (ii) uninterpretable features (e.g. Case) which were unable to be checked with any features contained in K and have therefore not been deleted and erased. In a fairly trivial way, the mere fact of being a trace in the chain of some element entails that feature-checking relating to that element has already taken place, thus eliminating at least some checking (hence movement) potential. Further, if the trace β does still have some checkable features of its own, there must then exist a target containing appropriate features with which it can enter into a checking relation. Such a target can never be K itself, since all features which can possibly be checked there by the element *will* have been checked there automatically as free riders (and subsequently erased, since the formal features of the target are always uninterpretable - see above). The Minimal Link Condition further reduces the possible targets for trace movement, as we will see below.

Let us firstly look at an actual example in which movement of β in a chain <α, β> is ruled out by the Last Resort part of the definition of Attract/Move, in the way just described abstractly. Such an example is in fact discussed in Chomsky 1995 and is given in (19):

(19) There_i seem t_i to be some books on the table

In *there*-constructions like (19), it has been argued that the associate of the expletive - here, *some books* (strictly speaking, its formal features) raises covertly to the matrix T(P), presumably to check Case and ϕ-features.[21] The question is, then, is it falsely predicted that the lower copy in the chain of *there*, in the embedded Spec-TP, should count with respect to movement and hence block the required movement of the associate's features up to the matrix T?

By Last Resort, the chain <*there*, *there*> in (19) is necessarily the result of feature-checking: in this case, of a strong feature of the matrix T. *There*, or more accurately, its two copies, will still possess their categorial feature, since these features are interpretable and cannot delete (see above). However, the strong feature of T itself will have been deleted and erased, meaning that no further element - including crucially the trace-copy of *there* - will possibly be able to check that particular feature. The head T may, and indeed does, contain further as yet unchecked uninterpretable features, namely a Nominative Case feature, as well as agreement features contained in a sublabel (V). However, this is immaterial here, since recall that when Attract/Move copies an element to some site, all features which can be checked in the process, are checked automatically as "free riders". This means that in this situation, if *there* was capable of checking the Case/ϕ-features of T, it would have checked them in the operation which formed the

[21] The reason concerns Binding theory. The associate of *there* seems to function for purposes of Binding and control as if it were in the matrix subject position. The same is apparently true of German *es* and Italian null expletives, but not of French *il* (see Chomsky 1995:274, although also Lasnik 1996 for a different view). For more about expletive-associate constructions, see chapter 5 below.

chain <*there*, *there*> in the first place. By hypothesis, *there* is not capable of checking Case and ϕ-features. Hence, neither of the two copies of its chain can enter into a checking relation with any sublabel of T, so that raising of *some books* is not in fact blocked by the trace of *there* - the necessary result. In this case, then, as is clear from Chomsky's (*op. cit.*:302) discussion, the very definition of Attract/Move itself correctly ensures that the trace will not be able to move to T itself, hence cannot count as blocking movement of any lower element there.

Now returning to the illustration (18), note that in addition to not being K itself, a possible target for Attraction/movement of β must be in such a position that it c-commands β but not α, since otherwise, attraction of β will be blocked by α itself, as the latter is automatically closer to the target by the Minimal Link Condition (see (16) above). This of course rules out J (or any head above it) as a target for β. The shaded area in (18) corresponds to sites to which movement of β is trivially and straightforwardly excluded by Last Resort and MLC, as described. What we are left with is the area underneath K: in the event that β has features which can enter into a checking relation with those on some element in this area (say L), is it predicted that the trace β can be attracted by these features?

There will in fact be more than one actual instantiation of this scenario, and in some circumstances, we shall see that it is advantageous for β to count for the purposes of Attract/Move (see Chapters 4, 5 and 6 of this thesis; also Chomsky 1995). But for the time being, let us concentrate on cases in which the trace ought not to count for Attract/Move. The main situation which this corresponds to involves A-movement (i.e. Case-checking) - and as we shall see in the next section, it is indeed traces of A-movement which Chomsky (1995) proposes must be stipulated to be "immobile".[22]

[22] Recall incidentally that in the 1993/1994 model outlined in chapter 1, which differed from the current model in certain respects, it was crucially assumed that the trace of the (VP-internal) subject (=β in (17)) *does* count as blocking movement of the object (= γ) to AGR_o (= L), if the verb has not raised to AGR_o (Holmberg's Generalization).

Let us suppose that <α, β> is the chain of a subject, K is T(P), and γ is an object, as depicted in (20) below. The required outcome is that the object is attracted by features, notably Accusative Case, contained in the complex formed by raising of V to the light verb *v*. The subject, i.e. both copies of its chain, will still possess interpretable categorial and *ϕ*-features which can potentially enter into a checking relation with features contained in the verbal complex. The question we need to ask is the following: is the trace of the subject falsely predicted to get in the way of attraction of the object to the verbal complex? It is easy to see that the answer is "no", due to the Minimal Link Condition, repeated here.

(16) "K [= a target for movement] attracts F if F is the closest feature that can enter into a checking relation with a sublabel of K" (Chomsky *op.cit.*, p. 297)

where a feature β counts as *closer* to the target K than a feature α iff:

(17) a. β c-commands α, and
b. β is not in the same Minimal Domain as (i) τ or (ii)α

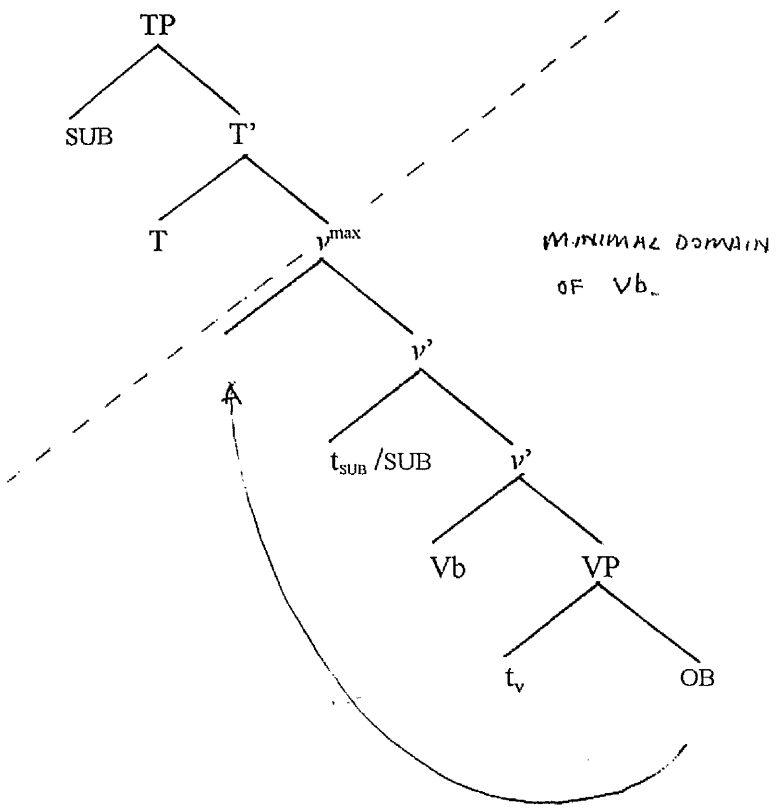
According to the Equidistance clause of the MLC as in (17b), the subject trace does not qualify as closer than the object to Spec-*v*, since the subject trace and Spec-*v* are in the same minimal domain, i.e that of "Vb". The critical point to be noted here is that the notion of Equidistance is required for independent reasons, namely to ensure that object movement is not blocked by the presence of the subject itself in Spec-VP, as required for overt Object Shift languages such as Icelandic. Hence in this case too, unwanted participation in movement by a trace is ruled out by general properties of the Attract/Move operation itself which are uncontroversially necessary in any case.[23, 24]

---

[23] On the further question of why the object does not check its features with the trace of V directly, see §5.1 below.

[24] In connection with this point, one should note that an alternative "stacking" view of clause structure has lately been advocated (see Koizumi 1993, 1995, Bobaljik 1995), according to which subject and object originate in separate verbal projections associated with their own functional structure. Such an assumption about clause structure would obviate the need for a notion of Equidistance, at least with respect to the situation at hand (i.e. the respective Case-movements of

(20)



Summing up this section, it has been shown how general properties of the operation Attract/Move itself are such that the copy β in a chain <α, β> will to a very great extent be irrelevant for subsequent operations of Attract/Move. The same point will apply to any of the non-head copies in a chain, should there exist chains with more than two members. The conclusion reached here suggests that it is safe to assume that all copies of a chain are active in the computational system, i.e. the Copy Hypothesis stated in (13) above: theoretically at least, there is no need to assume anything more

subjects and objects), and concomitantly render the issue of trace mobility null and void, again with respect to the situation at hand. The relevance of the trace mobility issue as a whole remains. However, the loss of motivation for Equidistance in the subject/object case needs consideration, since it is relied on at a later stage in the thesis (see Chapter 6). Thanks to Elabbas Benmamoun and Ian Roberts for drawing my attention to this point.

complicated.[25] In case this conclusion seems trivial, we should consider a proposal about traces and movement put forward by Chomsky 1995 which goes more or less directly against the Copy Hypothesis.

4 Potential arguments against the Copy Hypothesis

In the view of Chomsky 1995, it is necessary to state that certain traces, namely those of A-movement, are "immobile", i.e. unable to participate in movement operations: essentially the antithesis of the Copy Hypothesis. In this section, however, it will be seen that Chomsky's proposal is not well motivated.

4.1 Chomsky's (1995) proposal

Addressing the issue of traces potentially being active in the computational system, Chomsky makes a number of proposals which are collected together in (21) and (22):

(21) a. "raising of α heading the chain (α, *t*) deletes the trace formed by this operation that is, marks it invisible at LF...at LF, then, all that is "seen" is the chain...which satisfies the Chain Condition" (p. 301)

b. "The intermediate trace *t* of an argument cannot be attracted; hence, *t* does not prevent attraction of an element that it c-commands" (p. 301)

(22) a. "The operation Attract/Move can "see" only the head of a chain, not its second or later members"

b. "Only the head of a chain CH enters into the operation Attract/Move"

c. "Trace is immobile" (p.304)

Chomsky claims that the statements in (21) derive from independent considerations, in particular, an LF principle: the Chain Condition. We shall see shortly that this idea is

[25] In fact, I have not quite exhausted the relevant issues at this point - in particular, I have not discussed the question of traces as targets of movement, something which is also allowed in principle given the Copy Hypothesis. This is discussed in its own section, § 5 below.

somewhat unsatisfactory, its finer details and also, I think, in its general conception. The wider statements in (22), on the other hand, are described as "natural extensions" of (21); they do not seem to derive from any other principle.[26] As such, we would expect to find some good empirical motivation for adopting them. Unfortunately, the motivation which Chomsky provides is sparse, as well as problematic in a number of respects.

4.2 Problems with Chomsky's (1995) proposal

A first problem is found upon closely examining the motivation for the statements in (21). This involves considerations of LF well-formedness and goes as follows. Chomsky considers the case of "successive-cyclic" movement of an argument, as illustrated in (23) (adapted from his example (88)):

(23) John is likely [t_3 to be forced [t_2 to [t_1 submit his thesis]]]

(23) is potentially problematic due to the fact that all but one of the sub-chains relating to the element *John*, that is, <t_2, t_1> and <t_3, t_1> (but not <*John*, t_1>) ,violate an LF principle, the Chain Condition, which requires that "every argument chain must be headed by a Case position and must terminate in a θ-position" (Chomsky and Lasnik 1993). One way of getting around this difficulty would be to assume that the problematic traces, i.e. the intermediate traces t_2 and t_3, are for some reason exempt from the Chain Condition at LF. Recall (from §3.1 above) that Chomsky assumes a deletion operation which renders an element "invisible at LF but accessible to the computation" (1995:280). Making use of this notion he thus proposes (21a) - repeated here - and the problem involving the Chain

[26] Of course, I have been attempting to show (§ 3) that trace immobility *does* to a very great extent follow from independent considerations. From that perspective, it would make sense to think of (22) as independently motivated rather than stipulative. However, the important point about Chomsky's proposal, if I am not mistaken, is his conception that (22) must be separately stated. Chomsky does not appear to systematically pursue the possibility that something like (22) (subsuming (21)) could fall out from general constraints on movement.

Condition is thereby taken to be solved:[27]

(21a) "raising of α heading the chain (α, *t*) deletes the trace formed by this operation - that it, marks it invisible at LF...at LF, then, all that is 'seen' is the chain...which satisfies the Chain Condition."

The next step in the argument is that deleted material is "erased when possible", that is, it becomes invisible to the computational system, where *when possible* means basically when the material has no intepretation at LF and is not a term of the structure. Given this, "a formal feature of an intermediate trace may erase, and indeed must erase if possible" (*op.cit.*:301).On the basis of this, Chomsky concludes (21b), repeated here:

(21b) "The intermediate trace *t* of an argument cannot be attracted; hence, *t* does not prevent attraction of an element that it c-commands" (p. 301)

Yet here a problem seems to arise. Remember that among their formal features, arguments will always have some which are interpretable, for instance categorial and ϕ-features. Remember also that according to Chomsky (*op. cit.*, p.280), "interpretable features cannot delete even if checked. Hence, the question of erasure arises only for an uninterpretable feature F". Given this, the uninterpretable features of intermediate traces may erase, but

---

[27] Even if intermediate traces are invisible at LF as far as the Chain Condition is concerned, they are definitely present for some aspects of interpretation. This is so for both A'- and A-movement:

(i) Which pictures of himself$_{i/j}$ does John$_i$ think that Bill$_j$ likes?
(ii) John$_i$ seems to Fred$_j$ [t$_i$' to be likelier than him$_j$ [t$_i$ to leave]]

In the familiar reconstruction case (I), *himself* may take either *John* or *Bill* as antecedent. If the former, it seems that the *wh*-phrase has "reconstructed" to its intermediate position, Spec of the embedded CP. In (ii) (taken from van de Koot 1995; example attributed to M. Brody), *him* may be coreferential with *Fred*, indicating the presence of a trace of *John* in the highest of the two embedded clauses. It clearly needs to be explained how the notion of "invisibility" with respect to the Chain Condition which Chomsky invokes for intermediate traces fits in with the fact that they are taking an active role in Binding, which also applies "at LF" in some as yet to be clarified way. Chomsky addresses this issue briefly in a footnote (op. cit., p.387-388).

those that are interpretable will never be able to. And given that, it is difficult to see how the inability of intermediate traces to be attracted can derive from (21a) and ultimately, from the Chain Condition on LF.

A more general problem, I think, concerns the idea that a principle like (21a), which is a condition on operations of the *computational system*, can be motivated by, or have access to, purely LF-related considerations such as the Chain Condition at all. (21a) presumably has access to LF information, since it must be aware that certain traces, although they are formed by "raising of α heading the chain (α, *t*)" must be spared from deletion - namely, Case-marked traces left by *wh*-movement. These of course have to be visible so that the Chain Condition is *not* violated. Even if the principle with the effect of (21a) turns out to be necessary, it seems undesirable, if not actually unfeasible, to allow a constraint on syntactic operations to have the capacity to take LF wellformedness into consideration before applying.[28]

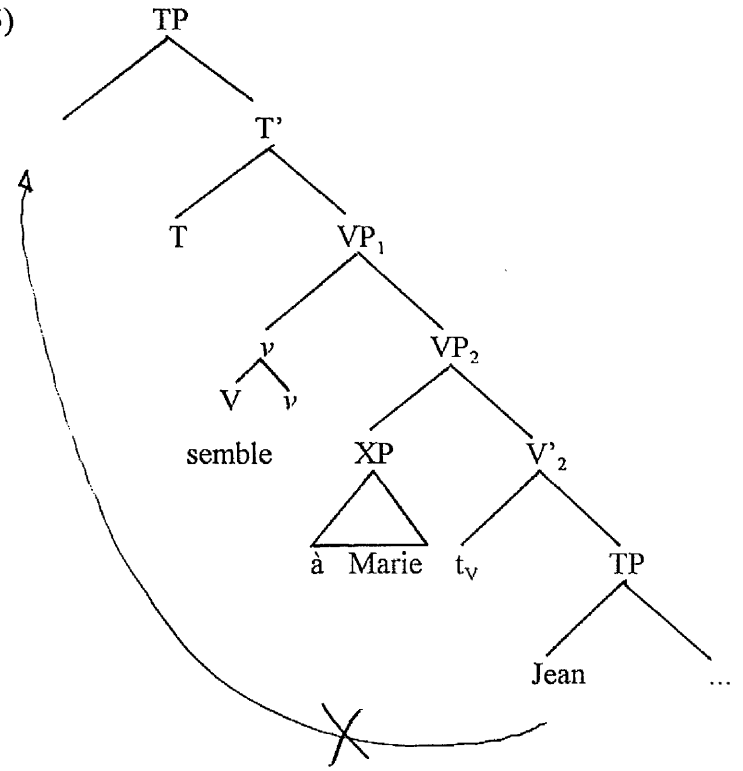
Another criticism which could be made of the statements in both (21) and (22). is that there seem to be very few actual cases where it is actually necessary to invoke particular principles to prevent traces from participating in movement, since, as was demonstrated in the previous section, the same result falls out from the definition of Attract/Move itself, incorporating Last Resort and MLC. It seems that there is not a single case in which it can be directly demonstrated that traces of A-movement must be specially "immobilized". The one possible candidate involves some interesting data from French which Chomsky discusses. Consider the grammaticality contrast in (24); the construction in question is analyzed as "a Larsonian shell, with [*semble*] raising to the light verb *v*" (*op. cit.*, p.305) as illustrated in (25) below.

(24) a. * Jean$_i$ semble à Marie [t$_i$ avoir du talent]
J. seems to M. to-have talent
b. Jean$_i$ lui$_j$ semble t$_j$ [t$_i$ avoir du talent]
J. to-her seems to-have talent

[28] This aspect of the issue is discussed in Pettiward 1995.

Chomsky assumes, crucially, that the ungrammatical (24a) involves an MLC violation: *Jean* cannot legitmately be attracted to the matrix T due to the intervention of *Marie*, which is a rival candidate for attraction to that site since it possesses features (e.g. a categorial feature) which could enter into a checking relation with a sublabel of T (Last Resort). But because *Marie* is not able to check the Case feature of T, its own Case property being presumably satisfied in some relation with the element *à*, the derivation cannot converge.[29] (24a), in other words, is seen as a kind of SuperRaising scenario.[30]

(25)



Chomsky then suggests that the grammaticality contrast between (24a) and (24b) hinges on the fact that in the latter, a *trace* - belonging to the Dative clitic *lui* - is in the position which *(à) Marie* occupies in (24a). Since, by Chomsky's hypothesis, traces are immobile,

[29] This incidentally is a good illustration of the absolute nature of the 1995 MLC: it cannot be overridden for convergence.

[30] See Rizzi 1990. The classical case of SuperRaising is * *John seems that it is likely to win.* As discussed briefly in Chapter 1 (§4), Chomsky (1993) proposes to account for such violations in terms of the Minimal Link Condition.

the trace of *lui* does not block raising of *Jean*, accounting for the grammaticality of (24b).

There are at least two reasons to question this explanation of the (24a)/(24b) contrast. The first is simply the fact that the equivalent of (24a) in English is grammatical, as Chomsky himself points out:

(26) John seems to Mary to have talent

It seems more likely that there is some particular property of French which disallows a full PP complement with the verb *sembler* 'seem', rather than it being the case that English for some reason permits violation of the MLC in certain situations.

A second and perhaps more interesting indication that MLC is not the relevant factor in the contrast between (24a) and (24b) emerges if we consider *sembler* constructions with the expletive *il*. It is assumed by Chomsky (*op. cit.*) that in constructions with *il*, no element raises to the position of the expletive.[31]

(27) a. * Il semble à Marie que Jean a du talent
b. Il lui semble que Jean a du talent

Yet interestingly, (27a) and (27b) exhibit the same contrast as (24a) and (24b). This suggests strongly that raising, or constraints relating to it, is not the factor responsible for the grammaticality difference in (24) (although of course it still need to be explained why elements claimed by Chomsky to be specifiers of the second VP - e.g. *Mary* in (26), *Marie* in (24a), (24b) - do not count as blocking movement of another element from the embedded to the matrix clause, despite the fact that elements in this specifier position

[31] See Chomsky *op. cit.*:274, for an argument that expletive *il* is able to check Case and ϕ-features, based on binding facts. For more on the syntax of *il*, see also this thesis, chapter 5. Thanks to Jeanne Cornillon for judgements on (27).

apparently c-command into the embedded clause).[32, 33]

In summary, it does not seem that the French data in (24) provides convincing motivation for Chomsky's proposal that trace is immobile. In other words, keeping to the null hypothesis (13) above, i.e. "all copies of a chain are active in the computational system", will not lead to any problems with respect to this data.

Apart from the French examples just discussed, Chomsky does not provide any clearcut empirical arguments for stipulating that trace is immobile. There are a couple of arguments in which this assumption makes an appearance; however, instead of going into the minute details of these arguments, which are fairly complex and involve rather esoteric issues,[34] I think it will be more enlightening to consider one straightforward property

---

[32] That this is the case is indicated by (i), ungrammatical on the reading where *him* is coreferential with *John*, suggesting that the latter is c-commanded by the former (Principle C). Note incidentally that it seems that *to him* must not be a regular Prepositional Phrase headed by *to*, since in that case, *him* as the complement of *to* should not c-command out of the PP. (cf. *I talked to John about himself*).

(i) * It seems to him_i that John_i's theory is unpopular.

[33] One might perhaps assume that in (24b) the clitic *lui* adjoins to *semble*, or otherwise to T, to which *semble* itself is adjoined. Consequently one could assume further that if there were any features in T (or one of its sublabels) which could enter into a checking relation with features of *lui* (and hence, with features of the trace of *lui*), then these features would have been deleted and erased at the point when the clitic adjoined (recall that all formal features of checking domains are uninterpretable). Hence, Last Resort would ensure that the clitic - or more precisely, any of the copies of the clitic's chain - would not be capable of entering into any further checking relation with T, because T would just lack any suitable features. Under this hypothetical analysis, then, the trace of *lui* would be predicted not to count as intervening between *Jean* in the embedded clause and the matrix T due to Last Resort (cf. the case of *there seem to be some books on the shelf* discussed in §3 above), so that again, no independent stipulation about trace immobility would be required.

[34] See e.g. Chomsky 1995: 365, in which the aim is to derive the supposition that expletives cannot appear in Spec-*v* at LF. The argument relates to Transitive Expletive constructions in Icelandic, and relies on assuming that the trace of a subject in the inner Spec-*v* cannot be attracted

which seems to characterize all the arguments in which trace immobility plays a role, since this highlights a general inadequacy in Chomsky's system. It is noticable that that the traces for which immobility is taken to be crucial are generally terminal rather than intermediate traces of A-movement. This highlights the fact that the prohibition on movement of *intermediate* traces of A-movement ((21) above) is to all intents and purposes an inactive assumption -- in reality, it is the "natural extensions" of (21), i.e. the statements in (22), which are relevant. But (22) is apparently not derivable from any independent consideration; in particular, it does not seem possible to derive it from the Chain Condition, as Chomsky maintains with respect to (21) (see above); if anything, (22) would seem to hold in spite of the Chain Condition. That is, unlike intermediate traces, terminal traces cannot possibly be taken to be immobile because they have undergone a process which makes them invisible at LF, since these traces (unlike intermediate ones) occupy θ-positions, meaning that if they were invisible at LF, violation of the Chain Condition would ensue, according to the premise upon which (21) is based (see above). To get round this problem, it could perhaps be assumed that it is only the formal features and not the entire trace which erases (as Chomsky 1995:303 suggests), but in that case, it is difficult to see in what sense the trace would be invisible at LF for the Chain Condition in the situations where this is actually required (intermediate traces).

4.3 Section summary

In this section, I have examined Chomsky's (1995) proposal that certain copies, namely the traces of A-movement, should be "immobile". We have seen that this proposal is problematic on a number of counts.

To begin with, the proposal that traces are immobile, as it stands, is essentially a stipulation. Of course there will always be occasions when stipulations are unavoidable, and resorting to them is a necessary strategy. However, in the particular situation at hand, it seems that the assumption that traces are immobile is not even fulfilling any genuine role

---

to the expletive.

since, as shown in § 3, independent general principles already tightly restrict the activity of traces. In addition, it was argued that (21) is not sufficient, since (i) it does not seem to prevent *terminal* traces from moving (as would be required in e.g. the French "super-raising" example discussed above), and (ii) it seemingly fails to predict that even intermediate traces of A-movement will be immobile, if these have interpretable features, which by assumption are unable to delete and erase. Add to these problems the fact that (21)/(22), as they do not seem to genuinely derive from any other principle, amount to relinquishing the notion that movement applies freely, and it seems to me that there is a strong case in favour of rejecting Chomsky's (1995) proposal, assuming instead that all copies in a chain participate in movement (or more generally, in computational operations) - i.e. what I have called the Copy Hypothesis. The case for the Copy Hypothesis would of course be reinforced if actual empirical uses could be found; as it turns out, there are a number of these, as will be demonstrated in Chapters 4, 5 and 6 below. Before this, I address two further theoretical issues connected with the Copy Hypothesis.

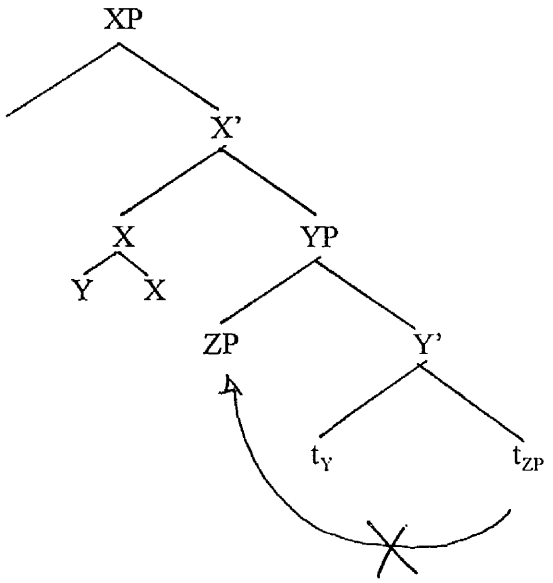
5 Movement to copies

So far in this chapter, I have concentrated exclusively on movement *of* traces, and not addressed the question of movement *to* (i.e. attraction by) a trace. Given the Copy Hypothesis, such an event is entirely possible in principle. Is this a problem?

In considering this question, one might begin by observing that *overt* attraction of an element by (features of) a trace would generally seem to be excluded. This is because traces are by definition embedded in a structure, and overt movement of an element to such a position should therefore be excluded by the Extension Condition (see Chapter 1, §4.2.3), which presumably is, or derives from, independent considerations. To be more precise, in the framework of Chomsky 1995, Strong features must be eliminated as soon as they appear in the structure, otherwise the derivation terminates: "we...define a strong feature as one that a derivation "cannot tolerate": a derivation D is canceled if Σ [= the syntactic object formed by the derivation] contains a strong feature...A strong feature thus

triggers a rule that eliminates it...Cyclicity follows at once" (1995:233). Hence a situation in which a strong feature is covered up by further structure-building cannot arise, and as long as all overt movement is movement to check a strong feature, then there should be no possibiltity of overt attraction of an element by a trace.[35]

(28)



by Extension Condition (Strict Cycle)

This being so, the situation which must be considered is the case of attraction of pure Features by a trace (i.e. covert movement to a trace). This means that the relevant evidence, should there be any, will involve Binding or other interpretive facts which reveal the structure of LF. What instantly comes to mind is evidence due originally to Postal 1974, more recently reconstituted by Lasnik and Saito 1991 (see also Branigan 1992, Vanden Wyngaerd 1989) that direct objects appear to c-command into VP-adjoined elements, in English at least. This implies that at LF, their formal features (at least) are

[35] In fact, I shall be assuming that in certain situations, *non-strong* features may be checked by overt movement (see Part II below). In that case, overt movement to check features belonging to a trace would not be excluded in the way just described in the text. This assumption will be instrumental in my analysis of local object-scrambling in Icelandic, German and Dutch (Chapter 6, § 2).

outside of VP, in other words, checking by the objects is not implemented by adjunction of their features to the trace of V itself, since the latter is obviously VP-internal and would not c-command VP-adjuncts, given standard assumptions.[36] As an example of this evidence, consider the following (taken from Lasnik and Saito 1991):

(29) a. The DA accused the defendants during each other's trials
b. The DA proved the defendants to be guilty during each other's trials
c. * The DA proved that the defendants were guilty during each other's trials

(30) a. * Joan believes him_i even more fervently than Bob_i's mother does
b. * Joan believes him_i to be a genius even more fervently than Bob_i's mother does
c. Joan believes that he_i is a genius even more fervently than Bob_i's mother does

In (29a,b), the reciprocal inside the adverbial *during each others trials* (which is presumably a VP-adjunct) is bound by *the defendants*, implying that the object is, or can be, in a position to c-command into the adverbial, i.e. it cannot be in situ, on standard assumptions (that is, unless the VP adverbial itself was a kind of complement of V itself (cf. Chomsky 1995: 332)). (30) (taken from Branigan 1992) demonstrates that objects and ECM subjects *must* be outside of the VP at LF (Principle C).[37]

These interpretive facts are immediately predicted if one assumes that the trace of V cannot be moved to: with movement to traces barred outright, as is the case under Chomsky's (1995) theory, the formal features of the object would necessarily end up outside of the VP: either they adjoin to T after V has itself adjoined there (Chomsky 1995) or they adjoin to V which then raises to T - in either case, the features will be in a position to c-command VP adjuncts. Given the object-related binding facts illustrated in (29) and

[36] This data is of course also compatible with *overt* Object Shift, if one further assumes that verbs also move overtly in English (Pesetsky 1989, Koizumi 1995). However, my assumption in this thesis is that English has no overt Object Shift (see Chapter 3 below for arguments).

[37] The reader is referred to the two references cited for many similar examples involving Negative Polarity Item and Binominal *each* licensing, "bound variable" pronouns as well as regular A-binding facts.

(30), it would seem obligatory to assume that traces cannot attract, which of course goes against the Copy Hypothesis. Unless, that is, the facts in question can be attributed to something else.

With respect to the particular issue about objects moving out of the VP, which is the only immediately apparent problem for the Copy Hypothesis in connection with movement to traces, notice that there is at least a way to side-step it. It is possible that Accusative Case is not directly connected with the verb itself, but in fact, with a higher functional head, either AGR (see e.g. Chomsky 1993, Ferguson & Groat 1995), or *v* (Chomsky 1995), or, more interestingly, perhaps a semantically contentful head associated closely with the verb, such as Aspect in some form (see e.g. Borer 1993, Diesing & Jelinek 1993) In the first place, this would do the job of accounting for the fact that formal features of objects do not seem to adjoin to the VP-internal trace of V, as seems to be necessitated by facts such as (31) and (32): this would be so simply because the trace would not have the Case features. Yet this approach is not entirely *ad hoc;* it would also account for the fact that objects generally do not move to Spec-VP when they raise overtly for Case, but to a position outside of VP,[38] something which the "trace is immobile" theory does not itself make direct predictions about. From a more general perspective, the view of Accusative Case as associated with a functional head builds into the system the fact that all elements have to move for formal licensing; in particular, the "unified theory" of Nominative and Accusative Case (now standard in the Principles and Parameters approach) would arguably be more contentfully unified, since Nominative is of course (standardly) associated with the functional head Tense, rather than verbs themselves.[39]

---

[38] In Icelandic, for instance, shifted objects precede VP-adjoined elements like *ekki* 'not' (see Jonas & Bobaljik 1992).

[39] An alternative approach to the movement-to-traces issue would be to assume (as in Williams 1994) that there is no X^0 -movement in the INFL complex (i.e. no short verb movement) at all. Given this, there would be no X^0 chains and therefore, no issue of attraction by the traces in such chains. Yet another possible alternative way to account for the object binding facts would be to adopt a Larson (1988) VP-shell structure in which VP-adverbial elements could actually be c-

I conclude tentatively that any problems concerning attraction of elements by traces can be overcome, and continue to assume that such operations are possible in principle, as entailed by the Copy Hypothesis. This conclusion will find empirical justification in Chapter 6 below, where the assumption that traces can attract will play an important role in the analysis of a range of object-related optionality data in Germanic.

6 Copies, derivations, and representations

In this final section, let us consider the issue discussed in this chapter within a different context: the question of whether the grammar is a derivational or a representational system, which Chomsky and Lasnik (1993) take to be "at root, a question of truth and falsity". It is worth noting that the potential problems concerning the copy theory which I have discussed here may have some bearing on the derivational/representational issue, since these problems by their very nature arise only in theories which are derivational, in a non-trivial sense. The implication is that if these problems for derivational theories (of which the standard Minimalist framework of Chomsky (1993 etc.) is obviously the relevant example)[40] - turned out to be insurmountable without the help of stipulations such as "trace is immobile", this would lend support to a theory which was non-derivational in the appropriate sense. A theory which fits this description is the "Radically Minimalist" theory of Brody 1994, 1995.

Brody's alternative minimalist model resembles Chomsky's in that there are only two levels of representation, LF and PF. However, in the former, unlike the latter, there is claimed to be no syntactic derivation linking the numeration to LF, other than a trivial one which instantaneously creates the representations out of pre-formed chains: "chain construction" itself, Brody maintains, takes place as follows: "suppose...that the chain...is

---

commanded by *in situ* objects. Such a position is argued against by Branigan 1992 in the course of arguing for covert movement of objects out of VP for Case reasons.

[40] See Introduction, § 3.1.

formed before LLF [i.e. LF] is projected, that is, it is formed presyntactically" (1995:41). Projection (in the sense of X-bar theory) where applicable, is also a "presyntactic" operation.

In Brody's model, there is one structural point alone at which Spell-Out can (successfully) apply, and that is LF itself (see p.20). Unlike in Chomsky's model, then, there is no overlapping at all between the operations Merge and Move/chain formation. The arguments which Brody advances in favour of his chain-based model are mainly of the conceptual economy type: "[a]ssuming that chains and Move α express the same type of relation, a theory that contains both concepts is redundant...the concept of chains is independently motivated by the principle of Full Interpretation and by the condition that determines the distribution of the set of thematic positions" (1995:5).

In common with Chomsky, it is necessary for Brody to subscribe to the copy theory, since his model, being minimalist (in the basic sense described in chapter 1), is equally reliant upon copies for dealing with reconstruction phenomena: as in the standard Minimalist framework of Chomsky, there is absolutely no possibility of appealing to different levels of representation, while lowering/undoing operations are still more resolutely ruled out, since there are no successive operations generally. It is in this last respect that it is interesting to compare the Radically Minimalist and standard Minimalist models, since it is clear that only in Chomsky's model does the potential problem of copies and derivations arise: only in this model is it actually conceivable that any part of a chain already formed might undergo further syntactic operations. This realization lends a further dimension of relevance to the discussion in this chapter. If the copy-movement problem were indeed such that it required the statements proposed by Chomsky to immobilize certain copies, this would provide a good argument in favour of the more representational model of Brody, since in this model it is obvious that no such stipulations would be needed. If the choice between the two types of model is itself of any significance (which I assume that it is), it is also significant to have shown that the stipulations needed in the derivational model which would, all other things being equal, tip the balance in favour of the representational model are not in fact necessary, as I have argued at length in this

chapter.[41]

6 Summary and conclusion

In this chapter, I have proposed that all copies (both "heads" and "traces" of chains) are free to undergo computational operations, as opposed to some being "immobile" and some not. I called this the Copy Hypothesis. Admittedly, this goes against the assumptions of the Minimalist framework in its present instantiation, specifically contradicting Chomsky's (1995) proposal that certain traces should be specifically prevented from moving. At the same time, it is clear that allowing movement (or indeed other) operations to apply to any element is theoretically the best option. For this reason alone, the Minimalist model should incorporate the Copy Hypothesis, if its theory of movement is to compare favourably to that of the GB model. It was argued indeed that general conditions on movement as currently assumed in Minimalism are in fact sufficient to properly constrain the behaviour of traces.[42] Complementary to this, it was argued that Chomsky's (1995) proposal to immobilize traces by stipulation is unsatisfactory in various ways, being, in a nutshell, both unnecessary and insufficient.

In the light of the above considerations, my conclusion is that there is good reason to adopt the Copy Hypothesis, although such an assumption is not standard within Minimalism.[43] The Copy Hypothesis will be made use of repeatedly during the rest of this thesis.

---

[41] We come back to the derivational/representational question in Chapter 7 below.

[42] Notice that the fact that certain constraints on movement are now incorporated into the definition of the operation itself does not mean that the notion of movement applying freely, in the sense under discussion here, ceases to have any content.

[43] It should however be noted that some authors have made use of the idea of trace movement (e.g. Kitahara 1994, Poole 1996) in analyses of certain phenomena, although as far as I am aware, the attendant theoretical issues have not been discussed (with the exception of Pettiward 1995, where a preliminary discussion can be found).

Chapter 3

When does the English Object Shift?

1 Introduction

The remainder of Part I of this thesis is devoted to preliminary demonstration of the empirical advantages of the Copy Hypothesis proposed in Chapter 2 (all copies in a chain are active in the computational system). The arguments, which for the most part involve *wh*-phrases in English - specifically direct objects and the associates of expletive *there* - are presented in Chapter 4. One important assumption made here is that neither objects nor associates of *there* raise overtly to check Case/agreement features. With respect to associates of *there*, this assumption is not really controversial, since the relevant position is occupied by the expletive itself. For objects, though, the situation is much less clearcut: although it is often assumed that they too check Case only at LF in English (see e.g. Chomsky 1993, Branigan 1992, Bobaljik 1995, among others), this viewpoint has not been argued for in a particularly comprehensive way, and moreover, a number of authors have argued in some depth that objects move *overtly* for Case (Johnson 1991, Koizumi 1993. 1995, among others). I take it that the question of when objects move for Case is basically still unresolved with respect to English, and in particular that an assumption that

objects move covertly must be actively argued for. This is what I intend to do in this chapter, which thus functions as a kind of auxiliary chapter in the context of the thesis. Before beginning, a couple of points need to be clarified, in order to avoid confusion and/or irritation. Firstly, since I am working within the Minimalist framework, I assume that objects do move at some derivational stage for Case reasons.[1] Hence the question is taken to be not whether, but only when, such movement takes place, and for this reason alone, I shall tend to write loosely of evidence against overt Object Shift as being supportive of *covert* Object Shift. It is of course recognized that evidence against overt Object Shift does not in and of itself consitute positive evidence for covert Object Shift. Secondly, the arguments I will present in this chapter against overt/for covert Object Shift primarily involve placement facts. Interpretive evidence was not considered, since at the time of writing, I was working under the assumption that such evidence only tells us something about LF (this last point is addressed in note 37 at the conclusion of this chapter).

The chapter is structured as follows. In Section 2, as a background, I give an overview of the main characteristics of object behaviour in English which any account should address, then review and compare two recent Minimalist proposals, one which claims that objects move overtly to check Case features (Koizumi 1993, 1995), and another which maintains that they move covertly (Chomsky 1995). We shall see that each of these positions seems to have empirical arguments in its favour, though positive arguments for covert object movement seem to be more scarce that those for the overt view. However, by examining the situation more closely (Section 3), it will be seen that the evidence given on behalf of overt Object Shift (which for the most part involves ECM constructions) is unconvincing. I advocate the view that Accusative Case checking is implemented covertly in English (Branigan 1992, Chomsky 1995, Bobaljik 1995), i.e. by pure feature-adjunction in terms of Chomsky's (1995) framework. In Section 4, I consider the phenomenon often known as the Case Adjacency Effect (Stowell 1981, Pesetsky 1989). In the P&P framework, the ability to derive the Case Adjacency Effect from conditions on Case assignment is more or less taken for granted as a necessity, and it is

[1] See Chapter 1 (§ 4.4.1) and Chapter 2 (§ 5).

arguable that overt Object Shift theories, despite other shortcomings, fares better in this particular respect, with the possible implication that it should still be preferred. However, I suggest that a range of further adjacency effects, involving elements which most definitely do not need Accusative Case, e.g. measure phrases, "quotative inversion" subjects, and adverbs (data which to the best of my knowledge has not been discussed widely, if at all, in this connection) should make us inclined to reconsider the relevance of the standard Case Adjacency effects to the Object Shift issue. Section 5 concludes.

2 Current views on objects in English

2.1 Preamble

As we saw in the introductory chapter, given the assumption that all syntactic licensing consists of feature-checking between functional and lexical elements, the Case of transitive objects, i.e. Accusative, is no longer taken to be assigned by V directly to the object *in situ*; objects, like subjects, must move out of the VP for Case, with the involvement (directly or indirectly) of a functional head of some sort: AGR_o (Chomsky 1991, 1993), some type of contentless affix such as Pesetsky's (1989) element μ, Chomsky;s (1995) light verb (see Ch.1, § 3.3), or possibly ASP(ect) as proposed by Borer 1993.[2]

On this view, the expectation is that we will find cross-linguistic variation with respect to the properties of this object A-movement, and in particular, variation as to whether it takes place overtly or covertly.[3] Recent research indeed suggests that object

---

[2] For generally related earlier proposals and/or discussion of this theory of Accusative Case, see e.g. Pesetsky 1989, Déprez 1989, Sportiche 1990, Johnson 1991, Chomsky 1991, Branigan 1992.

[3] In fact, although I am talking here about variation in the timing of Case-related movement, it ought to be noted that in the system of Chomsky 1995, it is explicitly assumed that Case features, and indeed operator-features, do not in themselves have the property of inducing "overt movement". That property belongs only to Strong features. No real explanation for overt movement is discernible in this system, as Chomsky himself admits. (Overtness of movement has been claimed to relate to

A-movement does vary cross-linguistically, along the same lines as A-movement of subjects,[4] and A'-movement generally.[5] Various instances of so-called object scrambling in various languages have been re-analyzed as overt movement to a Case position (see Déprez *op. cit.*, Mahajan 1990, among many others). Languages which it has, in one way or another, been claimed exhibit overt object-movement for Case include Hindi (Mahajan 1990), Japanese, Zarma (Koizumi 1995), Early Modern English (Roberts 1994), Dutch (Vanden Wyngaerd 1989, Zwart 1993). Contrastingly it has been argued that French has covert object shift (Branigan 1992, Koizumi 1995). In other languages, there seems to be some optionality with respect to when Object Shift occurs - Icelandic is one such language (see e.g. Collins and Thráinsson 1994, Kitahara 1994)).[6] As for A-movement of objects in English, disagreement prevails, as we shall see shortly.

As the Minimalist system has no level of S-structure, it is naturally more difficult to decide which elements move when, particularly when the movement in question lands in (what will ultimately be) a clause-internal position, as is the case with Object Shift. For example, interpretive facts concerning Binding, Negative Polarity Item licensing and so on are not precisely informative about the position of arguments when Spell-Out applies, because the conditions regulating binding etc. (if syntactic and representational at all) can only apply at LF, at which stage, on standard Minimalist assumptions, elements may well

---

"clausal typing" (Cheng 1991) in case of *wh*-movement, although this idea does not seem to provide any further insight into why the movement in question should be overt as opposed to covert.

[4] Cf. Koopman and Sportiche 1991 who suggest that there is parametric variation as to whether subjects are obliged to move to Spec-IP for Case (English, French), or can be assigned Case *in situ* under government by I (so-called VSO languages, e.g. Welsh, Arabic).

[5] For example, English has overt *wh*-movement, in Chinese and Japanese, *wh*-phrases are commonly assumed to move covertly; while in French, it appears that *wh*-movement is optionally overt or covert in matrix clauses, though obligatorily overt in embedded clauses.

[6] From now on, I use the term *Object Shift* (or OS) to refer to overt Accusative Case-related movement of an object. See Jonas and Bobaljik 1992 for discussion of this term and its different uses.

have gone elsewhere.[7] This all means that other diagnostics must now be relied upon. One well-established method is to look at the position and behaviour of elements which are assumed to occupy a fixed position - the prime examples being adverbials, as well as certain negative elements (as in e.g. Icelandic, see Jonas and Bobaljik 1992). Provided that we have some idea about where these fixed elements are situated - adjoined to VP, for example - then the overt position of other elements relative to this can be deduced. Another popular diagnostic is the position of so-called Floating Quantifiers (Dowty and Brodie 1984, Sportiche 1988; Bobaljik 1995 - see Chapter 4 for discussion). Inevitably, given the increased importance of adverb/floating quantifier placements as diagnostics for the overt position of elements, their reliability and interpretation becomes a more contentious issue in its own right, and this will be kept in mind.

As mentioned in the introduction, I am taking it for granted that objects move out of the VP at some stage, as is now more or less standard in the P&P approach - see Lasnik and Saito 1991, Branigan 1992 for arguments that objects c-command VP-adjoined elements based on the accepted tests involving Binding principles, pronominal variable binding and the licensing of Negative Polarity Items and binominal *each*.[8]

2.2 The standard object phenomena

When considering the syntax of objects in English, one comes across a well-known cluster of problems involving both differences between the behaviour of objects and other elements within English, and differences between the behaviour of objects in English and objects in other languages. Firstly, there is the apparent inseparability of main verbs and their direct objects,[9] a phenomenon sometimes known as the Case Adjacency effect.

---

[7] However, see note 40 below.

[8] See this thesis, Chapter 2, § 5 for examples.

[9] With the exception of when the object is "heavy" or "complex" (see Ross 1967) as in (i). Compare this with (ii).

(i) a. Please check carefully the details on this page

Consider (2):

(2) a. * John eats often lemons
b. John often eats lemons
c. John eats lemons often

As illustrated in (2), an adverb may not intervene between a main verb and its direct object. The same facts obtain with any adverbial element - *quickly*, *with his bare hands*; *every day*; *regrettably*; *allegedly*; and so on. The term *Case Adjacency* stems from an influential analysis of the phenomenon by Stowell 1981,[10] who accounted for the illformedness of such examples as (2a) by proposing that (Accusative) Case must effectively be assigned under strict adjacency between assigner and assignee; in (2a), the adverb *often* intervenes between the Case-assigner *eats* and the assignee *lemons*. This particular theory has empirical and conceptual drawbacks which are well-documented (see Johnson 1991, Koizumi 1993, 1995). The main conceptual objection concerns the use of Adjacency, which is thought not to be a legitimate syntactic notion (but more appropriate in phonology or morphology); it has however continued to be standardly assumed that Case has a hand in the effect, since the verb and various non-DP complements (e.g. PP or CP) which do not need Case, are allowed to be separated by adverbs, as illustrated in (3) (examples taken from Pesetsky 1989). The Case hypothesis is further corroborated by the fact that adverbs are also prohibited from appearing between an Exceptional Case Marking (ECM) verb and its IP complement, the subject of

---

b. We are investigating thoroughly the circumstances of that decision

(ii) a. * Please check carefully the details
b. ?* We are investigating thoroughly the decision

The phenomenon in (i) is usually known as Heavy NP Shift, and has sometimes been analysed as involving rightward movement of the "heavy" element (for an alternative analysis using leftward movement, see Larson 1988). More discussion of heaviness phenomena and adjacency effects will be found in § 4 below. On Heavy NP Shift as an apparent case of optional movement, see Part II, Chapter 6, § 5, note 39.

[10] I will continue to use the term *Case Adjacency* for ease of reference, following frequent practice in the literature.

which is standardly assumed to receive Case in some relation with the ECM verb - (4):

(3) a. Sue looked carefully at him
b. Harry relies frequently on it
c. Sue thinks sometimes that Harry is foolish

(4) * Bill believed often John to have eaten lemons

The facts about constructions with two DP Complements (Double Object constructions) and with a DP and a PP complement are a little more complicated: in both types of construction, adjacency is required between the verb and the linearly first complement (receiving the Goal θ-role) - (5a,b) - which fits into the pattern as long as it assumed that this DP checks an Accusative or at least structural Case. Further, adverbs are illicit between the two DPs of a Double Object construction (5c), but allowed between a DP and PP (5d) (although this seems to be restricted to manner adverbs). The second (Theme) object in the latter has Accusative Case, as shown by (5e):

(5) a. * Bill gave often John lemons
b. * John put often the lemons on the shelf
c. ?* Bill gave John often/quickly lemons
d. John put the lemons quickly on the shelf
e. Bill gave him them/*they

The adjacency phenomenon is one basic object-related phenomenon in English which needs explaining, and for the reasons just mentioned, most (if not all) Principles and Parameters accounts have sought to attribute it in one way or another to Case theory (Pesetsky 1989, 1995, Johnson 1991, Branigan and Collins 1993, Koizumi 1993, 1995, Chomsky 1995, Costa 1995, Stowell *op. cit*).[11]

Related and equally well-known is the fact that English shows a contrast with French - and likewise with European Portuguese, Italian and Spanish - where the

[11] Although in § 4 below, the correctness of deriving the Case Adjacency phenomenon from conditions on Case assignment/checking will be questioned.

equivalent of (3a) is grammatical; in these Romance languages, it seems that there is "more space" between the verb and its object when Spell-Out occurs, within which adverbs may be found:[12]

French

(6) Jean mange souvent des citrons
J. eats often lemons
'Jean often eats lemons'

European Portuguese (example from Costa 1995)

(7) O Paulo beija frequentemente a Maria
the P. kisses often the Maria
'Paulo often kisses Maria'

Italian (example from Stowell 1981:114):

(8) Mario legge spesso dei libri
M. reads often books
'Mario often reads books'

Spanish

(9) María leyo cuidadosamente los documentos
M. read carefully the documents
'Maria carefully read the documents'

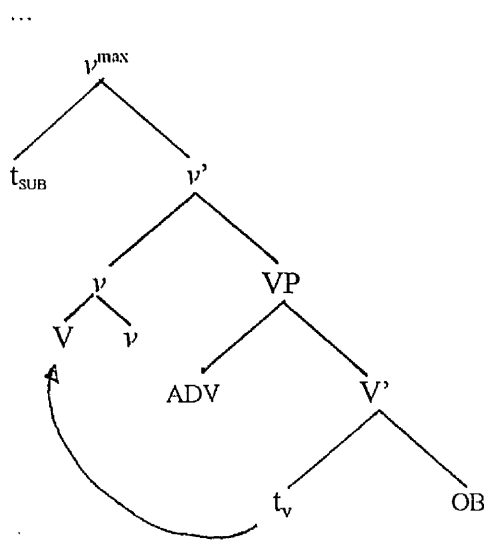
Having reviewed the most basic facts which a theory about the syntax of objects in English should deal with, I now outline and compare two specific recent views: Chomsky 1995 and Koizumi 1993/1995.

[12] Examples from French and Spanish provided by Jeanne Cornillon and Carmen Zoraquiain respectively. In French, the negative element *pas* may also intervene between main verb and object, as in *Jean ne mange pas des citrons* 'Jean eats not lemons'. See Pollock 1989.

2.3 Chomsky 1995

Adopting a type of "VP-shell" structure inspired by Larson (1988) (see Chapter 1, § 3.3), Chomsky 1993, 1995 argues that objects move covertly for Case-checking in English, while main verbs undertake a short overt movement to adjoin to a higher light verb, though they do not move overtly out of the VP complex. Chomsky's analysis is illustrated in (10) (regular lines represent overt movement and broken lines, covert movement):

(10) English (Chomsky 1995:331)

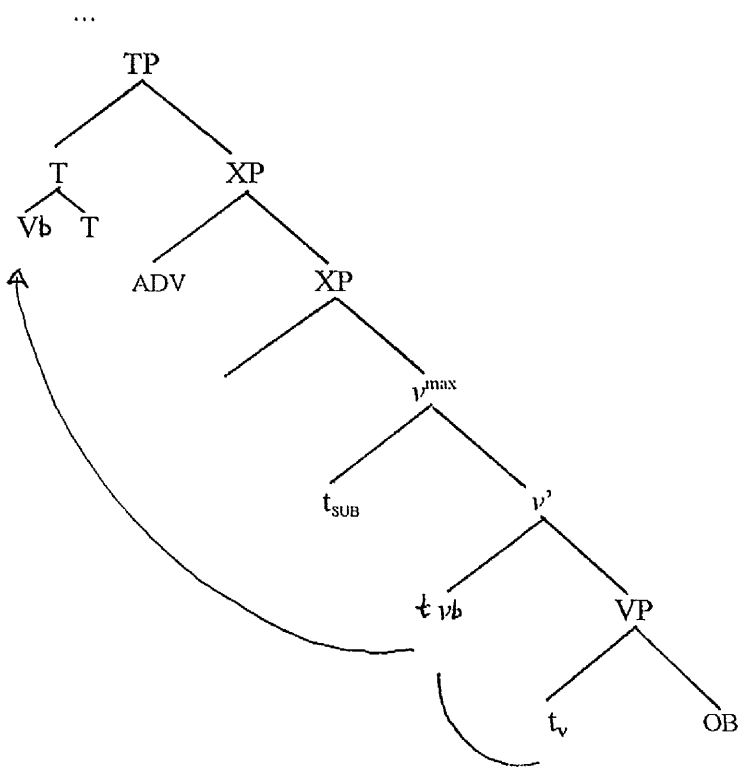


With respect to the Case Adjacency effect, Chomsky suggests that "[t]here is a Case solution, but it does not involve adjacency" (1995:332). The explanation is as follows: adverbs may in principle be generated as the Specifier of the VP complement of *v*, as shown in (10). However, the adverb blocks movement of the object to the outer Spec-*v* for Case by the Minimal Link Condition, given what he describes as the "plausible though not obvious" assumption (*op. cit.*:390, note 104) that the adverb actually counts as a candidate for attraction by the Case features. This analysis would seem to be quite easily extendable to the case of ECM (4). As for elements other than DP-complements, which

examples such as (3) above show are not subject to the Adjacency effect - the locality problem with adverbs does not arise, because non-DP complements do not need to move for Case, hence an adverb can occupy the VP along with them without any problems arising.

Although Chomsky does not go into details on the topic, the fact that French (and Spanish etc.) lacks the Adjacency effect may simply be captured by assuming firstly that main verbs raise overtly beyond VP (as in the earlier proposals of Emonds 1976, Pollock 1989), perhaps to T (see Koizumi 1993, 1995) and further, that there are potential adjunction-sites below TP and above the verbal projections, perhaps adjoined to the latter (presumably adverbs may not occupy the Spec of VP in Romance (or for that matter in any language) for the locality reasons just discussed in connection with the English facts).

(11) <u>French</u>

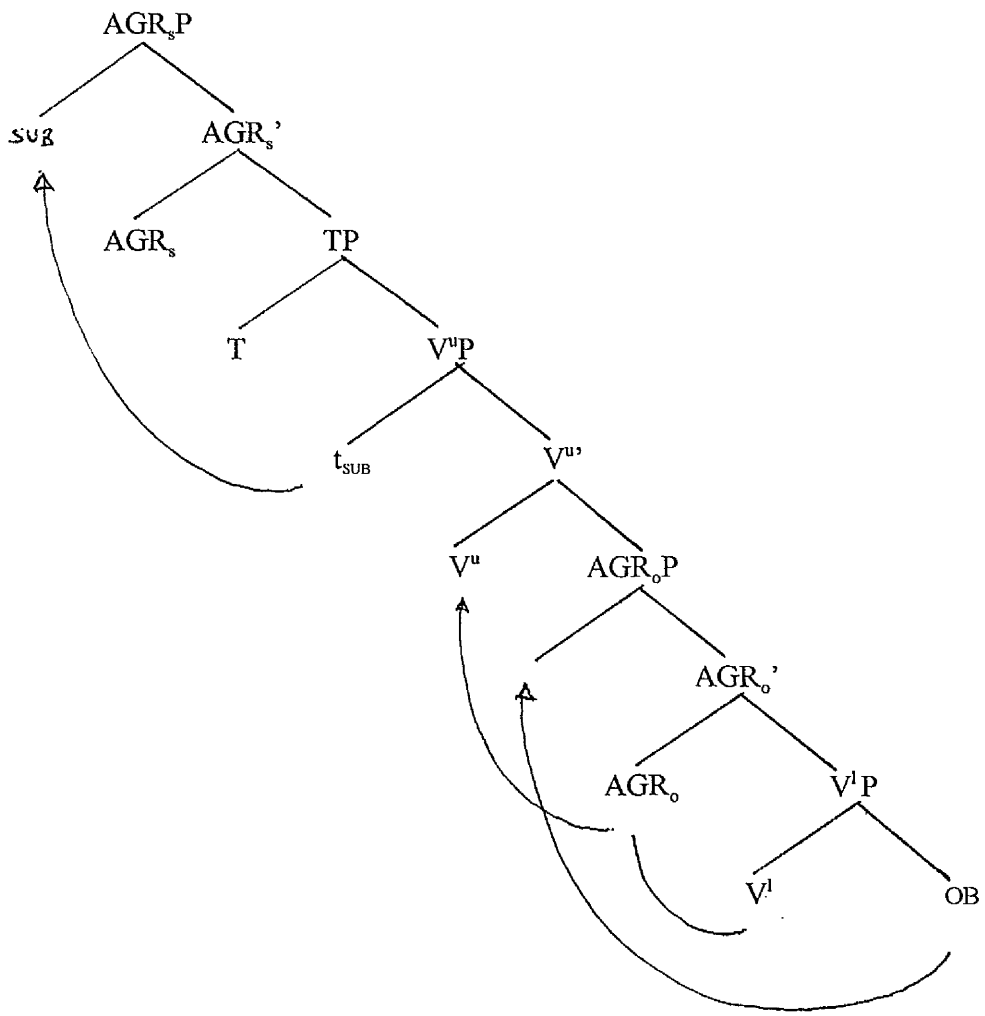


2.4 Koizumi 1993/1995

A contrasting approach to the placement of objects relative to verbs and other elements in English is developed by Koizumi 1993, 1995. Koizumi introduces an alternative clausal architecture, the main characteristic of which is that subjects and objects originate in separate VP + functional projection complexes. This "Split VP Hypothesis" is illustrated in (12) (Koizumi 1995:102), where V^u indicates "upper V" and V^l, "lower V". In addition to the Split VP Hypothesis, Koizumi proposes that both objects and verbs move overtly in English, the former to Spec-AGR_oP, and the latter making a shortish movement to V^u via AGR_o (this recalls earlier analyses by Pesetsky 1989, Johnson 1991, and see also Costa 1995 for a similar proposal). The Case Adjacency effect is derived with the addition of the following assumption: "AGR is a 'pure' functional category, it has no lexical semantic content. Thus adverbs may not adjoin to projections of AGR such as AGR_oP" (1995:28). Coupled with the assumption that the verb is in V^u and the object in Spec-AGR_oP, then, the possibility of Verb-Adverb-Object order is eliminated, in keeping with the facts.[13]

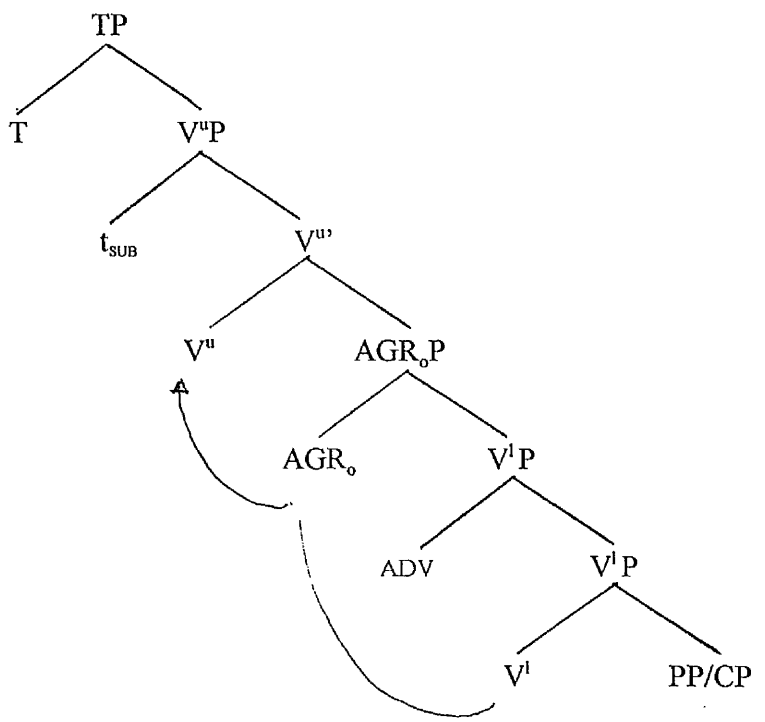
[13] Notice that technically Koizumi's assumption about adjunction becomes infelicitous within the most recent version of the Minimalist framework since it presupposes the existence of purely empty heads (see Chapter 1, § 3.3).

(12) English (Koizumi 1995)



In common with Chomsky's, Koizumi's system derives the adverb-placement contrast between DP and non-DP complements from Case-theoretic considerations: PPs and CPs do not need Case and therefore do not raise to Spec-AGR$_o$P, overtly or otherwise. Hence it is possible that a VP-adjoined adverb may show up between the verb (which recall raises to V^u) and non-DP complements which it has stranded inside the VP, as in (13) (cf. (3) above):

(13)



To derive the fact that adverbs cannot appear between the two objects of a double object construction, nor between the verb and the linearly first element in either Double Object or DP PP constructions, Koizumi assumes that the indirect object moves overtly to the spec of $AGR_{io}P$ which is positioned immediately above AGR_oP.

To account for the lack of a Case Adjacency effect in French, Koizumi simply assumes that in French, objects move covertly, while main verbs raise overtly, landing in a site higher than the upper V (see Koizumi *op. cit.*:41). Hence objects in French find themselves in a similar position to non-NP complements in English when Spell-Out occurs: the verb is outside of V^lP, so that adverbs attached at this projection will appear between the verb and the object. Koizumi (*op. cit.*:44) motivates his claim about French objects with data involving past participle agreement (cf. Kayne 1989); French is

discussed at length in chapter 5 below.[14]

3 Arguments and counterarguments for (c)overt Object Shift

In the last section I outlined two different theories about English: according to one - Chomsky 1995- objects raise covertly for Case checking; according to the other - Koizumi 1993, 1995 - this operation is overt.[15] Presumably only one of these views on the timing of object shift is correct, yet confusingly there seems to be evidence for both. In this subsection, I collect together and assess the evidence for the two views, concluding ultimately that the evidence goes against overt Object Shift.

---

[14] Another Minimalist variation on the Case adjacency theme is Branigan and Collins 1993 who, in the spirit of Pesetsky 1989, propose that AGR_o in English has optionally strong [N] features, these [N] features having to be checked by the verb. To derive the adjacency contrast between direct objects and other complements, they claim that "if the verb has a direct object and the [N] features of AGR_o are strong, there will be no convergent derivation." (p.3). This is because the strong [N] feature will also force the object to raise overtly, a situation which they take (following Chomsky 1993) to be excluded in English due to the Minimal Link Condition, the Strict Cycle (Extension Condition) and the fact that Spec-TP is unavailable as a landing-site (see this thesis, Chapter 1). In the case of verbs without a direct object, selecting AGR_o with strong [N] features obviously does not have this problem - hence verbs may raise overtly, giving rise to Verb-Adverb-Complement order. With respect to these latter constructions, a shortcoming of Branigan and Collins's account appears to be that they are forced to assume that verbs which do not check Accusative nevertheless check an N-feature on AGR_o. On verb-raising approaches to the adjacency contrasts see Chomsky 1995, who for a variety of reasons takes the view that "in general, it is doubtful that raising has anything to do with the relevant paradigms" (p. 331). For a critique of the notion of optional strong features, see Chapter 6, § 3.2.

[15] Note of course that the question of clausal architechture is separate from the question of what moves when. The reader is referred to Bobaljik 1995 for a detailed comparison between Chomskian and Koizumi views of clausal architecture. Bobaljik gives arguments in favour of the Split VP Hypothesis, but not for Koizumi's specific proposals about English. He explicitly claims in fact that objects do not raise overtly in English (p. 76).

As we shall see, the conflict between the evidence for overt and covert object shift in English coincides with the difference between recipients of Accusative Case in ECM constructions, such as *John* in (14a) - usually known as *ECM subjects,* since they tend to be analysed as originating as subjects of the IP complement of V (see Chomsky 1981), as in (14a) - and on the other hand, regular objects, exemplified by *the lemons* in (13b):

(14) a. Bill believes John to like eating
b. John ate the lemons

(15) shows that ECM subjects and regular objects check the same Case, Accusative:

(15) a. John ate them/*they
b. John expects them/*they to like fruit

As was recognized in the GB model, the null hypothesis is that Accusative Case licensing takes place in the same way in both environments. In the Minimalist framework, Chomsky 1993 suggests that "[e]xceptional Case marking by V is now interpreted as raising of NP to the Spec position of the Agr-phrase dominating V", i.e. ECM subjects check Case in the same way as regular objects. With respect to whether this movement is overt or covert, naturally the null assumption eliminates the possibility that Case is obligatorily checked overtly for ECM subjects but covertly for regular objects - or vice versa. Accusative Case must be checked either covertly or overtly across the board. There is then but one option for dealing with the conflicting evidence which we are about to see: some of it must be reinterpreted. As we will see, the arguments for overt OS involving ECM constructions are all somewhat weak, while at least one fairly strong argument exists for covert OS.[16]

[16] At this stage, I am omitting the Case Adjacency effect from the equation; this is fully discussed in § 4 below.

3.1 Adverb construal arguments

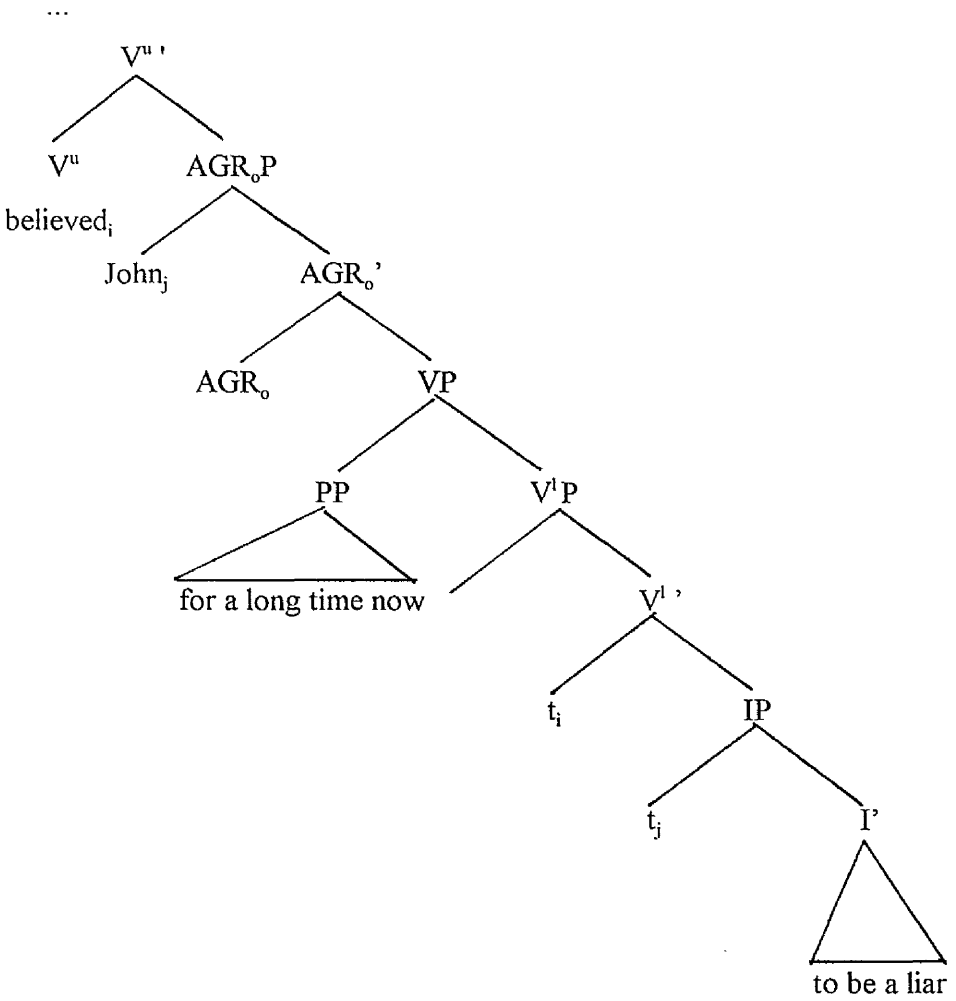
3.1.1 For overt Object Shift

The first piece of evidence given by Koizumi (1995) in favour of overt raising of ECM subjects comes from some facts concerning adverb construal. Citing Postal 1974, Kayne 1984, Authier 1991 and others, Koizumi notes that in (16), it is possible to construe the adverbial phrase *for a long time now* as modifying the matrix VP:

(16) I've believed John [for a long time now] to be a liar

Given this, plus the assumption that adverbs cannot be attached lower than the phrase they modify, Koizumi concludes that the matrix verb *believed* must once have occupied a position lower than *for a long time now* -- i.e. it has raised, into the higher V, given the clause structure which Koizumi assumes:

(17)



If the original position of the verb is below the adverbial, than the object to the left of the adverbial cannot be *in situ* itself: presumably it has moved (from the subject position of the IP complement of V, given the standard analysis of ECM constructions) - as illustrated in (17) above.

3.1.2 Counterargument

In fact, it seems that the situation with ECM and adverb construal is less conclusive than Koizumi's single example (16) would lead one to expect. As reported by Johnson 1991,

not all adverbs are in fact able to occur in the position of *for a long time now* in (16) and receive a matrix construal, as the examples in (18) (from Johnson 1991:587) illustrate:

(18) a. * Gary believes Mikey sincerely to be intelligent
b. * Sam considers Mittie strongly to have been misled

Stowell 1981, arguing against "Raising-to-Object" analyses, makes a similar point, maintaining that "the postverbal NP [i.e. the ECM subject - AMP] and the rest of the infinitival complement form a unit at S-structure, as shown by the fact that the clausal integrity of the complement does not allow other material in the matrix VP to intervene between the subject and the rest of the clause." (p.191). Stowell's examples are in (19) (judgements his own):

(19) a. * I never expected them at all to arrive so soon
b. * Paul discovered the pills by himself to be powerful
c. * Jim exposed Susan to us to be a liar

It cannot perhaps be said that the sentences in (19) or in (18) are worthy of a full "*", and there is at least a contrast generally between ECM and embedded finite clauses with respect to adverb construal: (16), repeated, contrasts with (20), in which it is impossible to interpret *for a long time now* as modifying *believed*; and there is a similar contrast between (18)/(19) and (21) (judgements in (21) are relativized to Stowell's in (19)):

(16) I've believed John [for a long time now] to be a liar

(20) * I've believed that John for a long time now has been a liar

(21) a. ** I never expected that they at all would arrive so soon
b. ** Paul discovered that the pills by himself were powerful
c. ** Jim exposed that Susan to us was a liar

However, it is quite possible that these particular contrasts could be attributed to some

deeper cause than simply the placement of the adverb relative to its (potential) modifiees,[17] and in view of the fact that in many, perhaps most, cases, the order Verb-ECM subject-Adverbial with the adverbial modifying the verb is at the very least marginal, it cannot be said that the adverb construal fact cited by Koizumi makes any conclusive contribution towards resolving the question of whether Case movement of the ECM subject is overt or not.

From a different perspective, the following counterargument to the alleged adverb construal evidence for overt OS is made by Chomsky 1995, who it will be recalled is a proponent of covert OS. His suggestion is that the overt adverb placement evidence which to some (not only Koizumi *op.cit.*, but also earlier "Raising-to-Object" authors such as Postal 1974, Bowers 1993) suggests overt movement of the ECM subject in fact "involves the kind of "rearrangement" that has been called "extraposition" in the past, but that may not belong at all within the framework of principles we are considering..." (p.33).

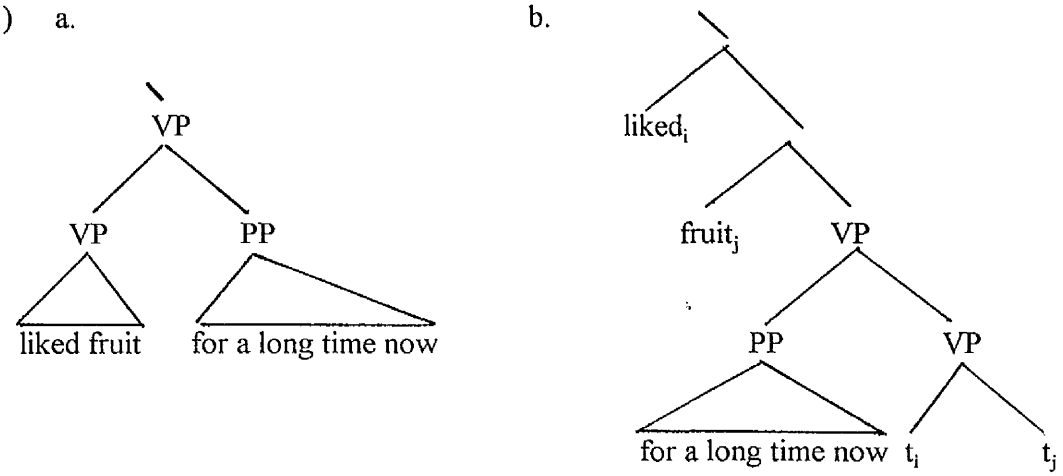
3.1.3 For covert Object Shift

It may be difficult to use adverb placement to draw conclusions about the overt position of ECM subjects, but it is still more difficult with regular objects. In (22), for example, the adverbial expression *for a long time now* modifies the verb, yet, unlike in the case of the ECM construction (16) discussed above, this fact fails to guarantee that the verb (and object) are outside of the VP - as in (23b), since the adverbial might equally be adjoined at the right edge of VP, as in (23a):

(22) John has liked fruit for a long time now

[17] In particular, some kind of ultimately semantic explanation suggests itself. For example, in examples (20) and (21) - unlike in the ECM examples (16), (18) and (19) - the finite embedded clause is specified for its own tense, which perhaps in some way could inhibit adverbs from being construed in the matrix clause.

(23) a. b.



However, a different type of adverb-construal argument can be constructed which turns out to favour the theory represented in (23a) over that in (23b) - i.e. the theory that regular objects do *not* move overtly. Pesetsky 1989, citing Andrews 1983, uses the relative scope among "stacked"adverbs as evidence that *verbs* (may) move overtly in English.[18] Firstly, consider (24):

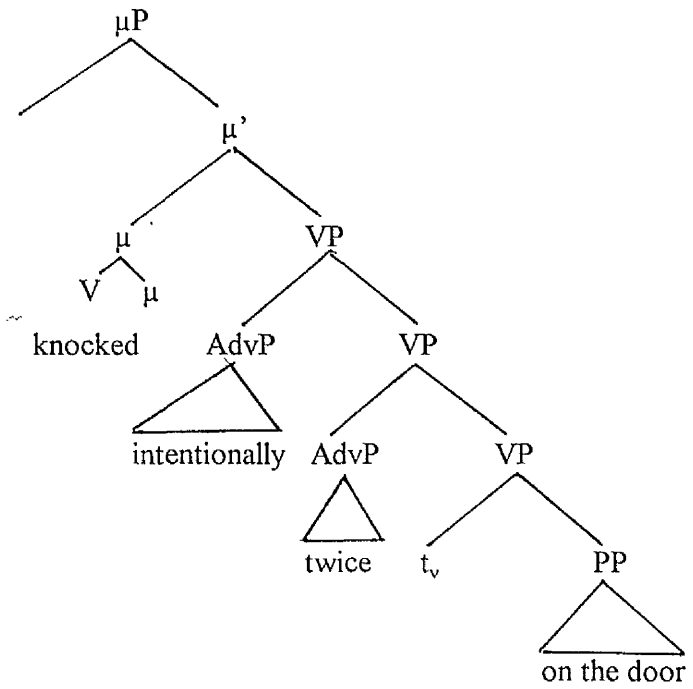
(24) John knocked intentionally twice on the door

Here, *intentionally* may be interpreted as taking scope over *twice*; i.e there is a single intention to knock twice on the door (indeed Pesetsky claims that this is the only possible reading, once the possibility of Heavy Shift of the PP complement is removed: *as for Mary, Bill relied intentionally twice on her*). Assuming, uncontroversially, that the syntactic reflex of the scope relation is c-command, it is deduced that *intentionally* c-commands *twice* in (24). If this is so, then the adverbs must be left-adjoined to the VP;

[18] Recall that it is assumed that adverbial elements (unless *wh*-phrases) do not move, which makes construal facts relating to them suitable for use in the diagnosis of overt positions of other elements. The question of where adverbs are actually situated in the first place is more controversial. In the theory of Kayne 1993, for example, there is no right-adjunction at all, and only one adjunct per head. The former of these assumptions is potentially relevant in the present discussion - see the next subsection.

and if they are left-adjoined to the VP, then the verb itself has moved, adjoining to the contentless affix μ in Pesetsky's own system, as illustrated in (25).

(25)



In fact, Pesetsky's scope argument is adopted by Koizumi 1995 in support of his own similar claim that the verb makes a short overt movement in English, so let us follow him in assuming that it is valid. Recall now that Koizumi's theory proposes the following difference between *John hit Bill* and *John knocked on the door*. In *John hit Bill*, all constituents move overtly out of VP, while in *John knocked on the door*, one constituent, namely *on the door*, is still within the VP when Spell-Out applies. Koizumi's hypothesis leads to a prediction: if there is a stack of adverbs to the right of a direct object, there should be a scope ambiguity - and in particular, the reading in which the leftmost adverb takes wide scope should be present - whereas if there is a stack of adverbs to the right of another sort of complement, e.g. a PP, then only one interpretation should be possible, namely the one in which the righthand adverb has wide scope.The examples we need to consider are (26) and (27):

(26) John hit Bill intentionally twice

(27) John knocked on the door intentionally twice

In (27), as reported by Pesetsky (*op. cit.*) (my (27) = his (66a), p.31), "*twice* unambiguously has scope over *intentionally*: the sentence can only refer to two events of intentional knocking", with the conclusion that the adverbs can only be right-adjoined to VP.

The interesting fact is that, contrary to the predictions of Koizumi's overt Object Shift hypothesis, the interpretation of the adverbs in (26) is exactly the same as it is in (27): *twice* has scope over *intentionally*, as unambiguously as it does in (27). But if the direct object is outside of the VP as Koizumi alleges, then left or right stacking of the VP-adverbs should surely be possible, leading to ambiguity - ambiguity which in reality does not exist. It is obvious that this state of affairs is compatible with the covert OS theory but not with the overt OS theory. We therefore have some support for the former.[19]

3.1.4 Counterargument

Just as one could object to the adverb construal evidence for overt OS in ECM constructions on the grounds that it is inconclusive, a defendant of the overt OS view might well make a similar point about the Pesetskyan adverb-stacking evidence which appears to argue for covert OS. Judgements involving scope ambiguities tend to be unstable and subjective, easier to dispute than Binding facts or purely syntactic placement facts.

---

[19] Koizumi in fact claims (note 13) that in (26), the scope of the adverbs is ambiguous. To my mind, wide scope for the lefthand adverb is more or less impossible to get, unless some heavy stress is applied to it - cf Mary hit John INTENTIONALLY twice, not ACCIDENTALLY twice). But the more important fact is that there is no interpretive difference between (26) and (27), which is definitely unexpected on Koizumi's analysis.

A diffferent type of objection would be to assume that right-adjunction is prohibited (Kayne 1993), in which case it would no longer be possible to assume such a simple relation between c-command and scope. It is not obvious that the conclusions reached on the basis of scope facts in the last subsection would be replicated after converting to a strictly Kaynian system. So far then, the adverb construal examples we have looked at in this section, although they are suggestive and one might resort to using them were no other evidence to be available, do not in themselves seem sufficient to resolve the Object Shift issue one way or another - so we put them aside and move on to some different evidence.

3.2 Particle construction argument

3.2.1 For overt Object Shift

Koizumi presents a second piece of evidence that ECM subjects are outside of the embedded clause at the point of Spell-Out, this time involving constructions with so-called Particle Verbs (on which see e.g. Kayne 1984, Johnson 1991, den Dikken 1995 and references therein).[20] Examples of Particle Verbs are *cut up*, *put down*, *throw out*, and so on. Certain of these verbs can also appear in ECM-like constructions, e.g. *make out* in the sense of 'imply' or 'suggest'. Koizumi initially assumes on the basis of (28) (= his (41a&b), 1995:35) that the Particle in Particle verb constructions cannot in general appear within a clausal complement of the verb:

(28) a. They made out that John was a liar
b. * They made that John out was a liar

[20] My use here of the term *Particle Verb* is not intended to imply any particular allegiance to the view that the verb and particle are some kind of unit lexically (a view argued for in e.g. Johnson 1991). This view is argued against extensively in Den Dikken 1992 (pp. 37-38), who treats particles as Preposition-like, ergative elements which head the predicate of a Small Clause. Den Dikken's account is discussed shortly.

This being so, Koizumi concludes that the particle in the "ECM" complex particle construction (29b) below must not be in the embedded clause; therefore, the ECM subject *John*, which shows up to the left of the particle, must not be in the embedded clause either - it would seem to have raised overtly.

(29) a. (?) They made out John to be a liar
b. They made John out to be a liar

3.2.2 Counterargument

Koizumi's argument for overt OS from complex particle constructions depends crucially on the assumption that the verb and particle cannot appear in different clauses. Unlike in the case of ordinary ECM constructions, there is no single standardly accepted analysis of complex particle constructions such as (29). But unfortunately Koizumi does not explicitly commit himself to any particular analysis, which makes it difficult to assess his argument. However, he does seem to implicitly reject one possible and plausible analysis, according to which particle constructions of both the simplex and complex type involve a Small Clause (SC) structure, headed (in some sense or other) by the particle itself.[21] An early version of this approach is Kayne 1984. Under such an analysis, contrary to what Koizumi supposes, the particle is by definition in a different clause to the verb associated with it,[22] so in theory, an element which occurs to the particle's left might be situated in the matrix or the embedded clause - in other words, under the SC theory of particle constructions, the fact that a DP can appear to the left of a particle is not really evidence that it has moved overtly out of its original clause, any more than is the fact that a regular

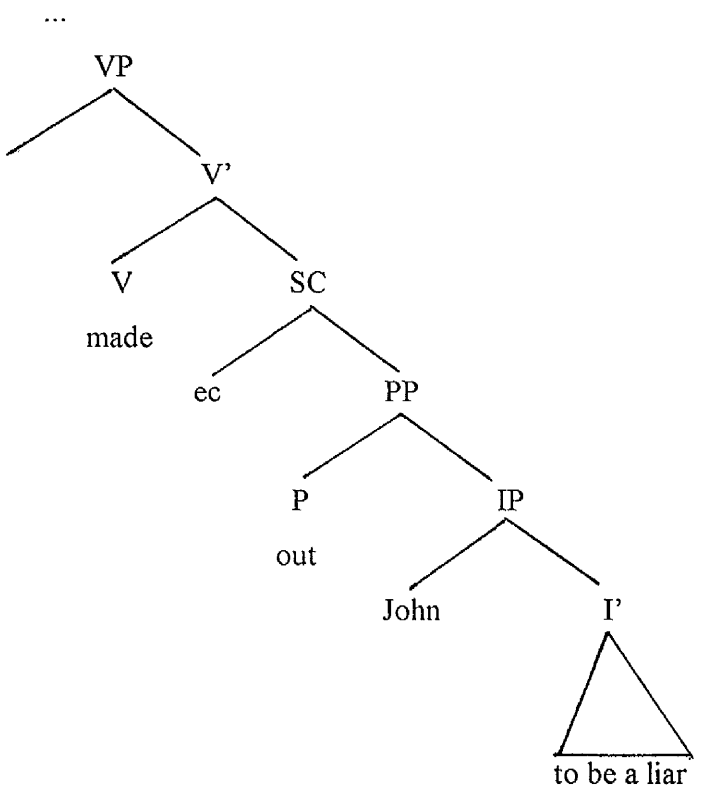
[21] There is some disagreement generally as to how to label the SC itself. In Stowell 1981, the SC is simply a maximal projection of the head of the SC predicate; in the particle cases, for example, it would be PP. There are various problems with this (see Williams 1994 for discussion) which seems to cause the question to frequently be avoided, as in for intance Den Dikken's account discussed in the text.

[22] Contra e.g. Johnson 1991; see note 19 above.

ECM subject can appear to the left of *to*.

A recent version of the SC analysis is developed by Den Dikken 1992, who proposes that the complex particle construction "involves a SC-in-SC substructure, the verb selecting a SC headed by an ergative particle which in its turn takes a second SC containing the additional secondary predicate" (p.65). Along the same lines, Den Dikken (p.52) proposes the underlying structure in (30) for the relevant part of (29b) (ec = empty category):

(30)



As for why the DP ends (or may end) up in front of the particle, there are a number of ways of describing this, not necessarily related to Case. Den Dikken himself (operating within basically a GB framework) claims that the DP moves in order to be assigned Case by the matrix verb in the SC subject position. This is not compatible with the Minimalist theory of Case, and so can be discounted here. A non-Case account of pre-particle

fronting is proposed by Svenonius 1995, who espouses a version of the SC analysis of particle constructions in which the SC is a Pred(ication)P (following Bowers 1993) with the particle-headed PP its complement. Svenonius suggests that there is an EPP-type feature (i.e. a strong D(P) feature) associated with Pred, which necessitates overt movement of the SC subject to check it.[23]

What is clear, in any event, is that Koizumi's argument for overt Object Shift from complex particle constructions is not compelling, since there is an obvious alternative (the SC analysis) to the assumption on which it crucially depends. Before leaving the topic of particles, we should look briefly at simplex particle constructions, as in (31) (from Kayne 1984:101). At first sight, it might be felt necessary to construe the object's pre-particle position as evidence for overt A-movement of regular objects.

(31) a. John looked the information up
b. John looked up the information

But again, if the particle *up* heads a small clause, independently of the verb, with *the information* the SC subject rather than the verb's object, - instead of the verb and particle originating as a unit - then the position of *the information* in (31a) does not necessarily imply that it has undergone overt object shift to check Case associated with the matrix verb. To conclude this section, I think that there is no strong evidence for overt OS from particle constructions, which brings us unfortunately no nearer to a resolution of the overt/covert Object Shift dilemma as yet. Let us now consider a third argument which Koizumi offers for overt Object Shift in Exceptional Case Marking constructions.

---

[23] Unlike these two authors, Kayne 1984 generates his SC subjects in Spec of the SC. In complex cases like (29b), the IP is the subject of the SC headed by the particle, which forces Kayne to postulate rightward movement of the *to*-constituent of the IP (i.e. leaving just its own subject behind) in order to obtain the correct order. Kayne also uses rightward movement of the SC subject in simplex particle constructions, to derive the optional Verb-Particle-Object order. See Den Dikken (*op. cit.*) for some arguments against Kayne's analysis. See below for some discussion of simplex particle constructions.

3.3 Deletion argument

3.3.1 For overt Object Shift

The third of Koizumi's three arguments for overt OS in ECM constructions concerns deletion, and is as follows. Citing Lobeck 1990, Saito and Murasugi 1990, he observes that only a constituent which is the complement of an agreement-inducing functional head (in the sense of Fukui and Speas 1986) may be deleted; hence we find contrasts as in (32) (= his (43)), on the assumption that +*wh* C is an agreement-inducing head - presumably reflected in the fact that something (the *wh*-phrase) checks features in its Spec - but *whether* (and similarly *if* and *that*) are not.

(32) a. We want to invite someone, but we don't know who [$_{IP}$ e]
b. * We thought Sue wanted to be invited, but we weren't sure whether [$_{IP}$ e]

The relevant element for the purposes of Koizumi's argument is *to*: this, it seems, is an agreement-inducing head in some situations but not others, as indicated by the fact that deletion of its complement is sometimes but not always possible. Following Martin 1992, Koizumi suggests that *to* is an agreement-inducing head in Control environments, something which he crucially attributes to the fact that *to* (i.e T) in Control infinitives assigns/checks Null Case with PRO, obviously making it an agreement-inducing head in the relevant sense (cf. also Chomsky and Lasnik 1993, Watanabe 1993). Koizumi's data is as follows:

<u>Control:</u>

(33) a. John wasn't sure he'd win the race, but he tried [PRO to [$_{VP}$ e]
b. (?) John convinced Bill to come to the party and Sarah convinced Mary [PRO to [$_{VP}$ e]

<u>Raising/ECM:</u>

(34) a. ?* Mary claims to not like baseball, but she appears to [$_{VP}$ e]
b. * John considers Mary to be clever, and Mike considers Sally to [$_{VP}$ e]

The fact that *to* in Raising/ECM environments apparently fails to behave like an

agreement-inducing head, i.e. no element is in its Spec, as deduced from the impossibility of deletion in the examples in (34), leads Koizumi to envisage the following problem: "...what forces the ECM subject to move from its base position to the pre-*to* position in overt syntax?...here we are faced with a paradox: *to* does not have features to attract the ECM subject, yet it appears to attract it to the pre-*to* position. Our overt raising analysis can provide a natural solution for this problem" (1995:36-37).

3.3.2 Counterargument

It could be argued that Koizumi's deletion argument for overt OS is on the wrong track, because contrary to the impression conveyed by his own data, the possibility of deletion of the complement of *to* does not in fact seem to relate to the Control versus raising/ECM distinction. Deletion after *to* is not necessarily possible with a Control infinitive, as shown in (35), and not necessarily impossible with a "raising" infinitive, be it in a regular raising to subject construction as in (36a), or in ECM constructions as in (36b,c) (to me, the illformed examples here are about as bad as deletion after *whether*; cf. (32b):[24]

<u>Control:</u>

(35) a. * John wasn't sure he'd be the winner, but he tried [PRO to [$_{VP}$ e]]
b. ?* Bill convinced John to be more confident, but Sarah couldn't convince Mary [PRO to [$_{VP}$ e]]

<u>Raising/ECM:</u>

(36) a. John claims to like baseball, but he didn't seem to [$_{VP}$ e] when I met him
b. Mary signed up for the baseball team, although John didn't want her to [$_{VP}$ e]
c. Mary and John's team won, though we hadn't expected them to [$_{VP}$ e]

---

[24] Thanks to M. Cobb, H. Gregory, K. Hayward, B. Howard, S.Lappin, A. Simpson, J. Watkins and G. Williams for their judgements on this data. The raising example (34a) which Koizumi claims to be ungrammatical is not. His example corresponding to my (34b) was also allocated a "?" by one informant.

In the light of (35) and (36), it appears that it is not after all the "agreement-inducingness" of *to* which is implicated in the contrast. Put in a more relevant way, since deletion *is* possible in raising/ECM contexts, it would have to be assumed that *to* is in fact an agreement-inducing head here, if the Fukui and Speas/Lobeck/Murasugi and Saito theory is to be upheld. But if *to* has to be assumed to be an agreement-inducing head in ECM constructions then, even if the identity of the features associated with *to* is unknown (perhaps the "EPP" is somehow involved - see Lasnik 1995), it is no longer any paradox that the ECM subject is attracted to the pre-*to* position, and there is certainly no need to assume that overt Case-related object movement must be responsible.[25]

I conclude that the deletion facts involving infinitivals, once we consider a wider range of them, cannot validly be argued from in favour of overt OS of ECM subjects, as Koizumi attempts to do. In the last three subsections, we have seen that three out of three of the arguments put forward by Koizumi 1995 in favour of overt OS in ECM constructions are dubious and/or inconclusive. I assume therefore that they can be put aside, and proceed to examine some further evidence.

---

[25] Looking informally at the data, what seems to be relevant to the possibility of deletion after *to* is whether or not the infinitival clause contains a main verb or *be* (cf. Akmajian & Wasow 1975): all the ungrammatical cases, namely Koizumi's (34b) (ECM) and my (35) (Control) involve *be*, while the good examples, namely Koizumi's (33) (Control) and my (36) (Raising and ECM) have main verbs: *win*, *come*, *like*, *sign up*. In discussing these deletion cases it should also be noted that there seem to be focus/contrast-like properties involved which may also affect well-formedness. Koizumi's example (34a), for instance, which he claims to be ungrammatical, is perfect if heavy stress is placed on *appears* (or with the additional of a contrastive adverbial like *nevertheless* - Ian Roberts p.c.) and slightly strange otherwise. The properly illformed examples with *be* in the infinitival clause cannot be improved in this way.

3.4 Extraction argument

3.4.1 For covert Object Shift

Branigan 1992 offers the following argument that English has covert as opposed to overt Object Shift, concerning the possibility of extraction out of complex DPs. Citing Bresnan 1972, he notes that there is a well-known asymmetry in the behaviour of subjects (and also DPs receiving objective Case in ECM, and Small Clause type constructions) on the one hand, and regular objects on the other: extraction is only possible out of the latter, as is well-known - with one theoretical construal of this being in terms of violation of Huang's (1982) Condition on Extraction Domain (CED).[26] The data is the following:

(37) a. * Who$_i$ did [pictures of t_i] offend Bill?
b. * Who$_i$ did John expect pictures of t_i to be on sale?
c. * What$_i$ did John paint pictures of t_i red?
d. * Who$_i$ did John look a reference about t_i up?
e. Who$_i$ did John see pictures of t_i?

If regular objects such as the *picture* DP in (37e) raised overtly to an argumental Spec position, we would expect any subsequent extraction out of them to be illicit, parallel to (37a-d) - yet this expectation is not met, since (37e) is perfectly formed.

3.4.2 Counterargument

Koizumi 1995 discusses the above argument of Branigan's (see also Lasnik 1995 for brief discussion) and dismisses it on the grounds that there are independent reasons for the illformedness of extraction from an ECM *picture*-DP. He suggests that the grammaticality contrast between examples like (37e) and (37b) should be attributed to what he calls "a

[26] Note of course that no corresponding argument for *overt* shift of ECM subjects could be made from the fact that extraction from these is *im*possible, since under the standard assumption that phrase structure is strictly binary branching (Kayne 1984, 1993; also Chomsky 1994) these must originate in a Spec position however the construction is analysed, even if one had them as the object of a "complex predicate" similarly to Larson's (1988) structure for secondary predication.

minimalist version of Huang's 1982 CED" (Condition on Extraction Domain). In cases involving a regular object, like (37e), Koizumi proposes that "the picture noun is in the minimal domain of its θ-role assigner. Suppose that a domain X is transparent with respect to extraction if (i) there is a head H that selects X, and (ii) X is in the minimal domain of H; it is opaque otherwise" (1995:37; note 22).

Technically, this explanation seems to work, although it could perhaps be seen as non-Minimalist in spirit, in the sense that by forcing the assumption that Spec-AGR_oP must count as being in the minimal domain of V, in order to allow the derived object to be extracted from, there is a similarity with the earlier GB theory of Case, in which the domains for the assigninment of Accusative Case and the internal θ-role were one and the same (specifically, government by V). In the Minimalist framework, recall that there is total complementary between complement domains and checking domains.

The question of whether (non)extractability from DP is a valid diagnostic for when that DP moves for Case-checking is an important topic in itself, so I propose to leave it open for the time being.[27] To conclude this section, let us look at what the placement of Floating Quantifiers can tell us about the timing of Object Shift in English.

3.5 Floating Quantifier arguments

3.5.1 For overt Object Shift

Bowers 1993 offers the following argument for overt object movement in ECM constructions, based on Floating Quantifier distribution. As well as being able to appear in its canonical position, as in (38a), he notes that the quantifier *all* can also appear detached or "floated" from its associated DP, either as in (38b) or (38c); the latter is the crucial case here:

---

[27] In Chapter 4 (§ 4), however, we shall see some evidence against an account of CED phenomena along the lines suggested by Koizumi, involving extraction asymmetries between overtly A-moved DPs (both subject and object) in German.

(38) a. John expects all his friends to like fruit
b. John expects his friends to all like fruit
c. John expects his friends all to like fruit

According to the most widely-accepted P&P analysis of the phenomenon, namely Sportiche 1988, Floating Quantifiers are "underlyingly" part of a constituent with the element with which they are semantically associated - *his friends* in (38) - and the quantifier becomes detached from its associate by leftward movement of the rest of the phrase. Given this analysis, (38c) suggests that the ECM subject is indeed outside of the embedded IP at Spell-Out, as Bowers concludes.[28]

3.5.2 Counterargument

There is a problem with the Floating Quantifier evidence for overt raising of ECM subjects which becomes apparent on considering the further example (39):

(39) What John wants more than anything is [$_{CP}$ for [$_{IP}$ his friends all to like fruit]]

If we suppose, quite reasonably, that *for* in (39) heads a CP taking an IP complement, it would seem that *his friends* must be within this IP, in spite of the fact that it appears to the left of a Floating Quantifier associated with it, and to the left of *to*. Whatever the explanation for this fact (either there must be a functional projection above the projection of *to* but still part of the IP complex - or possibly FQs are, or can sometimes be, adverbials[29]), it clearly indicates that the DP-FQ-*to* order is a possibility even when DP is

[28] Bowers in fact uses the Floating Quantifier facts as evidence for a "Raising-to-Object" analysis (cf. Postal 1974), developed within a somewhat different framework to that adopted in this thesis. However, as he notes, the Floating Quantifier evidence in itself favours an Overt-OS-to-AGRP analysis (e.g. Koizumi's) as much as it does a Raising-to-Object analysis, although he does then offer some further arguments in favour of the latter position (1993:620-621).

[29] As proposed recently by Bobaljik 1995. Competing analyses of Floating Quantifiers are discussed in Chapter 4 (§ 3) below).

unambiguously within its own clause. Given this, the possibility of this same order of elements in the ECM example (38c) cannot be construed as evidence that the ECM subject has moved overtly into the matrix clause.

3.5.3 For covert Object Shift

Floating Quantifier placement strongly suggests that regular objects do *not* raise overtly in English: they fail to license Floating Quantifiers, as shown in (40), thereby contrasting with ECM subjects (cf. (38b,c) above) and indeed with other subjects of various persuasions - (41b&c):

(40) a. John ate all the lemons
b. * John ate the lemons all

(41) a. All John's friends like fruit
b. John's friends all like fruit
c. John's friends were all given apples for christmas

The ungrammaticality of (40b) is problematic for overt object shift analyses of English - and this is so whether one assumes a stranding analysis of Q-Float, as in Sportiche 1988 (see above), or some kind of VP-adverb analysis (see Chapter 4, § 3.3.2).[30] Note that

---

[30] There is an apparent problem here, in that object-oriented Floating Quantifiers *are* possible in English just in case the object is a weak pronoun:

(i) John ate them all

However, note from (ii) that the pronoun and quantifier are inseparable by a VP-adverb, perhaps indicating that (i) is not best analyzed as the pronoun undergoing overt Object Shift. Contrast the situation with Norwegian, which also allows object-oriented FQs exclusively with weak pronouns: here, negation may intervene between pronoun and FQ, as in (iii):

(ii) * John ate them quickly all

(iii) Jeg leste dem ikke all
I read them not all
'I didn't read them all'

One possibility is that weak pronouns like *them* in English are heads and cliticize onto V, as

with respect to object-oriented Floating Quantifiers, English behaves like French, a language in which it has been argued on independent grounds by Koizumi 1995 (see above) that objects shift covertly, due to lack of agreement on the past participle (Kayne 1989)

(42) Jean a repeint(*es) les tables
J. has repainted(fem.pl.) the tables(fem.pl)
'Jean repainted the tables'

3.5.4 Counterargument

Koizumi claims that Floating Quantifiers *are* licensed by objects in English, giving the example in (43):

(43) * John read the books all

Since no speaker of English I have consulted agrees with his judgement on this data, it seems fair to discount it. However, a counterargument to the Floating Quantifier evidence for covert OS could be constructed following a suggestion made by Costa 1995 in

---

proposed by Diesing & Jelinek 1993 (following Haegeman 1992, Muysken 1983). Such cliticization, which would obviously not be available to full DPs like *the books*, could conceivably strand a quantifier. On the assumption (Chomsky 1995) that main verbs in English raise overtly out of VP (to *v*), it would remain to explain the impossibility of adverbial intervention between the pronoun and the stranded quantifier; an approach in terms of Larson's (1988) Light Predicate Raising is a possibility, although this is not the place to pursue it. Another possibility is that (i) involves movement of the pronoun into the specifier of the quantificational phrase, as proposed by Shlonsky (1991) for Hebrew. A different complication relating to Floating Quantifiers is that overt A-movement is not always sufficient to license them, a good example being (iv) (see Déprez 1989, Bobaljik 1995):

(iv) * <u>The lemons</u> were eaten <u>all</u>

However, see Chapter 4 (§ 3.1) in which I give evidence from a range of languages to substantiate the claim (Sportiche 1988) that Floating Quantifiers are licensed by overt Object Shift.

connection with subject-oriented Floating Quantifiers, as follows.

In an account partly akin to Koizumi's, Costa proposes that English has overt main verb movement and overt Object Shift, specifically making the following claim: "Nominal objects move overtly to Spec-AgrO, and the verb moves overtly to the first functional projection above AgrOP" (1995:25). Although Costa devotes himself mainly to arguing for short overt verb movement in this work, he does make an auxiliary claim about subject-oriented Floating Quantifiers which, if valid, could be simply extended to explain away the alleged Q-Float evidence for covert OS. I outline this now.

Costa's framework faces the problem that, contrary to what it predicts, it is not possible for *subject*-oriented Floating Quantifiers to appear postverbally in English, as shown in (44) (= Costa's (16a), p.27), an example originally attributed to Pollock 1989:[31]

(44) * <u>My friends</u> love <u>all</u> Mary

In short, Costa's response to this problem is to propose that floating *all* cannot adjoin to VP, nor to AGR_oP (the latter suggested to be for semantic reasons, as in Koizumi 1995 - see above, and note 12), but only to some higher projection, thus immediately predicting that *all* is illicit in its position in (44). This theory would extend straightforwardly to the case of object-oriented FQs too: if objects are overtly situated in Spec-AGR_oP, and the next projection down is VP, as Costa assumes, then there is nowhere but VP to which *all* in the ungrammatical (40b) (*John ate the lemons all*) could attach - hence the ungrammaticality.

This potential counterargument to the Q-Float argument for Covert OS suffers from problems of its own, however, since there appears to be evidence against its crucial assumption that Floating Quantifiers cannot adjoin to VP. In Icelandic, unlike in English, a subject-oriented Floating Quantifier can show up not only to the right of the verb (which

[31] Notice that this problem does not arise for Koizumi, since subjects originate in the Spec of the upper VP, whose head is the host of short overt verb movement.

clearly raises overtly) but also to the right of an overtly shifted object, as in (45) (data from Bobaljik 1995). If the shifted object, *húsið* 'the house' is in Spec-AGR_oP, as presumably it would be under Costa's analysis, then the subject-oriented Floating Quantifier must be adjoined to VP - unless there is some extra functional projection in between AGR_oP and VP - an otherwise clearly unmotivated assumption.[32]

(45) a. Í gær máluðu strákarnir húsið allir rautt
yesterday painted the-boys the-house all red
'yesterday all the boys painted the house red'

It seems unlikely that the hypothesized inability of "Floating Quantifiers" to adjoin to VP would vary parametrically, if, as Costa seems to imply in a footnote (*op. cit.*, p.12) this inability is ultimately semantic in origin. Given (45), Costa's hypothesis that Floating Quantifiers such as *all* cannot adjoin to VP, and its implication that FQs are "unreliable to test movement of the verb [and of course other VP-internal elements] out of the VP" would seem to be untenable, and so cannot in fact form the basis of a reasonable objection to the above argument from FQ placement that English has covert Object Shift - and we find ourselves with one resilient argument:- in favour of covert OS.

3.6 Section summary

In this section, we have collected and examined a range of arguments in favour of overt and covert Object Shift in English. Arguments for overt OS initially appear to be more plentiful. We looked at four of these involving the behaviour of ECM subjects with respect to adverb construal, particle constructions, deletion and Floating Quantifiers. But upon closer examination, each of these arguments turns out to be problematic. By contrast, there does seem to be at least one good argument against overt OS - and hence,

[32] I take it there is no possibility that *allir* 'all' could be adjoined any lower than the VP from which its associated DP *strákarnir* 'the boys' originated - i.e., it could not be adjoined to some putative functional projection belonging with the Small Clause, *húsið rautt* 'the house red' (assuming this Secondary Predication example to have some kind of an SC structure).

for covert OS, with the usual proviso - involving Q-float possibilities with regular objects. On the assumption that all Accusative Case features have the same properties, regardless of the construction they are in, and so are checked uniformly either by overt or covert movement, we conclude, following Chomsky 1995, Branigan 1992, Bobaljik 1995, that Object Shift is a covert operation in English, whether a regular object or an ECM subject is involved.

There remains one further and important matter to address before it can be claimed with a clear conscience that English is a covert Object Shift language. This has to do with the Case Adjacency phenomenon mentioned earlier, in § 2.2.

4 On the relevance of "Case Adjacency" effects

Recall the Case Adjacency effect, which any theory of English syntax needs to explain. The kind of data typically used to illustrate the effect is repeated here as (46): .

(46)	a.	*	John eats often/quickly lemons	V-Adv-DP
	b.	*	Mary believed often/sincerely John to be a fanatic	V-Adv-DP
	c.		Sue looked carefully at him	V-Adv-PP
	d.		Harry relies frequently on it	V-Adv-PP
	e.		Sue thinks sometimes that Harry is foolish	V-Adv-CP

The generalization, apparently, is that DPs with Accusative Case must appear strictly adjacent to the verb with the Case features (46a,b), while elements which are not in a Case relation with the verb are free of this restriction (46c-e). To capture this generalization, the adjacency effect has generally been attributed in some way or another to conditions on the assignment of Accusative Case (see the references in §2.2 above).

Recall that Chomsky 1995 and Koizumi 1995, who argue for covert and overt Object Shift respectively, both derive the Case Adjacency data from Case considerations (see §2.2). One could easily argue that in this respect, Koizumi's account is superior to Chomsky's, which relies on the assumption that adverbs count as potential blocks for A-

movement and hence must possess the same Case and/or agreement and/or categorial features as DPs - an assumption which appears somewhat contrived. Since the Case Adjacency data is of central importance, the fact that Koizumi's overt OS theory accounts for it more plausibly might then be taken as grounds for overturning our earlier decision that OS is covert in English. However, I do not think this would be the right step, for reasons which I explain next.

4.1 Case, Adjacency, and a Misconception

There exist in the literature a number of arguments that Stowell's original (1981) account of the Case Adjacency phenomenon is unsatisfactory. A conceptual objection sometimes raised against Stowell's account concerns its use of the notion *adjacency:* in the words of Koizumi 1995, "the Adjacency condition is problematic on conceptual grounds, since most syntactic principles of grammar are stated in terms of hierarchical notions such as "domination" and "c-command" ... linearity and adjacency surely play important roles in PF, but they do not seem to be relevant in syntax proper" (p.18).

This conceptual objection is very reasonable, provided that an explanation of the Case Adjacency data in terms of syntax proper is indeed required - which is of course the standard assumption. However, I think there is interesting evidence to suggest that Case (nor indeed any pure syntactic consideration) is *not* the crucial factor, although at first sight it certainly appears to be heavily implicated. We will see presently that neither all nor only elements in need of Accusative Case are subject to the type of adjacency effects which have conventionally been explained in Case-theoretic terms. As a result, I would suggest that there is actually no direct connection between Case and the so-called Case Adjacency phenomena, and in fact, as we will see below, it seems that no purely syntactic factor is involved. In a sense, therefore, Koizumi's conceptual objection to Stowell 1981 (while valid in and of itself), is founded on a misconception; the problem concerns not so much adjacency, as Case. Furthermore, while adjacency itself may not be the exact notion, it certainly seems that the phenomenon requires a PF explanation of some sort. All this,

to be elaborated next, has clear implications for the Object Shift question.

4.1.1 Not only XPs needing Case are subject to Adjacency effects

(47)-(51) are examples of adjacency effects between the verb and an element with which it presumably has no Case relationship: in (47), an adverb; in (48), a PP; in (49), a locative phrase in (50) a Measure phrase (a quasi-argument, in the sense of Rizzi 1990);[33] in (51), an adjectival Small Clause; in (52), the subject of a quotative inversion construction:[34]

---

[33] On Measure Phrases, see Lee 1995 who proposes that these elements bear ADV(erbial) features (p.67). See also Kitahara 1994, who claims that "...quasi-arguments do not bear any Case-feature, hence they do not undergo any NP-MOVEMENT for Case-checking". As evidence, for this (citing Chomsky 1989) he points out the following contrast between quasi- and true arguments (p.128):

(i) * 150 pounds were weighed by John
(ii) The potatoes were weighed by John

Hornstein 1994 gives a different argument that Measure Phrases are adjuncts rather than complements, involving antecedent-contained ellipsis, the traditional QR analysis of which he proposes to replace with a Minimalist Case-theoretic account which exploits the fact that objective Case involves movement out of the VP. Consider (iii) (p.473, note 27):

(iii) John wants to weigh as much as Bill does $[_{VP}$ e]

Here, $[_{VP}$ e] can be interpreted as 'weigh' but not as 'wants to weigh'. Hornstein attributes this to the fact that the phrase containing the ellipsis, *as much as Bill does*, is not an argument, hence will not raise to a Case position (Hornstein assumes that *want* is a (LF) restructuring verb, so that the relevant postion would be Spec-AGR_oP dominating that verb.) Compare with (iv), with a true object, which has the second reading lacking in (iii).

(iv) John wants to weigh every potato that Bill does

[34] The adjacency effect in quotative inversion constructions is noted by Branigan and Collins 1993. They propose the following Case-related explanation: the verb in these constructions assigns Case to the subject in the latter's base position (i.e. Spec-VP), this assignment seemingly subject to some kind of adjacency condition since an adverb cannot intervene between verb and subject. Branigan and Collins motivate this proposal with cases like (i), in which they claim that

(47) * John runs often quickly (cf. John often runs quickly)
(48) * Mary goes often out (cf. Mary often goes out)
(49) * Sue lives now here (cf. Sue now lives here)
(50) * Bill weighed occasionally 10 stone (cf. Bill occasionally weighed 10 stone)
(51) * Jackie seems currently sick (cf. Jackie currently seems sick)
(52) * "Yes, I hear what you say," replied insincerely Dave (cf. "Yes, I hear what you say," replied the Dean insincerely)

The "non-Case Adjacency effects" in the (a) sentences are very much akin to the classic Case Adjacency violations in (46a&b) above. It could be objected at once that the non-Case Adjacency effects are not necessarily relevant, since their deviance could be due to independent factors; for example, in the case (47a) with the two adverbs, the problem might be due to some semantically-based constraint on the placement of different classes of adverb. However, it is interesting to observe that the unacceptable V-Adv-XP cases, though involving a wide range of different elements syntactically, all improve considerably, in some cases recovering fully, when the XP is made "heavy" - as of course do ordinary Case Adjacency violations (see next subsection):

(47') ? John runs often more quickly than Jim does
(48') Mary goes often to the pub at the end of the street
(49') Sue lives now in the room next to the Phonetics Lab
(50') Bill weighed occasionally more than his doctor had advised him was healthy
(51') Jackie seems currently a great deal more sick than the other patients
(52') ? "yes, I hear what you say," replied insincerely the Dean of the postgraduate school

Given this quite striking similarity between the Case and the non-Case Adjacency effects, it is surely desirable that they should have a unified explanation. Since obviously Case is not relevant in the non-Case effects, then presumably the explanation, whatever it is, will not be in Case terms. Moreover, given the variety of syntactically different types of element involved in the adjacency effects, it seems quite unlikely that the explanation will

---

Mary cannot get Case since this has been assigned to *Bill*. The problem with this account is that quotative inversion subjects clearly have Nominative Case - cf. (ii) - which is presumably not assigned by verbs.

(i) * "What is the exchange rate?" asked Bill Mary
(ii) "I don't know - why not go to the Bureau de change?" said she/*her

be in terms of "syntax proper" at all.

4.1.2 Not all XPs needing Case are subject to Adjacency effects

The non-Case Adjacency effects noted above become more interesting when juxtaposed with the well-known phenomenon of "Heavy NP Shift" (first discussed by Ross 1967). As is well-known, Case Adjacency violations improve a great deal if the DP is "heavy":

(53) John ate quickly [all the lemons he'd discovered in the fruitbowl]
(54) Please check carefully [the details on this page]

It cannot be concluded on these grounds alone that not all elements needing Case are subject to the Case Adjacency effect, since it has been argued that HNPS involves rightward movement of the heavy DP across the adverb, an operation which clearly could apply after Case is appropriately licensed. In order to show that the base order Verb-Adverb-DP can actually be grammatical, the HNPS explanation needs to be ruled out.

A widely-accepted way to test whether a constituent occupies a base or derived position is to find out whether it can be extracted from, for further extraction out of a moved element is assumed to be impossible (Wexler and Culicover 1980). For example, Costa 1995, investigating the position of Prepositional Phrases, shows that it is possible to extract from PPs following adverbs such as *carefully* (i.e. manner adverbs), as shown in (55):

(55) Which painter did Bill look carefully at the pictures of?

Costa concludes from this that the PP in (55) is not in a right-adjoined position, but rather in its base position. (55) contrasts significantly with (56): here we have an adverb *yesterday* which is independently known to be restricted to sentence-initial or sentence-

final position.[35] The order Verb-*yesterday*-PP can only be derived by rightward movement, and extraction from XP is illformed as expected:

(56) * Which painter did Bill look yesterday at pictures of?

Let us try now to apply the extraction test to heavy DPs. The procedure is partly complicated by the fact that the more heavy (or complex) the DP, and hence the more amenable it is to "HNPS" in the first place, the more likely it is that extraction from the DP is prevented on independent grounds, since a complex DP will contain more nodes, generally inducing Subjacency violations, as in (57b). This means that the fact that the DP is not extractable from when it occurs post-adverbially - (57c) - does not necessarily reveal the information we are looking for.

(57) a. John denied strenuously the claim that he'd been harassing Sue
b. * Who did John deny the claim that he'd been harassing?
c. * Who did John deny strenuously the claim that he'd been harassing?

However, if Subjacency-inducing environments are avoided, it is still interesting to note that extraction from a DP following an adverb can indeed be wellformed, or at least, as wellformed as its non-extracted-from counterpart; similar extraction data is discussed in Kayne 1994. (It seems incidentally that adverbs in general occur more happily in relative clauses than in questions, for reasons which may well be pragmatic in nature).

(58) The application forms, which I'd checked carefully every detail of,....
(59) The problem, which she explained fully only part of,...
(60) Mary, who I removed quickly all the unflattering photos of before she could get offended,...
(61) Peppermint tea I advised Bill to drink regularly at least two and a half litres of

As in the case of Costa's PP example discussed above, a contrast arises if the adverbs are replaced with *yesterday* or something similar:

---

[35] As demonstrated by (i) (cf. Costa 1995:26):

(i) (Yesterday) John (*yesterday) read the newspaper (yesterday)

(62) * The application forms, which I had checked yesterday all the details of,...
(63) * The problem, which she explained yesterday only part of,...
(64) * Mary, who I removed yesterday all the more unflattering photos of before she could get offended,...
(65) * Peppermint tea I advised Bill to drink every day at least two and a half litres of

Since extraction from heavy DPs can be licit, "underlying" V-Adv-DP order must be possible. That in turn implies that not all elements needing Case are subject to the Adjacency effect. Putting this together with our previous conclusion that not *only* elements needing Case are subject to the Adjacency effect, the implication is clear: Case has nothing to do with the "Case Adjacency" effect.

If the answer does not lie in Case theory, how can the Adjacency effects be explained? As "weight" is generally crucial in determining whether adjacency effects obtain, it seems probable that an account in terms of PF is needed, rather than something purely syntactic. That is, complement DPs (and other elements) are syntactically authorized to appear in the order Verb-Adv-DP, but are prevented from doing so by PF considerations when "light" or non-complex in some yet to be determined way.[36] To pursue such an account is beyond the scope of this thesis, but the basic relevance of the discussion in this section is clear: the success or failure of a theory to derive the "Case Adjacency effect" from conditions on Case assignment is not an issue, if the two are not linked. This removes our hypothetical reason for preferring Koizumi's (or Costa's) theory (overt Object Shift) to Chomsky's (covert Object Shift)

---

[36] As Ross 1967 (p.32) observes (citing Chomsky 1961), "'complex' [i.e. heavy] cannot be equated with 'long'. This is exemplified by (i)and (ii):

(i) * Please check carefully [the details] [on your day off] (cf. (54))
(ii) * John eats often [quickly] [when he is nervous] (cf. (47')

This does not suggest that a purely syntactic account of the phenomenon is mandatory, but does suggest that a PF explanation will be in terms of aspects of PF in which syntactic structure is relevant; for instance, Stress (cf. Cinque 1993).

5 Summary and conclusion

I have argued in this chapter that Object Shift is covert in English: when Spell-Out applies, objects must be *in situ*. This hypothesis finds support in Floating Quantifier facts (see below for more discussion of these). Counterarguments involving both placement facts from ECM constructions and also Case Adjacency phenomena were addressed and argued against in detail. Recall that covert movement is taken to involve adjunction of just the appropriate formal features to the appropriate head (see Chapter 2, § 3).[37]

---

[37] As noted in § 1 above, there is one potentially important factor which I have not considered here, and that is the relevance of interpretation-related facts (concerning Binding, NPI licensing, etc.) to the overt/covert movement question. Lasnik (1995, 1996) argues that overt as opposed to covert movement is important for these interpretive relations, and that given that ECM subjects (and to some extent, regular objects) seem from a variety of interpretive data (reciprocal binding &c.) to c-command matrix VP-adjuncts (Postal 1974), English must have overt Object Shift. This of course conflicts directly with the conclusion reached in this chapter. However, while Lasnik's arguments that overt movement is relevant for the various interpretive licensing relations are quite convincing, I think that there is some reason to question at least some of the interpretive data itself. Consider the following: (i) with its Principle C effect is taken to demonstrate that an ECM subject is high enough to c-command into the adjunct. Assuming that adjunct to be (right) adjoined to the matrix verb which it modifies, the ECM subject itself must be higher than the matrix VP:

(i) a. Joan believes he_i is a genius even more fervently than Bob_i does
b. * Joan believes him_i to be a genius even more fervently than Bob_i does

The ECM case (ib) contrasts straightforwardly with (ia), in which the subject is obviously within the embedded clause. But now consider the further data in (ii):

(ii) a. Joan believes everyone respect him_i even more fervently than Bob_i does
b. * Joan believes everyone to respect him_i even more fervently than Bob_i does

In (ii) there is the same contrast between the ECM and the non-ECM case, with a Principle C effect obtaining in the former - and yet here, the pronoun is deeply embedded within the lower IP, where it cannot by any stretch of the imagination c-command into the matrix VP-adjunct. The implication is that it is not, after all, the height of *him* in (ib) which explains the Principle C effect, but rather, a more general property of the construction in question. Although this topic needs much further investigation, it does seem that there may be more to the standard interpretive evidence for raising of ECM subjects than initially meets the eye.

Chapter 4

The Case of *wh*- objects and associates

1 Introduction

In Chapter 2, I proposed the Copy Hypothesis: all copies in a chain are active in the computational system. The arguments given for the the Copy Hypothesis were basically theoretical. To complement this, I show in the present chapter that the Copy Hypothesis has practical applications. The discussion centres around two types of *wh*-phrases in English - objects and associates of expletive *there* - and the means by which the Case of these phrases is licensed.

Recall from the Introductory chapter that under Minimalist assumptions, argument *wh*-phrases such as *who* and *how many computers* in (1) participate in more than one act of feature(s)-checking: (i) Case and agreement features, and (ii) some kind of operator-related features, forcing movement relations with two separate heads.

(1) a. Who does John like?
b. How many computers are there in the lab?

The fact that what we are accustomed to think of in terms of a single "element" may require licensing by movement to more than one position raises interesting questions about how derivational economy conditions work. In theory, it may happen that the different properties of the features to be checked induce conflicting movement requirements on the element which needs to check them. In English, this possibility is realized by the case of *wh*- objects and *wh*-phrase associates of expletive *there* (illustrated in (1a) and (1b) above respectively). In general, *wh*-movement is obligatorily overt in English,[1] driven by a strong C-related feature, but the Case-related movement of both objects (see Chapter 3) and associates of *there* is a *covert* operation. Hence the question arises: how to reconcile the necessity for the *wh*-phrases in (1) to move overtly and check the strong feature of C, on the one hand, and the requirement of Procrastinate that the non-strong Case/agreement features are *not* checked overtly, on the other? In this chapter, we shall see that the Copy Hypothesis provides a simple answer to this question - and moreover, an answer for which there seems to be some evidence.

This chapter is structured as follows. In Section 2, I explain the theoretical problem which examples of the type in (1) raise, proposing then that the Case/agreement of object *wh*-phrases and of *wh*-associates of *there* is checked covertly, like that of regular non-*wh* objects and associates, in compliance with Procrastinate. Specifically, I claim that the VP-internal copy (or perhaps just its formal features; Chomsky 1995) created by overt *wh*-movement to Spec-CP raises covertly to the relevant position for Case-checking. In Section 3, I present evidence for the proposed derivations concerning Floating Quantifier distribution. In section 4, drawing on work by Kitahara 1994, I provide further motivation for the proposal concerning extraction asymmetries between objects and associates of *there*, on the one hand, and subjects and adjuncts, on the other. In Section 5, concluding both this chapter and Part I, I show how the Copy Hypothesis makes it possible for the economy principle Procrastinate to operate in a somewhat more "refined" way than the standard framework allows for. This discussion prepares the ground for the theory of

---

[1] Except in the case of multiple *wh*-constructions (i) and so-called echo questions (ii):

(ii) Who bought what?
(iii) (John bought a laryngograph) He bought WHAT?

derivational optionality to be presented in Part II.

2 *Wh*- objects and associates of *there*

2.1 Theoretical questions

In Chapter 3 it was argued that Accusative Case is checked in English by covert movement. In other words, there must be no Strong feature associated with the relevant head *v* (or AGR_0, or Aspect, ...), and overt Object Shift, as in (2), is excluded by Procrastinate.

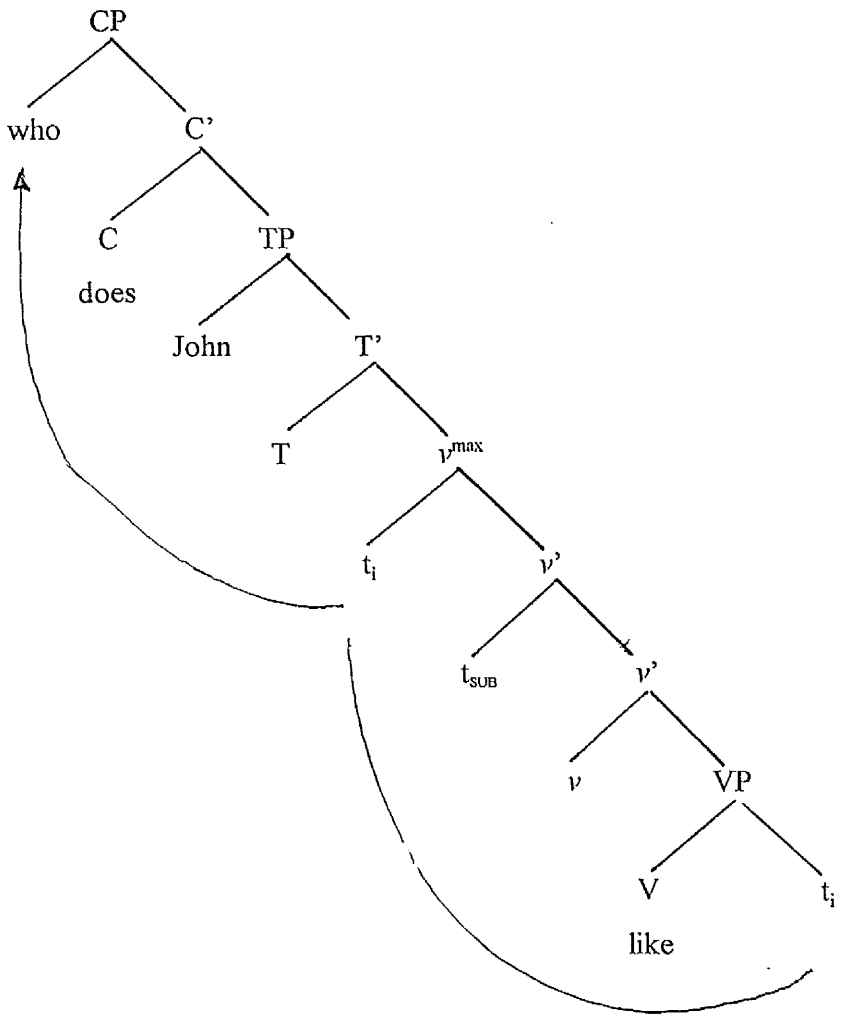
(2) * John Mary likes

However, consider a case in which the object is a *wh*-phrase, as in (3) (= (1a)):

(3) Who does John like?

Its obligatory clause-initial position clearly indicates that the object *who* needs to check a Strong feature of some kind, presumably on the head C. Furthermore, given that it is an argument, *who* must also move to check Case and Agreement features at some stage, in particular Accusative Case. We know from Chapter 3 that there is no Strong feature associated with these features (in contrast to Nominative, associated with Tense which also carries a strong D-feature). However, C(P), which contains the strong feature forcing overt movement of the *wh*-phrase, is obviously further from *who*'s initial position than are the Accusative Case and agreement features. As a result, it seems inevitable that in a construction like (3), there must be overt movement to/through the Case position Spec-*v*, i.e. overt Object Shift must be licensed, as in (4):

(4)



Whether one sees the movement in (4) as a classical step-by-step operation, or as an instantaneous application of Form Chain as in Chomsky 1993, Collins 1994,[2] it certainly appears that *who* has no option but to move overtly through its Case-checking position, if it is to land overtly in Spec-CP, as is required for convergence. Clearly, covert movement to the Case position "followed" by overt movement to Spec-CP is contradictory, at least in the standard Minimalist model in which overt/covertness is thought of as a difference in the timing of operations relative to Spell-Out.[3]

---

[2] See Chapter 5 for discussion of these two methods of movement.

[3] I assume that there is no possible derivation in which the *wh*-phrase moves to Spec-CP, with it or its relevant features later moving back down to check the Case and Agreement features,

But why and how should *wh*-phrase objects like *who* in (3) be licensed to move overtly to their Case position, while ordinary objects, like *Mary* in (2), are prevented from doing so? It does not seem plausible to suggest that the *wh*- and non-*wh*-constructions differ with respect to the properties of a Case or agreement-related feature, or that Procrastinate might apply in the latter but not the former. Therefore it seems necessary to assume that Procrastinate has the very global characteristic of being able to consider the properties of non-local elements when evaluating a particular step of movement. That is, in (3), Procrastinate will need to know that a *wh*-phrase must be in Spec-CP by Spell-Out, and accordingly, allow the phrase to move overtly to Spec-*v*. I take it that it would be preferable not to have to attribute this property to economy conditions. At the same time, as we have observed, there is apparently no alternative -- unless, I argue, we assume the Copy Hypothesis as proposed in Chapter 2.[4]

2.2 Proposal [5]

As I have just presented it, the question is how to rule *in* overt Object Shift in cases like (3), at the same time as ruling it *out* in cases like (2). However, given the Copy Hypothesis, repeated here as (5), this question - along with its problematic implications - need not arise at all.

(5) **The Copy Hypothesis** (Chapter 2, p.53)
All copies of a chain are active in the computational system

---

since if such "backwards" attraction were to be possible at all, it should be ruled out by the Minimal Link Condition due to the intervention of the subject in Spec-TP; the subject will have interpretable agreement features which should make it a closer candidate for attraction by the head bearing Accusative Case features.

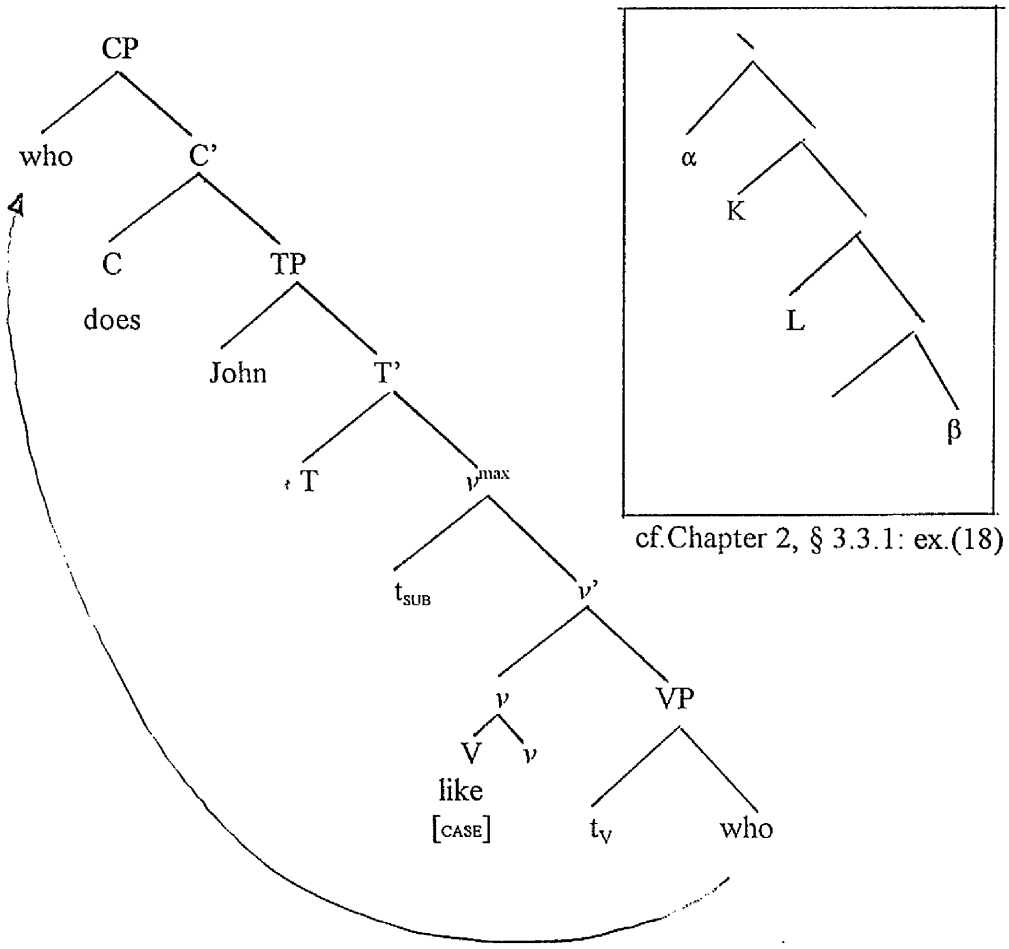
[4] The problem I describe was more complicated in the 1993/1994 Minimalist framework, in which overt Object Shift was excluded in English by the Minimal Link Condition (see Chapter 1). See Pettiward 1995 for discussion.

[5] The proposal of this section is based on Pettiward 1994.

With (5) at our disposal, there is naturally no theoretical necessity to assume that there is an overt Object Shift component to examples like (3), and as a result, no need to weaken any conditions which are supposed to ensure in general that Accusative Case is checked covertly. My hypothesis is that in no circumstances does movement to/through Spec-*v* take place overtly in English. Evidence for this will be presented shortly, but firstly let us elaborate the proposal.

The derivation I propose for object *wh*-movement in English (as illustrated by (3) above) is as follows: I assume firstly that the *wh*-phrase moves directly to Spec-CP, in order to eliminate the Strong feature of C; this step is shown in (5). Notice that this movement, in skipping the Case position, does not violate the Minimal Link Condition (Chomsky 1995), in that the latter is formulated from the point of view of the feature moved to: "K [= a target for movement] attracts F if F is the closest feature that can enter into a checking relation with a sublabel of K" (p.297). This formulation makes Chomsky's MLC spiritually similar to Rizzi's (1990) Relativized Minimality; an "intervening A-position" will have no relevance to an operation of A'-movement (or vice versa) as far as locality is concerned.

(6) a. Derivation for (3) - Step 1

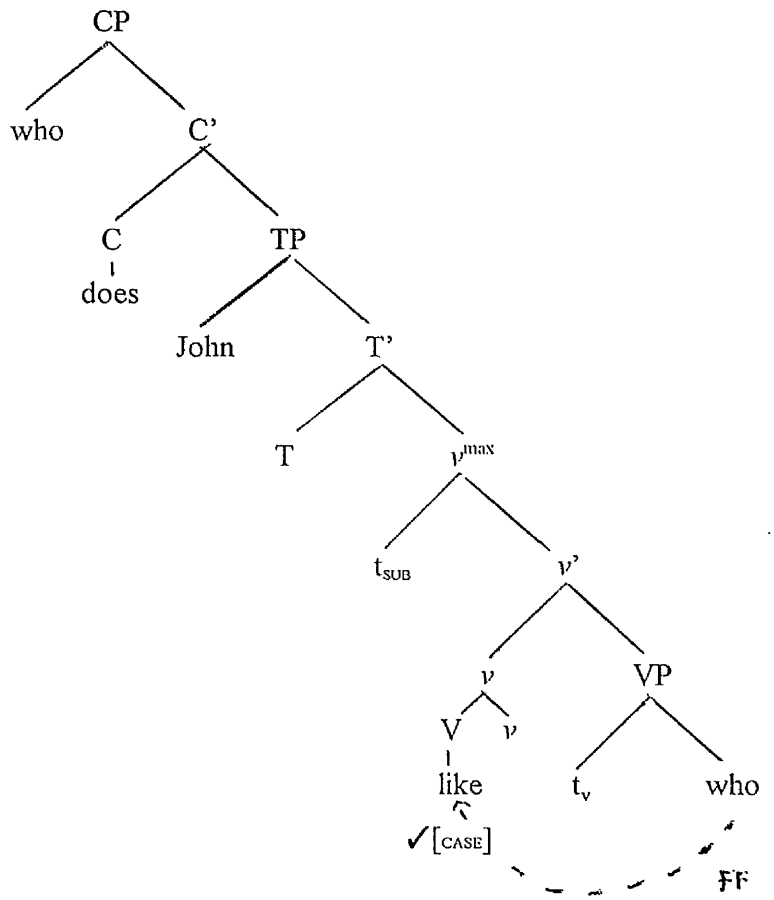


Given the copy theory, movement to Spec-CP as illustrated above leaves a copy of *who* in its initial position. When Spell-Out applies, note that Accusative Case and agreement features will be unchecked; these features are negligable at PF and will not cause the derivation to crash. However, the Case and agreement features are not negligible at LF, and must be checked eventually, otherwise the derivation *will* crash. At this point, it is clear that we have another instantiation of the situation considered in §3.3.1 of Chapter 2 when discussing the movement possibilities of traces (see (6a), inset): a chain of two copies - <*who*, *who*> - with a relevant target for movement - namely the V/*v* complex - intervening between them.

Following an earlier proposal (Pettiward 1995), I assume that checking of the

remaining Case and agreement features is implemented by raising of the lower copy of *who*, on the assumption that "apart from phonological aspects, a copy is...identical to its 'original', so that the copy left by *who* should itself be unchecked for Case" (p.196). This proposal may easily be modified in line with the recent (Chomsky 1995) conception of covert movement as raising of formal features alone, so that just the copy of the *feature* in question, rather than the copy of the whole phrase, adjoins to the complex head bearing the Case features.This operation is illustrated in (6b):[6]

(6) b.Derivation of (3) - Step 2



[6] It should be mentioned that Kitahara 1993, 1994 has independently proposed the same derivation for English object *wh*-movement as that proposed here (see § 4 below for further discussion of Kitahara's theory). In addition, Chomsky 1995 independently proposes to permit " the Case feature F in the [= trivial] argument chain headed by t [= trace left by *who's* movement to Spec-CP] to raise covertly for Case checking, which now deletes and erases it in both positions of the chain (F, t_F formed by the operation (and in the operator)" (1995:303) (see below for discussion).

A question immediately raised by this analysis is how, or when, does the copy of the *wh*-phrase in Spec-CP count as checked with respect to the Case feature which it, along with its copy left within VP, must possess? Evidently, it needs to be assumed that the copy in Spec-CP somehow "inherits" Case-checkedness when the Case feature in the *v*/V complex is eventually checked by raising of the lower copy's features. Notice that something along these lines would need to be assumed in any event, whether or not we countenance derivations involving movement of copies. For instance, if one assumed a more conventional step-wise derivation for (3), as pictured in (4) above, in which the object moves first to Spec-v, then on to Spec-CP, then the copy left in the VP-internal position by the first step must be counted as becoming checked for the operator-feature, as a result of the second step of movement to Spec-CP taking place. It could be generally assumed then that when some feature is checked by an element, all copies of that element (i.e., the chain) are considered checked for that feature (see Pettiward 1995 for discussion, and also Chomsky (1995), according to whom "the simplest assumption is that the features of a chain are considered a unit; if one is affected by an operation, all are" (p.381, note 12).[7, 8]

---

[7] For some discussion of the question of whether a derivation employing movement of a copy should be taken to violate LF wellformedness conditions (as suggested in Chomsky 1995a), see Pettiward 1995:195.

[8] It is probably worth considering whether the question about *wh*-/Case movement addressed here also arises in the alternative minimalist framework of O'Neill and Groat 1995 (O'N&G) mentioned in Chapter 2 (note 23). In O'N&G's theory, "covert movement" amounts to movement/copying of an element minus its phonological material, and must happen unless the feature to be checked is Strong. O'N&G's system, though different to Chomsky's in that Spell-Out applies to complete LFs only, resembles Chomsky's in that it has Procrastinate. The *wh*-movement question arises in essentially the same way as it does in Chomsky's framework. In O'N&G's theory, unless movement of/from copies is allowed, it has to be assumed that in cases like (3) (*who does John like?)* the *wh*-phrase takes its phonological material with it to the Case position (Spec-AGR_oP in their system), overriding/violating Procrastinate.

2.3 *Wh*-phrase associates of *there*

As mentioned in note 6 above, Chomsky (1995:303) independently suggests movement of features of a trace for the case of object *wh*-movement. Interestingly, he also points out a further situation in which it seems necessary to assume movement of traces, or their formal features: constructions featuring expletive (i.e. semantically null) *there* and a *wh*-phrase "associate". To see why, let us firstly consider some regular *there* constructions, illustrated in (7):

(7) a. There are three men sitting in the garden
b. John expected there to be a lot of people at the party
c. There arrived a man from the Inland Revenue
d. There were several people arrested that week

For a number of reasons, *there* is standardly assumed to be related to its associate (i.e. *five men*, *a lot of people*, etc) by covert movement of the latter.[9] Firstly, as is well-known, the verb shows agreement with the associate rather than the expletive:

(8) There are/*is several computers in the lab

Chomsky 1995 (p. 274) shows also that for the purposes of control, the associate of *there* behaves as if it is higher than its overt position: in (9a), for example, the adjunct *without* PRO *identifying themselves* requires subject control; this could not be achieved by *three men* in its *in situ* (VP-internal) position. The fact that (9a) is grammatical implies that the associate (or its formal features) raises covertly to a position, namely T(P), where it can control PRO in the adjunct, allowing the reflexive *themselves* be bound. A similar situation obtains in (9b) (from Uriagereka 1988 cited in Lasnik 1996):[10]

---

[9] See Chomsky 1991, Chomsky and Lasnik 1993 for different theories.

[10] Lasnik in fact argues that in general, covert movement does not alter binding (and other interpretive licensing) possibilities, although he does not dispute Chomsky's claim that expletive-associate constructions (or at least those involving *there*) do involve such movement, on the basis of the agreement facts noted above in the text.

(9) a. There arrived three men without identifying themselves
b. There arrived two knights on each others' horses

Chomsky (*op. cit.*) assumes that the motivation for raising of the associate is checking of Case and Agreement features: *there* is not capable of checking these features itself.[11]

Associates of *there* thus resemble objects in that both raise covertly for Case. Observe incidentally that associates pattern with objects - see (10) - rather than subjects and adjuncts - see (11) - with respect to CED-type phenomena (Huang 1982, Chomsky 1986): extraction from an associate DP causes no violation ((11a,b,f) are taken from Chomsky 1986).[12]

(10) a. The man who we saw [pictures of t] [objects]
b. Who did John see [pictures of t]?
c. The man who there was [a picture of t] recently in the papers [associates]
d. Who was there [a picture of t] in the papers?
e. Who was there thought to be [a picture of t] in the papers?

(11) a. * The man who [pictures of t] are on the table [subjects]
b. * The book that [reading t] would be fun
c. * Who did [pictures of t] upset John?
d. * Who were [pictures of t] believed to be in the papers?
e. ?* Who did we expect [pictures of t] to be on sale?
f. * To whom did they leave [before speaking t] [adjunct]

Coming back to the main point, (12) shows that the associate of *there* may be a *wh*-phrase ((12a) taken from Chomsky 1995:302):

(12) a. Guess what there is in the room?
b.. How many men did there seem to be in the garden?
d. How many men were there expected to be there?

---

[11] Unlike *it* and French expletive *il* - see Chapter 5 below.

[12] The extraction behaviour of associates of *there* is discussed in § 4 below.

Since there there can be little doubt that when the associate of *there* is *not* a *wh*-phrase, it does not raise overtly to the relevant Case position (T(P)), the question again arises, as with *wh*- objects, of how Case does get checked in examples like (12). In addition, unlike with *wh*- objects, it does not even seem likely from a technical perspective that *wh*- associates undergo exceptional overt movement to check Case, since the expletive itself already occupies Spec-TP.[13] Chomsky proposes that in cases like (12), Case and agreement features are checked by covert movement of the features of the *wh*-associate's copy, analogous to the object *wh*-movement derivation discussed above.

It is obvious that Chomsky's *wh*-associate data described here strengthens the case for allowing trace movement in general. Of course, at a conceptual level there is some difference between Chomsky's proposal and that of this thesis, which I shall now briefly discuss.

2.4 Comparison with Chomsky 1995

To conclude this section, let us briefly compare the current proposal with that of Chomsky 1995 with which it has similarities. Both in the current proposal and in Chomsky's, formal features of a trace (= lower copy in a chain) undergo a movement operation. However, it is significant to recall that Chomsky simultaneously (1995) proposes that in general, traces are resistant to movement operations, or "immobile", in his own terms (see discussion in Chapter 2, § 4). This latter feature of his theory forces Chomsky to introduce a qualification such that "...traces cannot attract and their features can be attracted only under narrow conditions..." (*op. cit.*, p.304).

Is this a purely stipulative addition to an already stipulative proposal? Chomsky

---

[13] One could possibly say that the *wh*-associate adjoins to the expletive, then excorporating and moving further to Spec-CP (E. Benmamoun p.c.). However, to assume this method of Case-checking for the associate might require some otherwise unmotivated relaxation of the definition of checking configuration.

suggests that it is not: "...formal features of trace are deleted (hence erased) if they are not necessary for the formation of legitimate LF objects that satisfy FI ... when *wh*-movement or some other form of operator raising takes place, the trace left behind heads an argument chain and must have the full complement of features: Interpretable features required for interpretation of the arguement at LF, and -Interpretable features that have not been checked (otherwise, the Case feature is never checked, remaining in the operator, and the derivation crashes. We conclude, then, that in A-movement the formal features of the trace are deleted and erased, but in *wh*-movement ... these features remain intact" (1995:303). This idea might perhaps be criticized on the grounds that it is undesirable to allow purely computational operations such as deletion and erasure to be directly sensitive to conditions on the wellformedness of LF (or PF, for that matter) (this is discussed in Chapter 2 (§ 4.2)). On the other hand, given the Copy Hypothesis proposed in Chapter 2, note that it is absolutely unnecessary to make any modifications of the above sort in order to account for the *wh*-movement cases in question, since movement of copies is freely available.

A different aspect of Chomsky's proposal about the Case-checking of *wh*- objects and associates is that he does not discuss any motivation for it, aside from its apparent theoretical necessity. This point brings us to the next two sections, in which I provide evidence for the existence of the copy-movement derivations just proposed. I begin by considering some Floating Quantifier facts which strongly suggest that object *wh*-phrases do not move overtly through their Case position in English.

3 Floating Quantifiers and *wh*- objects[14]

As we saw in Chapter 3, the overt position of elements is in general quite difficult to determine in the Minimalist system, and this difficulty only increases when the elements under investigation are themselves phonologically covert (i.e. traces). There is however one type of data which is frequently used to reveal the position of such elements, and this concerns the distribution of so-called Floating Quantifiers (hereafter FQs), commonly (though not universally - see below) believed to be licensed exclusively by (overt) A-movement (Sportiche 1988, Déprez 1989). After carefully collecting cross-linguistic evidence in support of the generalization that object-oriented FQs are possible if and only if a language has overt Object Shift (i.e. overt A-movement), drawing on data from German, Dutch, Japanese, Icelandic and French, I then turn to the behaviour of object *wh*-phrases in English. These, it is found, do not license FQs, a conclusion which is quite unexpected if *wh*-objects are taken to move overtly through Spec-*v*, but which on the other hand, very clearly supports the present proposal that Accusative Case-checking by object *wh*-phrases in English is covert, implemented by copy movement.

3.1 Sportiche's (1988) movement analysis

Quantificational elements, e.g. *all*, canonically make up a constituent with the phrase which constitutes their restriction, as in (13a). However, it is well-known that certain quantifiers - in English, *all*, *each* and *both* - may show up in an unexpected position, as does *all* in (13b-d), in which case they are commonly known as Floating Quantifiers:

[14] For judgements on the data in this section, I am grateful to Jeanne Cornillon, Agnès Dahan, Johanes Flieger, Bjarnir Magnusson, Lutz Marten, Taeko Maeda, Wilfried Meyer Viol, Stefan Ploch and Malte Zimmermann.

(13) a. All the children saw a film
b. The children all saw a film
c. The children have all seen a film
d. The children were all frightened by the film

Sportiche 1988 devised a highly influential syntactic account of this phenomenon. On the basis of data from French and English, he argued that quantifiers like *all* in (13b-d) are stranded by leftward movement of the associated material (*the children* in the above example): "any time a Q appears adjacent to an empty NP [i.e. NP trace]", Sportiche maintained, "the illusion of floating will be created" (*op. cit.*:435). Thus in the cases in (13), *all* is taken to be left behind in the subject's VP-internal position when the latter moves for Case (or perhaps EPP) reasons.

Since Sportiche's analysis will play a leading part in my FQ argument about *wh*-objects in English, it is worth taking some trouble to establish independently whether it seems to be on the right track, especially with respect to objects, which have not been so widely investigated in this connection. The reader may recall from Chapter 3 that in English, subjects but not objects license FQs. This is reillustrated in (14):

(14) * John saw the films all

However, it is important to note that in the present context, it is not legitimate to take this subject-object contrast in English as corroboration of Sportiche's analysis, since our main positive argument that objects do not move overtly in English (Chapter 3, § 3.5.3) was itself dependent on the assumption that Sportiche's analysis was correct. Moreover, even if we found out somehow that Sportiche's analysis *is* correct, English offers the additional complication that derived subjects, which clearly do move overtly, still for some reason fail to license a post-verbal FQ, as Sportiche himself notes:

(15) a. * The tables were repainted all

What this suggests is that that overt movement is not sufficient to license FQs, though it may well be necessary (perhaps independent factors of some kind rule out the stranding

of a quantifier in its VP-internal position).[15] Given (15), it is not safe to conclude from the ungrammaticality of (14) that overt Object Shift in English is prohibited. In order to establish a more stable link between overt movement and Floating Quantifier possibilities (and thereby reinforce the conclusion we will later reach on the partial basis of this assumption), it is necessary to examine situations where there is independent evidence for whether Object Shift is overt or covert, and then look at what kind of FQ patterns occur.

Let us start with French. Recall that regular objects in French have been argued not to check Accusative Case overtly, on the basis of the fact that past participles do not exhibit overt object agreement in such constructions, as in shown in (16) (see Koizumi 1995, also Kayne 1989). Note that just like English, French fails to license FQs with regular objects, as shown by (17a). Thus objects contrast with subjects, with which FQs are well-known to be licit (17b).

(16) Jean a repeint(*es) les tables
J. has repainted(fem.pl.) the tables(fem.pl)
'Jean repainted the tables'

(17) a. * Jean a vu <u>les films</u> <u>tous</u> [object]
J. has seen the films all

b. <u>Les enfants</u> ont <u>tous</u> vu le film [subject]
the children have all seen the film
'the children all saw the film'

Spanish, which I assume lacks overt Object Shift, like French, has a similar asymmetry in FQ-orientation: *todo* 'all' can be stranded by subjects but not by objects:

(18) <u>Los niños</u> han visto <u>todos</u> la pelicula [subject]
the children have seen all the film
'the children all saw the film'

(19) * Maria ha visto <u>las peliculas</u> <u>todas</u> [object]
M. has seen the films all

[15] Though this is certainly not the case universally - as shown by the case of Icelandic to be discussed below.

These facts obviously suggest a correlation between impossibility of overt Object Shift and impossibility of object-oriented FQs, thereby supporting Sportiche's theory. Note further that FQs are also not possible with the associate of an expletive *there* in English (see Bobaljik 1995), which we can also independently assume does not move overtly:[16]

(20) a. (?) There arrived all (of) the five men from Mars
b. * There all arrived the five men from Mars [associate]

Now let us look at some languages in which there is independent reason to believe that overt Object Shift *is* possible. Here it turns out that FQs associated with objects are indeed permitted; the striking subject/object Q-Float asymmetry found in English and French is missing. Firstly, consider the situation in German and Dutch, "SOV" Germanic languages with a type of short object "scrambling" which has been quite convincingly argued by several authors to be a case of A-movement (see e.g. Vanden Wyngaerd 1989, Zwart 1993 for Dutch, Déprez 1989, Mahajan 1990 for German).[17]

German

(21) a. ...weil die Kinder diesen Film alle gesehen haben [subject]
... because the children this film all seen have
'... because the children all saw this film'

b. ... weil Stefan die Filme gestern alle gesehen hat [object]
... because S. the films yesterday all seen has
'...because yesterday Stefan saw all the films'

[16] The existence of the so-called Definiteness Effect in such constructions - namely the fact that the associate must be "indefinite" - renders them slightly odd whether there is Q-Float or not, although I think that the relevant contrast does exist.

[17] For a lot more about this scrambling, in particular its apparent optionality, see Chapter 6.

Dutch

(22) a. De studenten zijn alle naar London gegaan [subject]
the students are all to London went
'the students all went to London'

b. Ik heb die mensen gisteren allen ontmoet [object]
I have those people yesterday all met
'yesterday I met all those people'

As (21b) and (22b) show, objects are able to strand quantifiers in both German and Dutch; following standard practice (e.g. Diesing 1990), the adverbs *gestern/gisteren* 'yesterday' are taken to mark the lefthand edge of VP, indicating that the objects have moved overtly.[18]

Next, consider Icelandic, an "SVO" Germanic language in which overt A-movement of full objects may take place in certain circumstances, essentially if [19] and only if there is overt verb movement.This is illustrated in (23) (from Bobaljik and Jonas 1994 cited in Bobaljik 1995:28); the position of the negative element *ekki* 'not' is taken to indicate the lefthand border of VP:

(23) a. Jólasveinarnir borðuðu bjúgun ekki
the-Christmas trolls ate the-sausages not
'the christmas trolls did not eat the sausages'

b. * Hann hefur bókina ekki lesið / ✓ ekki lesið bókina
he has the-book not read
'he has not read the book'

As in Dutch and German, not only subject-oriented but also object-oriented FQs are possible in Icelandic, provided the verb has raised (examples from Bobaljik 1995, citing

[18] See Merchant 1995 for a detailed movement analysis of Floating Quantifiers in German.

[19] Overt verb movement may not be quite sufficient, as it seems also that the shifted object must be "definite" (see Bobaljik 1995). On interpretive effects relating to Object Shift in Icelandic, and other languages, see Chapter 6, § 5.

Jonas and Bobaljik 1993 (24a), Bobaljik and Jonas 1994 (24b), Vikner 1991 (24d)).[20]

Icelandic

(24) a. Í gær máluðu strákarnir húsið allir rautt [subject]
yesterday painted the-boys the-house all red
'yesterday the boys all painted the house red'

b. það borðuðu margir strákar bjúgun ekki öll [shifted object]
there ate many boys the-sausages not all
'many boys didn't eat all of the sausages'

c. Á barnum drakk stúdentinn bjórinn studum allan [shifted object]
in the-bar drank the-students the-beer sometimes all
'the students sometimes drank all the beer, in the bar'

d. Hann las bækurnar eflaust ekki allar [shifted object]
he read the-books doubtlessly not all
'he undoubtedly didn't read all the books'

Moreover, if the conditions are such that the option of overt Object Shift is excluded, as in (25) (no overt movement of the main verb), object-oriented FQs cease to be possible ((25b is from Bobaljik 1995:246); (25a) was checked with an Icelandic informant)):

(25) a. * Á barnum hefur stúdentinn drukkið bjórinn allan [*in situ* object]
in the-bar has the-student drunk the-beer all

b. * Á barnum hefur stúdentinn allan drukkið bjórinn [*in situ* object]
in the-bar has the-student all drunk the-beer

Once again, we have evidence for Sportiche's theory linking (overt) movement and Q-Float. Finally, consider the situation in Japanese, where both subjects and objects, which have been argued on independent grounds to undergo overt A-movement (see e.g. Déprez 1989, Koizumi 1993, 1995)), may strand Numeral Quantifiers ((26) is taken from Koizumi 1995:108; Cl. = Classifier):[21]

---

[20] On the apparent dependence of overt Object Shift on overt verb raising (usually known as Holmberg's Generalization (after Holmberg 1986), see Chapter 1 and Chapter 6.

[21] Korean also has Floating Numerals with both subjects and objects (Jiyoung Shin p.c.).

Japanese

(26) a. Gakusei-ga 3-nin piza-o tabeta [subject]
Students-Nom. 3-Cl. Pizza-Acc. ate
'three students ate pizza'

b. John-ga piza-o Mary-ni 2-kire ageta [object]
J.-Nom. pizza-Acc. Mary-Dat. 2-Cl. gave
'John gave Mary two slices of pizza'

Summing up so far, I have outlined Sportiche's hypothesis that "any time a Q appears adjacent to an empty NP the illusion of floating will be created" (1988.:435), and examined the behaviour of object-oriented FQs in the light of this. Investigating cases where it is known on FQ-independent grounds whether or not objects move (overtly), i.e. languages other than English, we find that FQs indeed behave in a way which is highly supportive of Sportiche's theory. From considering a good range of data, no situations have been found in which there is overt Object shift but no possibility of object-oriented FQs (cf. German, Dutch, Icelandic examples (24b-d), Japanese), nor conversely where there are object-oriented FQs, but no overt Object Shift (French, Icelandic example (25), and along similar lines, the associate of *there* in English). On the basis of this, let us subscribe to Sportiche's analysis, in particular making the following assumption about objects:

(27) Object-oriented FQs are possible iff overt Object Shift is possible[22]

With this assumption in mind, let us now look at the interaction of FQs with *wh*-movement in English.

[22] Provided of course that a language allows FQs at all. Many if not most languages do, it seems. A few more examples are: Arabic, Hebrew, Italian, Modern Greek and Swedish. A language which apparently does not allow FQS is Haitian Creole (Ian Roberts p.c. citing M. de Graff p.c.).

3.2 *Wh*- objects and Floating Quantifiers in English (and other languages)

The fact is that in English, there are no FQs associated with object *wh*-phrases. The relevant data is given in (28) and (29). I have used a nonrestrictive relative clause and a case of topicalization (standardly assumed to involve A'-movement), since *wh*-phrases with another quantifier are in general slightly odd in questions, as can be seen from (30):[23]

(28) a. The children, all of whom John took to the cinema yesterday, ...
b. The children, who John (* all) took to the cinema (* all) yesterday,...

(29) a. All those films, John had wanted to see before he died
b. Those films, John had (*all) wanted to see (* all)before he died

(30) a. (?) All (of) which children did John take to the cinema yesterday?
b. Which children did John (* all) take (* all) to the cinema yesterday?

The assumption (27) stated above entails that if overt Object Shift is possible, then object FQs are possible. From (28)-(30) we may assume that object FQs are not possible with object *wh*-phrases. From here it is straightforward to conclude that overt A-movement is not possible for these *wh*- objects. This very clearly supports the proposal in Section 2 above with respect to object *wh*-movement: Accusative Case is checked not by "overt Object Shift" in such contexts, but covertly, i.e. by feature-raising from the copy of the *wh*-phrase left within VP by overt movement to Spec-CP.

3.3 Some potential counterarguments

3.3.1 Are Floating Quantifiers *not* stranded by movement?

The above argument against an overt A-movement component to object *wh*-movement in English relies heavily on Sportiche's (1988) analysis of Floating Quantifiers. There is however an alternative to Sportiche's analysis which claims that FQs are not stranded by movement of an associated DP, but rather adverbial elements adjoined to VP. This type

[23] The same data is used independently by Déprez 1989 in a different context - see below.

of analysis is exemplified by Dowty and Brodie 1984. A more recent version of the adverbial analysis is proposed by Bobaljik 1995, who draws attention to certain apparent shortcomings of Sportiche's (1988) theory. If the movement analysis of FQs is indeed incorrect, does this undermine the argument of the last section? As it turns out, it does not.

Bobaljik argues against Sportiche's (1988) stranding analysis of Floating Quantifiers on two main grounds. Firstly, he observes that FQs "do not surface in positions where one would posit DP traces, except when those positions are coextensive with independently motivated adverb positions" (p.242); secondly, he notes that "Floating Quantifiers surface in a healthy array of positions in which one would not wish to posit a subject trace, though these are clearly potential sites for adverbs" (*ibid.*). The point that FQs sometimes fail to surface where a DP trace should be is illustrated by the case of Passive in English, mentioned earlier ((31) = (15) above):

(31) a. * <u>The tables</u> were repainted <u>all</u>

To illustrate the converse fact that FQs may appear where there is no NP trace, Bobaljik gives examples of FQs appearing at the left of certain adjuncts, as in (32) (1995:212-213):

(32) a. <u>Larry. Darryl and Darryl</u> [sic] came into the café <u>all</u> very tired
b. <u>The magicians</u> disappeared <u>all</u> at the same time

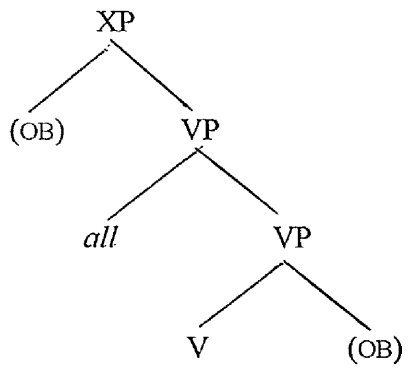
On the face of it, data like (31) and (32) go against Sportiche's theory. Bobaljik thus puts forward an alternative theory according to which "floating quantifiers like *all* do not directly modify a DP, but rather modify the predicate in a predictable manner with respect to some DP" (*op. cit.*:192).[24] This clearly allows for the adjunct cases like (32); it also predicts more or less that no FQ can appear to the right of *in situ* VP-internal elements - e.g. verbs or objects, explaining the impossibility of object-oriented FQs in

[24] Obviously Bobaljik assumes implicitly that this (or maybe all) adjunction must be to the left only, otherwise false predictions would ensue, for English at least.

covert Object Shift situations, and the impossibility of postverbal FQs in the case of Passive.

However, it will be noticed that with respect to *objects* (though not other elements, for instance, derived subjects), Bobaljik's and Sportiche's theories make basically the same prediction, as we can see from (33): if the object stays within VP, then a VP adverb will not show up to its left. The crucial assumption for our purposes, (27) - "Object-oriented FQs are possible iff overt OS is possible" is not really affected.

(33)



Having said that assumption (27) is not affected by adverbial/movement controversy, there is in fact one aspect of (27) for which this might at first appear to be false. Note that Bobaljik's theory falls short of Sportiche's in an important respect: it offers no obvious explanation for the fact that, descriptively speaking, FQs are licensed only by an *overt* A-chain (be it headed by an overt NP, or by a variable (see below, (36)/(37) for the latter); it seems that the DP associated with the FQ ("antecedent") must c-command the FQ "at Spell-Out" - see the examples in (34). Similarly, Floating Quantifiers tend to show agreement with the antecedent, as in the French example (35). Needless to say, these facts are not unexpected under Sportiche's movement analysis:

(34) a. * John <u>all</u> likes <u>the books</u>
b. * There <u>all</u> arrived <u>the five men</u>
c. * Into the bar <u>all</u> walked the students

(35) Les femmes ont toutes/*tous lu ces livres
the women(f.pl.) have all(f.pl.)/*all(m.pl.) read these books
'the women all read these books'

However, it is important to realize that these facts represent a *general* problem for the adverbial theory, since the overt c-command requirement on FQs pertains to all elements rather than exclusively to objects. Hence, if there is no adequate explanation of this, the implication is not that my FQ-related argument about objects is invalid, but rather that the entire adverbial approach is wrong. In summary then, it seems to me that the argument from Q-Float against overt OS given above remains unaffected by the movement/adverbial debate about Floating Quantifiers.

3.3.2 Are FQs *never* licensed by A'-movement?

Sportiche 1988 and Déprez 1989 argue that A'-movement fails to license Floating Quantifiers. In fact, Déprez uses examples equivalent to (28)-(30) as evidence for this.[25] Given this, one could object that lack of FQs in an A'-movement environment will not tell us anything one way or the other about whether *A*-movement is going on.

However, this objection is not a good one, since in the Minimalist framework, all Case-checking, and therefore Case-checking by *wh*-phrases, involves A-movement. And not surprisingly, FQs *can* be associated with an A'-moved phrase, just in case conditions are such that overt A-movement is permitted. Good examples of this are Dutch and German, in which overt Object Shift seems to be routinely allowed (see § 3.1). Here, Floating Quantifiers are possible with an A'-moved phrase, as shown in (36) and (37) (here again I have used nonrestrictive relative clauses, for the reasons noted in connection with the English examples earlier). Note that FQs can be stranded at the site of the

[25] In Déprez's system, developed within the GB framework, Case-related A-movement was taken to occur in some, but not all, languages - and not in English. Hence the assumption that A'-movement in English could not possibly have an A-movement component, and subsequent construal of the lack of FQs here as evidence that they are not licensed by A'-movement.

variable as well as in the object's base position:[26]

Dutch

(36) Die mensen, die ik allen gisteren allen ontmoet heb,...
the men, who I (all) yesterday (all) met have,...
'the men, all of whom I met yesterday,...'

German

(37) Die Filme, die Stefan (alle) tatsächlich (alle) gesehen hat,...
the films, which S. (all) indeed (all) seen has,...
'the films, which Stefan did indeed see all of,...'

Now compare this with languages known independently to lack overt Object Shift, such as Swedish and Norwegian. (38) illustrates the impossibility of overt Object Shift in Swedish (from Bobaljik and Jonas 1994, cited in Bobaljik 1995:28).[27]

(38) * Tomtarna åt korvarna inte
the-Christmas-trolls ate the-sausages not
'the christmas trolls did not eat the sausages'

As in English, FQs are impossible with object *wh*-phrases, as shown by (39) and (40)

---

[26] See also Hindi, which is argued to have (the possibility of) overt Object Shift, and where a Floating Quantifier is possible with an overtly Topicalized object, according to Mahajan 1990.

[27] In fact, overt Object Shift *is* possible in Norwegian and Swedish, but only with weak pronouns. Accordingly, FQs are licensed in this situation, as in (i) (Norwegian). An apparently similar phenomenon occurs in English, as in (ii) (this is discussed in Chapter 3, note 30). Cf. perhaps also the French example in (iii); an FQ is marginally possible with an object clitic (though not of course with a full NP).

(i) Jeg leste dem ikke all
I read them not all
'I didn't read all of them'

(ii) I read them (*those) all

(iii) Je les ai lu tous
I them have read all
'I read them all'

(from Déprez 1989):

Swedish:

(39) a. * Dessa flaskor vin har min kamrat alla druckit
these bottles wine has my friend all drunk

b. * Vilka flaskor vin har min kamrat alla druckit?
which bottles wine has my friend all drunk

Norwegian (from Déprez 1989)

(40) * boeken, att jeg ikke leste alla,...
the-books, that I not read all,...

We have seen in this subsection that *wh*-movement with an overt Object Shift component generally licenses Floating Quantifiers.[28] This reinforces our earlier conclusion that there

---

[28] The relevant facts in Icelandic are less clear. The expectation is that FQs should be licensed with *wh*-movement in just those environments in which the *wh*-phrase is permitted to undergo overt Object Shift prior to moving to an A'-position (i.e. when there is overt verb raising). However, consider the example in (i), Bobaljik 1995 (p.131), citing Déprez 1989, and interestingly used by both authors to argue that A'-movement fails to license FQs:

(i) * bækurnar, sem Jón keypti ekki allar,...
the-books, which J. bought not all,...

The verb *keypti* in the relative clause seems to have moved overtly, which should make overt Object Shift, and hence object-oriented FQs, possible. However, in the preliminary investigation of Icelandic which I have undertaken myself, it turned out that a quantifier may float from a "topicalized" object, as in (ii), in which it will be noted that the verb of the relevant clause is raised. On the other hand, consider (iii), where a similar-looking Floating Quantifier is possible even though the verb of the relevant clause is a past participle and presumably does not raise - i.e. it seems that overt Object Shift (and concomitantly object-oriented FQs) may be unexpectedly licensed in the context of A'-movement (cf. next section in text on French, and Chapter 5):

(ii) Dessar vínflöskur drakk vinur minn allar
these bottles of wine drank friend my all
'all these bottles of wine, my friend drank'

(iii) Dessar vínflöskur hefur vinur minn (*allar) drukkid allar
these bottles of wine has friend my (* all) drunk all

is no overt Object Shift subpart to object *wh*- movement in English, an environment in which object-oriented FQs are not possible at all. Before leaving the subject of FQs, let us briefly address an apparent problem from French.

3.3.3 FQs and A'-movement in French

Remember that in French, Object Shift has been argued to be covert on the basis of past particple agreement facts (Koizumi 1995, also Kayne 1989). If French is generally a covert Object Shift language, then given the proposal about English in § 2.2 above, French is expected to have the same derivation as English for object *wh*-movement, with covert Case-checking, and the two languages should behave alike with respect to Floating Quantifers in this environment. However, this does not seem to be the case at all: compare (41) with English (29)-(31) above:

(41) Les tables, que Paul a toutes repeintes,...
the tables, which P. has all repainted
'the tables, all of which Paul repainted,...'

So in French, it looks as if there is indeed an overt A-movement component to object *wh*-movement. Furthermore, there is independent evidence for this apparently unexpected situation, since overt past participle agreement, which is impossible with non-*wh*-objects, *may* appear with an overtly fronted *wh*-object (see Kayne 1989), as can also be seen in (41).

Furthermore, some sort of a Floating Quantifier does seem to be marginally possible with an object *wh*-phrase, as in (iv) (compare with Déprez/Bobaljik's (i) above):

(iv) Dessar vínflöskur, sem hún keypti allar, ...
these bottles of wine, which she bought all, ...

In view of the fact that the Icelandic counterpart of *all* can have readings in which it is not directly quantifying a nominal expression (as is also the case in English; see Bobaljik 1995, chapter 4), and until further research is undertaken, it is difficult to draw any definite conclusions from this data.

One way to deal with this apparent counterexample would be to adopt Déprez's (1989) proposal that cases like (41) do not involve a true Floating Quantifier, but rather are the product of a separate rule "L-*tous*" (proposed originally by Kayne 1975) which shifts the quantifier to the left. However, instead of taking this way out, I think it is possible to continue to analyse (41) a true case of Q-float (arguably the null hypothesis[29]) - and even turn it to the advantage of my own proposal. Notice firstly that past participle agreement with a moved *wh*-object, while possible, is not necessary, which, under the standard view that agreement is the reflex of (overt) movement of the agreed-with element, indicates that overt Object Shift of *wh*-phrases through Spec-*v* is *optional* in French - that is, it seems that a derivation involving copy movement is possible, though for some reason not obligatory, as it is in English.

(42) Les tables, que Paul a repeint,...
The tables(fem.pl.), which P. has repainted
'The tables, which Paul repainted,...'

The apparent optionality of overt Object Shift in *wh*-movement in French (and lack of this optionality in English) needs to be explained of course, but since this is a major topic in its own right, I delay addressing it until Part II (Chapter 5), where a full analysis is provided.

In conclusion of § 3, I have investigated the behaviour of Floating Quantifers and *wh*-objects in English and found evidence in support of my § 2 proposal that movement for Accusative checking in this environment takes place covertly, by movement of the copy. In the next section, I turn to some different evidence for the proposal, concerning extraction asymmetries.

---

[29] A Floating Quantifier can occur at exactly the same spot in French Passive. Notice also that something similar happens in English, as the translation of (i) shows.

(i) <u>Les tables</u> ont été <u>toutes</u> repeintes
the tables have been all repainted
'the tables were all repainted'

4 Approaching extraction asymmetries

In this section, building on a proposal by Kitahara 1994, I show how the copy-movement analysis of *wh*-movement outlined in § 2 above provides for an account of the well-known contrast in extractability between objects and associates of the expletive *there*, on the one hand, and subjects, adjuncts and quasi-arguments, on the other (contrasts accounted for in the GB model by the Empty Category Principle (ECP) - see below). Extending the investigation to Condition on Extraction Domain (CED) type phenomena (Huang 1982), it is suggested that a derivational approach along the lines of Kitahara's - crucially utilizing the notion of copy movement - is descriptively more satisfactory than a possible alternative account appealing to an extended notion of θ-government (as contemplated in Koizumi 1995, Lasnik 1995). Extraction data considered in this section comes from English, German, Dutch and French.[30]

4.1 A Minimalist dilemma

As is well-known, objects behave differently to other types of elements under both total and partial extraction. For example, an object can be extracted out of a strong island - a *wh*-island or a complex NP, say - with significantly less deviant results than can subjects, adjuncts, and quasi-arguments (e.g. measure phrases). (43) illustrates with *wh*-islands:

[30] Before embarking on this section, it should be noted that constraints upon extraction is still very much open within the Minimalist framework. Among other issues, there is the question of how to replace notions like *head government* which were crucial in GB theories of locality, but not technically tenable in Minimalism (see e.g. Roussou 1994 for a head government-free account of *that*-t effects). Important though these issues are, I shall not attempt to address them here, since this would be another thesis in itself. Thus, the discussion in this section takes place at quite an abstracted level.

(43) a. ? What did you wonder when John saw?
b. * Who did you wonder when saw John?
c. * Who did you wonder when was seen?
d. * How did you wonder when John killed Bill?
e. * How much did you wonder when John weighed?

All the examples in (43) involve some syntactic violation, conceived of in the GB framework in terms of the Subjacency condition,[31] with even (43a) being imperfect (compare with the perfectly well-formed *what did you think (that) John saw?*). The noticeably greater deviance of the (43b-e) examples was attributed to the fact that they also violate the Empty Category Principle (ECP) (see e.g. Chomsky 1981, Huang 1982, Rizzi 1990, Cinque 1990, Lasnik and Saito 1992). According to Rizzi's (1990) so-called conjunctive version of the ECP, traces "must be (i) properly head-governed (Formal Licensing)" and "(ii) antecedent-governed or Theta-governed (Identification)" (1990:32).[32] Under the assumptions of the GB model, within which Rizzi's theory is formulated, the contrast between (43a) and the rest is accounted for quite straightforwardly: the trace of *who* in (43a), assigned Case *in situ* under government by the verb, is both properly head-governed (by V) and θ-governed and hence obeys the ECP. Compare this with (43b-e), in which the relevant traces of the extracted *wh*-phrases are neither properly head-governed, nor θ-governed or antecedent-governed, which more suffices for the traces to violate the ECP.[33]

---

[31] See e.g. Chomsky 1986b.

[32] Rizzi defined head government as follows (1990:25): X head governs Y iff

(ii) a. X is a head
b.X m-commands Y
(ii) X = {[+/-V +/-N, Agr, T}
(iii) a. no barrier intervenes
b. Relativized Minimality is respected

Proper head government is head government "within the immediate projection" [i.e. c-command] (31). Clearly some of the notions involved in Rizzi's theory are not tenable in that form within Minimalist framework (see note 30 above).

[33] One might wonder, incidentally, why subject extraction (as in e.g. *who left?*) is ever permitted, given that subjects do not seem to be properly head-governed (here Rizzi's system contrasts with the earlier "disjunctive ECP" (Chomsky 1981), in which antecedent government alone

Even assuming that relevant technical notions such as head government can be appropriately replaced, the Minimalist framework raises a more general question concerning the treatment of extraction asymmetries: as Branigan 1992 points out, the Spec-head theory of structural Case appears to flatten out significant structural asymmetry between objects and other types of element. Given the unified Spec-Head theory of Case, the trace left by object extraction will itself be in the Spec of a functional projection (Spec-*v* in current terms, Spec-AGR_oP in Branigan's), and if extraction takes place across an island, the trace left in this Spec should violate the ECP, since it not properly head-governed. In order to preserve the essence of the ECP account within the Minimalist system, two basic types of strategy seem to be available: (i) somehow make Spec-*v* count as properly head-governed - and as θ-governed, if we are assuming Rizzi's conjunctive ECP (cf. Koizumi's (1995) suggestion with respect to CED phenomena), or (ii) somehow make extraction out of Spec-*v* not take place at all, at least in cases where object extraction is legitimate, as in the well-known case of English. An approach along the lines of (ii) is proposed by Kitahara 1994 in connection with extraction asymmetries of the type in (43) above. Kitahara's account, which I shall now outline, relies crucially upon assuming that object *wh*-movement in English does not involve the *wh*-phrase moving overtly through Spec-*v* - that, in turn, implying that Case is checked by raising of the copy. The account also extends quite attractively to the extraction behaviour of associates of *there* (not considered by Kitahara himself), which behave like objects.

4.2 Kitahara 1994

As mentioned earlier (note 6), Kitahara 1994 independently proposes an analysis of English object *wh*-movement which, like my own proposal in § 2.2 above, invokes covert movement of the VP-internal copy of the *wh*-phrase for Case-checking. On the basis of this, Kitahara proposes the following account of the object-subject/adjunct asymmetry illustrated in (43) above. The first step in object extraction, consisting of direct movement

---

was sufficient). In fact, Rizzi (*op. cit.*:60) discusses cross-linguistic subject-extraction data which suggests that this "surprising consequence" of his system may actually be to its advantage.

of the *wh*-phrase to Spec-CP, violates the Shortest Move Requirement,[34] since it must cross the embedded CP containing another *wh*-phrase in its Spec. However, the second, and final, step of the derivation, consisting of the VP-internal copy of the *wh*-phrase raising to check Case features, creates with no violation a chain with links at the three relevant positions (Spec-CP, the Case position and the base position). The idea behind this highly derivational account seems to be that the second movement can to some extent "repair" the Shortest Move violation initially incurred.

By contrast, this method of evading, or repairing, violation is not an option for *wh*-subjects, adjuncts or measure phrases should they find themselves within a *wh*-island. A *wh*-subject, for example (as in (43b,c)), cannot possibly move to Spec-CP skipping its Case postion, with a later, legitimate movement of the trace for Case-checking (as a *wh*-object can) simply because there is a Strong feature in T which cannot be skipped without causing the derivation to crash. This then forces the *wh*-subject to move overtly to Spec-TP, from where it must subsequently move across the island in order to reach Spec-CP. But this move violates the Shortest Move Requirement, and no legitimate chain is formed. *Wh*- adjuncts (e.g. *how* in (43d)) and measure phrases (e.g. *how much* in (43e)) are also unable to escape the "ECP" violation: they must move to the higher Spec-CP, crossing the *wh*-island, but unlike object *wh*-phrases, they get no second chance to form a chain without violation, since there is no other site at which they need to check further features.[35]

4.3 Extracting associates of *there*

Kitahara's account can also handle the extraction behaviour of associates of expletive *there*, and we shall see that this data gives us an initial reason to favour such an approach over a hypothetical alternative appealing to an extended notion of θ-government (i.e. a

---

[34] Kitahara's (*op. cit.*:61) version of Chomsky's Minimal Link Condition. The Shortest Movement Requirement states: "Minimize the length of each feature-checking movement".

[35] See Chapter 3 (note 33) below for references on Measure Phrases lacking Case features.

strategy of type (i), in the sense described above).

Recall that in § 2.3 above, following Chomsky 1995, I proposed that in constructions with *there* and a *wh*- phrase associate, Nominative Case (also agreement) is checked by raising the formal features of the copy left by overt movement of the *wh*-associate to Spec-CP. As we have already seen (§ 2.3 above), associates of *there* pattern with objects with respect to CED effects. Now as (44) shows, they also behave like objects - (45a) - rather than subjects (or adjuncts, measure phrases, etc.) - (45b) - under total extraction:

(44) ? What were you wondering when there would be?

(45) a. ? What were you wondering when John saw?
b.. * Who were you wondering when saw John?

A similar contrast between associates/objects and subjects/adjuncts is witnessed in extraction from *whether* clauses, with extraction of the former giving rise to minor violations, as opposed to much stronger ones in the case of the latter, as shown in (46). Similarly, A'-movement of the associate of *there* fails to give rise to the *that*-trace effects typical of subjects - (47).

(46) a. ? Who did you wonder whether John saw?
b. ? What did you wonder whether there might be in the garden?
c. * Who did you wonder whether saw John?
d. * Who did you wonder whether was seen?
e. * When did you wonder whether John saw Bill?

(47) a. How many people did you think that John would invite?
b. How many people did you think that there would be at the party?
c. * How many people did you think that would come to the party?
d. * How many people did you think that were offended by John's behaviour?

Kitahara's (1994) above-described account of the island behaviour of objects can also explain the ability of *wh*-phrase associates of *there* to evade "ECP" violation, as in (44): just like objects, there is no need for associates to check Case features via overt movement (recall that the expletive itself deals with the strong D(P) feature of T). Thus,

although *wh*-movement to the matrix CP violates the Shortest Movement, a licit chain can subsequently be formed by covert raising of the copy of the *wh*-associate (or copies of the relevant features) to the Case position. If we assume that the asymmetries in (46) and (47) are essentially of the same nature (see e.g. Rizzi 1990 for an ECP account of *that*-trace violations), then these too will be amenable to the same approach.

The important point in all of this is that Kitahara's approach to ECP-type extraction asymmetries depends on assuming derivations for object (and associate) *wh*-movement in which the *wh*-phrase does not move overtly through its Case position. To the extent that the account is successful, then, we have some motivation for adopting such an analysis of object (and associate) *wh*-movement, and, ultimately, motivation for the Copy Hypothesis as proposed in Chapter 2. Note that it would not be unreasonable to expect that a similar approach could be taken to CED phenomena,[36] since, as the examples in (48) will remind us, these too exhibit the object/associate versus subject/adjunct pattern, with proper head-government (or some "Minimalist" version of it - see below) clearly relevant:[37]

(48) a. Who did John see [pictures of t]?
b. Who was there recently [a picture of t] in the papers?
c. * Who did [pictures of t] upset John?
d. * Who were [pictures of t] believed to be in the papers?
e. * To whom did they leave [before speaking t]

Kitahara (*op. cit.*:92) does in fact develop a Minimalist account of CED asymmetries in terms of what he calls the Inner Minimal Domain Requirement (IMDR) on movement, which essentially comes down to a prohibition on extraction out of an element in the Spec of the head of a nontrivial chain. Although Kitahara's main concern is an extraction domain asymmetry between subjects in English and Icelandic (see note 41

[36] Condition on Extraction Domain (CED): Extraction out of a category K is possible only if K is properly governed (Huang 1982).

[37] See Rizzi *op. cit*, p.108: "the structural conditions on a trace and on the extraction domain are fundamentally homogeneous".

below), it is clear that in order for the IMDR system to achieve the correct result for objects and associates in English - i.e. that they are possible extraction domains), it must be assumed that objects do not move overtly to Spec-v - the very assumption which Kitahara's account of ECP asymmetries depends upon. From now on, let us call the Kitahara-style approach to ECP and CED extraction patterns *the Derivational Approach.*

At this stage it is relevant to compare the Derivational Approach with a hypothetical alternative which would make no appeal to copy movement. As mentioned very briefly in Chapter 3 above, Koizumi (1995) sketches what he calls a "Minimalist version of Huang's 1982 CED" as a method of explaining the contrasts exemplified in (48) above. Koizumi suggests that the Case position to which objects move (Spec-AGR_oP in his framwork) qualifies as θ-governed, due to the fact that the verb adjoins to its head. As a result, he claims, extraction from this position is legitimate.[38] Again, it is natural to wonder whether Koizumi's suggestion might be extended to account for the extraction asymmetries with respect to islands, given that the same factor, basically close relatedness to a lexical head, seems to be crucial throughout. Such a hypothetical extension of Koizumi's suggestion, which I shall refer to as *the Representational Approach*, exemplifies strategy (i) in the sense described above - in contrast to the Derivational Approach.

When it comes to the extraction contrasts between objects and subjects/adjuncts, there is no difference in the results of the Derivational and Representational Approaches. But here the associate extraction facts discussed in this section become relevant. For with associates, unlike with objects, there is obviously no possibility of explaining the relative

---

[38] "The picture noun is in the minimal domain of its θ-role assigner. Suppose that a domain X is transparent with respect to extraction if (i) there is a head H that selects X, and (ii) X is in the minimal domain of H; it is opaque otherwise" (Koizumi 1995:37; note 22). Koizumi's suggestion is made in the context of his argument for overt Object Shift in English, specifically to refute the argument of Branigan (1992) that the CED behaviour of objects is evidence that they do not move overtly (see Chapter 3 (§ 3.4.2) below). A more or less identical suggestion is made by Lasnik 1995, also as part of an argument on behalf of overt Object Shift and against Branigan *op. cit.*

legitimacy of extraction (either of or from the element) in terms of the Case position, Spec-TP, counting as θ-governed, since this would sacrifice the account of regular subject extraction violations. This then means that in the case of associates, one has to take the approach that it is lack of overt movement through the Case position which makes extraction grammatical -- an approach which requires a copy movement analysis of the derivations in question. However, if it is necessary to resort to a copy movement analysis to account for the associate extraction case, then it seems conceptually more economical to deal with the object extraction facts by the same means (which is of course possible), rather than bringing in an extended notion of θ-government for the benefit of this case alone.

The signs are, in other words, that the Derivational Approach seems more general in its coverage, a property which is in its favour. On the other hand, this is not exactly a knock-down argument for the Derivational Approach. Let us therefore look at some extraction patterns in other languages. As it turns out, these point towards the same conclusion.

4.4 Extraction asymmetries among objects: German, Dutch and French

The so-called Derivational and Representational Approaches to extraction asymmetries differ in their predictions with respect to subjects and objects (though not with respect to adjuncts). The Derivational Approach takes possibility of extraction to be linked to whether or not the DP must move overtly to its Case position prior to *wh*-extraction - and hence, linked ultimately to feature-strength properties of functional heads in a language. Thus, the Derivational Approach leads us to expect to find different patterns across (and within) languages with respect to extractability of subjects and objects. By comparison, the Representational Approach which I am also considering tends to predict that object extraction should be universally preferable to subject extraction, due to the θ-government-related difference between Spec-*v* and Spec-TP.

In order to put these predictions to a preliminary test, let us look at German and Dutch, both of which have (at least) two different positions for objects, an alternation often referred to as "scrambling"; (49) illustrates for German, (50) for Dutch (latter example from de Hoop 1992 cited in Reinhart 1995); the adverbs *gestern*, *gisteren* 'yesterday' are taken to mark the left edge of the VP:[39]

(49) a. ... weil Lutz gestern sein Handy verloren hat
... because Lutz yesterday his mobile lost has
b. ... weil Lutz sein Handy gestern verloren hat
... because Lutz his mobile yesterday lost has
'... because Lutz lost his mobile phone yesterday'

(50) a. ... dat de politie gisteren de taalkundigen opgepakt heeft
... that the police yesterday the linguists arrested have
b. ... dat de politie de taalkundigen gisteren opgepakt heeft
... that the police the linguists yesterday arrested have

'...that the police arrested the linguists yesterday'

It has recently been claimed with respect to both German and Dutch that object appearing outside of the VP, as in (49b), (50b), have undergone A-movement. For German, for example, Déprez 1989 (p.244) shows that a "scrambled" object can repair a Weak Cross Over violation, shifting the variable-site of the quantificational phrase *jeden Schuler* 'each student" to the left of the phrase containing the coindexed pronoun - (51a&b). Déprez also shows that the object movement in question is itself not subject to WCO - (51c) (and see also Mahajan 1990 for similar arguments concerning German):

(51) a. * ... weil ich seinem$_i$ Professor jeden Schuler$_i$ vorgestellt habe
... since I to his professor each student introduced have

[39] Both German and Dutch are so-called V2 languages, whose characteristic property is that the verb in declarative main clauses must be the second element. In embedded clauses, this constraint does not hold, which is why embedded clause examples are used here, following standard practice in the literature.

b. ... weil ich jeden Schuler$_i$ seinem$_i$ Professor vorgestellt habe
... since I each student to his professor introduced have

'... since I introduced each student to his professor'

c. ... weil Maria den Hans$_i$ seinem$_i$ Professor vorgestellt hat
... since M. the H. to his professor introduced has

'... since Maria introduced Hans to his professor'

In a similar vein, it is argued by Vanden Wyngaerd 1989, Zwart 1993 that the "scrambled" object in Dutch examples like (50b) has undergone overt Case-related A-movement. For example, Vanden Wyngaerd gives a variety of evidence that object scrambling creates new Binding configurations, illustrated in (52) (*op. cit.*, p.261), where a reciprocal *elkaar* 'each other' occurring inside a VP-internal phrase can be bound by a scrambled object, but not by an unscrambled one. Consider also (53), taken from Zwart 1993, showing that scrambling itself does not seem to create an operator-variable configuration: here, the object *Marie* finds itself to the left of the adjunct *volgens haar aanwijzingen* 'following her directions', yet may be interpreted as coreferential with the pronoun within that adjunct, the lack of WCO effect implying A-movement.

(52) a. * Ik heb aan elkaar$_i$ de jongens$_i$ voorgesteld
I have to each other the boys introduced

b. Ik heb de jongens$_i$ aan elkaar$_i$ voorgesteld
I have the boys to each other introduced

'I introduced the boys to each other'

(53) Jan heeft Marie$_i$ [volgens haar$_i$ aanwijzingen] gekust
J. has Mary following her directions kissed

'Jan kissed Mary according to her directions'

It may be concluded then that in German and Dutch, objects can undergo overt Object

Shift, presumably to Spec-*v*.[40] In the light of this information, let us have a look at the extraction behaviour of objects. Here I shall consider the case of extraction from, rather than of, the elements in question, since this obviously makes it easier to tell where the all-important attempted extraction site is.

Although German and Dutch do not have the equivalent of extraction from *picture*-DPs in English, both languages have what looks like some type of CED phenomenon involving 'what kind of-' DPs, in which either the entire DP is extracted - as in (54), (56) - or, under certain conditions, the operator moves alone, stranding the rest of the phrase. What are these conditions? Interestingly enough, extraction of just the operator - *was* in German, *wat* in Dutch - is possible only from objects which have not undergone overt Object Shift, as shown in (55) (adapted from Diesing 1992) and (57) (from W. Meyer Viol p.c.):

(54) [Was für Sonaten]$_i$ hat Lutz gespielt t_i ?
was for sonatas has L. played
'what kind of sonatas did Lutz play?'

(55) a. Was$_i$ hat Lutz immer [t_i für Sonaten] gespielt?
what has L. always for sonatas played
'what kind of sonatas did Lutz always play?'

b. * Was$_i$ hat Lutz [t_i für Sonaten] immer gespielt?
what has L. for sonatas always played

(56) [Wat voor boeken]$_i$ heeft Otto t_i geschreven?
what for books has Otto written
'what kind of books did Otto read?'

(57) a. Wat$_i$ heeft Otto altijd [t_i voor boeken] geschreven?
what has Otto always for books written
'what kind of books did Otto always read?'

---

[40] I am taking it that in (48a) and (49a), the objects are *in situ*, and thus that German and Dutch in some sense have optional overt Object Shift. Of course, this is not a straightforward assumption in Minimalism; see Part II for discussion of optionality, and in particular Chapter 6 for an analysis of the German and Dutch object "scrambling" alternations.

b. * Wat$_i$ heeft Otto [t$_i$ voor boeken] altijd geschreven?
what has Otto for books always written

Notice that the Derivational Approach to extraction asymmetries provides a straightforward explanation of these facts: the *wh*- objects in (55a), (57a) do not move as a whole to Spec-*v* (Case is presumably checked by pure feature-movement, analogously to the case of object *wh*-movement in English discussed earlier), so that extraction of the operator itself (*was*, *wat* obviously does not take place from Spec-*v* (cf. the legitimacy of extraction from object *picture*-DPs in English). By contrast, in (55b), (57b), the entire *was für/wat voor* DP undergoes overt Object Shift, placing it in Spec-*v* to the left of the adverbs *immer/altijd* 'always', from which position extraction of the operator alone is impossible (cf. the illegitimacy of extraction from subject *picture* DPs in English). On the other hand, the Representational Account, with its core assumption that extraction from Spec-AGR$_o$P/Spec-*v* is legitimate due to the position counting as θ-governed, would seem to predict that extraction from either of the two object positions is legitimate - a prediction that plainly does not correspond to the facts. Hence, if the leftmost of the two object positions under consideration is an A-position - as we have seen independent reason to believe that it is - then we have some indication that the Derivational Approach to extraction asymmetries is on the right track, rather than the Representational Approach. A clear link is seen in German and Dutch between extractability from objects and whether or not these undergo overt movement to the Spec of a functional projection for Case.[41]

---

[41] Notice that the extraction asymmetry seen among object *was-für* DPs seems to repeat itself with subjects which, according to Diesing (1992), may appear in Spec-VP or Spec-IP; the particle *denn* 'indeed' is taken to mark the left edge of VP. Extraction of *was* alone is possible only from subjects in VP:

(i) a. Was$_i$ haben denn [t$_i$ für Ameisen] einen Postbeamten gebissen?
what have indeed for ants a postman bitten
'what kind of ants have bitten a postman?'

b. * Was$_i$ haben [t$_i$ für Ameisen] denn einen Postbeamten gebissen?
what have for ants indeed a postman bitten

Something of a similar situation also obtains in Dutch, as shown in (ii) (from W. Meyer Viol p.c.); *er* is an expletive pronoun presumably in Spec-IP:

(ii) a. Wat$_i$ hebben er gisteren [t$_i$ voor mieren] een postbode gebeten?
what have there yesterday for ants a postman bitten

To conclude this section, let us briefly consider French, which has its own case of extraction out of DP involving *combien-de* 'how many' phrases (see Rizzi 1990): either the entire *combien de* DP fronts, as in (58a) and (59a), or, under certain conditions, the operator *combien* 'how many' may move on its own, leaving the remains of the DP behind. The extractability pattern we find here is again more compatible with the Derivational Approach than with the Representational Approach. Subjects in French obligatorily move overtly to Spec-TP, and extraction of *combien* from a subject is ungrammatical. Meanwhile, as noted in § 3.3.3 above, *wh*-objects seem to have an *option* but not an obligation to move overtly through their Case position; extraction of *combien* from an object is possible, as shown in (59b) - and note that in this situation, object agreement on the participle, taken to be the reflex of overt movement through the Case position (see Sportiche 1992, Branigan 1992, Chomsky 1991, 1995) is not permitted.[42]

(58) a.. Combien d'idiotes$_i$ t$_i$ ont repeint ces tables?
how-many of idiots have repainted these tables
'how many idiots repainted these tables?'
b. * Combien$_i$ ont [t$_i$ d'idiotes] repeint ces tables?

(59) a. Combien de tables$_i$ as-tu repeint(es) t$_i$?
how-many of tables(f.pl.) have-you repainted(f.pl.)
'how many tables did you repaint?'
b. Combien$_i$ as-tu repeint(*es) [t$_i$ de tables]?

---

b. * Wat$_i$ hebben [ti voor mieren] gisteren een postbode gebeten?
what have for ants yesterday a postman bitten

And in Icelandic, known to have a least two subject positions (Jonas & Bobaljik 1992, Chomsky 1995), Kitahara (1994) reports that subjects pattern with objects rather than adjuncts with respect to extraction from *picture*-DPs, in contrast to English:

(iii) a. hverjum$_i$ heldur þú að Jón hafi keypt [myndir af t$_i$]?
who think you that Jón has bought pictures of
b. ? hverjum$_i$ heldur þú að [myndir af t$_i$] séu til sölu?
who think you that pictures of are on sale

c. hvað$_i$ hefur Jón farið [eftir að hann keypti t$_i$]?
what has Jón left after that he bought

[42] See Chapter 5 below for a full account of why French has optionality between overt and covert Object Shift when the object is a *wh*-phrase, but obligatory covert Object Shift otherwise.

5 The Copy Hypothesis, economy and complexity

5.1 Preamble

To conclude this chapter and Part I, I investigate a different type of consequence of the Copy Hypothesis. Recall from the introductory chapter that derivational economy conditions - that is, those which pick "the optimal realization...of interface conditions" - are global, in the sense that determining the *optimal* derivation necessarily involves comparison between derivations. Conditions which have this property in the 1993-1994 framework are Last Resort, Procrastinate and MLC; and in the 1995 framework, Procrastinate and the Shortest Derivation Requirement (see Chapter 5 on the latter). In this respect, these conditions can be compared with the Principle of Full Interpretation, which is a non-global, i.e. absolute condition: a representation either obeys FI, or it does not. Greed in the 1993/94 framework is also an absolute condition, as are Last Resort/Greed and MLC in the 1995 framework.

As mentioned briefly in Chapter 1, conditions of the global type are associated with the property of computational complexity, in that it less simple to evaluate something with respect a global condition, such as Procrastinate, than it is with respect to an absolute condition like FI or Last Resort (1995). For example, to find out whether a representation conforms to FI, one needs only to examine properties of that particular representation itself. Likewise, an operation of Move will in and of itself either obey or violate Last Resort (1995) or the Minimal Link Condition (1995) - no further information needs to be taken into account. By contrast, in order to determine whether a movement is *optimal* by some condition, say Procrastinate, the circumstances of that movement must be taken into account; specifically, does the movement converge? Thus, Procrastinate, containing the notion *optimal* (cf. *most*, *many* etc.) is a "higher-order" condition which it is intuitively clear must operate in a more complex fashion than conditions like Full Interpretation and Last Resort, which are formulated in terms of first-order notions (*all*, *some* etc.) The complexity issue has become more prominent recently (see Chomsky 1995) and the high degree of complexity countenanced in the Minimalist model is seen by some authors as a serious shortcoming (see Johnson and Lappin (forthcoming)).

Its practical applications proposed and motivated in this chapter indicate that the Copy Hypothesis has implications for the computational complexity issue, specifically with respect to the condition Procrastinate. Recall that we started this chapter with the question of whether an element should be allowed, or forced, to move overtly through a site where there is no strong feature - normally a violation of Procrastinate - in special circumstances; namely, just in case the element will eventually need to check a strong feature in a more distant site. The question was, in other words, "can Procrastinate be overridden?" Although, as will become clear shortly, this is not the straightforward yes/no question it may seem - it should really be asked "to what extent can Procrastinate be overridden?" - I shall attempt to show that, given the Copy Hypothesis, it becomes possible to at least *reduce* in a contentful way the extent to which Procrastinate needs to be an overridable condition. Given the analysis which I have proposed in this chapter, itself dependent on the Copy Hypothesis of Chapter 2, it will be seen that there is no need to allow for "unforced violations" (in the sense of Chomsky 1995) of Procrastinate. I would mention here that the discussion in this section will be conducted on an intuitive level; whatever conclusion may be reached will be tentative.

5.2 Procrastinate, necessary and unnecessary degrees of violability

Let us firstly recap the status and role of Procrastinate in the Minimalist framework. As described in Chapter 1, accounting for obligatory word-ordering effects in the Minimalist framework is a two part procedure. Firstly, the features taken to be involved in the movement in question are designated a value: strong or weak: while features of any type cause the derivation to crash if they appear in the LF representation, strong features have the added property of PF uninterpretability. If one of these features is unchecked (i.e. still present) when the derivation is subject to Spell-Out, the derivation crashes. The second part of the account is the economy condition Procrastinate, repeated here:

(60) Procrastinate: "LF movement is 'cheaper' than overt movement...The system tries to reach PF 'as fast as possible', minimizing overt syntax." (Chomsky 1993:31)

Since a condition which actually distinguishes covert movement ("LF movement") from overt movement is, strictly speaking, inconsistent with "the Minimalist assumption that the computational procedure C_{HL} is uniform from N to LF" (Chomsky *op. cit.*, p.23 - see Chapter 1), it may be preferable to think of Procrastinate as a sub-case of Last Resort, as Wilder and Cavar (1993) suggest, rather than taking the words of (60) literally; this is possible, given the difference between Strong and weak features and the implications of this for convergence.[43]

At least one point about Procrastinate is clear: whether we think of it in terms of (60), or as an instance of Last Resort, as in Wilder and Cavar 1993, there is some sense in which the condition needs to be violable. This can be expressed either by saying that the condition holds among convergent derivations only, so that it fails to apply when there is a strong feature, since if it did, the derivation would not converge - cf. Chomsky 1994: "convergence requires that [Spec, d] be filled. Only one option exists: to raise there...we therefore select this option, *not violating Procrastinate, which does not arise*" (p.38). Alternatively, one may think of Procrastinate as "arising" but being violated when a strong feature is involved - cf. Chomsky 1995: "two *violations of Procrastinate* [is] the minimal number with two strong features" (sec.10, p.15); similarly, "a formal feature may or may not be strong, forcing overt movement that violates Procrastinate" (*op. cit.* 1995:232) (all emphases in the above quotes added - AMP). Whichever one of those perspectives one wants to see it from, the fact remains that Procrastinate is the type of condition which needs to know whether or not a step of movement will result in convergence before making a decision, rather than a condition which simply considers a movement on its own terms and rules it in or out. For convenience, I shall use the name *degree 1 violability* to refer to this necessary property of such conditions (e.g. Procrastinate; Last Resort and MLC in the 1993/1994 Minimalist framework, but not the 1995 framework). It is evident that even if all weak features are able to be checked by covert movement - which is allowed for, if copy movement is possible - Procrastinate cannot be free of degree 1 violability.

---

[43] Strong features are conceived of somewhat differently in the 1995 Minimalist framework, but the difference is not relevant to the present discussion.

However, depending on other theoretical assumptions which have been directly relevant in this chapter, it may or may not be necessary to attribute a further degree of violability to Procrastinate. Having said that it is "the type of condition which needs to know whether or not a step of movement will result in convergence", note that *will result in convergence* is somewhat ambiguous in that it could have the wider meaning of "ultimately", or the narrower meaning of "immediately". In this connection, let us consider the role of Procrastinate in a case like (61) (= (3) above):

(61) Who does John like?

On standard assumptions (i.e. with no derivations involving copy-movement), the *wh*-phrase in (61) moves overtly through the Case position Spec-*v*, although there are no strong features there. The reason that the *wh*-phrase undergoes this exceptional overt A-movement is that otherwise, the derivation will crash at PF, since the strong feature of C further up will not be able to be checked overtly. It is clear that in order to *force* overt movement to the Case position for reasons relating to the strong feature of C, one is obliged assume that Procrastinate can see further ahead than the consequences of the current step of the derivation when evaluating that step; that is, it has to know whether that step will *ultimately* result in convergence, and may overlook the immediate situation. Let us say that a principle which can work like this has the property of *degree-2 violability.*

There is no need to engage in any concrete calculations to see that conditions requiring degree-2 violablity in all probability entail greater computational complexity than those requiring only degree-1 violability; more possible situations need to be taken into account. In the light of this, notice that in a theory which allows for movement of copies, it is completely unnecessary to attribute degree-2 violability to Procrastinate (or for that matter to any economy condition), simply because there will *never* exist a situation in which a weak feature needs to be checked by overt movement as a result of the fact that the element also needs to check a strong feature at some higher position in the tree. To borrow Chomsky's (1995) terms, while in the standard Minimalist framework, "two

violations of Procrastinate [is] the minimal number with two strong features", in a version of the framework including the Copy Hypothesis, two violations of Procrastinate is the minimal, and the maximal, number with two strong features.

Summing up, it has been suggested that the Copy Hypothesis, which has the consequence that all non-strong features can be checked covertly by copy movement, allows Procrastinate to be seen as a less global constraint whose domain of application is limited to single steps of movement at a time. This, it was suggested, is a step towards reducing computational complexity in the model, presumably a positive result.[44]

6 Summary and conclusion

In this chapter I gave various arguments in favour of the Copy Hypothesis, focusing on *wh*-movement of objects and associates of expletive *there* in English. It was proposed that such *wh*-phrases check Case features (Accusative and Nominative respectively) covertly, just like their non-*wh* counterparts. This involves movement of the copy of the *wh*-phrase left by movement to Spec-CP. I gave empirical motivation for such an analysis pertaining to Floating Quantifier distribution, and also, more indirectly, to the treatment of certain extraction asymmetries. I suggested further that the availability of the proposed derivations might allow for the elimination of some (though not all) of the computational complexity entailed by the economy condition Procrastinate.

---

[44] In fact, it will become clear in Part II that the situation with economy and complexity is more complicated than implied in this section. While allowing for checking by copy movement does indeed mean that Procrastinate can apply in a more "local" way, as asserted in this section, the theory of syntactic optionality to be elaborated in Chapters 5 and 6 below will rely on the idea that in certain situations, a derivation which is *not* optimal by Procrastinate may nevertheless be permitted, just in case it is optimal by a different economy condition, Shortest Derivation. If such a seemingly "global" state of affairs is to be countenanced later on, then the suggestion made in the present section might appear to be debatable. However, I do not think that this is necessarily so; I delay discussion of the question until after the theory of optionality has been fully presented (see discussion in Chapter 6, § 2.1.3 on "modular economy").

In so far as these justifications are convincing, there is reason to believe that copy movement must be possible. Having arrived at this conclusion, we have a choice: either it can generally be maintained that traces are immobile (as proposed by Chomsky 1995), to which we then have to add the (further) stipulation that in certain situations - i.e. the cases of *wh*-movement discussed in this chapter - traces are *not* immobile. Or, it can be said that movement is allowed to operate on any element in principle, as expressed in the Copy Hypothesis proposed in this thesis. If, and only if, the latter choice is taken, then the availability of copy movement derivations and concomitant advantages come for free.

While the arguments of this chapter have mainly involved data from English (although other languages have certainly played a supporting role), we shall see in Part II (coming up shortly) that the Copy Hypothesis has further uses in dealing with data from other languages, not all of it involving *wh*-movement, and some of it involving *attraction by* rather than *movement of* copies, or their features.

Part II

Chapter 5

Syntactic optionality in theory and practice

1 Introduction

Remember that the rule Move-α of the GB model was in principle free to apply or not to apply, with necessary constraints provided independently. The inherent unconstrainedness of Move-α was naturally exploited to characterize phenomena which themselves seemed to be "unconstrained", namely various optionality and "free word order" data - so-called scrambling, to name but one example.

Unfortunately, as is well-known, such accounts are not possible in the Minimalist framework: since a tenet of the system is that movement is possible only if necessary, there would seem to be no place for movement which is possible and not necessary, i.e. "optional". So given the considerable quantities of data which apparently do exhibit the latter characteristic, the Minimalist framework potentially faces a large-scale empirical problem (see e.g. Fukui 1993, Marantz 1995 for general discussion). This is the point of departure of Part II, comprising this chapter and the next, in which I develop a Minimalist analysis of a number of phenomena in terms of optionality between overt/covert

movement; this mimics the forbidden optionality between moving/not moving, yet may in fact be permitted by derivational economy conditions, under particular, restricted circumstances.

I start out with two basic premises: firstly, that there exist at least some optionality phenomena which, in the absence of framework-specific considerations, can most plausibly be characterized in syntactic terms (as opposed to e.g. lexical or dialectal/parametric terms; all definitions are provided below). I also assume that a certain amount of syntactic optionality is at least logically possible within the derivational economy system (see e.g. Chomsky 1991). Given these premises (which I shall in fact elaborate on and in some cases argue for later), the objective is then to devise a system which permits syntactic optionality in an appropriately restricted way. The cases of optionality I investigate in this chapter and the next range from the optionality of phonologically overt agreement with object *wh*-phrases on past participles in French (classically analysed by Kayne 1989 in terms of an optionality in whether to move to Spec-CP via an agreement projection) through optional associate movement phenomena in expletive constructions in English and Swedish, to optional overt Object Shift in Icelandic, German and Dutch. Throughout Part II, I shall be concerned with related non-optionality phenomena, and how to integrate these into the account of optionality which I propose.

This chapter is organized as follows. Section 2 explains why, and to what extent, syntactic optionality is a problematic concept within the Minimalist framework. In Section 3, I examine the case of optional object agreement on past participles in certain syntactic environments in French, namely *wh*-movement and Accusative clitic constructions. I begin by giving arguments that this data needs to be treated as a case of syntactic optionality (in the spirit of Kayne's (1989) GB account), rather than as a case of lexical or dialectal optionality. More specifically, once again exploiting the Copy Hypothesis (see Part I, chapters 2 and 4), I argue that the optionality of agreement is the reflex of an optionality between overt and covert "Object Shift" (cf. Branigan 1992). Having thus characterized the agreeing and nonagreeing options, I go on to argue that their co-existence is due to the fact that neither qualifies as the unique "optimal realization of interface conditions" -

in fact, each of them is optimal: one with respect to Procrastinate, the other with respect to the Shortest Derivation condition (Chomsky 1995). For reasons which will become clear, this optionality is restricted to arise only in syntactic environments involving overt *wh*-movement and Accusative clitics. Among other characteristics of the analysis, we shall see that it assigns a non-trivial role to the concept of instantaneous Form Chain (Chomsky 1993, Collins 1994). In Section 4, I show how further cases in French in which participle agreement is *not* optional, but in fact obligatory or impossible, can be handled in a way which is properly integrated with the account of the optionality cases. In this section is included a unified account of the impossibility of agreement with both *wh*- and non-*wh*-associates of the expletive *il*, which I contrast with the proposal of Kayne 1989 in which this expletive data was first discussed. Finally, in Section 5, the theory is extended still further to a phenomenon which occurs in English and Swedish passive constructions with the expletives *there* and *det* respectively: optional partial A-movement of the associate, apparently to Spec-*v*. In exhibiting this optionality, English and Swedish contrast with French, in which the associate is obliged to remain *in situ* in such constructions. The contrast is attributed to the feature-wise different natures of English *there*/Swedish *det* and French *il* - for which there is independent evidence - in conjunction with principles of derivational economy.

2 Optionality is a Minimalist issue

Let us start with a general discussion of the status of syntactic optionality within the Minimalist system. To do this, I think it is useful to provide some definition of the concept at issue. This exercise may appear trivial, but we shall see that it has the advantage of highlighting instantly at least one very important question which surrounds the topic - namely the question of the role of so-called interpretive effects (see shortly below for prelimary discussion). It also helps to make explicit the domains in which optionality is *not* problematic, which in turn gives a clearer picture of the various alternative ways which exist for dealing with the phenomena in question.

2.1 An informal discussion of syntactic optionality

What exactly is this phenomenon which appears to be exist so widely in natural languages and causes such a problem in the Minimalist framework? Informally speaking, what we are interested in is cases of two (or more) sentences which have both the following properties: (i) they are made up of the same set of elements (i.e. have the same numeration) and (ii) they do not have necessarily different meanings.

This informal definition needs some elaboration. Characteristic (i) is straightforward enough: alternation in numeration is not in itself an interesting optionality, since determining an initial set of lexical items (by whatever means this may actually be done) is by definition a matter of choice. The meaning aspect (ii) is less straightforward. As is well-known, there is a lot of data, traditionally thought of in terms of syntactic optionality, where there exists a particular type of interpretive difference between the options, usually characterized in terms of discourse-related notions like *specificity*, *familiarity*, or *presuppositionality*; examples include object scrambling in Dutch and German (Diesing 1992) and participle agreement/lack of it with object *wh*-movement in French (Obenauer 1992). What typifies this phenomenon is that one option tends to have an obligatory specific/familiar interpretation, while the other *may* have this interpretation, but need not. It is in order to subsume such cases that (ii) says that the options have "not necessarily different meanings", rather than "the same meanings" - although in some instances, "the same meanings" is indeed what the options have.

But given the interpretive phenomena just described, one obviously has to ask whether it is appropriate to think of the relevant data as involving optionality at all. Instead, one might want to say that the relevant interpretive property is directly encoded in the form of features or formatives.[1] However, in addition to noting the fact that there are many cases of optionality in which there is no interpretive difference between the

[1] See e.g. Uriagereka 1995, Delfitto and Corver 1995.

options,[2] I take it that there is nothing necessary about this type of direct encoding approach,[3] and shall myself analyse some cases of optionality which *do* display interpretive effects as nevertheless having essentially the same LF structures, in a sense to be made precise in due course. An approach to the interpretive effects in terms of post-LF processes is offered in Chapter 6, after I have proposed a syntactic account of the relevant cases.

Having attempted to clarify what I mean by optionality, let us now examine a few pairs of sentences and see which of them might qualify. Consider (1):

(1) a. John hit Mary
b. John saw Mary

(1) complies with neither (i) nor (ii) - (1a) and (1b) have transparently different numerations and transparently different meanings - and therefore does not count as a case of optionality, a result which hopefully corresponds to our intuitions. Next, consider (2), (3) and (4):

(2) a. It seems that John has left
b. John seems to have left

(3) a. Bill was hit by Mary
b. Mary hit Bill

(4) a. Quand Jean est parti?
b. Quand Jean est-il parti?

The above three cases also fail to achieve optionality status: while they fulfil (ii), with the

---

[2] E.g. optional partial overt A-movement to Spec-v in expletive constructions in English and Swedish (see this chapter, § 5); scrambling in Japanese, described by Saito (1989) as "semantically vacuous".

[3] The proposal of Reinhart 1995 with respect to object scrambling in Dutch and German is one example of a method of dealing with interpretive effects without directly encoding the relevant property in syntax/LF.

(a) and (b) cases having the same meanings, these are clearly based on different lexical choices; (2a), for example, contains the expletive *it* which is missing in (2b).

In some sense the opposite to (2)-(4) is (5), which satisfies (i) but not (ii): the same set of lexical items is arranged in (5a) and (5b) in such a way that necessarily different meanings result:

(5) a. John saw Bill
b. Bill saw John

Let us now look at some pairs of sentences which do qualify as cases of optionality; a wide variety of these can be found in natural languages, even those with relatively fixed word order, such as English and French. What follows is far from a comprehensive list, including only the cases which I will provide analyses of in this chapter and the next.[4]

<u>Optional phonologically overt agreement in French</u>[5] (from Kayne 1989)

(6) a. Je me demande combien de tables Paul a repeint
I wonder how-many of tables(f.pl.) P. has repainted
b. Je me demande combien de tables Paul a repeintes
I wonder how-many of tables(f.pl.) P. has repainted(f.pl.)

'I wonder how many tables Paul repainted'

---

[4] A prominent omission here is the case of "free word order languages" such as Hungarian, Greek, Mohawk, Hindi and others. These also display discourse-type interpretive effects usually characterized in terms of notions like Focus and Topic, which has led some authors to posit functional heads encoding such properties (see e.g. Brody 1990 for focus in Hungarian).

[5] Classification of this group of examples as cases of syntactic optionality is based on the premise (Kayne 1989 among others) that phonologically overt agreement is triggered by movement. See § 3 of this chapter for discussion.

(7) a. (ces tables,) Paul les a repeint
(these tables(f.pl.),) P. them(f.pl.)-has repainted
b. (ces tables,) Paul les a repeintes
(these tables(f.pl.),) P. them(f.pl.) has repainted(f.pl.)

'(these tables,) Paul repainted them'

Optional overt Object Shift: Icelandic (from Collins & Thráinsson 1994)

(8) a. Jón las bækurnar ekki
J. read the-books not
b. Jón las ekki bækurnar
J. read not the-books

'Jón did not read the books'

Short object scrambling in Dutch (from de Hoop 1992 cited in Reinhart 1995)

(9) a. ...dat de politie gisteren de taalkundingen opgepakt heeft
....that the polcie yesterday the linguists arrested have
b. ...dat de politie de taalkundingen gisteren opgepakt heeft
...that the police the linguists yesterday arrested have

'...that the police arrested the linguists yesterday'

Short object scrambling in German

(10) a. ... weil Lutz gestern sein Handy verloren hat
... since Lutz yesterday his mobile lost has
b. ... weil Lutz sein Handy gestern verloren hat
... since Lutz his mobile yesterday lost has

'since Lutz lost his mobile yesterday'

"Heavy NP shift" in English

(11) a. Please check the details on this page carefully
b. Please check carefully the details on this page

Optional overt "partial A-movement" of associate in English and Swedish

(12) a. There were believed to be four men in the garden
b. There were four men believed to be in the garden

(13) a. There have only ever been given thirteen of these prizes
b. There have only ever been thirteen of these prizes given

(14) a. Det blev skrivet tre böcker
it was written three books
b. Det blev tre böcker skrivna[6]
it was three books written

'three books were written'

All the examples (6)-(14) seem to have the property of syntactic optionality as defined above: it is not obvious that different numerations are involved in the (a) and (b) versions; moreover, there is no necessary interpretive difference between them; no (a) version has an interpretation which is unavailable in its (b) version, nor vice versa - compare with (1) and (5). Certain of the cases display the type of discourse/interpretive effects mentioned earlier: for example, there is a "specificity" effect associated with overt Object Shift (or "scrambling") in German, Dutch and Icelandic, and also in the case of agreement-triggering *wh*-movement in French. In other cases, there is no interpretive difference between the options: in the case of optional agreement with clitics in the French example (7), *both* (a) and (b) versions must have a specific/familiar interpretation for the relevant phrase; while in the case of optional overt partial A-movement in English and Swedish - (12)-(14) - a specific/familiar interpretation of the relevant phrase is prohibited in both options (i.e. the so-called "Definiteness effect" which characterizes expletive constructions).

(6)-(14) thus resemble (2)-(4) above, in that there is no significant difference in meaning (contrast with (1)) - but unlike (2)-(4), it is not a question of presence/absence of expletive or "empty" elements in the numeration, but rather, *movements* which seem

[6] The participle in (14b) shows overt (plural) agreement with *tre böcker*, while the participle in (14a) does not. See the French example (6) and the associated footnote.

to have such properties. Why is such a concept problematic for the Minimalist framework?

2.2 Scope for syntactic optionality in the Minimalist system

Firstly, it cannot be disputed that given certain core Minimalist assumptions, there can be no simple optionality between moving or not moving. These (1989:26) words of Pesetsky sum up the situation: "[e]conomy prohibits spontaneous movement, unmotivated by any filter...". In more specific and up-to-date terms, recall that the Last Resort condition dictates that "a step in a derivation is legitimate only if it is necessary for convergence - had that step not been taken, the derivation would not have converged" (Chomsky 1993:32 - see Chapter 1). This consideration is sufficient to eliminate what in the GB framework was "the simple assumption that Move-α may fail" (Burzio 1986), and along with it, any analysis which countenances such an assumption, a good example being Kayne's (1989) account of the optionality of past participle agreement in French (which is reanalyzed in §3 below).

It can be agreed then that true choices in *whether* to move are nonexistent. In this sense, there is indeed no optional movement. However, consider a different aspect of movement, namely *when* it occurs in relation to Spell-Out. The overt/covert difference resembles or mimics the difference between movement/no movement, but in the former dimension, optionality is not excluded as profoundly as it is in the latter, which means that we might find phenomena which *appear* to involve an option in whether to move, but in reality involve an option in movement timing. Although in the case of strong features, movement is obligatorily overt (see Introductory chapter), consider the case of features which are not strong (i.e. "weak" features in the 1993 Minimalist framework; hereafter I will refer to these as *ordinary features*): these may convergently be checked by either overt or covert movement - in other words, the potential for optionality does exist here. In practice, in the standard Minimalist framework, the economy condition Procrastinate routinely eliminates this potential, since it requires that movement is overt only if convergence depends on this, that is, if the feature to be checked is strong. But in

principle, the scope for optionality is certainly present; to develop a system which exploits this is the objective of this chapter and the next.

Although it can be seen that the possibility for some sort of optionality in (or relating to) the computational system is not actually nonexistent under Minimalist assumptions, it is still obviously a lot more limited than in the GB model. As a result of these limitations, a number of alternative ways of explaining various optionality data have emerged in the literature, probably the most prominent of which is the *lexical approach*, as I shall call it. According to this, the "options" have different numerations, distinguished by presence/absence of an element, or by the strength property of a feature associated with some functional head (see e.g. Branigan 1992[7]). Other possible approaches are in terms of dialectal (i.e. parametric) variation and adverb placement (i.e. Merge) optionality. I shall look at specific instances of some of these approaches later on. In general, it seems likely that no single approach will cover all optionality phenomena. In what follows, I develop a theory of syntactic optionality which covers a series of cases involving objects. I look firstly at a case from French.

3 Optional agreement in French as optional overt Object Shift[8]

In French, phonologically overt agreement (hereafter *agreement)* on participles is optional[9] in a small range of syntactic environments: with overtly moved object *wh*-phrases, as in (15), and with Accusative clitics, as in (16) (= respectively (6) and (7) above):

---

[7] Branigan's proposal is reviewed in Chapter 6.

[8] This section is based on work presented at the LAGB Autumn Meeting, September 1995, and Going Romance 9, December 1995. Examples throughout this section are taken or adapted from Kayne 1989, unless otherwise stated.

[9] See note 5 above.

(15) a. Je me demande combien de tables Paul a repeint
I wonder how-many of tables(f.pl.) P. has repainted
b. Je me demande combien de tables Paul a repeintes
I wonder how-many of tables(f.pl.) P. has repainted(f.pl.)

'I wonder how many tables Paul repainted'

(16) a. (ces tables,) Paul les a repeint
(these tables(f.pl.),) P. them(f.pl.)-has repainted
b. (ces tables,) Paul les a repeintes
(these tables(f.pl.),) P. them(f.pl.) has repainted(f.pl.)

'(these tables,) Paul repainted them'

Optionality of agreement with Accusative clitics appears to be less well-known than with object *wh*-phrases; however, it is reported by Kayne 1975, Sportiche 1992 and Branigan 1992, and confirmed by informants.

3.1 The classical analysis: Kayne 1989

Probably the best-known Principles and Parameters account of participle agreement in French is that of Kayne 1989. Kayne attributes the optionality of agreement in cases like (15) and (16) to the fact that the potentially agreed-with phrase can take either of two different routes to its final landing-site, - for example, to Spec-CP in the *wh*-movement example (15). In particular, Kayne claimed that in the options where agreement shows up on the participle, the agreed-with phrase makes an intermediate landing in a site local to a functional projection containing the participle, Part(iciple)P, immediately above VP (cf. Chomsky's (1991) AGR_oP), thereby "triggering" agreement. In the options *without* agreement, this intermediate step is taken to have been missed out, i.e. the phrase moves directly to its final landing-site. In (15) for example, "two *wh*-movement sentences, one with and one without past participle agreement, will, even if otherwise identical, have category-wise different representations" (Kayne 1989:90). Kayne's account is formulated within the GB framework, predating the unified Spec-Head theory of Case. Accusative Case was thus taken to be assigned by V itself to a governee, and for this reason (alone),

the agreement-triggering position had to be assumed to be an A'-position.[10]

As established previously, the kind of optionality which Kayne envisaged for these cases is not possible in Minimalism. However, I think that his insight that syntactic optionality is involved is worth retaining, for reasons which I shall outline before developing a Minimalist analysis.

3.2 Arguments for a retaining a syntactic approach

By way of a rationale for the proposal to follow, I will begin by setting out my motives for treating (15) and (16) as cases of syntactic optionality. This is necessary since, as we shall see, quite apart from the theoretical problem already noted, the data itself has certain characteristics which might at first sight compel one to seek an explanation outside of "syntax proper". For one thing, there is the fact that "...past participle agreement with *avoir* in French is permeated with a great deal of artificiality" (Kayne 1975). In addition, there is more recently-discovered evidence (Obenauer 1992) that (some of) the agreement optionality data exhibits a "specificity" effect similar to that found with Scrambling of various sorts (e.g. in German and Dutch, see Diesing 1990). An examination of these characteristics will show, I think, that (15) and (16) are nevertheless most straightforwardly thought of as involving syntactic optionality (as in Kayne 1989).

3.2.1 Syntactic, lexical and dialectal approaches

Kayne's (1989) analysis of (15) and (16) is representative of what we can call a *syntactic optionality* approach. Its basic claim is that the agreeing and nonagreeing options do not

[10] The assumption that the agreement-triggering position was an A'-position had the apparent advantage of making available a straightforward account of the impossibility of participle agreement with *wh*-phrase associates of expletive *il*, in terms of improper movement. See § 4 for discussion and critique of Kayne's account of the agreement facts in *il* constructions.

differ from one another in any significant or "underlying" way. In contemporary (Minimalist) terms, this would seem to amount to claiming that the agreeing and nonagreeing versions in (15) and in (16) are essentially[11] equivalent at the points of input to, and output from, the computational system, with the observed difference between them due to derivational causes of some kind or another. As discussed in the last section, such a view of this kind is not easily taken in Minimalism. Here I shall look at some possible alternative approaches which could conceivably be applied to the case of (15) and (16); we shall see ultimately that they are inadequate. The first possibility to consider is what I call the *lexical optionality* approach. The gist of this would be to attribute the apparently syntactic optionality to some difference in underlying set of lexical items (numeration). On this approach, the difference between, say, the agreeing and nonagreeing versions of (15) and (16) would, in effect, be assimilated to the difference between, for example, *Paul a repeint les tables* 'Paul repainted the tables' and *Paul a repeint les chaises* 'Paul repainted the chairs' (though intuitively there is obviously a difference of some sort between the two cases - see below).

A version of the lexical approach has indeed been suggested for the French agreement cases by Branigan 1992. He proposes that "Agr [the functional head] can have either strong or weak features from the outset, freely" (1992:37). Thus Branigan maintains that the agreeing and nonagreeing versions of (equivalent examples to) (15) and (16) have different numerations, differentiated by the presence or absence of a strong feature.[12] On this account, there is of course no reason why agreeing and nonagreeing options should not coexist; like *Paul a repeint les tables* and *Paul a repeint les chaises*, they will not be in competition with respect to economy.

A second alternative which (for reasons to be discussed shortly) could be worth considering in the case of (15) and (16) is a *dialectal optionality* approach. According to

---

[11] The exact meaning of *essentially* will become apparent in § 3.3.1 below.

[12] An optional Strong feature account is also suggested by Collins and Thráinsson 1994 for the optionality of overt Object Shift in Icelandic. See Chapter 6, § 3.2 for a critique of this type of approach.

this, the optionality of agreement in (15) and (16) would be taken to reveal the availability of two different languages/dialects to the speakers in question: in one dialect, the participle agrees in the relevant situations, in the other dialect, it does not. Thus, while the hypothetical dialects and the difference(s) between them would need to be characterized, neither the fact that both exist, nor indeed the fact that individual speakers may have both to hand, would be problematic in and of itself (presumably explaining the factors responsible for which dialect is actuallly used when would be the responsibility of sociolinguists rather than syntacticians).[13]

Both the concepts of lexical and dialectal optionality - i.e. selecting a set of lexical items, and setting parameters respectively - are in themselves unproblematic, and trivially so - in no framework, Minimalism or otherwise, could such concepts fail to exist - so adopting an approach along one of these lines might initially appear to be a more attractive prospect than attempting to modify the framework to allow optionality "syntax internally". At the same time, one would of course hope to find some independent motivation for whichever approach one adopted. In the next subsection, I will show that in this last respect, a syntactic approach to (15) and (16) seems to have advantages over both the lexical and the dialectal approaches.

3.2.2 The syntactic roots of phonologically overt agreement

It is well-known that whether or not agreement on verbal elements is overtly realized relates strongly to syntactic factors of some sort (no matter what particular syntax of agreement is assumed). Consider for example a further fragment of the French participle agreement paradigm, with (15) and (16) repeated for comparison:

---

[13] Den Dikken 1992 speculates on an approach of this type to an apparent optionality in the ordering of elements in English complex particle constructions (e.g. ... *put the books down on the shelf* vs. ... *put down the books on the shelf*), suggesting that "speakers of English are bilingual ... disposing of two grammars, one in which V-Prt [=particle] reanalysis is obligatory and one in which it does not apply" (p.55, note 26).

(15) a. Je me demande combien de tables Paul a repeint (agreement
I wonder how-many of tables(f.pl.) P. has repainted optional)
b. Je me demande combien de tables Paul a repeintes
I wonder how-many of tables(f.pl.) P. has repainted(f.pl.)

'I wonder how many tables Paul repainted'

(16) a. (ces tables,) Paul les a repeint (optional)
(these tables(f.pl.),) P. them(f.pl.)-has repainted
b. (ces tables,) Paul les a repeintes
(these tables(f.pl.),) P. them(f.pl.) has repainted(f.pl.)

'(these tables,) Paul repainted them'

(17) Paul a repeint(*es) les tables (impossible)
Paul has repainted(*f.pl.) the tables(f. pl.)
'Paul repainted the tables'

(18) Il sera repeint(*es) beaucoup de tables cette année (impossible)
It will-be repainted(*f.pl.) many of tables(f. pl.) this year
'there'll be a lot of tables repainted this year'

(19) Les tables ont été repeint*(es) (obligatory)
The tables(f.pl.) have been repainted*(f.pl.)
'the tables were repainted'

(20) La femme de Paul est mort*(e) (obligatory)
The wife(f.) of Paul is died*(f.)
'Paul's wife died'

In the simplest terms, the appearance of agreement on the participle seems to relate to whether or not the agreed-with phrase precedes or follows the participle, i.e. something syntactic (in the wider sense) in nature. This pattern is not confined to French - consider for example the contrast from Swedish in (21) (= (14) above; from Holmberg 1994:218) - nor indeed to constructions involving participles and objects - a prominent example is the contrast between "agreement with subject versus object in nominative-accusative languages with the EPP" discussed in Chomsky 1995 (p.277).[14]

[14] Other cases which illustrate this pattern are Standard Arabic (Aoun, Benmamoun & Sportiche 1994), where the verb shows richer subject-agreement with preverbal subjects (person, gender, number) than with postverbal ones (person, gender). Also Franco-Provençal Valdôtain (Roberts 1994) where pre- and post-verbal clitics respectively do and do not show agreement with the

(21)[15] a. Det blev skrivet/*skrivna tre böcker (agreement impossible)
it was written three books(pl.)

b. Det blev tre böcker skrivna/*skrivet (obligatory)
it was three books(pl.) written(pl.)

'three books were written'

c. Tre böcker var skrivna/*skrivet (obligatory)
three books were written(pl.)/written
'three books were written'

No matter what it is that determines whether or not agreement features come to be overtly realized, the default assumption is of course that (15) and (16) conform to the pattern, rather than being exceptional cases where the usual rules cease to apply. And since such a wide range of data suggests that the overt realization of agreement features depends on the *syntactic* behaviour of the elements concerned, we should probably be inclined to assume that the constructions in (15) and (16) somehow possess the syntactic characteristics of both the cases where agreement is obligatory (cf. (19) and (20), *and* those where agreement is impossible (cf. (17) and (18),[16] i.e. (15) and (16) constitute microcosms of the overall agreement pattern. This is consistent with either a syntactic or a dialectal approach to the agreement optionality. By comparison, it would seem difficult to motivate an analysis of (15) and (16) in terms of some lexical difference between the options, given the clear implausibility of treating the agreement pattern as a whole in such terms. So, if we pay attention to the nature of agreement facts generally, it seems more promising to approach the optionality of (15) and (16) in syntactic or dialectal, rather than lexical, terms.

At the same time, there are other properties of the data which might at first appear

---

participle. Similarly Celtic languages (Ian Roberts, p.c.).

[15] An analysis of the optionality of (21a,b) is provided in § 5 below.

[16] As Kayne's (1989) paper made clear, this presupposes that the syntactic conditions underlying or "triggering" agreement must involve conditions more complex than precedence of the participle by the agreed-with element. See § 3.3.1 below for details.

to favour the opposite conclusion. Obenauer 1992 discovered that when the participle agrees with an object *wh*-phrase in French, the latter is obliged to receive what he called a specific interpretation.This specificity effect does not show up so clearly in (15) (at least in the absence of any context) since *combien-* ('how many') phrases are in themselves inherently ambiguous between specific and cardinality readings. However, Obenauer comprehensively demonstrates the specificity effect, both by using *wh*-phrases like *lequel* 'which', which are specific "de façon inhérente" 'inherently' (p.177), and by embedding ambiguous *wh*-phrases (with *combien*, as well as *quel* 'which') in disambiguating contexts. To give just one example, Obenauer shows that in *wh*-exclamative constructions, whose nature strongly disfavours a specific reading of the *wh*-phrase, agreement on the participle is unacceptable:[17]

(22) Quelle surprise elle m'a fait(*e)!
what surprise(f.) she me-has made(*f.)
'what a surprise she gave me!'

This specificity effect associated with agreement could easily be interpreted as compelling evidence for a lexical optionality approach to (15) and (16). Thus it could be claimed that the numeration of an option with agreement includes some extra item or feature encoding specificity (or whatever the interpretive property in question may turn out to be).[18]

However, although noone could reasonably disagree that Obenauer's specificity effect needs to be explained somehow, it is unclear to me how this could be done

---

[17] In my own tests on informants I found confimation of Obenauer's results. Several of his examples were interspersed with other agreement data, and in cases in which a specific reading of the *wh*-phrase is excluded (e.g. the exclamative example (21) in the text), even the informants generally most resistant to omitting agreement found it to be unacceptable here.

[18] See Delfitto and Corver's (1995) discussion of what they call the Syntax of Specificity. These authors in fact argue that the relevant property is Familiarity (in the sense used recently by Heim 1980). For a theory linking agreement (in the sense of the element AGR) with the property of specificity, see Mahajan 1991.

straightforwardly by means of a lexical approach.The problem, as I see it, is that it is *only* in the *wh*-movement cases that the forced specific interpretation of the relevant phrase, and the presence of agreement (or whichever syntactic conditions lead to its appearance - see below, §3.3.1) in fact coincide: as discussed from a different perspective a few paragraphs above, there is generally no direct link between the appearance of agreement on the participle and any interpretive property. In examples like (17) above (*Paul a repeint les tables*), where agreement is impossible, either the specific or the cardinality interpretation of the object is available. The same goes for the version of (15) without agreement ("...sans l'accord, la [*wh*-] phrase...est ambiguë." Obenauer 1992:177). And in cases like (19) and (20) above (*les tables ont été repeintes* etc), where agreement is obligatory, the same ambiguity obtains. An account which seeks to link the optionality of agreement with specificity will have to be complex enough to accommodate these facts. What is more, it will have to deal with the fact that in Accusative clitic cases like (16) (*Paul les a repeint(es)*), where agreement is optional as it is with *wh*-movement, there is no specificity-related difference between the options (Accusative clitics like *les* in (16) seem to be more or less invariably linked with a specific interpretation (see Sportiche 1992 and Uriagereka 1994).

So, given firstly the fact that the appearance of overt participle agreement in French seems generally to be ***dis***connected from the property of specificity, and secondly the fact that optionality of agreement in the clitic case ((16)) is accompanied by no interpretive alternation at all, it would surely be a difficult task to fully attribute the optionality of agreement to interpretive factors, be it in terms of a specificity-related element in the numeration (i.e. lexical optionality approach), or indeed by some other means.[19]

It is necessary to stress here that arguing against the type of hypothetical lexical optionality approach just described does not amount to denial of what is evidently a systematic link between agreement and specificity in the *wh*-movement cases, for it is

[19] One theory which implicates the specificity effect in the optionality of agreement in French is developed by Adger 1994. This proposal is reviewed in Chapter 6, § 3.3.

perfectly possible that the kind of interpretive effects in question might be associated with certain syntactic situations without this association taking the form of direct encoding in the system of morphosyntactic features.[20] Indeed, Obenauer's own account of the specificity effect, which bases itself on Kayne's (1989) syntactic analysis of the *wh*-movement cases, does not, as far as I can tell, involve any claim that the interpretive effect is *responsible* for the optionality. Having said all this, I shall not make any concrete proposals about how to deal with the specificity effect in the present chapter; the issue is addressed in Chapter 6, after I have dealt with the syntax of the French (and certain other) examples in which it arises.

In defence of a lexical optionality approach to (15) and (16), it might perhaps be argued that *not* taking the agreeing and nonagreeing options to have different numerations must entail a violation of Chomsky's (1994, 1995) *Inclusiveness* condition, whereby "any structure formed by the computation...is constituted of elements already present in the lexical items selected for N; no new objects are added in the course of computation" (1995:228). This is because in the French cases - in contrast to other typical optionality phenomena such as scrambling, Heavy NP Shift - the difference between the options seems to involve the presence or absence of some material, namely the overt ϕ-features on the participle, as opposed to a rearrangement of existing components. As such, the same numeration giving rise to both (15a) and (15b), the latter seemingly including an extra item, might appear to violate *Inclusiveness*. To avoid this, it would have to be said that the options have different numerations. However, this does not constitute a good argument for the lexical optionality approach, as long as the following two reasonable assumptions are made: (i) Chomsky's (1994, 1995) assumption that Inclusiveness applies only to the N --> LF computation, whereas "standard theories take it to be radically false

[20] For example, as mentioned briefly in note 3 above, Reinhart 1995 develops an account of German and Dutch object scrambling in which scrambled and non-scrambled options are both generated by the syntax (she in fact assumes that the word-order difference lies in adverb placement rather than movement). The interpretive difference between the options, also frequently characterized in terms of specificity (Diesing 1990), but which Reinhart sees in terms of Focus, is accounted for in an indirect way making use of Cinque's (1993) theory of stress and focus.

for the computation to PF" (228). (ii) the common view both in general and with respect to the French data (see e.g. Kayne 1989, Sportiche 1992) which is that phonologically overt agreement constitutes a "spelling out" of ϕ-features, triggered by certain syntactic conditions (see § 3.3.1 below). Given these two assumptions, it is clear that claiming that e.g. the two options of (15) and (16) stem from the same numeration is not incompatible with the Inclusiveness condition.

Having thus put aside a lexical optionality approach to (15) and (16), we are left with the syntactic and dialectal approaches. In deciding between these two, there is one further property of the data which is potentially relevant. As mentioned briefly above, speakers' judgements on agreement in cases such as (15) and (16) are affected by factors which are clearly non-linguistic, having to do with rules of prescriptive (as opposed to Universal) grammar. In this respect, (15) and (16) contrast sharply with examples like (17)-(20), where there is no such ambivalence about the presence/absence of agreement.

It is clear that these extra-linguistic factors do interfere with judgements on agreement in (15) and (16), and as Obenauer (1992:170) remarks, it may therefore be questionable whether an analysis of this data is a legitimate basis on which to form conclusions about UG - or, to be a bit more precise, conclusions about a particular instantiation of UG; is it really valid to conclude that the agreeing and nonagreeing options of (15) and (16) are the product of a single grammar, as entailed by a syntactic (or indeed lexical) approach to the data? On the other hand, the sociolinguistic effects could perhaps be construed as favouring a *dialectal* optionality approach to (15) and (16), whereby speakers for some reason have access to two dialects: in one, the participle must agree in the environments in question, while in the other, it must not (cf. note 14).

However, there is some reason to believe that the sociolinguistic factor, and the dialectal approach which it might inspire, are not the whole story about agreement optionality in French. As Obenauer (*op. cit.*) points out, the data which he discusses concerning specificity effects and agreement (see above) seems itself to constitute evidence that there is some other factor at the root of the agreement alternations, which

cuts across the sociolinguistic factor and could not be exclusively accounted for in such terms. In this connection, it should be noted that if agreement optionality in French were to be attributed to a dialectal alternation conditioned by nonlinguistic factors, then this case would constitute a strange exception to the following descriptive generalization discussed by Adger (1994:94): "there appears...to be a correlation between optionality in the...derivation, and obligatoriness of familiarity".[21] Of the various illustrations of this generalization which Adger mentions (for example, object scrambling in Dutch) to the best of my knowledge it is only the French agreement case in which there is any question that sociolinguistic factors might have a hand in the optionality. Presumably in all these other cases, the optionality in question must be taken to obtain within a single grammar. If, but only if, it is assumed that agreement optionality in French also reflects options within a single grammar (whether along the lines of a syntactic or a lexical approach) - as opposed to a nonlinguistically conditioned alternation between two grammars - the striking resemblance between French and the other cases discussed by Adger is straightforwardly captured.

As in the case of the interpretive effects discussed above, it should be stressed here that claiming that an approach in terms of dialectal variation is not wholly adequate to deal with the optionality in (15) and (16) should not be confused with denying the existence of the sociolinguistic effects, or the fact that one should be aware of their possible interference.

It has been argued in this section that, leaving aside framework-specific concerns about syntactic optionality, various properties of the data itself indicate that an account in these terms is needed. Such an account is developed in what follows.

---

[21] Adger's *familiarity* is extensionally equivalent to Obenauer's *specificity*. The exact nature of the interpretive effect at issue is not relevant to the current discussion. With respect to Adger's generalization, note that the correlation between interpretive effects and syntactic optionality is only one-way: if there exists an interpretive effect (of the type under discussion), then there is an optionality, but not vice versa, as was illustrated above by the Accusative clitic case, where agreement is optional as with *wh*-movement, yet no specificity effect obtains.

3.3 Characterizing the options

3.3.1 Preliminary assumptions

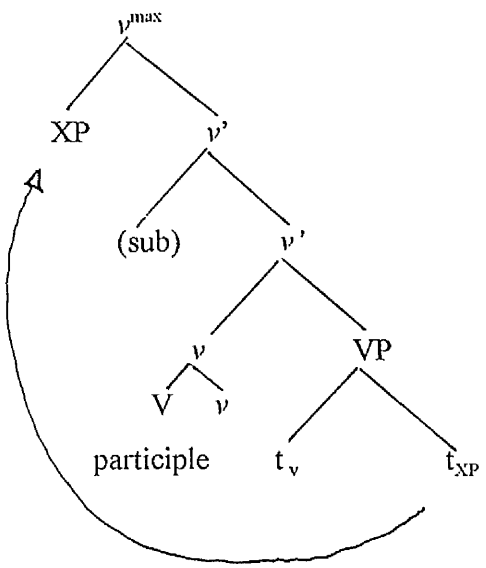
What exactly are the syntactic conditions which result in phonologically overt agreement? I assume, adopting Sportiche's (1992) adaptation of Kayne's original (1989) proposal, that agreement on the participle is not the result of the agreed-with element adjoining to a functional projection immediately above VP, but in fact reflects the presence of the agreed-with element in the Spec of this phrase, to whose head the participle itself is adjoined (overtly). I assume further, also following Sportiche 1992 (and most current work) that this is the Spec position where Accusative case is licensed. In contemporary terms, the projection in question is that of the "light verb" *v*, to whose head the "real verb" has adjoined. The configuration for agreement-triggering is depicted in (23) below.

What about when agreement fails to appear on the participle? As will be apparent from the brief discussion in § 2.2 above, Minimalist assumptions, and in particular *Last Resort*, prevent it from legitimately being assumed - as did both Kayne 1989 and Sportiche 1992, within pre-Minimalist models - that this state of affairs reflects actual "skipping" of an agreement-triggering position by the phrase in question. I will make the alternative assumption (contemplated though not adopted by Sportiche 1992, for a reason to be explained in the next subsection) that the distinction we are looking for is essentially that between overt and covert movement to check the appropriate features. Let us translate this into the framework of Chomsky 1995, in which it will be recalled that Move F (= formal feature) is the basic movement operation (see Chapter 2, § 3.2.1). Here, the difference between overt and covert movement corresponds, more or less, to the difference between movement which takes with it, i.e. pied-pipes, a full category, on the one hand, and movement of features alone, on the other.[22] Let us then make the following assumption about agreement-triggering conditions: agreement on the participle reflects

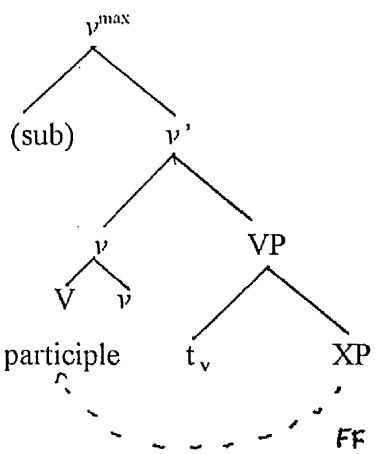
[22] The overt/covert -- category/feature correspondence is not ideal, since, while it must be the case that all covert movement is pure feature movement - due to the fact that full category movement is a reflex only of PF considerations - it need not be the case that all overt movement is movement of a full category.

a category of the chain of the relevant phrase in Spec-v, while no agreement reflects adjunction of just the relevant features of the phrase to the head of the verbal complex (cf. Chomsky 1995b:277), as shown in (24):

(23) Agreement (XP = agreed-with element)



(24) No agreement:



Given the possibility of both pure feature movement and full category movement, note that it becomes clear in general how two LF-structures might be "essentially the

same", and yet not strictly identical.[23] We shall see next how our preliminary assumptions about agreement as depicted in (23) and (24) provide the basis for a modern version of Kayne's (1989) proposal that "two *wh*-movement sentences, one with and one without past participle agreement, will, even if otherwise identical, have category-wise different representations".

3.3.2 Derivations

3.3.2.1 *Wh*-movement

Let us consider (15) again, in particular the option without agreement:

(15a) Je demande combien de tables Paul a repeint
I wonder how-many of tables(f.pl.) Paul has repainted

Here, *combien de tables*, the phrase which fails to be agreed with by the participle, overtly occupies a higher position than Spec-*v*, which I take to be Spec-CP. As Sportiche (1992) observes, this seems to rule out the assumption - which I myself made in the previous subsection - that the absence of agreement reflects covert movement to (as opposed to skipping of) (Spec) *v* (or Spec-AGR_oP, in Sportiche's own analysis) - unless the *wh*-phrase, having arrived in Spec-CP, then lowers there. But what we have here is in fact more or less the same scenario that we discussed in Chapter 4 with respect to *wh*-movement of objects in English: how can an element move *overtly* to its ultimate position yet *covertly* through an intermediate position? Given my proposal in Chapter 4, there is an obvious way of characterizing the agreement-less option of (15) as the product of a derivation in which movement of the phrase not agreed-with is covert, or, to be precise, movement of features alone. Crucially, this derivation will involve copy movement; we shall look at it shortly. I assume that like English, French is a covert Object Shift language, with no strong feature associated with the *v*/V complex (see Chapter 3).

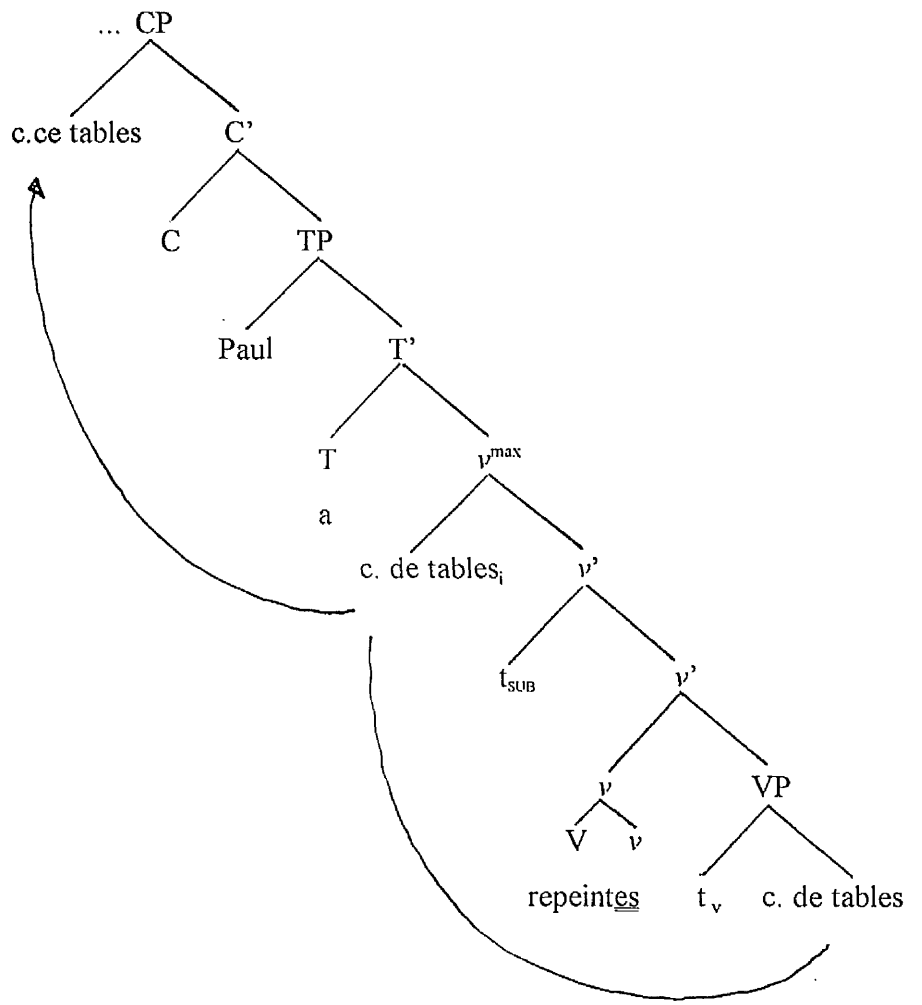
---

[23] This will provide the basis for an account of the interpretive effects mentioned above. See Chapter 6.

The two derivations for (15) can now be characterized as follows, beginning with the option in which the participle does agree, i.e. (15a) (for the time being, I shall simply describe the derivations; the question of economy is fully addressed in §3.4 below). In the relevant part of the derivation leading to this option, i.e. the derivation of the embedded clause - shown in (25) below - *combien de tables* moves overtly to Spec-v, checking Case and ϕ-features, then to Spec-CP, to check the strong feature of C. In principle, this may involve a step-wise movement of the *wh*-phrase, or an application of Form Chain (in the sense of Chomsky 1993); I will be more specific about this in §3.4; what is important at this point is the fact that the derivation results in agreement on the participle, since the *wh*-phrase has moved overtly to check features of the verbal complex containing the participle, which it will be recalled (cf. (23) above) corresponds to leaving a full categorial copy of itself in Spec-v.[24]

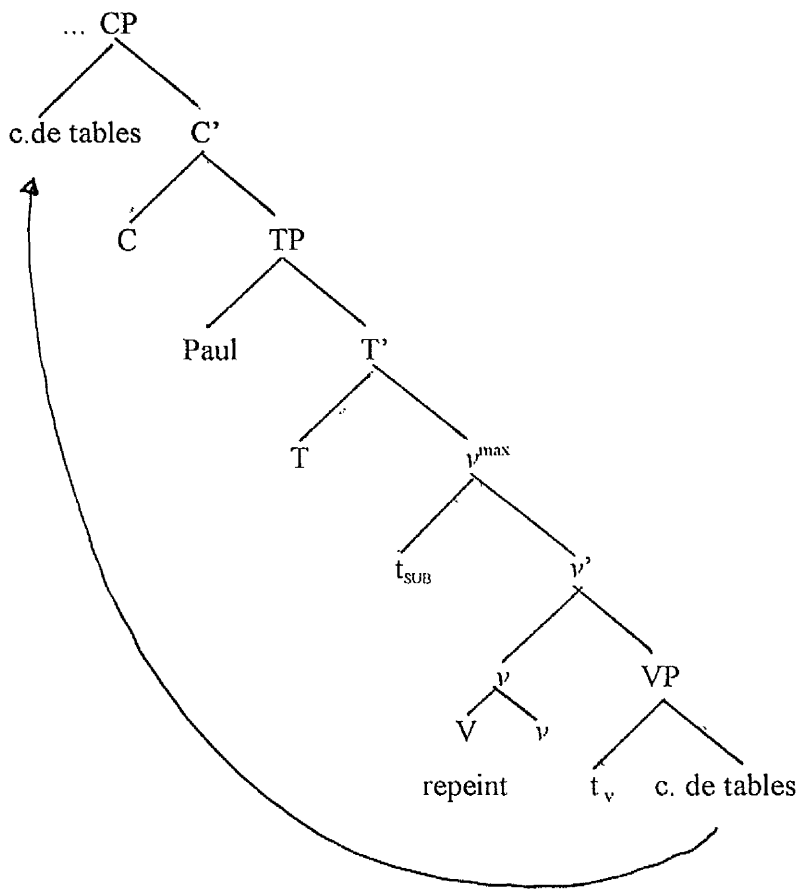
[24] Of course I take it that all movement creates copies, but for convenience and clarity, I only represent the copies for the elements whose copyhood is relevant..

(25)



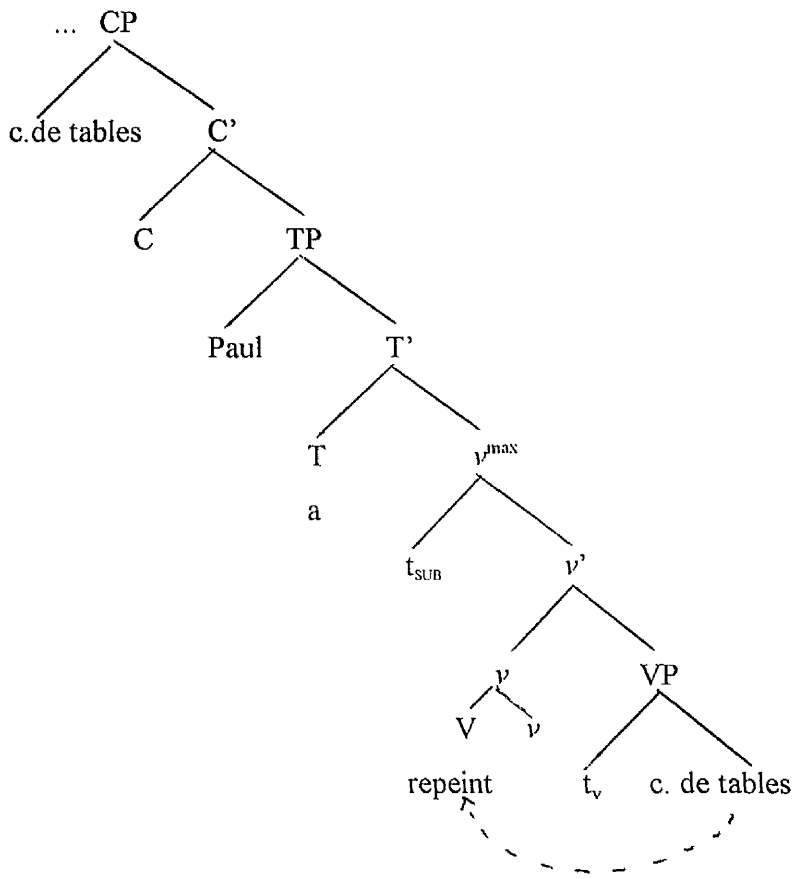
Now contrast (25) with the following derivation, which corresponds to the option of (15) *without* agreement on the participle. As before, the *wh*-phrase has to move overtly to Spec-CP in order to check the strong feature of C. In this case however, the *wh*-phrase moves *directly* to Spec-CP, as in (26a) below (this derivation exactly parallels the derivation proposed for object *wh*-movement in English in Chapter 4, § 2). Under the assumption made above that there is no strong feature associated with the Accusative Case system in French, the failure of *combien de tables* to check the Case and ϕ-features overtly will not cause the derivation to crash.

(26) a.



Let us now assume, following the proposal in Chapter 4 for object *wh*-movement in English, that the Case and ϕ-features associated with the participle, skipped by movement of the full *wh*-phrase, are checked by covert raising of the appropriate features of the copy which the *wh*-phrase left in its original VP-internal position, as shown in (26b).

(26) b.



In this derivation - in contrast with (25) - it is obviously not the full phrase which checks the Case and ϕ-features on the complex verbal head, but only its formal features, which adjoin to the head, resulting in a participle without agreement - cf. (24) above. This completes my characterization of the two options of the *wh*-movement example (15). Let us look next at the Accusative clitic case.

3.3.2.2 Accusative clitics

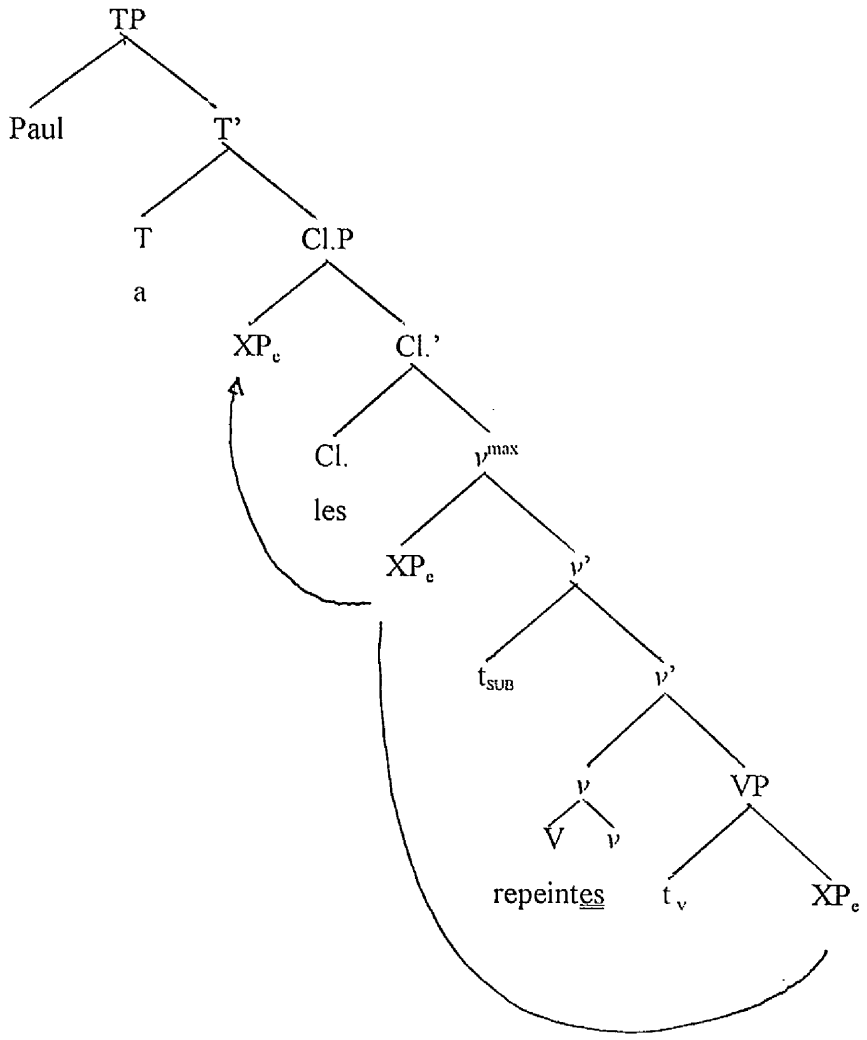
The case of (16) (*Paul les a repeint(es)*), in which the past participle optionally agrees with an Accusative clitic, responds to the same treatment as (15), if one adopts the theory of clitic constructions proposed by Sportiche 1992. According to this, the clitic (*les*

'them' in the case of (16)) is itself the head of a projection situated somewhere above what Sportiche labels AGR_oP (= Chomsky's complex verbal projection), but presumably within the IP system. In the relevant argument position is a phonetically empty phrase which Sportiche likens to *pro*. The latter must ultimately move to the Specifier of the clitic phrase (or *Clitic Voice*, as Sportiche's terms it) for licensing which relates in some way to an interpretive property which he identifies as specificity. At some stage in the derivation, the empty *pro*-like element, being an argument, will also need to check Case/agreement features, in (Spec-)AGR_oP, in Sportiche's own framework, or the complex verbal projection in the framework I am using. For reasons which are not directly relevant here, Sportiche maintains that movement of the silent phrase is overt in French.[25] Given this analysis, with the clitic construction seen as involving a kind of overt operator movement, (16) can be dealt with exactly analogously to the case of *wh*-movement (15). In one possible derivation, shown in a simplified form in (27) (cf. (25) above), the empty phrase moves overtly to Spec-*v*, thereby fulfilling the conditions for participle agreement, then on to the Spec of the clitic phrase. The alternative derivation, shown in (28) (cf. (26) above) is one in which the empty phrase moves directly to the Spec of the clitic phrase, with the relevant features of its copy then adjoining covertly to the V/*v* head, resulting in no agreement on the participle.[26]

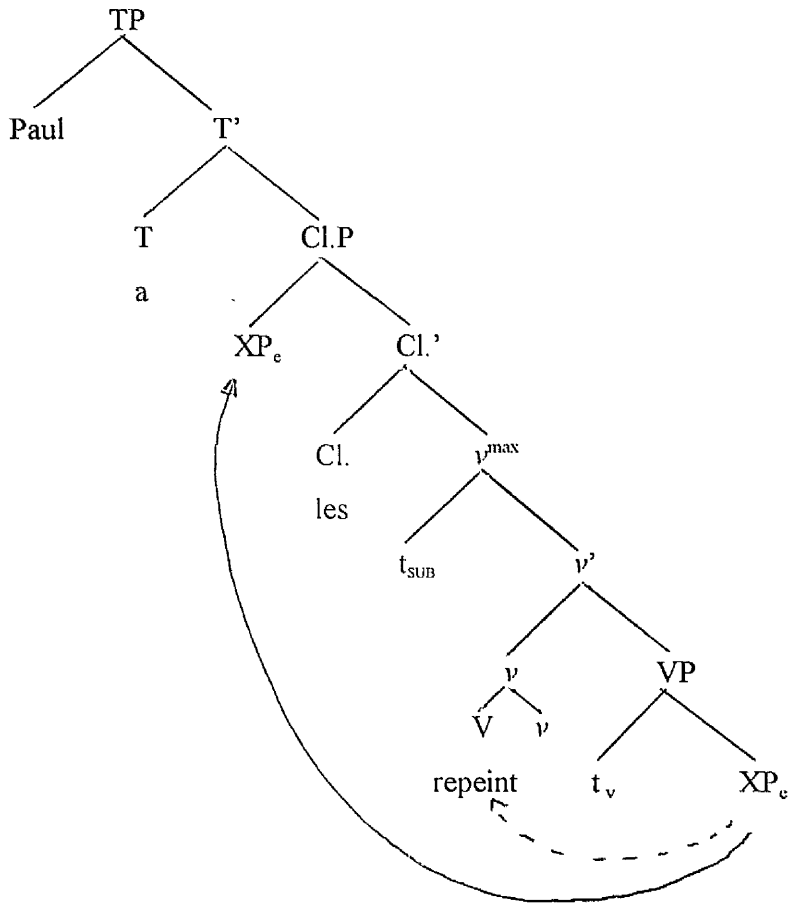
[25] See Sportiche 1992 for details.

[26] I shall assume for concreteness that the clitic raises to T, thereby deriving the observed order.

(27)


TP
Paul
T'
T
a
Cl.P
XP$_e$
Cl.'
Cl.
les
v^{max}
XP$_e$
v'
t_{SUB}
v'
v
V
v
repeintes
VP
t_v
XP$_e$

(28)



So far, we have half an account of the optionality of agreement in (15) and (16): it is a question of overt versus covert movement to check Case/ϕ-features, that is, of movement of an entire category to a specifier position versus adjunction of just formal features to an X^0. The remaining half of the account will consist of an answer to the following question: how is it that derivational economy, in particular the condition Procrastinate, allows both derivations to survive?

3.4 Shortest Derivations, Form Chain and anti-Procrastinate effects

As Chomsky remarks in connection with the French data we are discussing here, "[n]ote that we must assume the two derivations to be 'equally costly', each being 'minimal' by

successive-cyclic movement. This consideration would lead to a further refinement of the notion of 'cost'". (1991:435, note 30).

The aspect of movement which is relevant in connection with (15) and (16), given my analysis of them at least, is when it takes place in relation to the application of Spell-Out, i.e. is it overt or covert? Remember from our discussion in §2.2 above that while there can be absolutely no optionality with respect to whether something moves or not, optionality in the timing of movement is possible in theory. The stage relative to Spell-Out at which movement to check ordinary (i.e. not strong) features takes place has no relevance for convergence, since these features are not relevant at PF. As mentioned earlier, the economy condition Procrastinate in the standard Minimalist framework prohibits overt movement unless it is to check a strong feature. However, as mentioned at the beginning of this chapter, I assume that it is at least a theoretical possibility that the "optimal realization of interface conditions" as determined by derivational economy conditions need not invariably constitute a unique derivation - in principle there may turn out to be more than one.[27] Whether or not this possibility will actually be realized depends on the formulation of the economy conditions, as indicated in the quote from Chomsky above (see also Fukui 1993).Two ways in which to make overt/covert optionality a practical possibility[28] suggest themselves: one involves finding a further derivational economy condition which may, in contrast with Procrastinate, indirectly result in a preference for *overt* movement. Alternatively, one could substitute the standard version of Procrastinate with an "all-in-one" economy condition which can result in either overt or covert movement to be preferred, depending upon other variables, as proposed by Kitahara 1994 in his account of optional overt Object Shift in Icelandic (see note 27 below).

---

[27] Various works make this assumption either implicitly or explicitly; see e.g. Chomsky's (1991) account of optional infinitve movement in French; Kitahara's (1994) account of optional overt object shift in Icelandic (reviewed in Chapter 6); Svenonius (1995) on optional particle shift.

[28] Apart of course from taking a lexical approach in terms of optionally strong features, a strategy which I have chosen to avoid here.

Here, I shall explore the hypothesis that the factor responsible is the requirement that computational operations (in a sense to be clarified shortly) be minimized - a requirement which has surfaced in the form of Chomsky's (1991) Least Effort principle, Branigan's (1992) Global Economy, Kitahara's (1994) Shortest Derivation Requirement, among others. A condition of this type is also assumed by Chomsky 1995, under the name of Shortest Derivation (hereafter SD). SD states that "shorter derivations block longer ones" (1995:314). It may well be asked, If two derivations from some numeration are convergent and moreover obey Last Resort, in what sense can one of them be "shorter" than the other? Chomsky does not give a fully explicit definition of SD, but the answer seems to be that one *movement* operation may result in one or more *feature-checking* operations - as is explicit in the most recent theory of feature-checking (Chomsky 1995; see Chapter 2, § 3.2.1 for an outline). According to this theory, when one feature of some element moves for checking, all the other formal features belonging to that element get transported, or strictly speaking, copied, to the checking-site, and potentially checked there, for no extra cost. Recall that these subsidiary features are referred to as *free riders*. Given this conception of movement, one application of Move will result in the checking of at least one, but possibly more, features - hence two convergent derivations from the same numeration may indeed involve different numbers of Move operations, depending crucially on what moves where -- and when, as we shall see shortly. It can be assumed then that the Shortest Derivation is that in which all necessary feature-checking is completed by means of the fewest Move operations possible. This seems to be exactly what Chomsky (*op. cit.*) assumes in his discussion of Icelandic-like languages, which consists virtually the sole argument he offers in favour of the SD condition.[29]

In using the Shortest Derivation condition to potentially favour overt movement (counteracting Procrastinate) I shall make the crucial assumption that Form Chain operations, in the sense of Chomsky 1993 and Collins 1994, are possible. Form Chain creates a chain via a single "instantaneous" operation, rather than a sequence of steps, which makes a difference when it comes to chains of more than two members. The notion

---

[29] For more about Icelandic and on Chomsky's discussion of Shortest Derivation in this connection, see Chapter 6.

of Form Chain was introduced by Chomsky 1993 in order to avoid a theoretical conflict between the requirement for derivations to contain the fewest possible steps, on the one hand, and the Minimal Link Condition (MLC/Shortest Move), on the other: "If a derivation keeps to shortest moves, it will have more steps; if it reduces the number of steps, it will have longer moves" (1993:15). Chomsky's solution to this predicament was to assume that all chains are formed in by a single step, in an operation he called Form Chain. The idea of this was, in other words, that the sub-operations involved in creating a chain, although they may involve separate instances of feature-checking, do not count as separate operations for the purposes of calculating economy. Due to the fact that MLC is not an economy condition at all in the framework of Chomsky 1995,[30] this particular motivation for Form Chain operations ceases to exist. However, there appears to be no reason to believe that many-linked chains should exclusively either be formed by a step-wise series of movements, or, in Collins's (1994) words, "all at one time", and I shall assume that *both* these methods are available in principle.[31]

Stated in this way, this assumption sounds somewhat *ad hoc*. However, there is a way to think of the availability of both instantaneous and step-wise methods of chain-formation as resulting from principles which are already familiar. Recall that the Shortest Derivation condition requires all necessary feature-checking to be done via the fewest operations possible. Let us then assume that SD requires chains to be formed by instantaneous Form Chain, unless, as with Procrastinate, this would prevent the derivation from converging. When might such a situation arise? Note that if strength of a feature amounts to that feature needing to be checked before any further Merge operations take place (as in Chomsky 1995), then there will indeed be situations in which, if the derivation is to converge, a chain can only be formed in a successive, step-by-step way - namely, when more than one link of the chain involves the checking of a strong feature. This is exemplified by the case of *wh*- subjects in English, which must check two strong features: one ("EPP") in T, followed by another in C. Here it is obvious that the derivation will

---

[30] See Chapter 2, § 3.1.

[31] See Collins *op. cit.* for arguments for assuming Form Chain operations. These have to do with what he calls "chain interleaving" effects in English and Ewe.

crash if the first strong feature (i.e. the EPP feature) is not checked as and when it appears - i.e. *before*, rather than instantaneously with, the strong feature in C which comes into the structure at a later stage. The relevance of this discussion of Form Chain and the Shortest Derivation condition will become clear shortly.

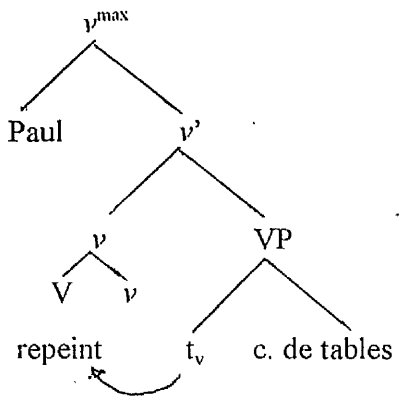
We are now equipped to explain how, in (15) and in (16), the Shortest Derivation condition can result indirectly in what we could call an *anti-Procrastinate effect* - that is, overt movement for the checking of an ordinary feature - counteracting the preference of Procrastinate itself for covert movement, and thereby making possible *either* overt *or* covert movement to/through the checking domain of the ordinary feature - i.e. an optionality between overt and covert movement.[32] The optionality data we are seeking to explain is repeated here:

(15) Je me demande combien de tables Paul a repeint(es)
I wonder how-many of tables(f.pl.) P. has repainted(f.pl.)

To illustrate clearly how I am claiming that optionality arises, let us consider the derivation/s of (15) "from the bottom up"; as before, I describe only the relevant part of the derivation, i.e. that involving the embedded +*wh* clause. Recall that there is just the one numeration involved in both options. In any event, a structure consisting of the participle itself (*repeint*), its internal argument (*combien de tables*) is first constructed, with the functional category *v* and the subject then Merged. Presumably, the participle raises overtly to *v* to form a complex head. At this stage we have the structure in (29).

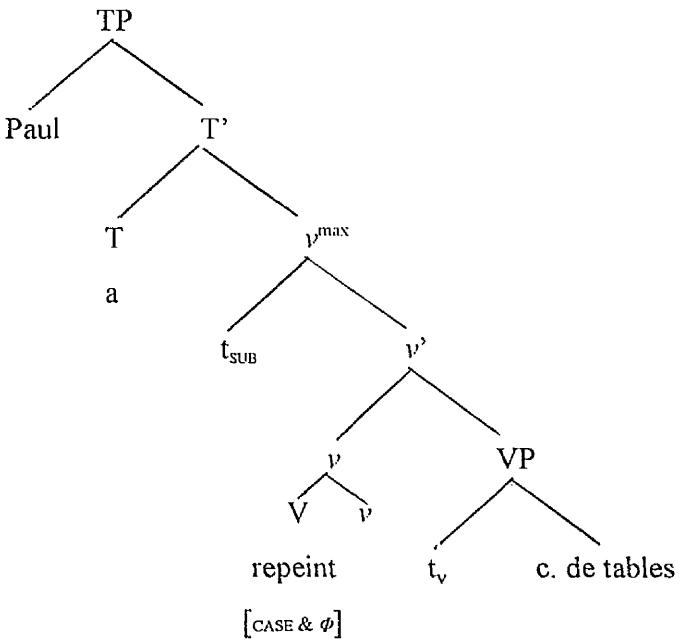
[32] Recall that Procrastinate does not actually make direct reference to overt or covert movement, i.e. it is not the case that overt operations are *inherently* more economical than covert ones, given the core Minimalist tenet that "the computational procedure C_{HL} is uniform from N to LF; any distinction between before and after Spell- Out is a reflex of other factors" (Chomsky 1995:327) (see Chapter 1).

(29)



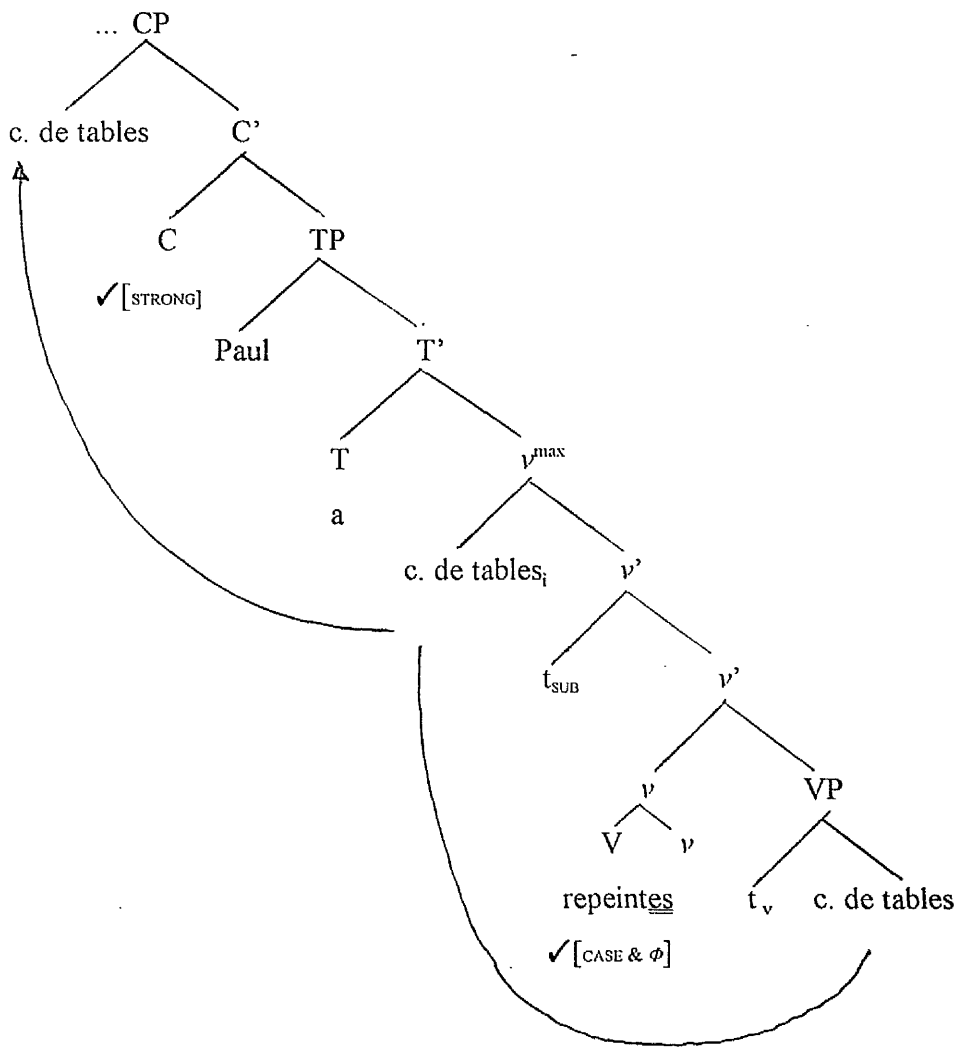
The complex head containing the participle contains Case and ϕ-features, but no strong feature (see above). Since it is strong features only which force overt movement, the Case and ϕ-features remain unchecked at this stage, and become embedded by further operations which construct the rest of the "IP" complex, with their prospective checker *combien de tables* remaining in situ for the time being - (30).

(30)



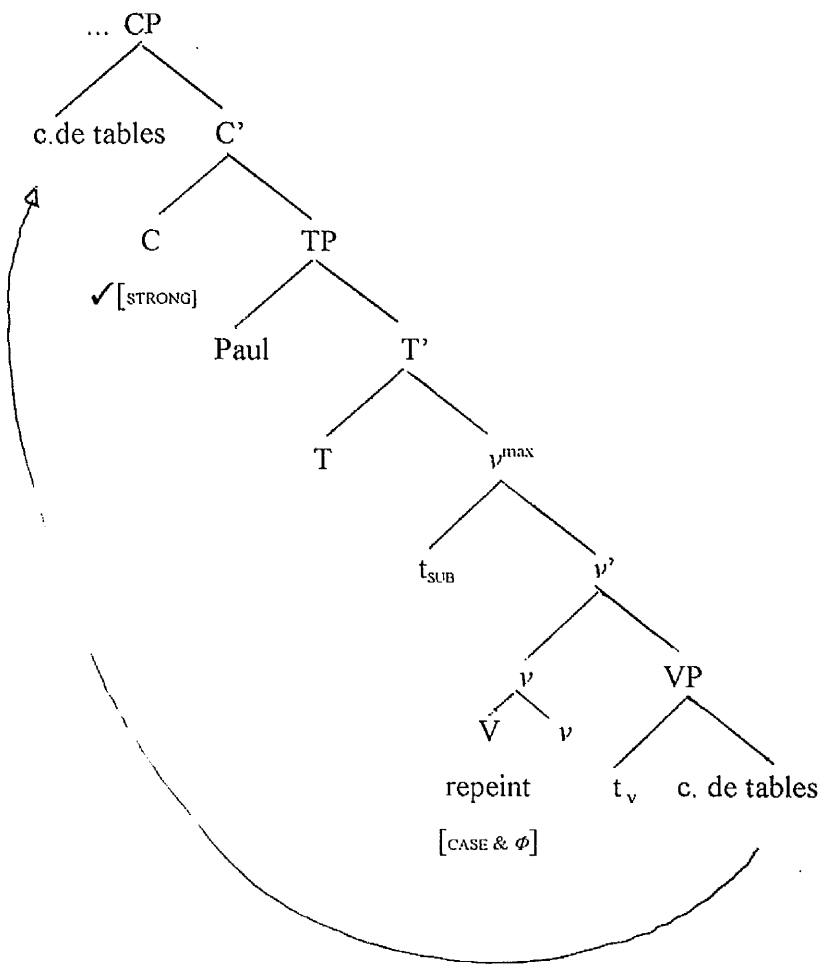
The next stage is for the IP to be merged with a C element, which does contain a strong feature requiring immediate movement of *combien de tables* to check it. It is at exactly this point that two convergent continuations of the derivation become possible with respect to the *wh*-phrase - viz. (25) and (26) above: either *combien de tables* checks all features relevant to it via an application of Form Chain, instantaneously creating a chain with links in Spec-*v* and Spec CP. This derivation results in agreement on the participle (cf. (23) above). Alternatively, *combien de tables* moves directly to Spec-CP, bypassing the Case/ϕ-features in the verbal complex, which are later checked by covert raising of the appropriate features of the *wh*-phrase's copy - resulting in no agreement on the participle (cf. (24) above).

(25) Derivation resulting in agreement



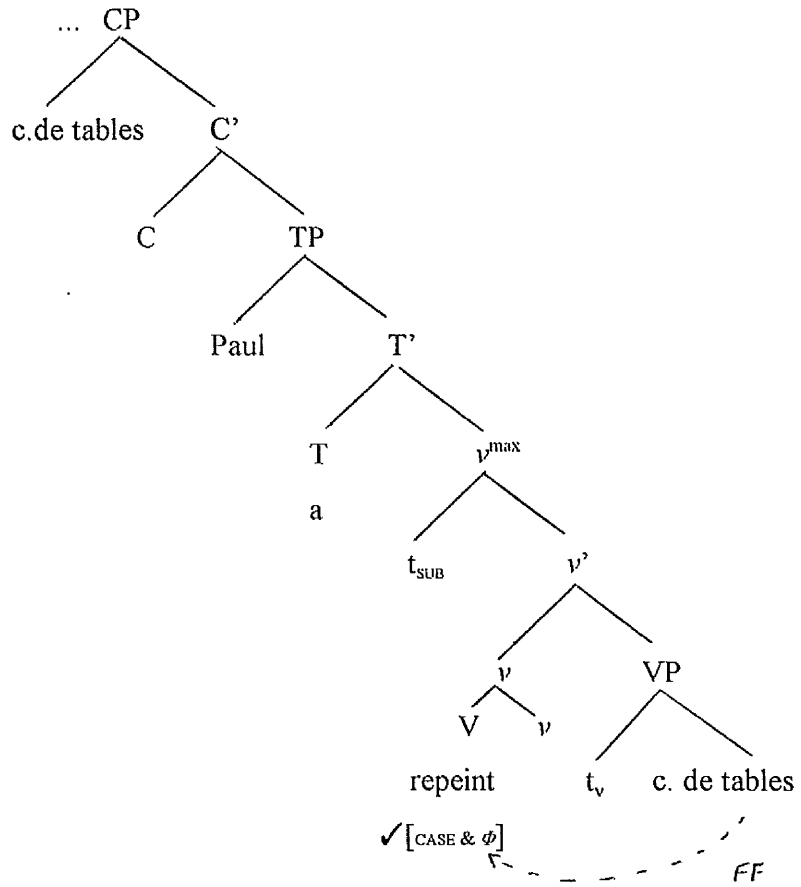
(26) <u>Derivation resulting in no agreement</u>

a.



(26) Derivation resulting in no agreement

b.



And finally, how do the two convergent derivations fare with respect to derivational economy? It is obvious that Procrastinate should prefer derivation (26) to derivation (25), since the latter involves overt checking of the Case and ϕ-features, which are not strong features and could legitimately have been checked by means of covert raising of the appropriate features. At the same time, however, the Shortest Derivation condition causes an anti-Procrastinate effect; it prefers derivation (25), since in this case, a single movement operation - i.e., an application of Form Chain - brings about checking of all the features (both the Case and ϕ-features in the verbal complex, and the features associated with C) associated with the *wh*-phrase, in contrast to derivation (26) in which, with one overt and one covert operation, the required feature-checking is necessarily

implemented by *two* separate operations. Both (25) and (26) count as the (or rather an) optimal route to convergence - one optimal for Shortest Derivation, the other optimal for Procrastinate - and I therefore assume that both are permitted. Recalling our characterization of the clitic case (16) in terms of Sportiche's (1992) theory, i.e. as involving overt A'-movement of a silent argument, the same account extends straightforwardly to the optionality in this case.

Summing up, my claim is that agreement optionality in (15) and (16) - itself by hypothesis the reflection of an optionality between overt and covert movement - is attributable to the fact that derivational economy conditions effectively fail to determine a unique optimal derivation. This situation in turn is linked to properties (strength/lack of it) of the features involved. In this last respect, the account of *optionality* in overt/covert movement is formulated in the same terms as the standard Minimalist account of movement which is obligatorily overt or obligatorily covert, a welcome characteristic.

It could be pointed out, of course, that it is easy enough to make "refinements" to the derivational economy system for the purposes of accounting for one or two cases of optionality. But it needs to be ensured that such an account does not predict too much optionality - we know, for instance, that in the majority of syntactic environments in French, participle agreement is not optional at all, but in fact impossible or obligatory (see examples (17)-(20) above). I address these facts next.[33]

---

[33] Under my analysis, notice that the question arises as to why overt Object Shift is not an option in *wh*-movement in English - as argued for in Chapter 4 (on the basis of e.g. the distribution of Floating Quantifiers). In an earlier version of the analysis, set within the 1993/1994 version of Minimalism, I took overt Object Shift in English to be completely excluded by the Shortest Move/MLC condition, due ultimately to the fact that the verb does not raise overtly (in contrast to French). See Pettiward 1995. It is not so obvious whether this solution is still available within the more recent version of Minimalism. I return to this important question in Chapter 6, § 2.2.4.

4 When agreement is not optional in French

As we saw in §2, (15) and (16) are in fact the only types of construction in French in which participle agreement is optional. In all other situations, agreement is either impossible or obligatory.

4.1 Two situations where agreement is impossible

Let us start with the case of (31), in which the participle cannot agree:

(31) Paul a repeint(*es) les tables
Paul has repainted(*f.pl.) the tables(f. pl.)
'Paul repainted the tables'

In this case it is transparent that the agreement-triggering configuration is not met: *les tables* is not in Spec-*v*. The real question is, why is this so? In other words, what explains (32)?

(32) * Paul a les tables repeint(es)
P. has the tables repainted

I have assumed throughout that in French, there is no strong feature associated with the functional category *v*. Staying with this assumption, overt movement of the object to Spec-*v* will be ruled out straightforwardly by Procrastinate. It is important to note that in this simple declarative case, the Shortest Derivation condition cannot possibly have an "anti-Procrastinate effect" on the object, because here - unlike in (15) and (16) - the object only has to check features at one site, with the result that only one application of movement can possibly be involved, be it overt or covert; the Shortest Derivation condition is immaterial here, with respect to the object.

We also need to account for why agreement on the higher participle *dit* is

impossible in an example such as (33) (example adapted from Branigan 1992):

(33) Les tables(f.pl.) qu'il a dit(*es) que Paul a repeint(es)
The tables that-he has said(*f.pl.) that Paul has repainted(f.pl.)
'the tables which he said Paul repainted'

This example is not problematic for any of the assumptions I have made. Since I have assumed, following Sportiche 1992 and Branigan 1992, that agreement is triggered by the presence of the agreed-with element in a Case, i.e. A-, position, it is clear that this configuration could not be met with respect to *les tables* (or more precisely, the associated operator) and the participle *dit* (assuming in the first place that this element even has relevant features), since this would entail "improper movement" of the operator from the embedded Spec-CP to Spec-*dit* [34] (see Chomsky 1995:325 for brief discussion of this case) .

4.2 Obligatory and impossible agreement in *être* environments

Let us now examine a further set of cases in which participle agreement is not optional, all of them involving the auxiliary *être*. In some of these cases, agreement is obligatory, in others, it impossible:

(34) Les tables ont été repeint*(es)
The tables(f.pl.) have been repainted*(pl.)
'the tables were repainted'

(35) La femme de Paul est mort*(e)
The wife(f.) of P. is died*(f.)
'Paul's wife died'

[34] The improper movement explanation presupposes that movement is forced through the embedded Spec-CP for locality reasons of some sort. If this is not assumed, one might alternatively rule out A-movement from Spec-*v* dominating *repeint* to Spec-*v* dominating *dit* (if indeed the latter exists) as a "superraising" violation, due to the presence of *Paul* in the embedded Spec-TP.

(36) Il sera repeint(*es) beaucoup de tables cette année
It will-be repainted(*f.pl.) many of tables(f. pl.) this year
'a lot of tables will be painted this year'

(37) Je me demande combien de tables il sera repeint(*es) cette année
I ask how-many of tables it will-be repainted(*f.pl.) this year
'I wonder how many tables will be repainted this year'

Let us start with (34) and (35), in which agreement is obligatorily triggered on the participle. In the context of the proposal in the previous section, these examples raise the following potential difficulty. Recall that feature properties were instrumental in accounting for the optionality in the *wh*-movement and Accusative clitic cases (15) and (16). With respect to features, notice that (34) and (35) bear some resemblance to (15) and (16), in that features must evidently be checked at two different sites by the objects. The two sets of examples are also alike in that in both, the ultimate checking-site contains a strong feature ("EPP" feature, in the NP movement cases). Yet despite these resemblances between (34)/(35) and (15)/(16), agreement is optional only in the latter cases. The question is, therefore, why is the derivation involving covert feature-raising from a copy not a possibility in the NP-movement cases, leading to a permissable alternative version without agreement? What *forces* the objects in (34) and (35) to move overtly through Spec-v?

There are various factors which could possibly be invoked to explain this property of the NP movement cases. One strategy would be to say that the objects are prohibited from skipping (Spec-)v on their way to the higher A-position for reasons of locality. This approach is suggested by Adger 1994, although it is not so clear whether this would be viable under the recent reformulation of locality (i.e. the Minimal Link Condition) in terms of the feature moved-to (Chomsky 1995).

Notice that another possible explanation might be in terms of "reconstruction", making use of the idea that only A'-movement results in a full copy of the moved element, as proposed by Chomsky 1993 - or indeed, that the formal features of the traces of A-movement are deleted and erased, as proposed by Chomsky 1995 but argued against in

Chapter 2 of this thesis. Under such an account, derivations involving raising of the features of a copy would obviously be unavailable in A-movement situations like (34) and (35); the only convergent course of action for the objects would be overt movement to Spec-TP via Spec-*v*, hence the obligatory agreement. This hypothetical explanation is particularly relevant, since if it is successful, it provides indirect evidence for Chomsky's (*op. cit.*) proposal that the formal features of the traces of A-movement delete and erase, and against the Copy Hypothesis proposed in Chapter 2 of this thesis. Below, we will in fact see that the "reconstruction" account is unsuccessful.

The third possible strategy for accounting for the obligatoriness of overt movement through Spec-*v* in (34) and (35), and perhaps, all other things being equal, the least attractive, is to propose some strong feature associated with the functional head *v*. An approach of this type is indeed suggested for the cases in question by Branigan 1992. It is clear that any of the above three approaches will be able to handle the simple cases (15) and (16). However, their predictions diverge when more complicated cases, in particular, (37) (see below) are considered; in particular, we shall see that the "reconstruction" approach loses out to the strong feature approach. Before explaining why this is so, let us firstly outline Branigan's (1992) version of the latter which I shall essentially be adopting here.

Noting that the constructions at issue all share a certain syntactic property, namely the presence of the auxiliary *être* ('be'), Branigan suggests that this is the factor ultimately responsible for obligatory agreement, an idea which he implements in the form of the following statement: "auxiliary BE selects an AGRP with strong N-features on the head" (1992:43).Under this assumption, the explanation for the obligatoriness of agreement in e.g. (34) is simple: due to the proposed strong feature in the AGR/V complex (= my *v*/V), the only convergent derivation is one in which *les tables* moves overtly through Spec-AGRP; if the strong feature fails to be checked overtly (the situation which would give rise to no agreement), the derivation crashes. The necessity of agreement on the participle in these cases is thereby accounted for, in a way which is not in itself fully explanatory (as Branigan admits), and yet is likely to be on the right track, since the cases

in question are independently known to share certain syntactic properties.[35]

It will however be noticed that there is an apparent counterexample to the Strong feature approach in (38):

(38) * Il sera beaucoup de tables repeintes cette année
It will-be many of tables repainted this year
'many tables were repainted'

Given the alleged strong feature associated with *v*, why is it that the object *beaucoup de tables* cannot move overtly to Spec-*v*? (38) in fact makes an interesting pair with example (36), repeated below, where agreement is impossible: how is the strong feature of *v* supposed to be checked here, since *beaucoup de tables* is overtly *in situ*?

(36) Il sera repeint(*es) beaucoup de tables cette année
It will-be repainted(*f.pl.) many of tables(f. pl.) this year

I think that these two problems - the impossibility of participle agreement in (36), and the apparent counterexample (38) to the strong feature approach to the NP movement cases - are amenable to a single solution. This involves derivational economy, and is based on Chomsky's (1994) analysis of some partly similar cases from English.[36] The major premise is that Merge (i.e. structure-building) operations, unlike movement, cost nothing for the purposes of derivational economy (Chomsky 1995).

Let us assume, following Chomsky 1995, that *il* is itself capable of checking ϕ-

---

[35] A useful place to look with a view to adding some content to the strong feature idea might be Kayne's (1993) theory of "auxiliary selection", in which he attributes a type of biclausal structure to constructions involving both *être* and *avoir*.

[36] The English examples are as follows; I return to them in § 5.

(i) There seems to be a man in the room
(ii) * There seems a man to be in the room

(and Case) features,[37] and is therefore not linked to the associate by movement at all. In the first place, there is the fact that the relevant verb does not show agreement with the associate, instead showing either no agreement, or default agreement (third person singular) - see (39). Further evidence involves the fact that the "associate" of *il* behaves for the purposes of Binding as if it were *in situ*, suggesting that it never raises covertly to the subject position (Spec-TP). (cf. expletive replacement) - (40), compare with (41). The implication of this is that the expletive itself checks all the relevant features. With respect to these properties, *il* contrasts with English *there*. In *there* constructions, the relevant verb shows subject agreement with the associate although its "surface subject" is the expletive - see (42); furthermore, the associate of *there* is able to bind as if from subject position. Chomsky therefore proposes that *there* can check only categorial features - see (43).[38]

(39) Il est/*sont arrivé trois hommes
It is(m.s.)/*are(m.pl.) arrived three men(m.pl.)
'there arrived three men'

(40) * Il est entré trois hommes sans s'annoncer
it is entered three men without themselves-to-identify
'there entered three men without identifying themselves'

(41) Trois hommes sont entrés sans s'annoncer
three men are entered without themselves-to-identify
'three men entered without identifying themselves'

(42) There are/*is three men in the garden

---

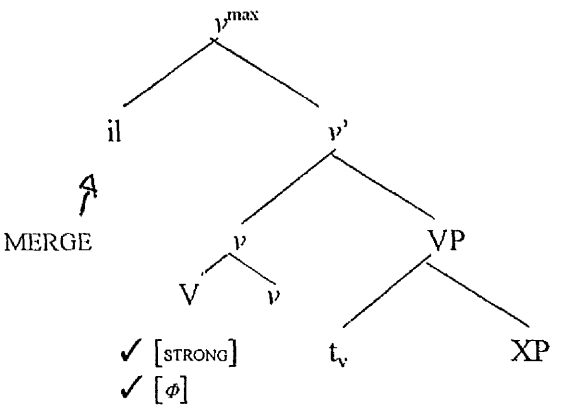
[37] In the examples at hand, all involving unaccusative constructions, there will be no Case features associated with *v*, though there is evidently agreement.

[38] In § 5 below, we shall see a consequence of this proposed feature difference between the French and English expletives. With respect to *there*, notice incidentally that its agreement behaviour appears to be unexpected in the light of my (§ 3.3.1 above) assumption that overt realization of agreement features on a verb-form reflects the presence of a full category of the chain of the agreed-with element in the Spec of phrase headed by the verb-form. Clearly, *the chain of the agreed-with element* has to be construed as including an expletive-associate chain.

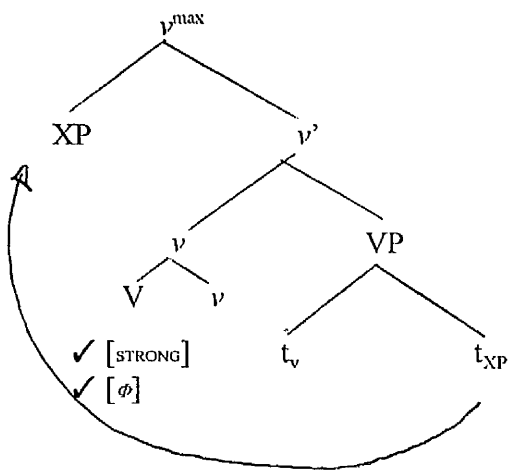
(43) a. There arrived three men without identifying themselves
b. There arrived with their own books three men from England

On the assumption that *il* is capable of checking agreement (and Case) features, it turns out that the Shortest Derivation condition (see § 3.4.2 above), will prefer a derivation in which *il* checks the strong feature of *v* via a Merge operation - viz. (36) - with the ϕ-features getting checked as "free riders". As desired, this blocks the alternative derivation - (38) - in which the strong feature of *v* gets checked via overt movement of the associate *beaucoup de tables*. The reason is that the derivation including overt movement of the associate incurs *one* movement operation to check the features associated with the *v*/V complex (viz. strong feature and agreement), as compared to the derivation where *il* implements the checking by Merge, which obviously involves *no* movement to do this. The situation is illustrated in (44) (XP = associate).

(44) a. <u>Preferred derivation (cf. (36)) (zero moves)</u>



b. Blocked derivation (cf. (38)) (1 move)



The important assumption is that *il* can check agreement features, so that in the preferred derivation ((36)), those associated with *v*/V get checked by *il* as free riders, resulting in lack of agreement/default agreement on the participle. Note that the same situation obtains with the auxiliary in this example, which is also prohibited from agreeing with the associate, instead displaying default agreement, like the participle.

(36') Il * seront/sera repeint beaucoup de tables
It will-be(*pl.) repainted many of chairs(pl.)
'a lot of chairs will be repainted'

Notice incidentally that there should be nothing problematic about the claim that *il* can check ϕ-features in Spec-*v* then move on for further checking. An analogous situation occurs in the non-expletive counterpart of (36), as in (45):

(45) Beaucoup de tables seront repeintes cette année
many of tables(f.pl.) will-be(f.pl.) repainted(f.pl.) ...

Here, both the participle and the auxiliary agree with *beaucoup de tables*. This indeed is expected if the ϕ-features on arguments are "interpretable" and hence may check repeatedly (Chomsky 1995b:284; see Chapter 2).

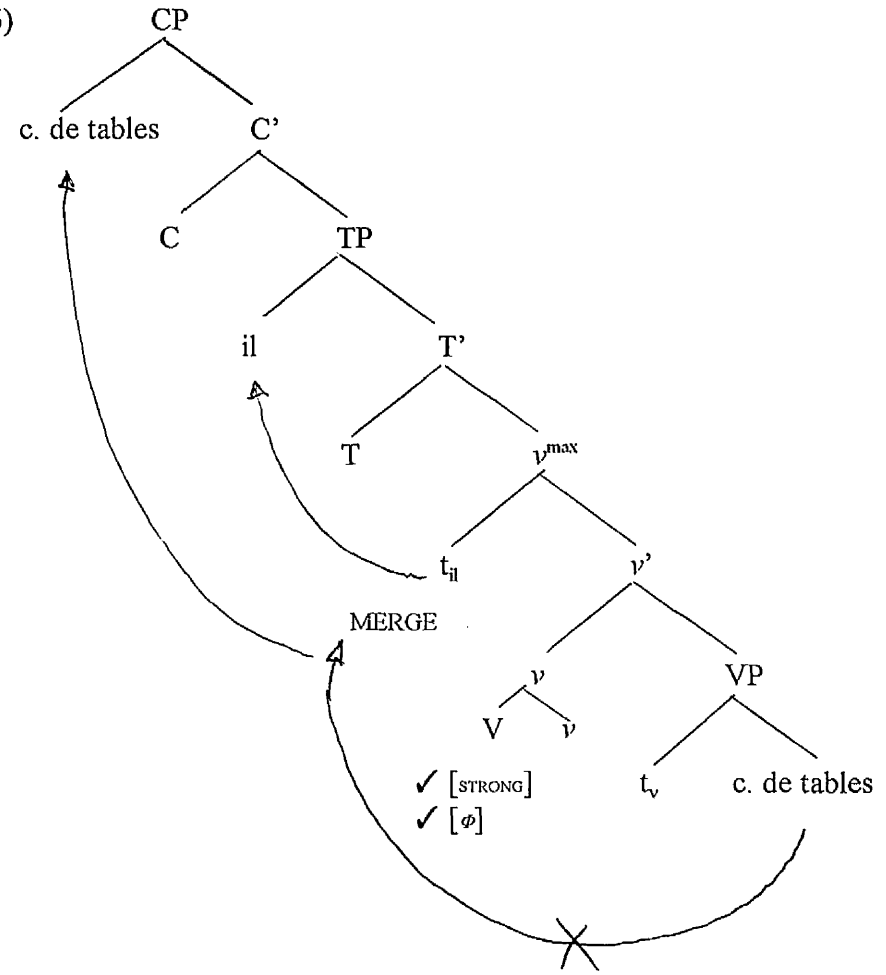
The conclusion is that (38) does not constitute a true counterexample to the strong feature approach to obligatory agreement in the NP movement cases. The example can be excluded by independently motivated assumptions, in a way which simultaneously accounts for the impossibility of participle agreement in (36).

Last but not least, we must consider (37), repeated below. This example is an important one, since we shall see that it distinguishes between the strong feature approach to obligatory agreement in (34)/(35), and the "reconstruction" approach briefly described earlier. In this construction, featuring *il* this time with a *wh*-phrase associate, Kayne reports that agreement on the participle is ungrammatical, analagously to the non-*wh*-associate example (36) (*il sera repeint beaucoup de tables...*).

(37) Je me demande combien de tables il sera repeint(*es) cette année
I wonder how-many of tables it will-be repainted(*f.pl.) this year

Why is the participle unable to agree with *combien de tables*? Let us see firstly how this can be explained using exactly the same assumptions as were used to explain the lack of agreement in (36) (giving an important advantage over Kayne's (1989) proposal - see § 4.3 below), and secondly, how the "reconstruction" approach by contrast fails to get the desired result. As in (36), the numeration corresponding to (37) contains *il*, and, by the Shortest Derivation condition, checking of the strong feature on *v* is most economically, and therefore necessarily, achieved by merging the expletive, which furthermore can and will also check the ϕ-features as free riders (resulting in no/default agreement on the participle). As a result of this, there will be no features left to be checked on *v* by the time *combien de tables* itself is forced to move, i.e. when C is merged into the structure: *combien de tables* will not be allowed to land in Spec-*v*, as there will be no checkable features there, and by Last Resort "Move raises feature F to target K only if F enters into a checking relation with a sublabel of K as a result of the operation" (Chomsky 1995:280).

(46)



Hence, the only possible move for *combien de tables* is straight to Spec-CP; it cannot form a chain with a link in Spec-*v* (giving rise to agreement) because this would violate Last Resort: there are simply no relevant checkable features on *v*/V. These have already been checked by the expletive, giving rise to no agreement (or default agreement) on the participle.

Finally, let us see how the example (37) argues against a "reconstruction" approach to the obligatoriness of participle agreement in "NP-movement" cases like (34) and (35) above, as I claimed earlier. Recall that the fundamental idea of the reconstruction approach is that traces of A-movement have their formal features deleted and erased (Chomsky 1995, refuted in Chapter 2), hence the only convergent derivation for the NPs in such cases as (34) and (35) would be overt movement to Spec-*v* then on to Spec-TP; direct movement straight to Spec-TP followed by covert feature-raising from the copy

would not be possible, since movement to Spec-TP is A-movement and hence results in deletion and erasure of the formal features of its trace. Although, as observed earlier, this "reconstruction" approach works adequately for the simpler cases (34) and (35), it is plain to see that when it comes to (37), this account can only work if supplemented with the strong feature assumption. On the other hand, the strong feature approach works independently of the "reconstruction" approach, for (34)/(35) as well as for (37), which makes the strong feature approach the superior of the two.[39]

To show that the "reconstruction" approach cannot work unsupplemented by the strong feature approach, let us see what happens if we assume (contrary to above) that there is no strong feature associated with *v* in constructions with *être*: just ordinary ϕ-features. Now reconsider the example in question, (37), for which the desired outcome is that agreement is impossible, i.e. that the associate, *combien de tables*, cannot move overtly through Spec-*v*. The ϕ-features in *v*/V, since they are by hypothesis not strong, will not need to be checked as soon as they enter the structure (be it by Merge of the expletive *il*, or by movement of the associate *combien de tables);* on the contrary, the ϕ-features.can still legitimately be unchecked by the time the element C is added to the structure (which will consist of TP with *il* as its Spec). But notice that in the relevant respect we now have a duplication of the scenario which obtains in the case of *wh*-movement without an expletive, as in (15) above: the *wh*-phrase may either undergo instantaneous Form Chain, creating links in Spec-*v* and Spec-CP (optimal by Shortest Derivation), or it may move straight to Spec-CP - crucially leaving a full copy, since this is A'-movement - with the ϕ-features of the verbal complex checked by covert raising of the corresponding features of this copy (optimal by Procrastinate). In short, the prediction is that we should get optional agreement on the participle. The prediction is false, since participle agreement is in fact impossible in (37). It is clear moreover that the faulty assumption is that there is no strong feature in *v*: if, contrary to this, we assume that there is indeed a strong feature in *v*, then we predict - as outlined above - that the expletive must

[39] Note of course that whichever of the two approaches we chose, there is no avoiding strong features, and so to favour the "reconstruction" approach on the grounds of shortcomings of the notion of strong features (e.g. stipulativeness) would not be valid.

check the feature by merge, resulting in obligatory default agreement/lack of agreement on the participle -- precisely the desired outcome.

4.3 Kayne, expletives and agreement: a short comparison

To conclude Section 4, I shall draw a small comparison between the above proposal on agreement facts in *il* constructions and that of Kayne 1989. Kayne attributed the impossibility of agreement in cases like (37), repeated yet again below, to improper movement of the associate from the agreement-triggering position - which recall that he took to be an adjoined, hence an A'-, position - to replace the expletive in Spec-IP. Indeed, one of the main virtues which Kayne attributed to his particular analysis of participle agreement was the fact that it made possible this improper movement account of (37).

(37) Je me demande combien de tables il sera repeint(*es) cette année
I ask how-many of tables it will-be repainted(f.pl.) this year
'I wonder how many tables will be repainted this year'

Of course, the improper movement explanation of (37) is unavailable in the context of my own analysis, in which the agreement-triggering position is taken to be an A-position. But leaving aside this framework-specific detail, note that from a more general perspective it is not clear to what extent Kayne's analysis allows for a properly unified explanation of the impossibility of agreement in cases with expletives generally, i.e. in cases with both regular and *wh*- associates. It seems that to account for all the expletive facts in the same way, i.e. in terms of improper movement, one would be forced to assume a somewhat unlikely difference between the expletive cases such as (36), (37), on the one hand, and NP-movement cases such as (34), (35), on the other.

Recall that in the NP-movement examples, participle agreement is possible, and indeed obligatory. It is therefore necessary to say that in this case, there is an A-position available to the associate, in which agreement can be triggered. However, to explain the

impossibility of agreement with the regular associate of an expletive (36) in the same way as the *wh*-associate case (37) - i.e. to explain it in terms of improper movement of the associate, according to Kayne's analysis - it would be necessary to claim that in (36), there is no A-position available in which the associate could land (if there was, it would be falsely predicted that agreement was possible with the associate). From a theoretical perspective, a unified account of the impossibility of agreement with both regular and *wh*-associates of *il* is obviously desirable; however, countenancing such a difference between constructions identical but for the presence of the expletive (as Kayne's analysis seems to require) is a high price to pay. By comparison, the economy-based analysis presented in this section deals with the expletive facts without the need to complicate basic assumptions about agreement-triggering configurations.[40]

5 Expletives and optional associate movement in English and Swedish

In the previous section, we saw that in French, the associate of an expletive in a passive construction is obligatorily *in situ:*

(47) a. Il sera repeint beaucoup de tables
b. * Il sera beaucoup de tables repeint

It is interesting to note that English *does* have the option for associate movement in what appears to be a similar type of construction involving the expletive *there*: this optionality is illustrated in (48)-(50):[41]

---

[40] On the other hand, it should be pointed out that my analysis involves some departure from Minimalist assumptions, since an expletive merging in a lower position then moving up is ruled out by a condition called "have an effect on output" (Chomsky 1995:294).

[41] In the context of discussing the fact that overt Object Shift is generally not licensed in English, Déprez (1989:244, note 73) independently notes the existence of overt partial A-movement examples like (49b); her example is *there was a man arrested.* Note that "heaviness" seems to be something of a factor here: ?? *there was arrested a man* vs. *there were arrested several men.*

(48) a. There have only ever been given thirteen of those awards
b. There have only ever been thirteen of those awards given

(49) a. (?) There were repainted many tables that year
b. There were many tables repainted that year

(50) a. There were believed to be four men in the garden
b. There were four men believed to be in the garden

Why should it be that English has the option for this partial movement of the associate, presumably to Spec-*v*, while the same movement is prohibited in French? After all, in environments similar but for having overt NP-movement rather than an expletive-associate relation, it seems that both languages do allow overt movement through Spec-*v*; consider the position of the Floating Quantifiers in (51) and (52):

(51) <u>The tables</u> were <u>all</u> repainted

(52) <u>Les tables</u> ont été <u>toutes</u> repeintes
the tables have been all repainted

I would like to propose that the relevant variable in this alternation between optionality/impossibility of partial associate movement lies in the featural properties of the expletives themselves. Recall that Chomsky (1995) argues that French *il* and English *there* differ in their feature-composition:*il* is able to check Case and agreement (ϕ) features, whereas *there* can check neither of these features.[42] In the previous section, I made crucial use of Chomsky's (1995) claim about the feature-properties of *il* (in conjunction with the Shortest Derivation Condition) to account for the fact that overt associate movement to Spec-*v*, and therefore participle agreement with the associate, is impossible in French, in constructions with both regular and *wh*-phrase associates. Building on this, I think it is also possible to link the different feature-wise nature of the expletive *there* with the optionality exhibited in (48)-(50). The intuition which I am aiming to implement in what follows is simply this: in contrast to the case of French *il*, the economic benefit of

[42] The evidence concerned Binding properties of the DPs associated with these expletives - see above.

merging *there* in Spec-*v* instead of associate will be offset by the fact that *there*, unlike the associate, is incapable of checking anything but the strong feature of *v*, hence necessitating a further operation to check the features which remain.

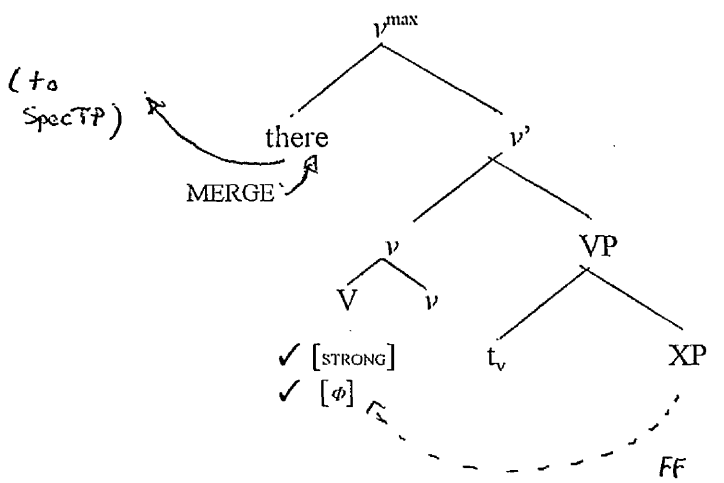
In addition to assuming that *there* is incapable of checking anything other than categorial features, let us also assume that in English, as in French, passive verb-forms are associated with a *v* bearing a strong feature[43] (cf. Branigan 1992, see above). Given these assumptions, along with derivational economy principles which are presumably universal, I suggest the following account of why *both* partial A-movement of the associate, as in (48b), (49b), (50b), *and* associate *in situ*, as in (48a), (49a), (50a) are permitted in English. Consider the derivation/s for (49): in all cases, a substructure is created by merging the passive participle *repainted* with its internal argument *many tables*, followed by the addition of the light verb *v* and subsequent adjunction of the participle to this element. By (above) hypothesis, the resulting complex head contains a strong feature, plus ϕ-features associated with the participle; recall that such features on a verbal element, unlike those on nominals, are uninterpretable and must be eliminated at some (in fact, any) time during the derivation (see Chapter 2). The strong feature, on the other hand, must be eliminated immediately it enters the structure. Given the presence of expletive *there* in the numeration, a derivational choice arises: either *there* itself checks the strong feature, by Merge - as illustrated in (53a) "derivation 1" below. Recall that unlike movement, Merge is a cost-free operation; however, in this case, it is also an operation which fails to achieve checking of all features at that site, since *there* is not capable of checking ϕ-features; the latter will then have to be checked by a later movement of the appropriate features of the associate *many tables*. That derivation obviously corresponds to the option of (49) where the associate is *in situ*.

The alternative derivation is as follows: the associate itself moves to check the strong feature of *v*, as illustrated in (53b) "derivation 2". This involves *one* operation of

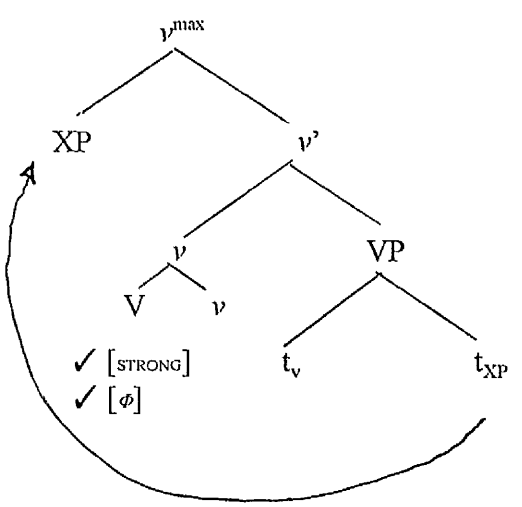
[43] This is of course a stipulation (see note 35 above). I am not convinced that there *is* necessarily a light verb *v* in this construction, but am using *v* to represent at least some functional head whose exact nature I will not dwell upon here.

Move, yet the associate, unlike *there*, is capable of checking not only the strong feature, but crucially also the ϕ-features: an operation of Move, though more costly than an operation of Merge, results in checking of all the features associated with the verbal complex.

(53) a. <u>Derivation 1 (cf. (49a)): optimal for Procrastinate, two applications of Move</u>



b. <u>Derivation 2 (cf. (49b)): nonoptimal for Procrastinate; 1 application of Move</u>



Why are (53a) and (53b) both optimal with respect to derivational economy conditions? In (53b), movement of the associate to check the strong feature of v violates

Procrastinate; more accurately, (53a), where the strong feature is checked by cost-free Merge of *there*, with later checking of Case and ϕ-features by the formal features of the associate, is preferable by Procrastinate. At the same time, however, (53a) must entail more operations than (53b), since the expletive must then *move* up to Spec-TP, whereas in (53b), in which the expletive has not been Merged in Spec-v, it simply Merges in Spec-TP. As a result, this derivation is optimal by SD. Each of the derivations is then optimal, so that each is permitted, bringing about what we think of as the "optionality" between them. This optionality hangs crucially on the feature properties of the expletive *there*: since this element can check D(P) features, but not Case and ϕ-features, Merging it in Spec-v fails to eliminate the need for an application of move to check these features against those of the associate. Compare this with the case of French discussed earlier, in which recall that associate *in situ* is the only option. Unlike *there*, *il* is able to check categorial, Case and ϕ-features, so that merging it in Spec-v is not offset by the necessity of further raising from the associate.[44]

While it is recognized that the above account is rather speculative and not fully

---

[44] In note 36 above, it was mentioned that my account of these (non)optionalities of overt associate movement in French and English are based on Chomsky's (1994) account of the following data:

(i) There seems to be a man in the room
(ii) * There seems a man to be in the room

The reader may notice an apparent incompatibility between Chomsky's account and my own. Chomsky assumes that (i) *blocks* (ii) as follows: Merge of *there* in the embedded Spec-TP is cheaper than moving *a man* to that position; the latter violates Procrastinate. Since my account of the English cases (48)-(50) seem to involve essentially the same situation, why is associate movement optional in one case, but not in the other? In fact, there is a simple solution, since there *is* a difference between (i)/(ii) and the optionality cases discussed in the text, and that is that in the former, the site in which we need it to be cheaper to Merge *there*, arguably has no Case or ϕ- features, but only a D(P) feature, since it is a tenseless clause. Hence - unlike the v/V complex in the text cases, which will contain ϕ- features from V - once the expletive has merged, no further covert movement there involving features of the associate will be necessary: as in the French case, though for different reasons, merging the expletive in the lower position is unequivocally the cheaper option.

worked out, note that Swedish provides support for the proposed link between expletive type and optionality of associate movement to Spec-*v*. This language (and apparently others of Mainland Scandinavia, apart from Norwegian) has what looks to be a similar kind of optionality to English in passive/expletive constructions - see (54) (from Holmberg 1994):[45]

(54) a. Det blev skrivet tre böcker
it was written three books(pl.)

b. Det blev tre böcker skrivna
it was three books(pl.) written(pl)
'three books were written/there were three books written

It is extremely interesting to note that with respect to Binding and control, the expletive *det* in Swedish (and likewise the expletive *där* 'there') seems to pattern with English *there* rather than French *il*, as reported by Chomsky 1995 (p. 384, note 46, citing Cardinaletti 1994). The relevant data is in (55); it should be noted that in general, Swedish does not have phonologically overt agreement with subjects, accounting for the superficial difference with English, in which the verb shows subject agreement with the associate of *there*).

(55) a. Tre män kom utan att identifiera sigsjälva
three men arrived without to identify themselves
'three men arrived without identifying themselves'

b. Det kom tre män utan att identifiera sigsjälva
it arrived three men without to identify themselves
'there arrived three men without identifying themselves'

The fact that Swedish *det* constructions pattern with English *there* constructions and against French *il* constructions with respect to both the existence of optional partial movement of the associate and the possibility of Binding by the associate provides some initial support for the view that there is some link between the two properties, most

[45] Thanks to Angelica Fransman for help with the Swedish data.

plausibly stemming from the feature properties of the expletives themselves.[46]

Turning lastly to a potential problem with the above set of analyses, concerning locality/globality of derivational economy. It seems that my analysis of French nonoptionality and English/Swedish optionality of associate movement in expletive constructions rely on incompatible assumptions about how "local" derivational economy must be. For the French case, it seemed necessary to assume that derivational economy applies "locally", in the sense of Chomsky (1995), i.e. no more than one step at a time is taken into account. If, contrary to this, the Shortest Derivation condition was to consider the *whole* of the two derivations in the case in question, then apparently they would measure the same with respect to SD: in the derivation in which the Strong feature is checked by Merge of *il* rather than movement of the associate - claimed here to be the preferred derivation - the saving incurred by this would in fact be a false economy, since the expletive would subsequently have to move up to Spec-TP - a move which would be unnecessary in the alternative derivation where the associate checks the strong features. It did not seem necessary to make this "locality" assumption in connection with my account of the optionality cases. Hence, there seems to be some inconsistency. Although much further thought is needed to make this issue precise and determine its relevance (if any), it might be worth considering using Procrastinate instead of SD to account for why the derivation of involving Merge of *il* (leading to default agreement on the participle) is preferred to overt movement of the associate. If Merge operations are not relevant for economy at all, it could just be said that overt associate movement is disfavoured because it violates Procrastinate (checking non-strongϕ-features overtly). On this account, the difference between the two derivations as to whether *il* has to move to Spec-TP or not would not be relevant. I hope to take up this question in further research.

While I have concentrated my efforts on seeing whether the "theory of optionality"

---

[46] One alternative approach to the cross-linguistic variation in partial associate movement might be to tie it to short movement (or lack of it) of the participle itself (cf. Pollock 1989). However, it would be difficult to link the differing behaviour of the participles with the feature-properties of the relevant expletives.

I have outlined can be made to work technically, there are obviously much wider issues raised by such a theory, not least concerning the notion of "multiple optimality" of derivations. This is returned to and addressed preliminarily in the next chapter (§ 2.1.3).

6 Summary and conclusion

In this chapter, I have demonstrated how syntactic optionality can arise in theory and in practice, and characterized some optionality and related non-optionality phenomena in this way. In particular, I outlined a Minimalist account of French agreement optionality in the case of *wh*-objects and Accusative clitics, and showed how this system fits in with further data from French in which agreement is not optional, also extending this system to the case of optional partial associate movement in English and Swedish. An attractive feature of the account proposed here is that it derives overt/covert optionality using the same means as one standardly derives fixed overt/covert effects, viz., the properties associated with functional categories, in conjuction with general economy principles. This account then allows for variation in optionality to be linked to properties of features. Cf. the widely-adopted lexical/functional parameterization hypothesis which holds that all crosslinguistic variation must be characterized in such terms (Borer 1984, Chomsky 1991).

The next chapter takes up a few of the questions raised by this account: can it be extended to further cases of optionality, in particular, those involving Object Shift? I show how this can be done straightforwardly for the cases of optional overt Object Shift in Icelandic, Dutch and German, with the assistance of a few independently motivated assumptions. I also consider how the account compares with existing alternatives within the Minimalist framework, e.g. Branigan's (1992) optional strong feature account. And how, if at all, does a syntactic account of optionality allow us to deal with the interpretive restrictions which arise in some of the cases I have discussed?

Chapter 6

Extensions

1 Introduction

Building on the proposal developed in Chapter 5, the present chapter extends the feature-strength-plus-economy conditions approach to a range of further object-related (non)optionality facts, from Icelandic, German and Dutch. I also compare my approach to optionality with three comparable accounts from the recent Minimalist literature. Finally, I address the important issue of how to account for the "interpretive effects" which arise in the Germanic and in the French *wh*-movement cases.

The chapter is structured as follows. In the next section (2), I give a characterization of optional overt Object Shift in Icelandic in terms of an anti-Procrastinate effect, following my analysis of French agreement optionality. For Icelandic, the optionality of Object Shift is crucially linked to obligatory overt verb-raising to check a strong feature of Tense (cf. Holmberg's (1986) Generalization). The same account is also applied to the case of short object scrambling in German and Dutch, which I shall also take to be a case of A-movement, following recent proposals in the literature (see below

for details). I further show why in Icelandic, optionality is precluded when the verb is a participle, but why in German and Dutch, optionality persists in such environments, suggesting that in the latter languages but not the former, participles do in fact move overtly out of the VP. In my account of the German and Dutch object facts, I make the assumption that both finite verbs and participles move overtly out of the VP, to the right (contra Kayne 1993; cf. Sabel 1996). This assumption is supported with evidence involving adverb placement, among other things. In § 3, I review three existing Minimalist proposals on optionality - Kitahara 1994, Branigan 1992 and Adger 1994 - and compare these critically with my own. In the last major section, § 4, I investigate the possibilities for dealing with the interpretive effects which are well known to be associated with both the French and the Germanic optionality data (see Obenauer 1992 on the former, and e.g. Diesing 1992 for the latter). I argue that my account of the phenomena in question in terms of syntactic optionality (overt versus covert movement) does not preclude an account of these systematic effects; in fact it provides a promising basis for an account in terms of post-LF mappings/processes, since the LF representations of the options are in fact distinct from one another.

2 Object Shift (non)optionalities in SVO and SOV Germanic

In this section, I consider a range of further cases of object-related optionality - and associated non-optionality facts - from Icelandic, German and Dutch. These cases are an interesting challenge, since, unlike in the French case studied in Chapter 5, the optionality seemingly involves a "whole movement" rather than a subpart of one.

2.1 Icelandic

2.1.1 Data and some background assumptions

As is well-known, Icelandic has the option for overt Object Shift provided that certain

conditions obtain, that is, if the main verb raises overtly - as in (1).[1, 2] If the main verb does not raise overtly, as in (2), the option for overt Object Shift disappears:

(1) a. Jón las bækurnar ekki
J. read the-books not
b. Jón las ekki bækurnar — overt OS optional
J. read not the-books

'Jón did not read the books'

(2) a. Hann hefur ekki lesið bækurnar
He has not read the-books
b. * Hann hefur bækurnar ekki lesið — overt OS impossible
He has the-books not read

'He has not read the books'

In this section, I shall attempt to link the *optionality* of overt Object Shift (as opposed to the possibility thereof, i.e. traditional Holmberg's (1986) Generalization) to overt V to T raising. Recall that for French, I claimed that the situation which "triggers" optionality between overt and covert Object Shift is the object itself undergoing overt A'-movement to check a strong feature associated with the operator-related head (either C, or the head of a Clitic projection). But in Icelandic, the situation seems different. Descriptively, as in the French case, there *is* a clear link between an operation of forced overt checking of a feature on a higher head. However, in Icelandic, it is a question of overt movement involving the *verb*, and presumably the strong feature at issue belongs to the head T. As in French, there is clearly a link between optionality and movement to check a higher strong feature, but unlike in French, the strong feature does not relate to

---

[1] It is reported in Bobaljik 1995 that the optional overt Object Shift in Icelandic has an associated interpretive effect analogous to that discussed for the French agreement data by Obenauer 1992, for German object scrambling by e.g. Diesing 1992, etc. This topic is treated in a later section dedicated to the discussion of these interpretive effects (§ 4).

[2] In Icelandic (and other Scandinavian languages), object *pronouns* actually move obligatorily when the Verb raises. I shall not address this here.

the object itself. The main question is, then, how to extend the French account in a way which implicates overt verb movement in overt/covert optionality relating to the object? In fact, it is possible to do this without compromising the account of French at all.

Let us begin by assuming, following e.g. Kitahara 1994, that (1) manifests optionality between overt and covert A-movement of the object - presumably to (Spec-)*v* - and the concomitant assumption that there is no strong D(P)-feature associated with *v* in Icelandic.[3] As noted above, optionality between overt and covert movement is the only possible type of optionality which can be assumed to exist in the computational system, given Last Resort. In this respect, then, the Icelandic case is characterized in the same terms as the case of agreement optionality in French *wh*-movement/clitic constructions investigated in chapter 5. I make the crucial further assumption that there is no strong V-feature associated with *v*; the "real" verb is not forced to raise to *v* overtly (though it must of course raise overtly to Tense - see shortly below). As evidence for the assumption that *v* contains no strong V-feature, note the fact that participles in Icelandic seemingly cannot move overtly out of VP, under the usual assumption that the adverbials *ekki* 'not' and *vandlega* 'carefully' mark its lefthand border (here, the features of Tense are clearly checked by the auxiliary *hafa* 'have'):

(3) a. * Margir stúdentar hafa lesið ekki t_v bókina
many students have read not the-book
b. * Margir stúdentar hafa lesið vandlega t_v bókina
many students have read carefully the-book

Let us now see how the above assumptions about feature-strength, in association as usual with the economy conditions Procrastinate and Shortest Derivation, account firstly for the optionality of overt Object Shift in cases exemplified by (1), and secondly, for the obligatory covertness of Object Shift with non-finite verbs, as illustrated in (2).

[3] If there was, Object Shift would of course be obligatorily overt.

2.1.2 Derivations and anti-Procrastinate effects

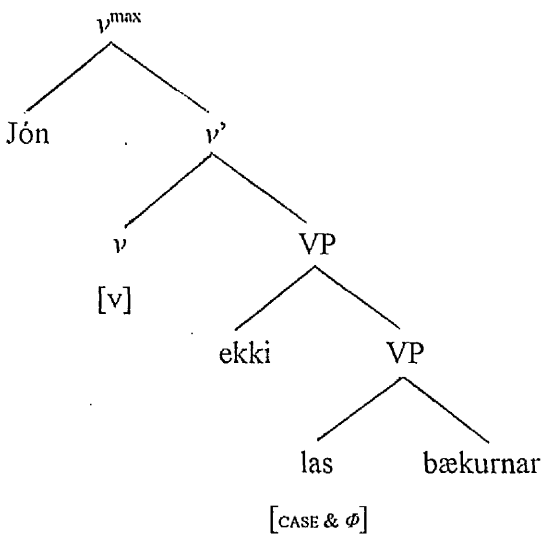
Firstly, let us consider the (relevant part of the) derivation/s of (1), repeated here:

(1) a. Jón las bækurnar ekki
J. read the-books not
b. Jón las ekki bækurnar — overt OS optional
J. read not the-books

'Jón did not read the books'

I assume of course that (1a) and (1b) derive from one and the same numeration. In all cases, the first step is the creation of a structure consisting of *las* 'read' and its internal argument *bækurnar* 'the books'. Let us then assume that the adverbial *ekki* adjoins to the VP, with the resulting VP then merged with the light verb *v*, and the subject *Jón*. At this stage, we have a structure as in (4):

(4)



As I have assumed, there is no strong D(P)-feature associated with *v*, and moreover, no strong V-feature. Hence, all features in *v*, viz. Case and agreement, and whatever features relate to the verb, will remain unchecked at this stage of the derivation,

with the verb and object themselves remaining within their VP. In the next stage of the derivation, a Tense element is combined with the existing structure. Tense contains a strong V-feature which must then be checked by immediate movement of the verb. It is at exactly this point that two convergent continuations of the derivation become possible with respect to the movement of the object.[4] I describe them first, and then explain why both qualify as optimal by derivational economy.

In one possible derivation, corresponding to overt Object Shift (1a), I assume that the verb checks *all* features relevant to it - that is, the strong V-feature of T, and further non-strong features associated with T (say tense) and with *v* (possibly aspectual) - via an application of Form Chain, instantaneously creating a chain with links at T, *v* and V. The object then raises overtly to Spec-*v*,[5] checking Case and ϕ features, which will be there to be checked since the "real" verb which possesses them has left a copy in this site. This two-stepped derivation is pictured in (5):[6]

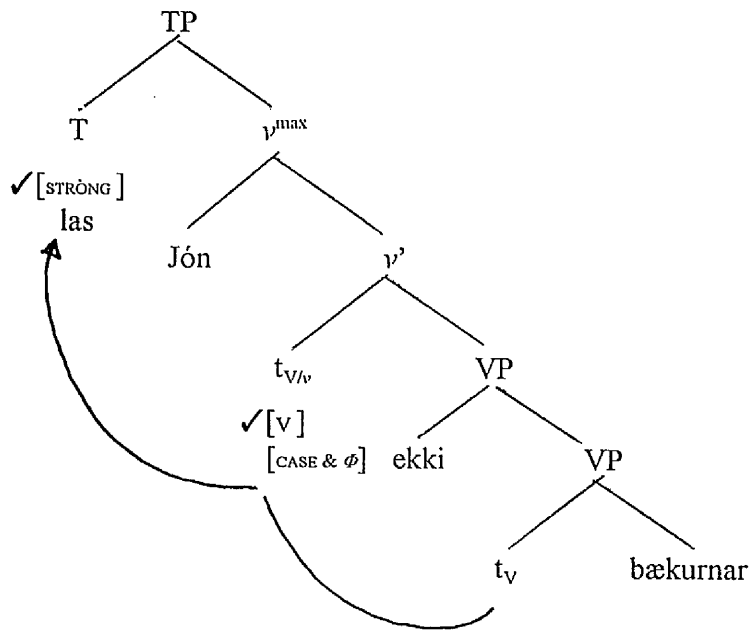
---

[4] Cf. the merging of strong-featured C in the French optionality examples of Chapter 5.

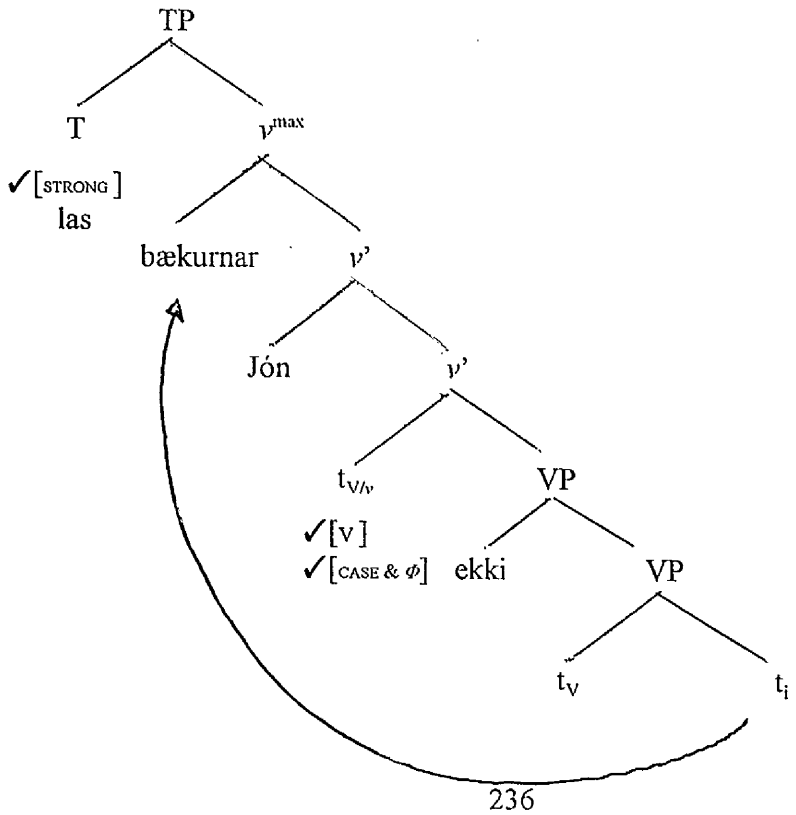
[5] This proposed step obviously appears consitute a violation of the Extension Condition; but see § 2.1.3 below.

[6] As Ian Roberts has pointed out to me (p.c.), the question arises as to why the strong feature of T cannot be checked by *v* rather than by V, the former being closer to T (it is clear that *v* alone cannot satisfy the strong feature, given the real verb's overt position). The obvious answer would be that *v*, contrary to what its name perhaps implies, is not itself of the category V, but rather a functional head encoding a distinct semantic property, e.g. Aspect (see Chapter 2, § 5).

(5) a. Overt Object Shift derivation: Step 1

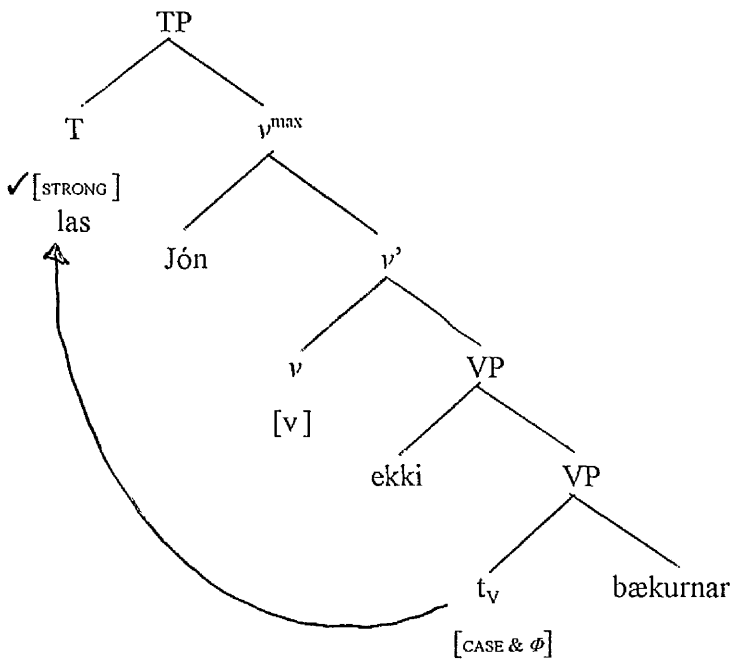


b. Overt Object Shift derivation: Step 2



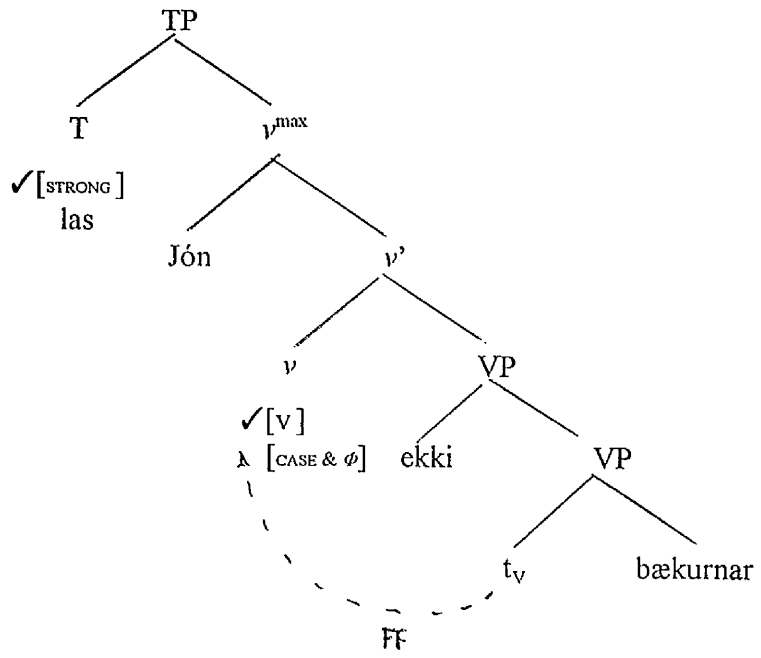
The alternative derivation, corresponding to covert Object Shift (1b), is as follows. Suppose that at the point when T is merged in the structure, requiring overt checking of its strong feature, the verb moves there *directly*, bypassing features relevant to it in *v*; remember that these particular features are not strong and therefore need not be checked overtly for convergence. This step is illustrated in (6a).[7] Next, these remaining non-strong V-features in *v* are checked by covert raising of the formal features of the VP-internal *copy* of the verb - (6b). Among these formal features of the verb are also Case and agreement features. These are checked by raising of the features of the object - (6c); note here that movement *to* a copy is assumed to take place (cf. Chapter 2, § 5):

(6) a. Covert Object Shift derivation: Step 1

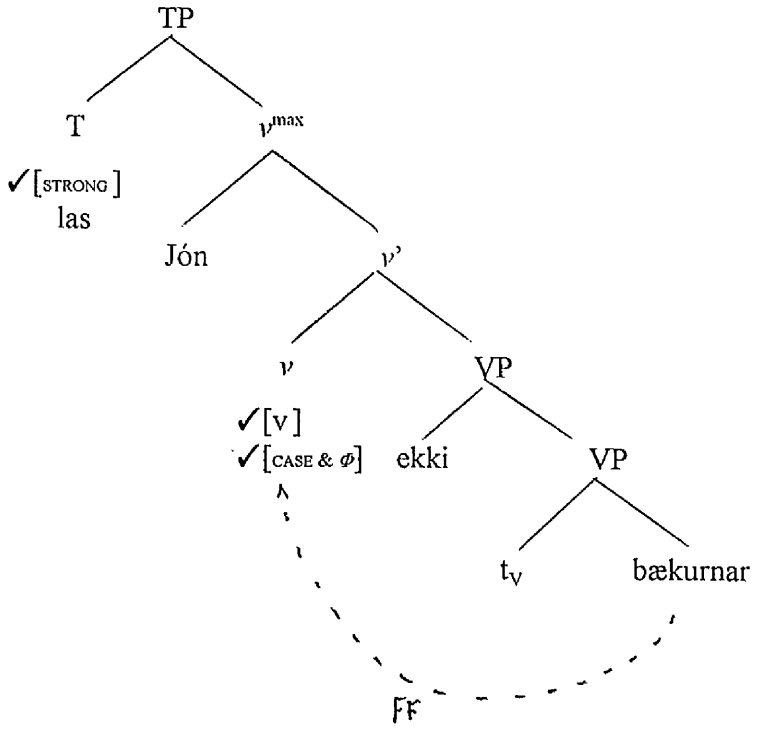


[7] Since this proposed movement skips *v*, one might ask if it ought to violate the Minimal Link Condition. But if we assume, as in note 6 above, that *v* itself does not have suitable features to check with those of T, then direct movement of V to T should not be blocked (recall that according to Chomsky (1995:297), "K [a target for movement] attracts F if F is the closest feature *that can enter into a checking relation with a sublabel of K*" [my emphasis - AMP]).

(6) b. Covert Object Shift derivation: Step 2



(6) c. Covert Object Shift derivation: Step 3



Such is my characterization of the derivations corresponding to overt and covert Object Shift options of (1). Why are both allowed by derivational economy? I suggest that the answer is essentially the same as the French optionality cases analysed in Chapter 5; in short, we have a case of "multiple optimality" - the derivation with covert Object Shift is preferred by Procrastinate, while the one with overt Object Shift counts as the Shortest Derivation. Let us elaborate on this. Consider the derivation in (6), corresponding to covert Object Shift (1b): here it is the case that *no* non-strong feature - relevantly, the DP and V-related features in the verbal complex - gets checked by overt movement. Hence naturally this derivation is optimal by Procrastinate: compare it with the derivation in (5), corresponding to overt Object Shift. Here, there are two instances of overt movement to check ordinary features, namely, the overt Object Shift itself, plus the overt checking of V-features in *v* by the verb, constituting part of its overt movement up to T.

At the same time, while the derivation in (5) is not preferred by Procrastinate, it is easy to see that it will count as a shorter - in fact, the shortest - derivation, and hence count as optimal for the Shortest Derivation condition. This is so since in (5), two operations are involved to check the features relating to the verb and the object, namely the operation on V which creates a chain linking it with *v* and T, and the operation which moves the object to Spec-*v*. Compare this against (6); here, no less than *three* operations are necessary to check these same features. This is crucially due to the fact that V has skipped relevant weak V-features in *v* *en route* to T, meaning that the operation to raise the formal features of V's copy necessarily counts as a different operation from that which raised V to T, since the two are "separated" by Spell-Out. In sum, derivations (5) and (6), corresponding respectively to (1a) overt Object Shift and (1b) covert Object Shift, are *each* optimal with respect to derivational economy, the former with respect to SD, and the latter, Procrastinate. I thus assume that both are permitted.[8] Having set out the account, I now briefly address some questions which it raises.

---

[8] Collins and Thráinsson 1994 suggest that in Icelandic, "the N features of AGR_o are optionally Strong. If they are Strong, object shift is obligatory, and if they are weak, no object shift is possible." See Section 3 for a critique of optional feature-strength theories of optionality.

2.1.3 "Modular economy" and other issues

In the first place, it may well be objected that the derivation in (5) involves a violation of the Extension Condition (Strict Cycle), since clearly it contains an overt movement into into an *embedded* position in the structure, i.e. the overt Object Shift. However, this should not necessarily be a problem, under the usual (and in Minimalism, obligatory) assumption that the cyclicity of overt movement is not itself ensured by a special constraint, but derives from more general considerations. Recall that the classic cases which cyclicity is needed to deal with all involve movement to check an embedded *strong* feature.[9] We could assume, following Chomsky 1995 (p.233) that the required cyclicity of overt movement is merely a reflection of the fact that *strong* features must be eliminated (i.e. checked) before any further structure-building occurs. Under this view of cyclicity, the overt Object Shift in (5) is not countercyclic in any meaningful way, because it does not involve checking any strong feature.

A further question we need to consider in connection with the above account of optionality concerns two other derivations which seem to be possible with respect to (1). Why exactly are these ruled out? Firstly, suppose that the derivation has proceeded to the stage shown in (6a), i.e. the verb has moved directly to T, not via *v*. The question is: why, instead of the continuation shown in (6b) and (6c), can the object not move overtly? In fact, I take it that this imaginary continuation of the derivation is not viable, for the following reason. It is plausible to assume that it is the "real" verb, rather than *v*, which carries at least ϕ-features (though probably not Case features).[10] Since, by hypothesis, the "real" verb has moved directly to T, leaving no trace in *v*, not all the features relevant to the object will actually be in *v* "in time" for the object to check them overtly. The object might move overtly in order to check the Case features, but a further covert operation would still be necessary to check the ϕ-features, when the relevant features of the copy of V have raised to *v*. But this derivation will be neither optimal by Procrastinate (since it involves overt Object Shift rather than covert Object Shift as in (6) itself), nor optimal

[9] See Chapter 1, § 4.2.3.

[10] But see Ferguson and Groat 1995.

by Shortest Derivation (since it takes *four* operations to check the relevant features). In other words, overt movement of the object to Spec-*v* will be prevented outright by Last Resort, since there are no suitable features for the object to check.[11]

More interestingly, the other potential derivation to consider with respect to (1) would be like (5), except that the object shifts *covertly*. To keep track of things, let us call this hypothetical derivation (5'). Notice that (5') does not differ from (5) in terms of the number of operations involved - each has only two (in relation to the relevant features), and so presumably both are optimal for the SD condition. However, notice that we might expect *Procrastinate* to prefer (5') to (5), since in (5'), object shift is covert. Does this mean that (5) should be ruled out, counterfactually predicting no option for overt Object Shift?

In order to prevent this unwanted prediction, we evidently need to make an assumption along the following lines:

(7) **"Modular Economy"**:

Each economy principle in the grammar recognizes only the "optimal" derivation/s with respect to *itself*.

That is to say, each condition recognizes its own optimal derivation, or possibly derivations, but in the event that more than one of these exists, the condition in question will not then differentiate between them on grounds of their virtues with respect to different conditions. For example, in our case (1), the grammar gives us (i) the optimal derivation for Procrastinate, which turns out to be the derivation in (6), and (ii) the

[11] Alternatively, one might assume that *v* itself had a D(P) feature, so that the object could in fact move to Spec-*v* as far as Last Resort was concerned. However, in this case, a further movement would later be needed, in order for the Case and ϕ-features to be checked. This further movement will have to happen after the relevant features of V's copy have raised to *v*. Thus, this derivation will still not be permitted: it is neither optimal by Procrastinate, having overt Object Shift rather than covert Object Shift as in (6) itself; nor optimal by Shortest Derivation, having *four* operations.

optimal derivation for SD, which turns out to be derivations (5) and (5'). What I think the grammar will *not* recognize is the concept of "optimal optimal derivations", e.g. it will not look at (5) and (5') in their capacity as the optimal derivations for SD, and then further differentiate between the two on the basis of other dimensions of economy.

I think that (7) is a fairly reasonable assumption about how economy conditions work. But does it exclude the possibility that one numeration can yield two optimal derivations, the central assumption in my account of optionality? It does not, since all that the account really entails is that to be permitted, a derivation has to be optimal by at least one economy condition.[12] This does not exclude the possibility of two derivations independently being optimal by different conditions (i.e. optionality), or of course the possibility that a single derivation turns out to be optimal with respect to more than one condition (i.e. nonoptionality).

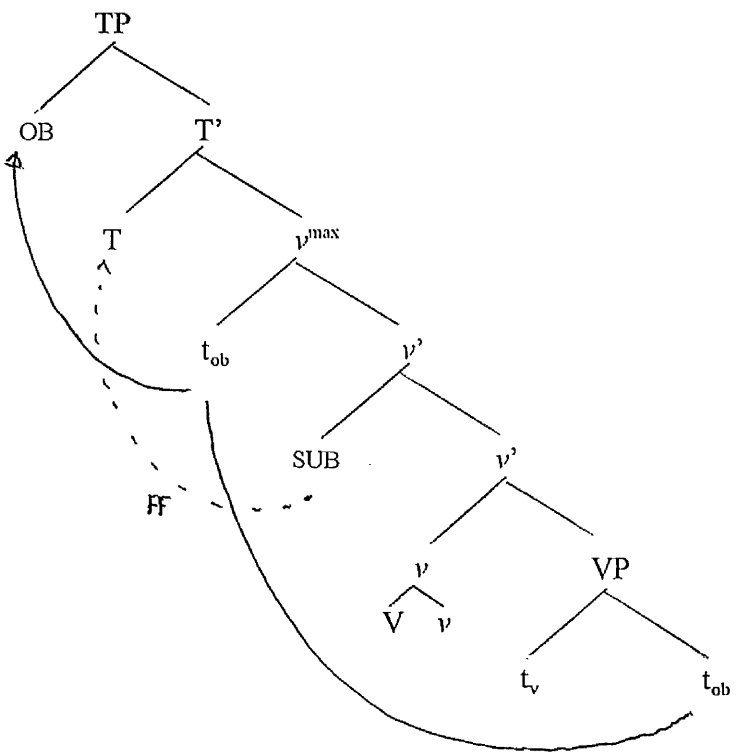
One final issue which I would like to address in connection with my account of the optionality in (1) (and indeed of the proposal on French in Chapter 5) has to do with Chomsky's (1995) use of the Shortest Derivation condition, which appears to be incompatible with my own. To see the problem, consider the argument which Chomsky (*op. cit.*, p.357-358) offers on behalf of an economy condition concerning derivation length, i.e. SD. This argument, virtually the only one he gives, concerns languages in which overt Object Shift is possible, and which also have the EPP - that is, Icelandic itself, among others. Chomsky points out that in the 1995 Minimalist framework, locality (MLC) does not prevent a derivation in which an object, once overtly shifted to Spec-*v*, subsequently moves up to Spec-TP to check the EPP feature, with Case and agreement features then checked by raising of the formal features of the subject.

Chomsky's argument is that this obviously unwanted derivation is excluded by the Shortest Derivation condition, since more steps are involved to check the same features

[12] It may be a shortcoming of my account that, with respect to Icelandic, the system redundantly yields two possible derivations with the same result, i.e. (6) and (5') for covert Object Shift.

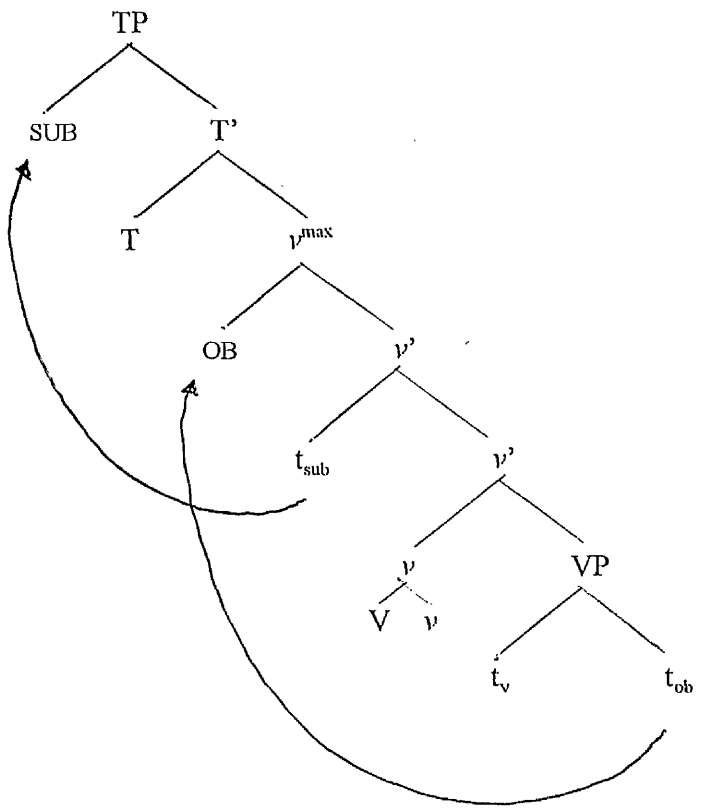
than an alternative derivation in which the object moves to Spec-v, then the subject, to Spec-TP, as illustrated below:

(8) <u>The unwanted derivation</u>



"...this derivation is blocked by economy conditions. It involves three raising operations...."

(9) <u>The correct derivation</u>



"...and two would suffice for convergence" - Chomsky 1995:357)

The basic idea here is that in the unwanted derivation - (8) - there is an operation, namely movement of the object to Spec-TP (allowed by Last Resort, since the object can check the "EPP" feature of T), which fails unnecessarily to get a feature-checking result; the object has already used up its Case feature, which means that an extra, separate operation, namely movement of the subject's formal features to T, is then required for convergence. This extra operation is obviously not necessary if the object stays where it is in Spec-v, letting the subject, which does have a Case feature to spare, move up to Spec-TP to check the strong EPP feature, as in (9). As (9) has two steps of movement in comparison to (8)'s *three*, Chomsky claims that (9) blocks (8) by SD.

The problem which arises for my own account in this connection is that it seems

to predict that the correct derivation and the unwanted derivation will involve *the same* number of steps, i.e. two. The reason for this is that, under my assumptions, the two overt movements undertaken by the object could, and would, count as a *single* operation of Form Chain. This is clearly the wrong result.[13] However, it does not seem unrealistic to suspect that the unwanted derivation (8) might be ruled out by considerations other than the Shortest Derivation. For example, in their discussion of the same case from Chomsky 1995, Johnson and Lappin 1996 suggest that movement of the object to Spec-TP could be ruled out on Case grounds. I shall not pursue this matter any further here, but take it that the problem is not insurmountable. Having set out an account of the optionality of overt Object Shift in Icelandic, let us now extend it to situations where Object Shift is obligatorily covert.[14]

2.1.4 When overt Object Shift is not an option in Icelandic: verbs *in situ*

Recall that in Icelandic, the overt/covert optionality for Object Shift disappears when the verb is nonfinite. The relevant data is repeated here:

(2) a. Hann hefur ekki lesið bækurnar
He has not read the-books

---

[13] Chomsky himself seems to handle the optionality of overt/covert Object Shift in Icelandic by means of the option for a strong feature in *v*. On this assumption, the strong feature of *v* must be checked *immediately* - it cannot remain unchecked in the tree until the next strong feature is added. Hence, the two strong features could not be checked by a simultaneous Form Chain operation. But the point is that my own account has no concept of "optional feature-strength", and relies on the assumption that (any case of) optional overt/covert movement must involve a non-strong feature.

[14] A serious potential problem which I have not addressed in this section is the following: why is there no option for overt Object Shift of non-*wh*- objects in French, given my analysis of Icelandic? I discuss this point in § 2.2.4 below, after dealing with a further range of Germanic optionality data.

b. * Hann hefur bækurnar ekki lesið
He has the books not read

'He has not read the books'

In the account of (1) proposed above, it was ultimately the necessity of overt verb movement to T which brought about overt/covert optionality for Object Shift. In Icelandic, nonfinite verbs, or at least participles, seem not to move overtly, as (10) (= (3) above) shows:

(10) a. * Margir stúdentar hafa lesið ekki bókina
many students have read not the-book
b. * Margir stúdentar hafa lesið vandlega bókina
many students have read carefully the-book

That is, as I have already explicitly assumed, there is no strong V-feature associated with *v* in Icelandic; in cases like (2) and (10), the features of Tense must be checked by the auxiliary verb. Presumably, the nonfinite, "real" verb only has to check features with one functional head, *v*. I suggest that this is the key to explaining why overt/covert optionality with respect to the object disappears in this context, with only covert Object Shift being allowed. Overt checking of the V-feature of *v*, which violates Procrastinate, does *not* - in contrast to the case of (1), incur any advantage relating to the Shortest Derivation condition: a derivation with overt Object Shift, which overt V-raising theoretically makes possible, will crucially involve the same number of steps as its covert Object Shift counterpart. Meanwhile, Procrastinate of course does differentiate between the two, preferring the derivation with covert V-raising and covert Object Shift.[15]

Let us sum up the proposal so far. I have extended the system introduced in Chapter 5 to cover the case of (non)optionality of overt Object Shift in Icelandic. In the next subsection, I address optional overt Object Shift in two other Germanic, but "SOV",

[15] This account is more or less parallel to that of the obligatory covert movement of French in non-operator environments discussed in chapter 5, § 4.1. On the important question of why overt Verb-movement to T is insufficent to license optional overt Object Shift in French, see § 4 below.

languages: German and Dutch. In the context of the foregoing account of Icelandic, these languages are interesting, since they appear to be straightforward counterexamples: while there appears to be no overt verb movement (except in matrix clauses, for independent reasons), there is optionality between overt and covert Object Shift. Moreover, this optionality obtains with both finite and nonfinite verbs.

2.2 Optional overt Object Shift in German and Dutch

2.2.1 Data

German and Dutch are two of many languages having so-called scrambling, i.e. the possibility for rearrangement of certain constituents - quite possibly the archetypal case of optionality. It is likely that the term *scrambling* in itself does not denote a unitary class of phenomena (see e.g. Grewendorf and Sternefeld 1990, Déprez 1989)). Here, I restrict my attention to a type of short object scrambling in German and Dutch which has been argued to be Case-related A-movement (see Déprez *op. cit.*, Mahajan 1990 for German; Vanden Wyngaerd 1989, Zwart 1993 for Dutch, see also Bobaljik 1995).[16] Short object scrambling in German and Dutch are illustrated in (11) and (12) respectively, with the (a) sentences the versions without scrambling and the (b)s, those with it. Like Icelandic *ekki*, the adverbial elements *grundsätzlich* 'always' and *gisteren* 'yesterday' are generally taken to mark the left edge of VP; note that manner-adverbs are also possible in these positions.[17]

---

[16] As in optional overt Object Shift in Icelandic, and as in optional overt Object Shift in the context of *wh*-movement in French, this short object scrambling in German and Dutch displays an interpretive alternation sometimes characterized in terms of specificity. Again, I delay all discussion of this matter until § 4 below.

[17] The German data and judgements I use in this section were provided by Lutz Marten (and thanks also to Stefan Ploch for judgements), and the Dutch, by Wilfried Meyer-Viol. In this discussion, following standard practice, I shall illustrate with embedded clauses in order to avoid the interference of the "verb second" effect which obtains in main clauses.

(11) a. ... weil Lutz grundsätzlich sein Handy verliert
... since Lutz always his mobile loses
b. ... weil Lutz sein Handy grundsätzlich verliert
... since Lutz his mobile always loses

'since Lutz always loses his mobile'

(12) a. ... dat de politie gisteren de taalkundigen verhoorde
... that the police yesterday the linguists interrogated
b. ... dat de politie de taalkundigen gisteren verhoorde
... that the police the linguists yesterday interrogated

'that the police interrogated the linguists yesterday'

It will be noticed that German and Dutch contrast with Icelandic in that the option for overt Object Shift persists in environments where the main verb is nonfinite, and appears not to raise overtly: thus compare (13) and (14) with Icelandic (2) above, repeated here:[18]

(13) a. ... weil Lutz gestern sein Handy verloren hat
... since Lutz yesterday his mobile lost has
b. ... weil Lutz sein Handy gestern verloren hat
... since Lutz his mobile yesterday lost has

'since Lutz lost his mobile yesterday'

(14) a. ...dat de politie gisteren de taalkundigen opgepakt heeft
....that the polcie yesterday the linguists arrested have
b. ...dat de politie de taalkundigen gisteren opgepakt heeft
...that the police the linguists yesterday arrested have

'...that the police arrested the linguists yesterday'

(2) a. Hann hefur ekki lesið bækurnar
He has not read the-books
b. * Hann hefur bækurnar ekki lesið
He has the-books not read

'He has not read the books'

[18] (14) is taken from Reinhart 1995 citing de Hoop 1992.

At first sight, the above data from German and Dutch appear to doubly counterexemplify the analysis proposed for the Object Shift facts of Icelandic. In the first place, optional overt Object Shift apparently arises in the absence of overt verb raising, which is unexpected given my account of Icelandic (1), and also of course a problem for any theory which links overt Object Shift with overt verb raising.[19] In the second place, there is no optionality-related contrast between finite and non-finite contexts, which is again unexpected in the light of what I proposed above.

A very obvious, though not uncontroversial, way to salvage my account of Icelandic (and also standard Holmberg's Generalization theories, although see § 2.3 below) would be simply to claim that there *is* overt verb raising out of VP in the relevant cases, but "string vacuously" to the *right*, as has indeed been argued recently for German by Sabel (1996). An approach of this type would be in line with traditional generative accounts of German and Dutch syntax, although not with recent proposals of Zwart (1993) on the latter. But is there any independent reason to believe that verbs raise overtly rightwards out of VP in any or both of these languages? In the next section, I suggest that there is - and more interestingly, that this is the case for both finite verbs and participles. This will then provide the basis for a straightforward extension of my French/Icelandic analysis to the German and Dutch optionality data in (11)-(14).

2.2.2 An old-fashioned solution: "covert" overt verb-raising

There is some reason to believe that even in embedded clauses in German and Dutch, both finite and nonfinite verbs *do* move overtly. I shall assume here that this movement is to the right and hence countenance an analysis of the Germanic SOV languages according to which they are "head-final" (at least with respect to some categories, namely V and I/T). Given the recent controversy over whether theoretical concepts such as "head final" and "rightward movement" should be permissible at all (see Kayne 1993 and

[19] I.e. Holmberg's Generalization. On the apparent problem caused by German and Dutch for theories which derive this generalization, see e.g. Van de Koot 1995, Bobaljik 1995).

subsequent related works), it is almost unnecessary to mention that the proposals made in this section are tentative.[20]

As a first step in arguing that verbs move overtly and to the right in German and Dutch, recall that it has been argued with respect to both languages that the type of short object movement at hand (see (11)-(14) above) is actually A-movement for Case reasons (presumably landing in Spec-v) - see e.g. Déprez 1989, Mahajan 1989 for German, Vanden Wyngaerd 1989, Zwart 1993 for Dutch. That is, German and Dutch seem to have optional overt Object Shift, like Icelandic. If object "scrambling" is Case-related A-movement, this tends to imply that the "unscrambled" position corresponds to the object's original, VP-internal position. If the object's unscrambled position is its base position, then the verb which follows it may either be *in situ* itself, or in some derived position to its right. In other words, the possibility at least exists that the verb moves overtly, in which case all that remains is to decide between these two options. As Koopman (1995) points out with respect to Dutch, this is not an easy task. However, I shall present a small collection of data concerning adverb placement which I think at least suggests that the movement analysis is preferable to the *in situ* analysis.[21]

Let us firstly examine the behaviour of finite verbs. Consider the position of the adverbial phrases *mit der neuen Schere* 'with the new scissors', *achtsam* 'carefully in (15) (German), and *voorzichtig* 'carefully' in (16) (Dutch):

---

[20] Although I shall assume here that V-movement is to the right in Dutch and German, which is definitely incompatible with the Kaynian approach, I do not think that it would be impossible to maintain that there is overt V-movement *leftwards* in these languages (for the purposes of upholding my own analysis of optional overt Object Shift, if nothing else, it is irrelevant in which direction the verb moves, as long as the movement is overt). To account for the final position of V, it would then be necessary to have wholesale movement of the rest of the clause to the Spec of a functional head higher than that occupied by V.

[21] See also independent arguments for rightward V-movement in German in Sabel (*op. cit.*). For arguments for overt V-movement (and related discussion of the Antisymmetry issue) in another SOV context, Japanese, see Koizumi 1995 (chapter 7).

(15) a. ... daß Lutz achtsam die Blumen mit der neuen Schere schneidet
... that L. carefully the flowers with the new scissors cuts

'... that Lutz carefully cuts the flowers with the new scissors'

b. ... daß Lutz im Gewächshaus die Blumen achtsam schneidet
... that L. in-the greenhouse the flowers carefully cuts

'... that Lutz carefully cuts the flowers in the greenhouse'

(16) ... dat de politie gisteren de taalkundigen voorzichtig verhoorde
... that the police yesterday the linguists carefully arrested

...'that the police carefully arrested the linguists yesterday'

My crucial assumption is that the adverbial elements *achtsam* 'carefully' in (15a), *im Gewächshaus* 'in the greenhouse' in (15b), and *gisteren* 'yesterday' in (16) are VP adjuncts. This is a standard assumption in the case of *gisteren*, and plausible with respect to the manner and place adverbials. The implication is that the objects *die Blumen* 'the flowers' in (15), and *de taalkundigen* 'the linguists' in (16) have not moved leftwards out of the VP. With this in mind, notice that it is perfectly possible for further such adverbials - *mit der neuen Schere* 'with the new scissors', *voorzichtig* 'carefully' - to appear in between the objects and the verbs.[22] Assuming that these adverbials are also VP-adjoined and so delineate its righthand border, it is clear that the only elements within the VP in the relevant examples are the objects: the verbs themselves must have moved out - to the right, and presumably to an inflectional functional head which I take to be Tense. The situation is illustrated abstractly in (17):

(17) $[_{TP}$ subject $[_{VP}$ adv. $[_{VP}$ Object t_V $]_{VP}$ adv. $]_{VP}$ ... $]_{T'}$ T/V $]_{TP}$

At this stage it becomes possible to extend the "anti-Procrastinate" account proposed for optional overt/covert Object Shift in Icelandic to the cases of similar

[22] I am assuming here that adverbs are not generated within a VP shell structure as in Larson 1988 and recently suggested by Chomsky 1995.

optionality in finite verb contexts in German and Dutch. The data is repeated here:

(11) a. ... weil Lutz grundsätzlich sein Handy verliert
... since Lutz always his mobile loses
b. ... weil Lutz sein Handy grundsätzlich verliert
... since Lutz his mobile always loses

'... since Lutz always loses his mobile'

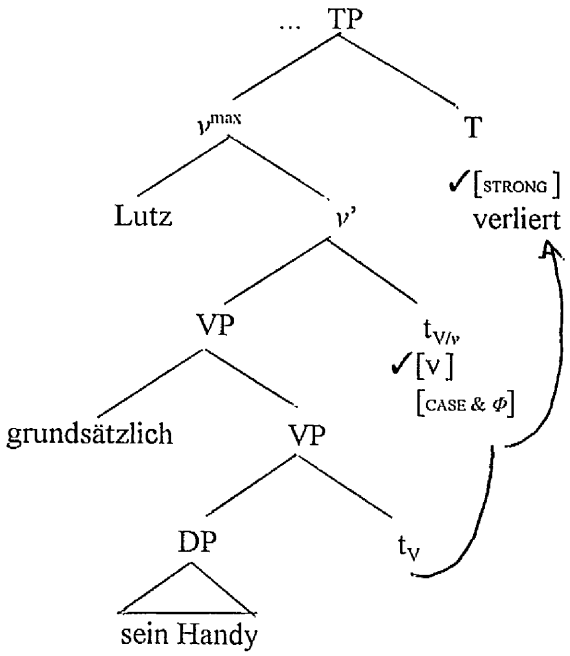
(12) a. ... dat de politie gisteren de taalkundigen verhoorde
... that the police yesterday the linguists interrogated
b. ... dat de politie de taalkundigen gisteren verhoorde
... that the police the linguists yesterday interrogated

'... that the police interrogated the linguists yesterday'

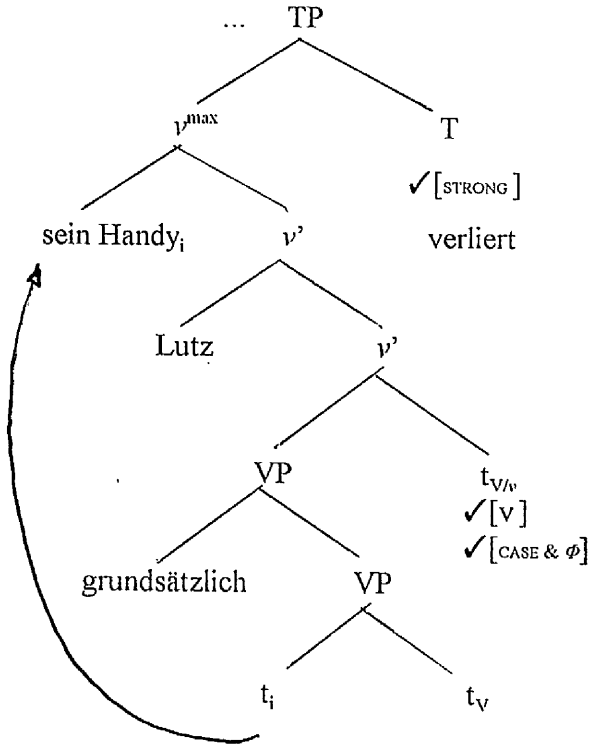
The account is simply as follows. Since the finite verb *does* move overtly in German and Dutch, although the effects may be "covert", or "string vacuous", due to the head final property of verbal and inflectional projections in these languages, the relevant syntactic conditions now obtain to induce overt/covert optionality with respect to Object Shift; German and Dutch are like Icelandic in the relevant way. Let us illustrate for the German example (11).

In one possible derivation, corresponding to overt Object Shift (11b), the verb *verliert* checks *all* features relevant to it - that is, the strong V-feature of T, and further non-strong features associated with T (say tense) and with v (possibly aspectual) - via an application of Form Chain, instantaneously creating a chain with links at T, v and V, as shown in (18a) below. In the second step, shown in (18b), the object then raises overtly to Spec-v, checking Case and ϕ features, which will be there to be checked since the "real" verb which possesses them has left a link in this site.

(18) a. Overt Object Shift derivation: Step 1

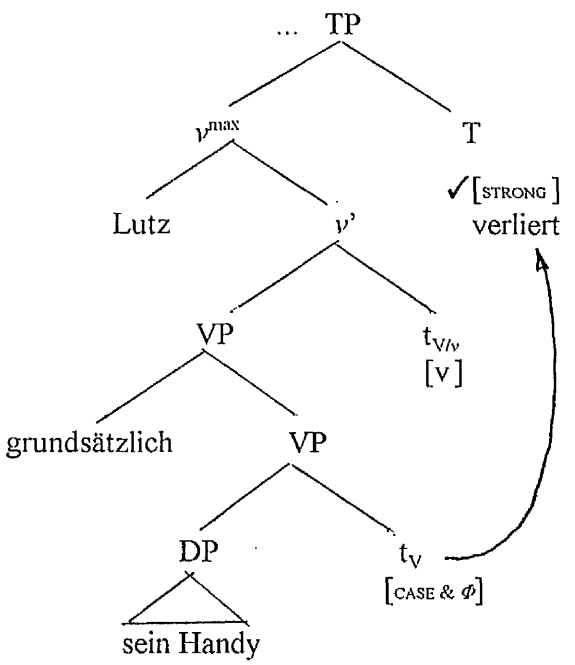


b. Overt Object Shift derivation: Step 2

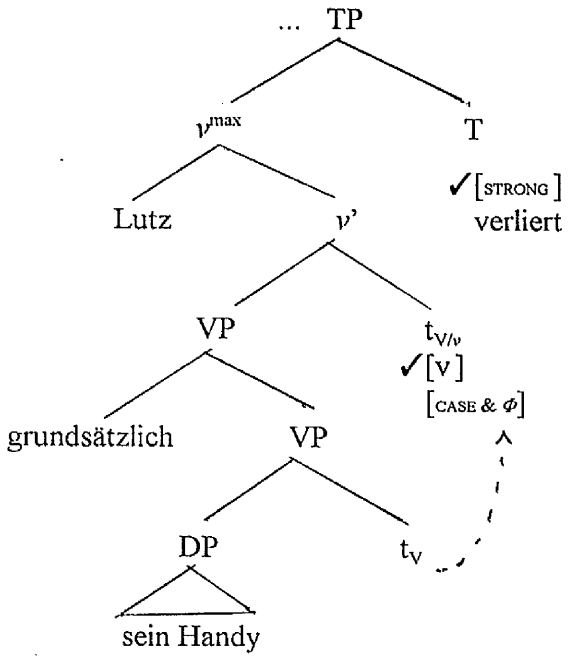


The alternative derivation, corresponding to covert Object Shift (no scrambling) (11a), is as follows. At the point when T is merged in the structure, requiring overt checking of its strong feature, the verb moves there *directly*, bypassing features relevant to it in *v*. I assume that these features are not strong and therefore need not be checked overtly for convergence. This step is illustrated in (19a) below. Next, these remaining non-strong v-features in *v* are checked by covert raising of the formal features of the VP-internal copy of the verb, as in (19b). Among these formal features of the verb are also Case and agreement features. These are then checked by raising of the formal features of the object, as in (19c). Note again that movement both *of* and *to* a copy is involved.

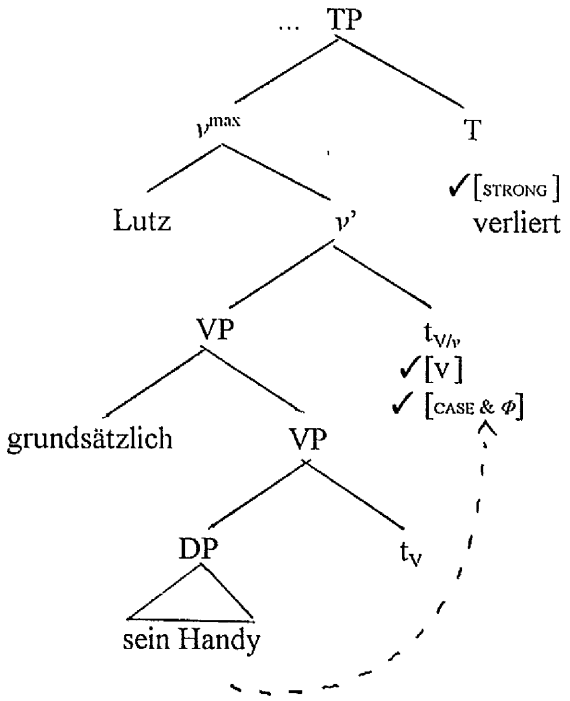
(19) a. Covert Object Shift derivation: Step 1



(19) b. Covert Object Shift derivation: Step 2



(19) c. Covert Object Shift derivation: Step 3



As with the Icelandic case (1), we can explain why both derivations (18) and (19) are optimal with respect to economy. The derivation in (19) - covert Object Shift - is optimal by Procrastinate, since no non-strong feature - relevantly, the DP and v-features in the verbal complex - gets checked by overt movement. Compare this with the derivation in (18) - overt Object Shift: here, there are two instances of overt movement to check non-strong features, namely, the overt Object Shift itself, plus the overt checking of v-features in the head *v* by the verb, as part of its overt movement to T.

On the other hand, while the derivation in (18) is not preferred by Procrastinate, it does count as the shortest derivation, and hence is optimal by the SD condition. This is so since in (18), two operations are involved to check the features relating to the verb and the object, namely the operation on V which instantaneously creates a chain linking it with *v* and T, and the operation which moves the object to Spec-*v*. Compare this against (19): here, no less than *three* operations are necessary to check these same features. This is crucially due to the fact that in this derivation, V skips relevant non-strong v-features in *v* when it moves to T, meaning that checking these v-features must count as a separate operation involving raising of the formal features of V's copy. In sum, derivations (18) and (19) are *each* optimal with respect to derivational economy, the former with respect to SD, and the latter, Procrastinate. I thus assume that both are permitted - hence optional overt Object Shift. The same analysis applies to the Dutch case (12) above.

2.2.3 Why optionality persists with non-finite verbs

We now need to return to the question of why the option for Object Shift persists in German and Dutch when the verb is a participle, in notable contrast to Icelandic. The relevant data is repeated here:

(20) a. ... weil Lutz gestern sein Handy verloren hat
... since Lutz yesterday his mobile lost has
b. ... weil Lutz sein Handy gestern verloren hat
... since Lutz his mobile yesterday lost has

'since Lutz lost his mobile yesterday'

(21) a. ...dat de politie gisteren de taalkundigen opgepakt heeft
....that the polcie yesterday the linguists arrested have
b. ...dat de politie de taalkundigen gisteren opgepakt heeft
...that the police the linguists yesterday arrested have

'...that the police arrested the linguists yesterday'

In § 2.1.4, I linked the disappearance of Object Shift optionality in Icelandic with the fact that the nonfinite verb does not move overtly out of VP. In the present context, it is therefore interesting to note that nonfinite verbs in German and Dutch behave differently from their Icelandic counterparts. Whereas in Icelandic, it is quite clear that the participle cannot move overtly out of VP, as shown by (22), in the SOV languages, adverbial elements of various kinds are freely able to intervene between presumably *unscrambled* objects and participles (just as is the case with finite verbs - see (15) and (16) above);

(22) a. * Margir stúdentar hafa lesið ekki bókina
many students have read not the-book
b. * Margir stúdentar hafa lesið vandlega bókina
many students have read carefully the-book

(23) ... weil Lutz gestern sein Handy unvorsichtigerweise verloren hat
... since L. yesterday his mobile phone carelessly lost has
'since Lutz carelessly lost his mobile phone yesterday'

(24) ... dat de politie gisteren de taalkundigen voorzichtig opgepakt heeft
... that the police yesterday the linguists carefully arrested have
'... that the police carefully arrested the linguists yesterday'

On the basis of this data, we might tentatively hypothesize that in German and Dutch, not only finite verbs, but also participles, move out of the VP to a higher functional head, the nature and location of which I return to shortly. The tentative hypothesis

receives some support from the fact that in neither German nor Dutch can distinct elements of any kind intervene between the participle and the auxiliary:[23] this is illustrated in (25). Meanwhile, Icelandic displays the opposite behaviour: here, as expected if participles fail to raise overtly, adverbial elements such as *ekki* 'not' can of course show up between a participle and an auxiliary in T - (27):

(25) *... daß Lutz die Blumen geschnitten achtsam/gestern hat
... that L. the flowers cut carefully/yesterday has

(26) * ... dat de politie de taalkundigen opgepakt voorzichtig/gisteren heeft
... that the police the linguists arrested carefully/yesterday have

(27) Margir stúdentar hafa ekki lesið bókina

If participles do raise overtly out of the VP in German and Dutch, this puts us in a position to explain why Object Shift is optionally overt or covert in this situation. One potential complication here is that for my account to work, it has to be the case that the participles are forced to raise to a head beyond *v* (because my accounts of optionality in French, Icelandic and German and Dutch finite clauses have all worked off the fact that an element is forced to move to a strong-featured head situated above another head with relevant but non-strong features). Presumably, this head is not T itself, since this contains the auxiliary, but it is likely to be almost as high as T, given the type of adjacency effect which obtains between auxiliary and participle. I leave this for further research.

2.2.4 The Germanic-French difference

Given the above claim that overt verb-raising toTense plays an instrumental role in bringing about optionality of overt Object Shift in Icelandic, German and Dutch, a serious question arises: why is there no optionality of overt Object Shift in declarative environments in French, given that French is standardly thought to have overt verb-raising (see e.g. Pollock 1989)? The lack of overt Object Shift in French is illustrated in (28): I

[23] Apart from parantheticals.

assume that the subject-oriented Floating Quantifier *tous* marks the left edge of VP:

(28) a. * Les hommes aiment Paul tous
the men like Paul all
b. Les hommes aiment tous Paul
the men like all Paul

Of course, this is a question for all theories which derive a connection between overt V-raising and overt Object Shift. It has been suggested by Jonas and Bobaljik 1992, Branigan 1992 that while overt verb-raising to T is necessary for overt Object Shift, it is not sufficient; in addition, Spec-TP must be available as a landing-site if overt Object Shift is to be possible.[24] This is because the availibility, or lack of it, of Spec-TP has implications concerning locality and cyclicity. Under the assumptions of the 1993/1994 Minimalist framework, in an ordinary declarative environment, if an object moves overtly to its Case position, then it must do this *before* the subject, because of the Extension Condition. If the object moved *after* the subject, and overtly, a violation of this condition would ensue.[25] Hence, if the object is to move overtly, it must move before the subject. But this then raises the further possibility that the object in its derived position above the base position of the subject will subsequently block the subject from moving to its own Case position. Significantly, the only way in which the derived object will not count as closer than the subject is if there is an *equidistant* potential landing-site which, in the 1993 Minimalist framework assumed by Jonas and Bobaljik, meant Spec-TP but crucially not Spec-AGR_sP. Thus Spec-TP had to be available as a landing-site for overt Object Shift to be possible.

This idea could be utilized to account straightforwardly for why there is no option for overt Object Shift in French: Spec-TP is not a possible landing-site. It then remains to explain why the restriction is obviated when the object is a *wh*-phrase. For this, a proposal of Branigan's (1992) can be adopted. When, but only when, the object is a *wh*-phrase, it

[24] Whether or not Spec-TP is available is taken to reduces ultimately to a parametrically varying property of the head T itself.

[25] See Chapter 1 on the Extension Condition.

can move *after* the subject, because its ultimate landing in Spec-CP will extend the structure, in compliance with cyclicity.

The one potential problem with this approach to the French/Icelandic variation is that it fails to translate into the 1995 Minimalist framework, in which there are fewer functional heads, and multiple Specs. However, nothing which has been proposed in this thesis is dependent on this recent revision of clause architecture, and some authors, e.g. Jonas 1996, have moreover argued for retaining the earlier (Chomsky 1991, 1993) view. If this is correct, the potential problem disappears. In addition, we get a solution for another cross-linguistic problem noted in Chapter 5 (note 32): given my account of optional agreement with *wh*-objects in French, the problem was to explain why English lacks an option for overt Object Shift in the environment of *wh*-movement, as shown by the Floating Quantifier data discussed back in Chapter 4; compare (29) and (30):

(29) * The tables, which John all repainted,...

(30) Les tables, que Jean a toutes repeint(es),...

Crucially, given the clause structure assumed in the 1993/1994 Minimalist framework, English contrasts with French in that in the former, overt Object Shift is completely excluded by the Minimal Link Condition, for reasons ultimately relating to the fact that in English, the verb fails to raise overtly (see Branigan 1992 for detailed discussion). This would then be one possible way of capturing the English-French contrast - how satisfactory a way, it remains to be seen; I leave this for future research.

2.3 Section summary

In this section, I have shown how my optionality system can cover a range of overt/covert (non)optionality phenomena found in Icelandic, German and Dutch. In what is essentially a variation on the classical "Holmberg's Generalization" approach - i.e. overt Object Shift is contingent on overt V-raising - I attributed overt/covert Object Shift optionality to

overt V-raising to Tense, or, in the case of participle environments in German and Dutch, overt V-raising to a functional head above *v* but below Tense, whose exact identity is yet to be determined. I made the important auxiliary assumption that in German and Dutch, both tensed verbs and participles move overtly out of the VP, to the right (though perhaps not necessarily so - see note 20 above). Although some details remain to be worked out, this assumption is not entirely lacking in independent motivation, as I attempted to show from certain adverb placement facts.

Having noted that my account is a variation on Holmberg's Generalization accounts, there is a difference which should be noted. In the 1993/1994 Minimalist framework, as we saw in Chapter 1 (§ 4.2.3), Holmberg's Generalization itself was derived from the Minimal Link Condition, but in the 1995 framework, with its fewer-headed clause-structure, this can no longer be the case: "note that we have lost Holmberg's generalization and other effects of V-raising on extension of chains; that is a consequence of excluding chains from the definition of "closeness". Such generalizations, if valid, would now have to be stated in terms of a property of Vb [= the *v*/V complex - AMP] ... it can have a second outer Spec only if it is a trace. There is no obvious reason why this should be so" (Chomsky 1995, p.358).

My slightly altered perspective on Holmberg's Generalization is that overt V-movement is linked to *optionality* of overt Object Shift, rather than *possibility* of overt Object Shift, and locality is not directly involved. In at least one way, this difference seems to be advantageous: in Scots Gaelic, there are syntactic environments in which obligatory overt Object Shift occurs in the absence of overt verb raising, as shown in (31) (taken from Adger 1994:93):

(31) Feumaidh Daibhidh cat a bhualadh
must David cat agr. strike
'David must hit a cat'

(31) seems to counterexemplify a theory in which overt Object Shift is contingent on overt verb raising, although tests would need to be carried out to see if the verb is in fact

in situ. On the other hand, (31) does fit in with my weaker version of Holmberg's Generalization, namely that *optionality* of overt Object Shift depends on overt verb raising, for in (31), there is no optionality with respect to the object: it is obliged to move overtly. Again, this is a topic for future research.

3 Three comparisons

In this section, I critically review three Minimalism-based theories which are comparable in various ways to the account of optionality which I have presented in this and the previous chapter. The first of these theories is by Kitahara 1994, who develops a fairly similar derivational economy-based account of optional overt Object Shift in Icelandic. The second is by Branigan 1992, who suggests an account of the French agreement optionality data (see Chapter 5) in terms of optionally strong features. The third is by Adger 1994, who addresses the French data and also Dutch object scrambling, as well as a number of other cases of optionality. I shall suggest that my theory has advantages over all three of these accounts.

3.1 Kitahara 1994

Mine is not the only attempt to account for a case of optionality in terms of more than one optimal realization of interface conditions; the first such account, appropriately enough, was contemplated by Chomsky 1991, for the case of optional infinitive movement in French. Within a more contemporary Chomskyan framework, Kitahara 1994 has proposed an account of the case of optional overt Object Shift in Icelandic which is similar in many ways to my own proposal. The relevant data is repeated here as (32):

(32) a. Jón las bækurnar ekki
J. read the-books not
b. Jón las ekki bækurnar
J. read not the-books

'Jón did not read the books'

In common with my proposal, Kitahara claims that (his example equivalent to) (32) involves an alternation between overt and covert Object Shift, and that both derivations are permitted because both qualify as optimal with respect to derivational economy. Though similar to my account both in spirit and in end product, Kitahara's system differs in certain intensional properties in a way which, in my view, is to its disadvantage. To illustrate, let us summarize the account.

In Kitahara's theory, in contrast to my own, optionality between covert and overt movement in (32) involves a single economy condition, the Shortest Derivation Requirement (SDR), which states: "minimize the number of operations necessary for convergence" (1994:32). On the question of what counts as an operation, Kitahara commits himself to two crucial assumptions which are not standard in Minimalism, and which do not have any place in my own account. The first of these assumptions is that removal or "deletion" of the phonological matrix of a copy, i.e. a Spell-Out type process, counts as an operation for the SDR; since overt movement entails more such operations than covert movement, it is therefore calculated as more costly by the SDR. The second assumption is that covert movement, of phrasal elements at least, entails an operation to construct an empty position prior to and separate from the movement itself; overt movement does not involve these separate position-construction operations, and on this basis is calculated as *less* costly than covert movement by the SDR. Given the SDR incorporating these two assumptions about what counts as an operation for its purposes, there is, modulo strong features, almost a complete balance between overt and covert movement with respect to economy.[26] Let us now see how this system predicts

[26] Almost, because in the case of X^0-movement, covert movement (if possible) will *always* be cheaper than overt movement, on the assumption that no empty position has to be created.

overt/covert optionality of Object Shift in Icelandic, as in (32), but obligatory covert Object Shift in English.

The crucial background assumptions to this account are firstly that in general, overt Object Shift is possible only if there is overt movement of the main verb, which Kitahara takes to stem from locality conditions;[27] and secondly, that overt main verb movement (at least of finite verbs) is obligatory in Icelandic - due to a strong feature - but not in English, where presumably the relevant feature is weak. The account of the optionality of overt Object Shift in Icelandic is then as follows: overt Object Shift, possible due to overt verb raising, incurs an operation to delete the phonological matrix of one of the copies of the overt chain (so that only one copy "spells out".). In this respect, the overt OS derivation is more costly than its covert counterpart. At the same time, covert OS necessitates the prior construction of an empty Spec position - Spec-AGR_oP in the framework Kitahara uses - an operation which is not necessary if OS is overt. Thus, "these two 'competing' derivations ... employed the same number of Target α applications to converge ... these derivations each satisfy the SDR; i.e. they are equally most economical" (1994:45).

In Icelandic, recall that overt verb movement is forced by a strong feature and as such is not a factor in determining economy, since obviously it must occur in all derivations which are convergent. This is the source of the difference between Icelandic and English: why in the latter there is no option for overt Object Shift. As mentioned, if there is to be overt OS in English at all, the verb also has to raise overtly - i.e. there is no convergent possibility involving overt OS but covert verb movement. But a derivation with overt verb movement and overt Object Shift is of course more costly than one with covert verb movement and covert Object Shift, since in the former, an extra operation of

[27] In keeping with views prevalent at that time, Kitahara attributes the dependence of overt Object Shift on overt verb movement (Holmberg's Generalization) to the Shortest Movement Requirement, i.e. his version of Chomsky's Minimal Link Condition (locality). About this topic, see Chapter 1 of this thesis.

phonological matrix deletion is required to apply to the overtly-formed V-chain.[28] Thus, covert OS is obligatory in English, due to the SDR.

It can be seen that Kitahara's theory is quite similar to my own, since it derives (non)optionality facts from feature-strength, subject to variation, in conjunction with general economy conditions. But notice that although Kitahara's system captures both the optionality of overt Object Shift in Icelandic, and the obligatorily covert nature of the same operation in English, it is arguable that some of the crucial assumptions upon which the theory depends are questionable. In the first place, as mentioned above, Kitahara claims that the Shortest Derivation Requirement counts deletion of the phonological matrix of elements as a computational operation. Although Kitahara subsumes such phonological deletion and "regular"syntactic operations like Merge and Move under the same title, "Target α", it does appear odd to lump these together, in that phonological matrix deletion is presumably part of the Spell-Out component which converts syntactic structures to PF representations, traditionally and contemporarily thought to involve a very different type of "derivation" to that which is involved in constructing LF structures. (see Chomsky 1994). It seems to me undesirable to have a condition which encompasses two such different systems; this seems to entail a level of globalness which goes beyond merely comparing derivations within the syntax (which of course is necessary in Kitahara's and my own account, and in the Minimalist framework generally). In comparison, the account of the French data which I developed in chapter 5 does not assume that considerations of the Spell-Out component have any direct relevance to conditions which regulate the derivation of syntactic structures.

In addition to this, Kitahara's Shortest Derivation Requirement has another problematic characteristic, and this concerns his assumption that empty positions are constructed to be moved into at a later, separate stage. The problem with this is that in the Minimalist framework,[29] it is assumed that a position is created when and only when

---

[28] A derivation with overt verb-movement and covert OS is also less economical than one with covert verb-movement and covert OS, for obvious reasons.

[29] Specifically Chomsky 1994 (*Bare Phrase Structure*)..

an element actually moves "to" it; in such a framework, it is not particularly plausible to maintain any meaningful distinction between the creation of a position and movement into that position. In the framework of Chomsky 1995, where covert movement basically corresponds to pure feature-adjunction to an X^0, this point is clearcut. Yet without the assumption that covert movement entails an extra prior structure-building operation, Kitahara's system of deriving overt/covert optionality collapses. In comparison, the superficially similar system which I used to derive overt/covert optionality for the French data depends on no such assumption, which is surely to its advantage.[30]

In comparing Kitahara's syntactic optionality system with my own, I have suggested that the former possesses some specific conceptual shortcomings which the latter does not. The two accounts are likely to be similar in empirical coverage, although note that Kitahara's theory, where timing effects all derived from a single condition (the SDR), seems to predict that *all* XP-movement to check weak features is optionally overt/covert in principle. In comparison, my own account, with its use of two separate economy conditions, seems likely to predict timing optionality in more restricted circumstances.

---

[30] As a further minor criticism of Kitahara's theory, one might consider his assertion that the Shortest Derivation Requirement is an improvement on standard Minimalism because it eliminates Procrastinate as a separate condition: "notice that Procrastinate, which stipulates a cost-distinction between overt and LF operations, plays absolutely no role in this analysis." (p.43). To me, this does not seem to be an authentic advantage of Kitahara's system, since Procrastinate (nor any condition) does not (cannot) make direct reference to c/overtness of movement (see chapter 5), although it is possible to phrase it in such a way. (Cf. Chomsky 1993's original version (see Chapter 1), or Kitahara's own rendition: "LF operations are cheaper than (i.e. less costly) than overt operations" (p.24).

3.2 Branigan 1992

In this section, I look at a Minimalist account of the French agreement optionality data developed by Branigan 1992, which I have already briefly referred to (see Chapter 5). Branigan's theory belongs to the category of what I called "lexical approaches" to optionality phenomena, and in fact he asserts, in the form of a rationale for his proposal, that "optionality is, in general, a property of lexical insertion, and not of syntactic processes" (p.37). Insead of positing a formative or some particular features with interpretive import for the options with agreement - a strategy which I discussed and argued against earlier in connection with the lexical approach generally - Branigan's theory is characterized by what Marantz 1995 describes as "...the imaginative use of some of the functional heads...for example, suppose that all the N features of T and AGR are potentially weak ... Then no DP need raise from the VP...prior to Spell-Out. Now if some of these N-features are optionally strong, movement of some of the constituents from VP but not others could be forced..." (378).[31]

Branigan assumes, as also in Sportiche 1992 and the present thesis, that the agreement-triggering position is an A-position where Accusative Case is checked - Spec-AGR_oP in his particular system. His basic suggestion is the following: "Agr [[32]] can have either strong or weak features from the outset, freely" (1992:37). That is, Branigan claims - in contrast to my own account - that the sentences containing agreeing and nonagreeing participles have different numerations, distinguished by the presence or absence of a strong feature.[33] The account works like this: if the strong feature option is selected, then the object is forced to move there overtly, triggering agreement. If the object is a non-*wh*-phrase, then this eventuality is actually blocked by the combined efforts of Relativized Minimality (i.e. MLC) and the Extension Condtion/Strict Cycle; if

---

[31] The actual example of optionality which Marantz discusses is scrambling in Japanese, but the comment is general.

[32] Agr = The functional head which I am thinking of as light *v*.

[33] An optional Strong feature account is also suggested by Collins and Thráinsson 1994 for the optionality of overt Object Shift in Icelandic.

the object is a *wh*-phrase, this prohibition on overt movement to Spec-AGR_oP is claimed by Branigan to be obviated due to the fact that it constitutes a subpart of movement to Spec-CP, which (unlike movement to Spec-AGR_oP) extends the structure.[34] It is not quite clear what derivation Branigan envisages for the *wh*-movement option without agreement. In any case, on his lexical account, there is obviously no reason why the agreeing and non-agreeing options should not coexist. Economy conditions compare only derivations from the same numeration (see Chapter 1), so, like for example *Paul a repeint les tables* and *Paul a repeint les chaises*, or *Quand est parti Jean?* and *Quand Jean est-il parti?* (see §x.x..x), the agreeing and non-agreeing options will not be in competition with each other.

However, although this optional feature-strength account does work, in some sense of the word, I think that in another way it is unsatisfactory. It seems essentially that the theory is unfalsifiable; it is evident *a priori* that it can deal with any case of suspected movement optionality one might discover. The vacuity of the optional strong feature approach can be illustrated in the following way: if a suspected case of syntactic optionality (e.g. the case of French agreement under discussion) is explained in terms of the existence in the lexicon of both strong and weak features associated with a certain functional element, then consider how this explanation interacts with related non-optionality data (e.g. the environments in French in which participle agreement is either obligatory or impossible). To explain these cases, it will then be necessary to either (i) state that in these constructions without the optionality, the lexical choice of strong or weak features dissapears - a solution which is explicitly *ad hoc* and perhaps unworthy of consideration at all, or failing that, (ii) assume that the feature-strength optionality exists in general, with the unwanted options thereby predicted to be ruled out by means of some separate constraint or constraints. That is, it seems that the lexical approach to optionality in terms of optional feature-strength, as exemplified by Branigan's (1992) proposal

[34] Note incidentally that Branigan's proposal would not be tenable in the framework of Chomsky 1995, since in that framework, a strong feature is that which must be checked before any further structure is added. Hence, even a *wh*-phrase ultimately on its way to Spec-CP would, under the assumptions Branigan is using, violate the Extension Condition in the same way as a non-*wh*-object would.

discussed here, is inherently incompatible with any contentfully integrated account of optionality and non-optionality phenomena. To give a concrete example: Branigan (*op.cit.*) explains agreement optionality cases like *Quelles tables as-tu repeint(es)?* in terms of optional feature-strength; then, to deal with a case where participle agreement is not optional but impossible, e.g. *Paul a repeint(*es) les tables* - he employs completely different means, viz. Relativized Minimality and the Strict Cycle.

If there were no conceivable alternative approach to optionality phenomena, then there might perhaps be a case for resorting to the optional feature-strength strategy despite its essentially *ad hoc* character. However, in the approach to optionality which I have outlined for the French and Germanic cases - and indeed in Kitahara's analysis of Icelandic reviewed above - it is at least true in principle that cases of optionality and non-optionality can be made to fall out from the same, more general, principles. In summary, there is good reason to reject the optional feature-strength account in view of both its inherent and relative disadvantages.

3.3 Adger 1994

I lastly consider a very different approach to optionality. Adger (1994a,b,c) develops a theory in which syntactic optionality is causally linked with interpretive effects of the type mentioned briefly in chapter 5. Though ostensibly set within the Minimalist framework, we will see that Adger's system relies on a fairly non-Minimalist premise: that considerations other than feature-checking - in particular, considerations of an extra-syntactic nature - are involved in sanctioning syntactic operations (movement).[35] In the above-mentioned works, Adger deals with several cases of optionality, including the cases involving object scrambling in Dutch and French agreement which I have dealt with in this thesis. For concreteness, I discuss his theory with reference to the French data which I discussed in the previous chapter.

---

[35] Cf. Reinhart's (1993) Minimalist account of certain facts concerning Superiority and *wh-in-situ*.

What figures crucially in the optionality as far as Adger is concerned is the interpretive phenomenon which arises in the case of *wh*-movement, first observed by Obenauer 1992. Recall that in the *wh*-movement option with agreement, the agreed-with phrase has to be interpreted as specific - or familiar (in the sense of Heim 1980), as Adger argues the relevant interpretive property to be. He claims firstly that elements in the Specifier of what he calls AgrP (= to all intents and purposes, Spec-*v*) are interpreted as familiar. In this theory (as compared to Diesing's (1992) Mapping Hypothesis which Adger critically discusses, *op. cit.*, p.86), the link between syntactic conditions and the discourse property is only a partial one, in the sense that having a chain-link in Agr is sufficient but not necessary for a DP to be interpreted as familiar; that is, if a DP has a link in Agr (triggering agreement), then it will be interpreted as familiar, but a DP may be also be interpreted as familiar without having a link in Spec-Agr, by non-syntactic means i.e. pragmatically. Even this weak correlation between syntactic form and interpretation, which is fundamental to Adger's account, encounters what seem to be serious counterexamples, which we will discuss shortly. However, it does capture the following facts: firstly, that in the case of optional agreement with a *wh*-phrase in French, the *wh*-phrase is interpreted as familiar in the version with agreement, whilst in syntactic contexts *without* participle agreement, be it the agreement-less option in the *wh*-movement case, or the case of a regular object as in (33), the non-agreed-with phrase can be interpreted as *either* familiar *or* unfamiliar (cardinality reading), as informally indicated in the translation:

(33) Paul a repeint(*es) cinq tables
Paul has repainted(*f.pl.) five tables(f. pl.)

'Paul repainted five (of the) tables'

But, as mentioned above, further facts appear cause problems for Adger's weak Agr-familiarity correlation, for although it states that if a DP-chain has a link in Agr, that DP will be interpreted as familiar, the fact is that, as mentioned in the last chapter, it is not only constructions where agreement is impossible (as in (33)), but also those in which it is *obligatory* - exemplified by (34) below - where the relevant DP can in fact be

interpreted either as familiar or unfamiliar.

(34) Cinq tables ont été repeint*(es)
five tables-fem.pl. have been repainted-fem.pl.

'five (of the) tables were repainted'

Importantly, as Adger points out, in cases of syntactic optionality, we find interpretive obligatoriness, whilst in cases of syntactic obligatoriness, we find interpretive ambiguity. Given that this state of affairs clearly fails to correspond perfectly to the statement "if a DP-chain has a link in Agr, then it is interpreted as familiar", some additional assumptions are needed. This is exactly where the explanation of syntactic optionality comes in.

Adger proposes to account for the correlation between interpretive effects and the optionality of agreement in the following way. He assumes firstly that in a DP-chain with a link in Agr (which recall suffices for a familiar interpretation of that DP), the link in Agr may in principle be deleted, with *deletion* here crucially referring to syntax/LF, rather than PF, yielding a non-familiar reading of the DP. Now, in environments in which movement through Spec-AgrP is for independent reasons syntactically obligatory, namely NP movement cases such as (34) above,[36] the claim is that either the link in Agr remains, or it deletes, resulting in the observed ambiguity with respect to familiarity. But why should such deletion of the Agr-link *not* be possible in the case of the *wh*-movement option with agreement, as must be so given the *wh*-phrase's obligatory familiar interpretation? Adger's claim is that it is its very status *as an option* which rules out a deletion operation on the chain: for its sister option *without* agreement, assumed by Adger to reflect lack of movement through Spec-AgrP, itself results in a nonfamiliar interpretation of the DP, due to its having no link in Agr.[37] Derivational economy then rules out the possibility of achieving the same interpretation by means of forming a chain with a link in Agr, then

[36] Adger attributes the obligatoriness of movement through AgrP to Relativized Minimality - see previous chapter.

[37] Incidentally, Adger does not give an account of how the phrase gets Case in this option, although he explicitly identifies his Spec-AgrP with Accusative Case.

deleting this link. By contrast, in the case where movement through Spec-AgrP is for independent reasons obligatory, economy does not rule out the deletion method of obtaining the nonfamiliar reading, since this is the only possible way to achieve this. The major assumption in this set-up is that a syntactic operation can be licensed in virtue of its resulting in an otherwise unavailable (syntactically-encoded) interpretation.

I think that the latter idea is at the root of several problems with Adger's account. Notably, the fact that it entails the claim that derivations are not "driven by the narrow mechanical requirement of feature checking only" (Chomsky 1993:33) goes against the standard Minimalist assumption. Whilst that alone is not necessarily an insurmountable problem, one must question a system in which information of an extra-syntactic nature (presumably in addition to morphosyntactic requirements) participates in determining the class of operations permitted by the syntax. As discussed earlier in connection with Kitahara's (1994) account of optionality in Icelandic, I assume that on conceptual (and probably also computational) grounds, it is preferable to avoid such direct interaction between independent systems if at all possible.[38] Of course, depending upon how successful a system of the type just described was at accounting for the facts, one might have to reconsider these "conceptual" reservations. But as it happens, Adger's system is not without empirical problems, since, contrary to its predictions, there do exist cases of syntactic optionality which seem to be associated with no interpretive differences of the relevant type.

One especially salient example of this is the case of optional agreement in Accusative clitic constructions in French. The data, which Adger does not discuss, is repeated here:

---

[38] It is important to dissociate these reservations from the perhaps similar-sounding claim that discourse-related properties should not be encoded in syntax (e.g. via morphosyntactic features, see Delfitto and Corver 1995). As far as I can see, there is nothing particularly wrong with the latter claim in conceptual terms, although as indicated in the previous chapter (§ 3.2.2), I think that in the case of the French agreement facts and familiarity, an account in these terms would face problems.

(35) (les tables,) Paul les a repeint(es)
(the tables,) P. them(f.pl.) has repainted(f.pl.)
'(the tables,) Paul repainted them'

Here, in contrast to the *wh*-movement case, there is no interpretive difference at all between agreeing and nonagreeing options, as already noted in Chapter 5 (§ 3.2.2). Moreover, this is is not the only case of syntactic optionality in which there is no interpretive difference between the options; consider for example the case of optional associate movement discussed earlier in Chapter 5, or the optional movement of infinitives in French, discussed by Pollock 1989 and Chomsky 1991, illustrated in (36):[39]

(36) a. Ne pas être heureux
. ne not to-be happy
b. N'être pas heureux
ne to-be not happy

'to not be happy'

Last but not least, Adger's account has a technical problem in that it countenances, and in fact requires, the possibility of syntactic deletion of the Case-marked link of a chain, which should on standard assumptions lead to violation of the Chain Condition (see Chomsky and Lasnik 1993).

Unlike Adger's account, the system which I have developed in this thesis deals with optionality in purely syntactic terms (feature properties plus economy conditions). It does not assign any role to interpretive considerations and hence avoids the concomitant conceptual and empirical problems. On the other hand and by the same token, it remains to be demonstrated exactly how my account can deal with the interpretive effects which obtain in the *wh*-movement case. A positive aspect of Adger's (1994) account is the important insight that there is a systematic connection between optionality and interpretive effects, although, as indicated above, I think that this connection needs to be characterized in terms other than those suggested by Adger himself. I turn to this in the next section.

---

[39] Chomsky 1991 gives a syntactic optionality account of this data, in terms of the now obsolete economy condition Least Effort (see Chapter 1).

4 On the status of "interpretive effects"

Let us finally turn to the interpretive effects which have been mentioned from time to time throughout this chapter and Chapter 5. My objective here is merely to describe the facts in question, to explain why the syntactic account I gave of the French agreement and Germanic object movement optionality definitely does *not* preclude the existence of some systematic link between aspects of the syntax and the interpretive effects, and to engage in some speculation as to a possible account.

In Part II, I have looked at the cases of optionality of overt Object Shift in Icelandic, German and Dutch, and optionality of phonologically overt participle agreement with *wh* and clitic objects in French; recall that the latter was also treated in terms of optional overt Object Shift. Now in the Icelandic, German, Dutch cases, and in the case of optional agreement with French *wh*-movement, the following phenomenon occurs: in the options *with* overt Object Shift, there is a restriction on the interpretation of that object such that only one of its inherently possible interpretations is permitted. In the options *without* overt Object Shift, on the other hand, this interpretive restriction does not hold. The interpretive property in question has been characterized in a number of different ways - in terms of specificity (Enç 1991; Obenauer 1992 for French); presuppositionality (Diesing 1992); familiarity (in the sense of Heim 1981; Adger 1994, Delfitto and Corver 1995). For the present purposes, the exact nature of the interpretive property is not relevant, but for convenience, I shall use the term *familiarity*.

In some sense, it might be said that reordering of constituents will in and of itself tend to affect interpretation; after all, this arises from different ordering of entire sentences, and indeed of non-linguistic stimuli. However, it seems that something of a more systematic nature is going on, as far as the familiarity effects are concerned. For one thing, in the case of French *wh*-movement, there is no apparent "reordering" of constituents; the difference between the options concerns presence versus absence of overt agreement. More generally, as noted in § 3.3 above, Adger 1994 (p.94) notes that "there appears...to be a correlation between optionality in the...derivation, and

obligatoriness of familiarity". In short, situations of syntactic optionality tend to give rise to the interpretive restriction on one of the options, described above, while related situations where optionality in the syntax disappears - e.g. in the context of nonfinite verbs in Icelandic, or in the case of NP-movement in French - interpretive ambiguity obtains. This seems to be an important generalization about the relation between syntax and interpretation, and I assume that a theory which lacks the means to capture it is probably on the wrong track. So let's see how my theory fares.

I characterized optionality in the French and Germanic cases in terms of overt versus covert Object Shift. Let us look at the account in representational terms. It is clear that in all the cases, our alternating derivations give us different chains for the element involved, as illustrated abstractly in (37) (FF = formal features):

(37)	a.	Chain: <XP, XP>	overt Object Shift
	b.	Chain: <FF(X), XP>	covert Object Shift

As (37) plainly shows, there is a quantitative but not a qualitative difference between the (a) and (b) object chains, and hence between the LF structures which contain them - but a difference nonetheless: obviously, the chains formed by overt movement contain a full extra category, on the assumption that overt movement, but not covert movement, forces "pied piping" of the entire category (Chomsky 1995). Since the options in the Object Shift cases will indeed involve different LFs, and since LF is of course the input to interpretive processes, we clearly have the basis for an account of the systematic interpretive effects arising in the optionality cases.

Notice that in predicting quantitatively different LF-structures for the alternating options, the account generally replicates Kayne's (1989) analysis of the French *wh*-movement case: recall that he maintained that "two *wh*-movement sentences, one with and one without past participle agreement, will, even if otherwise identical, have category-wise different representations" (p.90). In fact, building on Kayne's analysis, Obenauer 1992 attributes the obligation for a familiar (in his terms, specific) interpretation of *wh*-object in the structures with participle agreement to the fact that the chain of the *wh*-phrase

cannot be a true operator-variable construction, and must instead involve a pro-like element which he claims is akin to resumptive pronouns. However, Obenauer's conclusion that the chain in this instance cannot be an operator-variable construction is not compatible with my own account, since it is dependent on the assumption that the agreement-triggering position is an A'-position (Kayne 1989). In addition, it is not clear how Obenauer's account could be extended straightforwardly to the Germanic Object Shift cases, since these do not seem to involve A'-movement at all. Let us investigate an alternative idea.[40]

According to my account, the syntax itself provides two "essentially the same" yet not identical LF structures in the optionality cases. Now in general, it must be assumed that there exist various post-syntactic processes which operate on the LF structures which are the output of the computational system, to create more complete representations which include contextually-supplied information. It has even been proposed, by Brody 1995 (contra Chomsky 1993), that the formation of actual operator-variable constructions is a post-syntactic process.[41] Following this, I shall make the assumption that, in the same way as A'-chains must be turned into actual operator-variable constructions for the purposes of interpretation - it is not permitted simply to "ignore" superfluous material - overt A-chains must be subject to a similar type of removal: all but the interpretable features must be deleted from all but one copy of the A-chain (extra material which is after all present as a by-product of PF requirements). As in Brody's proposal on operator-variable constructions, I assume that this removal is a post-LF process. Further to this, I assume that the post-LF rules/processes, including those which contribute to establishing the propositional form of the utterance, are themselves subject to economy considerations of some sort, perhaps Relevance, in the sense of Sperber and Wilson 1986. There is a good deal of independent motivation for the latter assumption.

---

[40] The suggestion I shall make here owes a lot to discussions with Richard Breheny, which were themselves partly inspired by some comments by Misi Brody in one of his 1996 syntax groups at UCL, on another interpretation and LF-related topic.

[41] Brody motivates this assumption with certain facts involving reconstruction.

With the above assumptions in mind, I propose the following account of the obligatory familiar reading in the overt Object Shift options. These options, unlike those with covert Object Shift, contain a chain as in (37a) above, and as such require an operation to remove the superfluous material from one of the two categorial copies. By economy considerations of some kind (though not derivational economy conditions, obviously), the extra operation to implement this removal is permitted only if it is offset by some corresponding interpretive benefit, namely, that the element in question is construed as recoverable from the existing context, i.e. "familiar". In cases where the syntax yields only one possible LF, as in e.g. the case of subjects in English, non-*wh*-objects in French, then this particular economy issue will obviously not arise.

Of course, this proposal on the link between syntactic optionality and interpretive effects is speculative and informal, merely suggesting a direction for further research. After all, the issue in question goes beyond the realm of syntax, and as such, this thesis is not the place to pursue it in depth (this situation can be seen as an LF analogue of the "Case Adjacency" effects and PF discussed in Chapter 3). What I do hope to have achieved is to demonstrate that the syntactic account of optionality phenomena which I have proposed does not necessarily preclude an account of the interpretive phenomenon. Should the account which I outlined be on the right track to any extent, notice that in contrast to Adger's (1994) theory, there will be no need to assume any unmotivated deletion operations in the syntax in order to predict the correct outcome with respect to interpretation in the nonoptionality cases. It still remains to deal with the cases of syntactic optionality which seem to be accompanied by no interpretive effect, e.g. optional agreement with object clitics in French; I leave this for future research.

5 Summary and conclusion

In this chapter, I have further developed the theory of optionality introduced in Chapter 5. I extended the system to a further range of cases of Object Shift optionality found in Icelandic, German and Dutch. I then reviewed some comparable proposals on optionality from the Minimalist literature, which highlighted both advantages and potential difficulties with my own account. Consideration of Branigan's (1992) account, which uses the notion of "optionally strong features", showed an approach along these lines to be inherently *ad hoc* in a way in which, in my opinion, my own approach is not. On the other hand, a review of Kitahara's (1994) theory, which is similar to mine in deriving optionality from derivational economy conditions, illustrates a tendency of such approaches to be *ad hoc* in practice (it is noticeable that many such syntactic optionality accounts deal with one or two isolated cases of optionality: Object Shift in the case of Kitahara 1994; optional raising of infinitives in French (Chomsky 1991); optional particle shift in English and Norwegian (Svenonious 1995)). I think that my account achieved some improvement in this respect, covering a range of similar optionality (and associated non-optionality) facts from six different languages - French, English, Swedish, Icelandic, German and Dutch - in a reasonably integrated fashion. On the other hand, numerous further cases of optionality exist, and it remains to be seen whether my account could be extended to any of these.

Finally in this chapter, I suggested a possible way in which to account for the interpretive effects arising in some of the optionality cases. I suggested that Adger's (1994) generalization that syntactic optionality correlates with an interpretive restriction on one of the options should be dealt with in terms of economy considerations - but those connected with post-syntactic (or "interface") processes, rather than the derivational economy conditions taken to characterize the syntactic computational system.

Chapter 7
Concluding remarks

This thesis investigated two facets of the characterization of movement phenomena within the Minimalist framework, taking as its point of departure a comparison with the earlier GB model. I first considered a theoretical issue: is it possible to allow *Move* to apply to anything - to be a "blind" computational procedure in the spirit of the GB rule *Move-α?* Secondly, I investigated an apparently major empirical problem relating to movement in Minimalism: how to deal with data previously characterized in terms of optional movement - a type of analysis which now seems unavailable, given the Minimalist assumption that derivations are constrained by economy conditions.

In Part I, I set out to discover whether the theoretically optimal notion of free application of movement operations can be upheld in the Minimalist model. That it can is not a foregone conclusion, since fundamental properties of the model, ending ultimately in the introduction of the copy theory of movement (Chomsky 1993), appear to necessitate the imposition of special conditions upon certain "traces", in order to prevent them from undergoing *Move* (Chomsky 1995). However, I argued in detail against such a position, suggesting instead that "all copies in a chain are active in the computational

system" (the Copy Hypothesis; Chapter 2), and showing how general conditions on movement are sufficient to properly regulate the activity of traces. In the remainder of Part I, I gave a preliminary demonstration of the practical applications of the Copy Hypothesis, involving primarily the behaviour of *wh-* objects and associates of *there* in English (Chapter 4). The Copy Hypothesis furthermore turned out to be quite pervasive in Part II of the thesis.

In Part II I explored the idea that there is in principle some potential for syntactic optionality within the Minimalist system - namely optionality in the occurrence of movement relative to the application of the mapping to PF (Spell-Out) - and attempted to develop a theory which yields such optionality in appropriately restricted circumstances. I proposed that overt/covert optionality - in common with obligatory overt/covert movement effects - can be characterized in terms of the feature-strength properties of functional heads in conjunction with derivational economy principles, in particular Procrastinate and the Shortest Derivation condition. This system was first applied to the optionality of phonologically overt agreement on participles in French which occurs in the case of object A'-movement (object *wh*-phrases and clitics) (Chapter 5). It was further shown how non-optionality of agreement is predicted in the appropriate syntactic contexts. Building on that, I then proposed an analysis of optional overt movement of associates of *there* in English and *det* in Swedish, attributing the nonexistence of this optionality in French to independently-motivated differences in the feature-properties of the expletives. In Chapter 6, I extended my feature-strength+economy account to object-related optionality and related non-optionality data from Icelandic, Dutch and German. It will be noticed that the analysis of optionality developed here depends heavily on the assumption that there is covert movement for feature-checking. To the extent that the analysis is successful, then, it provides motivation of a sort for assuming the existence of such covert operations; recall from the discussion of objects in English in Chapter 3 (see especially note 37) that it is doubtful to what extent facts involving Binding and so on can be taken as evidence for covert movement (Lasnik 1996).

Finally, let me return very briefly to the question of whether the grammar is a

derivational or a representational system (discussed in Chapter 2 (§ 6) in connection with the Copy Hypothesis). It may be significant to note that as well as seeming to be formulable only within a nontrivially derivational model, the theory of optionality developed in Part II also relies critically on the Chomskyan notion that Spell-Out operates at any point during the syntactic derivation, rather than applying simply to LF only (cf. O'Neill and Groat 1996). If successful, this theory of optionality would then provide some unusual support for a model of the Chomskyan type.

References

Abney, Steven (1987) The English Noun Phrase in its sentential aspect. Ph.D thesis, MIT.

Adger, David (1994a) Functional Heads and Interpretation. Ph.D thesis, University of Edinburgh

Adger, David (1994b) Movement, Meaning and Economy. In *Proceedings of WCCFL 13.*

Åfarli, T. (1995) Dimensions of phrase structure: the representation of adverbial phrases. Abstract from Workshop on Optionality, OTS/Utrecht University.

Aoun, Joseph, Elabbas Benmamoun and Dominique Sportiche (1994) Agreement and conjunction in some varieties of Arabic. *Linguistic Inquiry* 25, 195-220.

Baker, Mark (1991) On some Subject/Object Non-Asymmetries in Mohawk. *Natural Language and Linguistic Theory* 9:4, 537-576.

Barss, Andrew (1986) Chains and anaphoric dependence: on Reconstruction and its implications. Ph.D. thesis, MIT.

Barss, Andrew (1994) Derivations and Reconstruction. In *Studies in the Linguistic Sciences*, vol. 24, no.s 1/2. University of Maryland.

Belletti, Adriana and Luigi Rizzi (1988) Psych-Verbs and θ-theory. *Natural Language and LinguisticTheory* 6:3; 291-352.

Bennis, Hans (1986) *Gaps and Dummies.* Foris: Dordrecht.

Bobaljik, Jonathan (1995) Morphosyntax: The Syntax of Verbal Inflection. Ph.D thesis, MIT.

Borer, Hagit (1984) *Parametric Syntax: case studies in Semitic and Romance languages.* Foris, Dordrecht.

Borer, Hagit (1993) The Projection of arguments. ms. UMass.

Bowers, John (1993) The Syntax of Predication. *Linguistic Inquiry.*

Branigan, Phillip (1991) Variables, Case and economy. In L.Cheng and H. Demirdache, eds., *MIT Working Papers in Linguistics 15: More Papers on Wh-movement.*

Branigan, Phillip (1992) Subjects and Complementizers. Doctoral dissertation, MIT.

Branigan, Philip and Chris Collins (1993) Verb movement and the quotative construction in English. In *MIT Working Papers in Linguistics* 18: *Papers on Case and Agreement* 1.

Brody, Michael (1990) Some Remarks on the Focus Field in Hungarian. In J. Harris, ed. *UCL Working Papers in Linguistics 2*.

Brody, Michael (1992) A note on the Organization of the Grammar. In *UCL Working Papers in Linguistics 4*, H. van de Koot, ed.

Brody, Michael (1993) Θ-Theory and Arguments. *Linguistic Inquiry*

Brody, Michael (1994) Lexico-Logical Form: A Radically Minimalist Theory. Ms, University College London.

Brody, Michael (1994) Phrase Structure and Dependence. In *UCL Working Papers in Linguistics 6*, J. Harris, ed.

Brody, Michael (1995) Perfect Syntax. ms., UCL.

Brody, Michael (1996) Class lecture notes, UCL.

Burzio, Luigi (1986) *Italian Syntax*. Dordrecht: Reidel.

Cheng, Lisa L.-S. (1991) On the typology of *wh*-questions. Ph.D. thesis, MIT.

Chomsky, Noam (1965) *Aspects of the Theory of Syntax*. MIT Press, Cambridge, Mass.

Chomsky, Noam (1973) Conditions on transformations. In *A festschrift for Morris Halle*, S.Anderson and P. Kiparsky, eds., 232-286. New York: Holt, Rinehart and Winston.

Chomsky, Noam (1980) *Rules and Representations*. New York: Columbia University Press.

Chomsky, Noam (1981) *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, Noam (1982) *Some Concepts and Consequences of the Theory of Government and Binding*. MIT Press, Cambridge, Mass.

Chomsky, Noam (1986a) *Knowledge of Language*. New York: Praeger.

Chomsky, Noam (1986b) *Barriers*. Cambridge: MIT Press.

Chomsky, Noam (1991) Some Notes on Economy of Derivation and Representation. In R.Freidin, ed., *Principles and Parameters in Comparative Grammar*. MIT Press, Cambridge, Mass.

Chomsky, Noam (1993) A Minimalist Program for Linguistic Theory. In K. Hale and S.J. Keyser, eds., *TheView from Building 20.* MIT Press, Cambridge, Mass.

Chomsky, Noam (1994) Bare Phrase Structure. In G. Webelhuth, ed. (1995), *Government and BindingTheory and the Minimalist Program*.Oxford: Blackwell.

Chomsky, Noam (1995) Categories and Transformations. Ms, MIT.

Chomsky, Noam (1995) *The Minimalist Program*. Cambridge: MIT Press.

Cinque, Guglielmo (1990) *Types of A '-dependencies*. Cambridge: MIT Press.

Cinque, Guglielmo (1993) A null theory of phrase and compound stress. *Linguistic Inquiry* 24.2, 239-297.

Collins, Chris (1994) Economy of Derivation and the Generalized Proper Binding Condition. *Linguistic Inquiry* 25, 45-61.

Collins, Chris and HüskuldurThráinsson (1994) Object Shift in Double Object constructions and the Theory of Case. In C. Phillips, ed. *MIT Working Papers in Linguistics* 19: *Papers on Case and Agreement II*, 131-174.

Costa, Joao (1995) Adverb positioning and V-movement in English: some more evidence. *Studia Linguistica.*

Delfitto, Denis and Norbert Corver (1995) Feature Primitives and the Syntax of Specificity. ms. OTS/University of Utrecht & CLS/Tilburg University.

Den Dikken, Marcel (1992) *Particles*. Dordrecht.

Den Dikken, Marcel (1994) Auxiliaries and Participles. In M. Gonzalez, ed., *Proceedings of NELS* 24 Vol. 1, 65-79.

Déprez, Viviane (1989) On the Typology of Syntactic Positions and the Nature of Chains: Move α to the Specifier of Functional Projections. Ph.D. thesis, MIT.

Diesing, Molly (1992) *Indefinites*. Cambridge: MIT Press.

Diesing, Molly and Eloise Jelinek (1993) Distributing arguments. Ms.

Dowty, David and Belinda Brodie (1984) A semantic analysis of "floated" quantifiers in a transformationless grammar. In M. Cobler, S. MacKaye and M.T.Wescoat, eds.,*Proceedings of WCCFL,* Stanford Linguistics Association.

Enç, Mürvet (1991) The semantics of specificity. *Linguistic Inquiry* 22,1-25.

Ferguson, K. Scott and Erich Groat (1995) Redefining Shortest Move. Ms., Harvard University.

Fukui, Naoki (1993) Parameters and Optionality. *Linguistic Inquiry* 24, 399-420.

Grewendorf, Günther and Wolfgang Sternefeld, eds. (1990) *Scrambling and Barriers*. Amsterdam: John Benjamins.

Haegeman, Liliane (1994) *Introduction to Government and Binding Theory*. Oxford: Blackwell.

Heim, Irene (1980) File change semantics and the familiarity theory of definiteness. In R. Bäuerle, C. Schwarze and A. von Stechow, eds. *Meaning, Use and Interpretation of Language*. Berlin: de Gruyter.

Holmberg, Anders (1994) The Pros and Cons of Agreement in Scandinavian Impersonals. In *Paths Towards Universal Grammar. Studies in Honor of Richard S. Kayne*, G.Cinque et al (eds.), 217-236. Georgetown University Press.

Hornstein, Norbert (1994) An argument for Minimalism: the case of Antecedent Contained Deletion. *Linguistic Inquiry* 25, 455-480.

Huang, James C.T. (1982) Logical relations in Chinese and the theory of grammar. Ph.D. thesis, MIT.

Huang, James C.T. (1993) Reconstruction and the structure of VP: some theoretical consequences. *Linguistic Inquiry* 24, 69-102.

Johnson, Kyle (1991) Object Posititions. *Natural Language and Linguistic Theory* 9, 577-636.

Johnson, David and Shalom Lappin (forthcoming) A Critique of the Minimalist Program. Ms., SOAS.

Jonas, Diane and Jonathan Bobaljik (1992) Specs for Subjects: The Role of TP in Icelandic. Ms,MIT.

Kayne, Richard (1975) *French Syntax: The Transformational Cycle*. Cambridge: MIT press.

Kayne, Richard. (1984) *Connectedness and Binary Branching*. Foris, Dordrecht.

Kayne, Richard (1984) Principles of Particle Constructions. In *Grammatical Representations*, J. Guéron et al (eds.). Foris: Dordrecht.

Kayne, Richard (1989) Facets of Romance Past Participle Agreement. In *Dialect Variation in the Theory of Grammar*, Paola Benincà (ed.) Foris, Dordrecht.

Kayne, Richard (1993) Toward a modular theory of auxiliary selection. *Studia Linguistica* 47, 3-31.

Kayne, Richard (1994) *The Antisymmetry of Syntax*. MIT Press, Cambridge, Mass.

Kitahara, Hitatsugu (1993) Checking theory and scope interpretation without Quantifier Raising. ms., Harvard University.

Kitahara, Hisatsugu (1994a) *Target α: A theory of movement and structure-building.* Ph.D.thesis, Harvard University.

Kitahara, Hisatsugu (1994b) Target α: Deducing strict cyclicity from principles of derivational economy. *Linguistic Inquiry*.

Kitahara, Hisatsugu (1994c) A Minimalist analysis of cross-linguistically variant CED phenomena. In M. Gonzalez, ed., *Proceedings of NELS 24* Vol. 1, 241-253.

Koizumi, Masatoshi (1993) Object Agreement Phrases and the Split VP Hypothesis. Ms. MIT.

Koizumi, Masatoshi (1994) Layered Specifiers. In M. Gonzalez, ed., *Proceedings of NELS 24* Vol. 1, 255-269.

Koizumi, Masatoshi (1995) Phrase Structure in Minimalist Syntax. Ph.D. thesis, MIT.

Koopman, Hilda (1995) On verbs that fail to undergo V-second. *Linguistic Inquiry* 26.1, 137-163

Koopman, Hilda and Dominique Sportiche (1991) The Position of Subjects. *Lingua* 85: 211-258.

Koot, Hans van de (1994) Class handout, University College London.

Koot, Hans van de (1994) On the status of the Projection Principle in the Minimalist Program. In J. Harris, ed. *UCL Working Papers in Linguistics* 6.

Kuroda, S.Yuki (1988) Whether We Agree or Not: A Comparative Syntax of English and Japanese.*Linguisticae Investigationes* 12: 1-47.

Larson, Richard (1988a) On the Double Object Construction. *Linguistic Inquiry* 19.3, 335-392.

Larson, Richard (1988b) Light Predicate Raising. Ms.

Lasnik, Howard (1995) A note on pseudogapping. In *MIT Working Papers in Linguistics* 27. 143-163.

Lasnik, Howard (1995) Case and expletives revisited: on Greed and other human failings. *Linguistic Inquiry* 26.1, 615-633.

Lasnik, Howard (1996) Levels of representation and the elements of anaphora. ms University of Connecticut.

Lasnik, Howard and Mamoru Saito (1991) On the subject of infinitives. In Lise Dobrin et al, eds. *CLS 27: Papers from the 27th Regional Meeting of the Chicago Linguistic Society*. Chicago Linguistic Society.

Lasnik, Howard and Mamoru Saito (1992) *Move α.: Conditions on Its Application and Output*. Cambridge: MIT Press.

Lasnik, Howard and Tim Stowell (1991) Weakest Crossover. *Linguistic Inquiry*

Lee, R.K. (1995) Economy of Representation. Ph.D. thesis, U.Conn.

Li, Yafei (1990) Conditions on X^0-movement. Ph.D thesis, MIT.

Maeda, Taeko (1995) Numeral Quantifiers. ms., SOAS.

Mahajan, Anoop (1989) Agreement and Agreement Phrases. In I.Laka and A.Mahajan, eds. *MIT Working Papers in Linguistics* 10: *Functional Heads and Clause Structure.*

Mahajan, Anoop (1990) The A/A' Movement Distinction and Movement Theory. Ph.D thesis, MIT.

Mahajan, Anoop (1991) Operator movement, agreement and referentiality. In L.Cheng and H. Demirdache, eds., *MIT Working Papers in Linguistics 15: More Papers on Wh-movement.*

Manzini, M. Rita (1994) *Locality*. MIT Press, Cambridge, Mass.

Manzini, M.Rita (1995) UCL Working paper.

Marantz, Alec (1995) Minimalist Program. In G. Webelhuth, ed., *Government and Binding Theory and the Minimalist Program*. Oxford: Blackwell.

May, Robert (1985) *Logical Form*. Cambridge: MIT Press.

Merchant, Jason (1995) Floating Quantifiers paper. Proceedings of *NELS*.

Müller, Gereon (1995) Economy, Optionality, and the concept of Form Chain. Abstract from Workshop on Optionality, OTS/Utrecht University.

Obenauer, Hans-Georg (1992) L'interprétation des structures *wh* et l'accord du participe passé. In Hans-Georg Obenauer and Anne Zribi-Hertz (eds.) *Structure de la Phrase et Théorie du Liage.* Editions CNRS, Paris.

O'Neill, John and Erich Groat (1995) Unifying Spell-Out and the LF interface in the Minimalist framework. Abstract, *GLOW* Newsletter..

Pesetsky, David (1989) Language particular processes and the Earliness principle. Ms., MIT.

Pesetsky, David (1995) *Zero Syntax*. MIT Press, Cambridge, Mass.

Pettiward, Anna (1994) Two-step chain formation: the Case of *wh*-trace. Paper presented at Girona International Summer School in Linguistics Workshop, Girona.

Pettiward, Anna (1995) *Wh*-movement and (C)overt Object Shift. In Sean Jensen (ed.) *SOAS Working Papers in Linguistics 5*

Pettiward, Anna (1996) Anti-Procrastinate effects and optional agreement in French. In S. Ploch and A. Simpson (eds.) *SOAS Working Papers in Linguistics 6.*

Pollock, Jean-Yves (1989) Verb movement, Universal Grammar and the structure of IP. *Linguistic Inquiry* 20, 365-424.

Postal, Paul (1974) *On Raising*. Cambridge: MIT Press.

Reinhart, Tanya (1993) *Wh*-in-situ in the framework of the Minimalist Program. ms., Tel Aviv University.

Reinhart, Tanya (1995) *Interface Strategies*. University of Utrecht.

Riemsdijk, Henk van and Edwin Williams (1986) *Introduction to the Theory of Grammar*. Cambridge: MIT Press.

Rizzi, Luigi (1990a) *Relativized Minimality*. MIT Press, Cambridge, Mass.

Rizzi, Luigi (1990b) Speculation on verb second. In Mascaró, J. and M. Nespor, eds. *GLOW Essays for Henk van Riemsdijk*. Dordrecht: Foris.

Roberts, Ian (1994) Object movement and verb movement in Early Modern English. Ms., University of Wales.

Roberts, Ian (1994) Agreement and object clitics in Franco-ProvençalValdôtain. In *PathsTowards Universal Grammar. Studies in Honor of Richard S. Kayne*, G.Cinque et al (eds.), 217-236. Georgetown University Press.

Ross, John R. (1967/1896) *Infinite Syntax!* Ablex: New Jersey.

Roussou, Anna (1994) *That*-t phenomena: a non-head government approach. In *Langues et Grammaire, Actes du Premier Colloque*, L. Nash and G. Tsoulas, eds.

Rouveret, Alain and Jean-Roger Vergnaud (1981) Specifying reference to the subject: French causatives and conditions on representations. *Linguistic Inquiry* 11, 97-202.

Sabel, Joachim (1996) Vacuous V-to-Infl in German. Ms., University of Frankfurt.

Saito, Mamoru (1989) Scrambling as semantically vacuous A'-movement. In M. Baltin and A. Kroch, eds., *Alternative conceptions of phrase structure*. Oxford University Press.

Shlonsky, Ur (1991) Quantifiers as functional heads: a study of quantifier float in Hebrew. *Lingua* 84, 159-180.

Speas, Margaret (1990) *Phrase Structure in natural language*. Kluwer, Dordrecht.

Sperber, Dan and Deirdre Wilson (1986/1995) *Relevance: communication and cognition*. Oxford: Blackwell.

Sportiche, Dominique (1988) A Theory of Floating Quantifiers and its Corollaries for Constituent Structure. *Lingiustic Inquiry* 19: 425-450.

Sportiche, Dominique (1992) Clitic Constructions. ms, UCLA.

Stowell, Tim (1981) Origins of Phrase Structure. Ph.D. thesis, MIT.

Svenonius, Peter (1995) The economy of Particle Shift. Abstract from Workshop on Optionality, OTS/Utrecht University.

Vanden Wyngaerd, Guido (1989) Object Shift as an A-Movement Rule. In P. Branigan et al, eds. MIT Working Papers in Linguistics Volume 11: Papers from the Student Conference in Linguistics 1989.

Watanabe, Akira (1993) Agr-based Case theory and its interaction with the A-bar system. Ph.D dissertation, MIT.

Webelhuth, Gurt (1992) *Principles and Parameters of Syntactic Saturation*. Oxford University Press.

Wilder, Chris and Damir Cavar (1992) Word order variation, verb-movement and economy principles. In *Sprachwissenschaft in Frankfurt* 10.

Williams, Edwin (198?) Against Small Clauses. *Linguistic Inquiry* 14.2, 287-308..

Williams, Edwin (1994) *Thematic Structure in Syntax*. MIT Press, Cambridge, Mass.

Zubizarreta, Maria-Luisa (1994) Word order, prosody and focus. Ms, University of Southern California.

Zwart, Cornelius Jan-Wouter (1993) Dutch Syntax: a Minimalist approach. Groningen Dissertations in Linguistics.