ON THE TYPOLOGY OF SYNTACTIC POSITIONS AND THE NATURE OF CHAINS: Move α to the specifier of functional projections.

by

Viviane M. Déprez

B.A University of Strasbourg, FRANCE (1981)
M.A., Syracuse University (1983)

SUBMITTED TO THE DEPARTMENT OF
LINGUISTICS AND PHILOSOPHY
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

December 1989

c Viviane Déprez 1989

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

Signature of Author Viviane Déprez
Department of Linguistics and Philosophy
December 1989

Certified by
David M. Pesetsky
Thesis Supervisor

Accepted by
Wayne O Neil
Chairman, Departmental Committee

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUN 12 1990

LIBRARIES

HUM.

ON THE TYPOLOGY OF SYNTACTIC POSITIONS AND THE NATURE OF CHAINS:
Move α to the specifier of functional projections

by

Viviane M. Déprez

Submitted to the Department of Linguistics and Philosophy in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in Linguistics

ABSTRACT

This thesis proposes a redefinition of the A/A' dichotomy, a concept central to the Principle and Parameter model of syntax developed by Chomsky (1981,82,85,86,89) and others. It is argued that in a model which incorporates recent hypotheses concerning the basic structure of clauses, such as the VP-internal subject hypothesis and the Split-INFL hypothesis, the A/A' dichotomy is inadequate to properly account for the distribution and properties of chains created by Move α. The central claim of this thesis (Chapter II) is that the dichotomy relevant to the identification of chains types should be expressed in terms of the distinction between [+HR] and [-HR] positions which are defined as follows:

A. α is a Head Related position ([+ HR]) iff
α is a sister to X° or to X', i.e, α is a specifier or
a complement in the X' theoretic sense.

B. α is a non-Head-Related position ([-HR]) otherwise

Chapter I is a brief overview of the evidence for the VP-internal subject hypothesis and the Split-INFL hypothesis, providing the theoretical background against which this thesis is set. Chapter II discusses the various properties of syntactic positions and the clusters of properties which identify two distinct types of chains. The central claim of the thesis is motivated in this chapter.

Chapter III is an in depth crosslinguistic study of the movement of objects (scrambling and Object-shift). It is shown that the chains created by object movement have properties characteristic of [+HR] chains. Object movement is analyzed as movement to the Spec of the various functional projections made available under the Split-INFL hypothesis. It is shown that object movement to [+HR] positions is constrained by the movement of verbal heads, a constraint which is argued to follow from the ECP given a dynamic notion of Minimality. A consideration of properties of the German/Dutch Mittelfeld scrambling leads us to suggest that a binary partition of positions and chains may

be insufficient. We identify a third type of chain which has mixed properties and suggest that this particular type of chain is created by movement to [+HR, -Case] positions.

Chapter IV considers WH-movement ([-HR] chains) and the theory of the ECP. It is argued that contrary to recent assumptions, WH-movement is not always movement to (or through) the Spec of CP (which in our view is a [+HR] position comparable to the Spec of other functional projections). The Spec of CP is available only for subject extractions; other WH-extractions are adjunctions to CP. This proposal accounts for the fact that the que/qui alternation of the French complementizer (and similar alternations in other languages) is restricted to subject extractions and is impossible otherwise. The last part of Chapter IV is devoted to Subjacency. The difference between ECP and Subjacency violations is derived from a recursive application of the ECP after trace deletion. That is ECP applies twice: first at S-structure and then at LF, after trace deletion. The intuitive idea behind this proposal is that chains which violate the ECP both at S-structure and at LF lead to stronger jugements of ungrammaticality (i.e. , standard ECP violations). Chains which violate the ECP only at S-structure lead to mild jugements of ungrammaticality (i.e., standard Subjacency violations). Further distinctions (in this case, between weak and strong Subjacency violations) are accounted for in terms of a calculus of barrier force. We follow Chomsky (1989) in assumming that obligatory trace deletion in non-homogeneous operator-variable chains is due to the Principle of Full Interpretation. This hypothesis is shown to have empirical consequences for the distribution of floating quantifiers.

Thesis Supervisor: David M.Pesetsky
Title: Associate Professor of Linguistics

ACKNOWLEDGMENTS

My deepest gratitude goes to my committee: Noam Chomsky, Howard Lasnik and David Pesetsky; without them this thesis would have never been written. The opportunity to work with such an exceptional group of people has been the most exciting intellectual challenge I have ever had. David helped me transcend the rigidity of my own ideas and explore new territories. Howard provided me with the rigor which transforms shapeless intuitions into concrete proposals. He always understood, inspite of my faulty expression, what is was that I meant to say. (He even understood that I meant "right" when I had written "left" and vice-versa.) My intellectual debt to Noam is beyond what words can express. (Hopefully, it will be apparent throughout this dissertation).

Another member of my (unofficial) committee that I wish to thank warmly is Richie Kayne. I started working with Ritchie back in Paris and he is the very reason why I am still a linguist. I will never forget a certain appointment during which I told him I could not continue to work in linguistics; his solution to my problem was to suggest that I apply to MIT... Richie's monumental work on comparative grammar has influenced my thinking at every step.

Thinking about Paris, I take the opportunity to express my gratitude to the European linguists whom I have had the priviledge to know and learn from: Jean Yves Pollock, Hans Obenauer, Alain Rouveret, Jacqueline Guéron, Luigi Rizzi, Pino Logombardi, Adriana Belletti, Tarald Taraldsen, Manuella Ambar, Eduardo Raposo, Alexandra Giorgi, Andrea Morro ...

My warmest thanks also go to the faculty members at MIT whose door was always open for me: Ken Hale, Richard Larson, Irena Heim, Donca Steriade, Jim Higgibotham, Wayne O'Neil, Ken Wexler. Morris Halle has never failed to support me and to let me understand that he believed in me sometimes more than I did myself. His unfailing support has been a great help.

MIT is such a wonderful place because there are always linguists around who are willing to talk for hours. I have learned so much from discussions with many visiting scholars and students that it is impossible for me to mention them all. Some discussions have had more bearing on this dissertation: I would like to thank particularly Mamoru Saito, who carefully read my Japanese section, Joe Emonds, Eric Reuland, Jane Grimshaw and Irena Heim.

To all my colleagues and friends,(past, present and future generations) who have entertained me as well as educated me, many thanks. To Li Yafei whose wise advice on life and linguistics have helped me get through critical moments, to Kyle Johnson for such good laughs, to Peggy Speas, Maggy Browning, Juan Uriageraka, Fernande Dupuis, Marie Odile Junker, Chris Collins, Betsy Klipple, Tova Rappaport, Pierre Pica, Andy Barss, Kelly Sloan, Eulalia Bonet, Sabine

Iatridou, the Magazine street trio, Kumiko Murasugi, Harry Leder, Beatrice Santorini, Anoop Mahajan, Utpal Lahiri, TROM, Phil Brannigan, John Frampton, Michael Heggarty, Janis Melvold, Scott Meredith, John Lumsden, Doug Saddy. A special thanks goes to Mikka Hoffman for help in proofreading, to the Japanese community of linguists at MIT for 24 hour jugement and comment service, to Arild Hestvik and Gudrun Gunnarsdotir for Scandinavian data, and to Amy Pierce for last minute survival help.

Thanks also to the Athena Language project for summer support in a very friendly work environment.

Finally, my gratitude and love goes to the friends who have enjoyed / suffer with me every moments of this dissertation writting process, even though they did not really know what it was all about, and who have helped me escape with music and poetry, à Véronique et Suzanne, Eric et Isabelle, Pierre et Henri, un grand merci.

This thesis is dedicated to my family.

TABLE OF CONTENTS

CHAPTER I -- THEORETICAL BACKGROUND

1.0.1 Introduction.....9

1.1 The VP internal subject hypothesis
1.1.1 Theoretical arguments for the VPS.....10
1.1.2 Variants of the VPS.....12
1.1.3 Empirical arguments for the VPS.....14
1.1.2.1 INFL as a raising category.....18
1.1.2.2 Languages with S-structure VP subjects.....19
1.1.2.3 Floating Quantifiers.....21

1.2 The Split-INFL hypothesis.....30

1.3 Conclusion of the chapter and overview of the thesis 40

CHAPTER II -- PROPERTIES OF POSITIONS AND CHAINS

2.1 Introduction.....47

2.2 Properties of positions.....54
2.2.1 The syntactic relevance of positions.....54
2.2.1.1 θ-marked positions.....59
2.2.1.2 Case-marked positions.....60
2.2.1.3 Base generated and derived positions 61
2.2.1.4 Adjunct structure.....67
2.2.1.5 Adjunct at S-structure.....69
2.2.1.6 Base generated adjuncts.....71
2.2.1.7 Summary.....76

2.3 Properties of chains.....76
2.3.1 Standard properties.....77
2.3.2 Floating Quantifiers.....86
2.3.2.1 Summary.....76

2.4 Towards a definition of the relevant dichotomy.....96
2.4.1 Properties of the VP external subject position 96
2.4.2 Properties of intermediate specifiers.....106

2.5 Summary of Chapter II.....116

CHAPTER III -- MOVEMENT TO INTERMEDIATE SPECIFIERS

3.0.1 Introduction.....119

3.1 Hindi Scrambling.....121
3.1.1 Word order in Hindi.....123
3.1.2 Sentence internal scrambling.....127
3.1.3 Long distance scrambling.....132

3.1.3.1 Summary.....138

3.2 Japanese Scrambling.....143
3.2.1 WCO in Japanese.....143
3.2.1.1 Word order in Japanese.....143
3.2.1.2 The disapearance of WCO effects.....144
3.3.1.3 Against a PG analysis of WCO repair 148

3.2.2 Scrambling as an instance of AN-chain.....153
3.2.2.1 Vbl-chains vs AN-chains and reconstruction.....153
3.2.2.2 Vbl-chains vs AN-chains : reconstruction and proper binding.....163
3.2.2.3 Anaphoric binding.....173
3.2.2.4 Numeral quantifiers.....175
3.3.3.5 Functional projections in Japanese..185

3.3 Object Shift.....189
3.3.1 Object Shift in Mainland Scandinavian.....190
3.3.2 Object Shift in Icelandic.....200
3.3.3 Evidence against the Head Determinaticn Hypothesis.....208

3.3.4 Object Shift as movement to a [+HR] position..212
3.3.4.1 Verb movement in Scandinavian.....212
3.3.4.2 Universal properties of verb movement to C.....219
3.3.4.3 Clausal structure in Scandinavian...222

3.3.5 Restrictions on Object Shift.....227
3.3.5.1 Dynamic Minimality.....227
3.3.5.2 Pronouns vs NPs.....236
3.3.6 Summary.....242

3.4 German and Dutch Mittelfeld Scrambling.....244
3.4.0.1 Scrambling vs Base generation.....245
3.4.1 Vbl-chain properties of German scrambling.....251
3.4.2 AN-chain properties of German scrambling.....257
3.4.3 Towards a ternary partition of positions.....263
3.4.4 On the nature of the positions which head chains with mixed properties.....269
3.4.5 Functional projections in German.....281
3.4.6 Bare quantifiers in French.....286
3.4.7 Conclusion of chapter III.....300

CHAPTER IV -- THE SPEC OF COMP

4.1 The problem.....305
4.1.0.1 Introduction.....305
4.1.1 Short distance WH and NP movement.....312
4.1.2 Long distance WH and NP movement.....321
4.1.2.1 Long distance WH movement and intermediate traces.....322
4.1.2.2 Weak Cross over.....324

4.1.2.3 Super Raising.....325

4.2 The ECP: a first approach.....328
4.2.1 What counts as a barrier?.....329
4.2.1.1 The barrierhood of IP.....329
4.2.1.2 The barrierhood of VP.....335
4.2.2 What can be adjoined to? Towards an elimination of the ban on CP adjunction.....336
4.2.2.1 WH-islands.....338
4.2.2.2 Adjunct islands.....347
4.2.2.3 Subject islands.....349
4.2.2.4 CNPC islands.....358
4.2.2.5 Conclusion.....361

4.3 Towards a definition of the ECP.....364
4.3.1 C-command vs M-command.....366
4.3.2 Relativized Minimality vs Dynamic Minimality..373

4.4 Case study.....380
4.4.1 Extraction out of sentential complement.....381
4.4.2 WH-islands.....394
4.4.3 Adjunct islands.....397
4.4.4 Subject islands.....406
4.4.5 CNPC islands.....410

4.5 Consequences.....415
4.5.1 Solutions to 4.1.2.....415
4.5.2 Super Raising.....418
4.5.3 Agreement in CP across languages.....425

4.6 Case marking into CP.....430
4.6.1 Infinitival relatives.....430
4.6.2 Variables and Case-marking.....447

4.7 Subjacency.....464
4.7.1 Trace deletion.....466
4.7.1.1 Referential indices.....469
4.7.1.2 Quantifier Float and Operator variable chains.....477
4.7.2 Subjacency.....481
4.7.2.1 ECP vs Subjacency.....481
4.7.2.2 Weak vs Strong Subjacency.....494
4.8 Conclusions.....498

REFERENCES.....500

In the "principle and parameters" model of syntactic theory, it is standardly assumed that the structure of a clause is as in (1):

(1) [IP NP [I' I [VP V NP]]]

Given this structure, and the general assumption that in UG most relations have a very local character two questions arise:

1) How is the NP/IP related to the verb which assigns its θ-role?
2) How is the verb associated with INFL, the category which contains its inflectional features

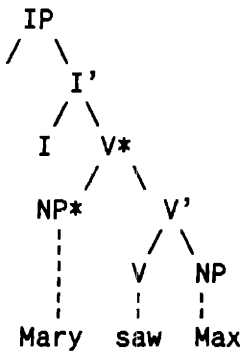
Various answers have been given to these two questions within the principle and parameter framework. As we will see, both the VP-internal-Subject hypothesis and the Split INFL hypothesis are in fact new answers to these two old problems.

1.1 The VP internal-subject hypothesis

A number of proposals in the syntactic literature of recent years converge on the following hypothesis: the sentential subject is generated at D-structure within the maximal projection of the verb (Kuroda (1986), Kitagawa (1986), Fukui and Speas (1986), Koopman (1988), Koopman and Sportiche (1985),(1988) among others.). According to this hypothesis, the D-structure representation of the sentence in (2) is as in (3):

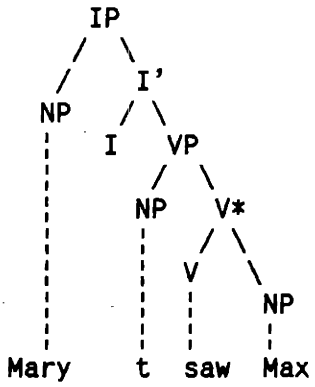
(2) Mary saw Max

(3)



A basic tenet of the VP-internal hypothesis (VPS) is that subjects, like objects are θ-marked within the maximal projection of the verb. Thus NP* is a theta-position. It is assumed by most proponant of the (VPS) however, that in languages such as English and French, NP* is not a Case marked position. A consequence of this assumption is that at S-structure the NP* must raise to the Spec IP position where it receives Case from Agr in INFL. Under this view the S-strucure of (2) is an in (4):

(4)



1

---

1. The spirit of this "raising" analysis can be traced back to Filmore's (1968) in the framework of Case grammar and to Mc Cawley's proposal that English has an underlying V-S-O order.

Various empirical arguments have been put forth in support of this hypothesis some of which we review in what follows. But first and foremost the VP-internal-subject hypothesis has been argued for on the basis of the conceptual simplification it allows with respect to various modules of the theory such as Theta-theory, and X' theory.

1.1.1 Theoretical arguments for VPS

As argued by Chomsky (1986) and many others, the canonical relation of theta assignment between X and Y requires sisterhood. Although this requirement is straightforwardly met in the case of a verb and its complement, the notion of sisterhood must be extended to account for theta-assignment to the subject under the standard clause structure (1), so as to ignore intermediate projections of INFL. Chomsksy (1986)b. proposes the following extension of sisterhood:

(5) α and β are <u>sisters</u> if and only if they are dominated by the same <u>lexical</u> projections

The requirement that the projection be "lexical" allows to bypass INFL in simple sentences. But since INFL can contain lexical material (such as the auxiliaries or the modals in English), the needed extension of the notion of sisterhood cannot straightforwardly rely on differences between lexical and functional categories. Aside this particular problem, it has been argued by Fukui (1986) that such an extension of the notion of sisterhood leads to some empirical problems. It is thus

quite undesirable. [2] Under the VPS, no such extension of the notion of sisterhood is required and theta-assignment can be generally assumed to proceed under strict sisterhood and/or under government. Thus the VPS avoids a unneeded complication in the definition of "sisterhood" and allows a simplification in the module of θ-theory.

It has been argued by Chomsky (1981) and Marantz (1984) among others, that there is an asymmetry between θ-role assignment to a complement and θ-role assignment to a subject: the θ role of a subject is determined "compositionally" by the Verb and its internal argument. This asymmetry has been argued to support the distinction established in the GB model between direct θ-marking (θ-marking of a complement by its head) and indirect θ-marking (θ-marking by an XP). Such an asymmetry however can be easily transposed to a model incorporating the VSP: suffices it to assume that the subject θ-role is assigned by some projection (segment) of V which includes the complement but excludes the subject[3].[4]. Ceteris paribus, the VPS appears to be simpler and to cause no loss of significant distinctions.

Another simplicity argument which has been advanced in support of the VPS involves considerations relative to X' theory (Kuroda 1986). The

---

2. We refer the reader to Fukui (1986) for a detailed discussion.

3. The terms exclusion here is not to be understood in the technical sense given to it by Chomsky (1986)b.

4. This would be accomplished simply by assuming that the requirement for θ-marking is c-command. Perhaps, more generally, the notion of government should be defined in terms of c-command rather than m-command. See Chapter IV section 4.3.1 for a discussion of this issue.

leading idea behind X' theory is the proposal that all categories have essentially the same structure. But although it is assumed that most categories such as IP, CP, AP and NP have specifiers, no role is given in the theory to the specifier of VP[5]. Thus in some sense VP appears to be a "defective" category. The VPS corrects this defectiveness and makes VP similar to other projections. With respect to X'theory, the assumption that there exists a specifier of the VP projection is in fact the null hypothesis.

The general structure of the simplicity arguments for VPS is the following: the VPS allows some simplification in some module of the theory. While simplicity arguments are never strong, since principles should be as complex as they need be (notewitstanding considerations of elegance and simplicity whose metrics are often arbitrary), it is quite clear that the arguments given above have at least the effect of shifting the burden of proof to advocates of the standard clause structure given in (1). Given this situation, one can in fact assume that unless there are evidence to the contrary, the VPS is the null hypothesis.

1.1.2 Variants of the VPS

There are several variants of the VPS which differ essentially on the question of how constituents are organized within the VP projection.

---

5. The notion of specifier is to be understood here as a position which is the sister of an X' projection.

Kuroda (1986) and Kitagawa (1986) assume that the subject is the specifier of VP and thus that in (4) V* = V'. Sportiche and Koopman (1985) (1988) on the other hand assume that the subject is adjoined to the maximal projection of VP and thus that in (4) V* = VP. Finally, Fukui and Speas propose that lexical categories differ from functional categories in never projecting up to the X" level. In their view, V* = V' and there is no VP projection. In the chapters to come (see in particular Chapter III section 3.3.5.1), we will argue that VP is a barrier for certain types of objects movement. Since the notion of barriers is defined in terms of maximal projections (X" level), we must assume that V projects a maximal projection. Thus we do not adopt Fukui and Speas's proposal.

Koopman and Sportiche's (1988) proposal that the subject is generated outside the VP projection, is based on an analysis of object extractions in Dutch.They argue that objects move to the specifiers of VP, a position in which it is assigned Case. We discuss this issue in Chapter III, where we show that the data can be reanalyzed in terms of movement of objects to the specifier of a functional projection. Given this proposal, the idea that in transitive VP, the Spec of VP must be preserved as a possible landing site for movement of the object is no longer necessary[6]. The simplest view is then to assume that the subject is generated as the specifier of VP, a proposal we adopt[7].

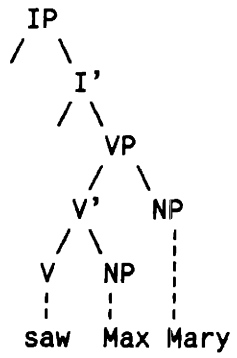
---

6. In fact, the assumption that the VP spec is available for movement in transitive and intransitive constructions is problematic in our view, since it would permit an "escape hatch" for object movement.

7. In chapter II, we return to a discussion of argument projection and of the notion of specifier. We will adopt the view of argument projection proposed by Richard Larson (1988).

Variants of the VPS also differ with respect to the direction in which the subject is assumed to be generated within the VP projection. For Koopman and Sportiche, the direction of the VP subject remains unspecified and is assumed to be determined parametrically by independent principles of linearisation. Kitagawa (1986) proposes on the contrary that the position of the subject is determined by the directionality of theta-marking, a principle argued for independently by Stowell (1981) and Travis (1984). Since the directionality of θ-marking is to the right in English as shown by the position of the object, a consequence of Kitagawa's proposal is that VP-internal subjects in English are generated as right branching sisters of V'. Thus for Kitagawa, the D-structure of (2) is as in (6):

(6)



Support for this hypothesis is drawn from an analysis of so called "extraposed" clauses in English which show overtly the assumed D-structure order. In Kitagawa's view, "extraposed" clauses are in fact generated in the Spec VP position and are allowed to remain there at S-structure because, contrary to NP, clauses do not require Case (Stowell 1981).

Several linguists have adopted Kitagawa's hypothesis that subjects can be generated as a right branching sister of the VP to account for inverted subjects in Romance languages (Bonet (1988), Sportiche and

Koopman (1985) (1988), Contreras (1987). In our own work (Deprez (1988) we have defended the position that, in French, the subject is based generated as a leftward specifier of the VP and that the rightward position of the subject in such contructions as Stylistic Inversion is derived in two distinct ways, either by verb movement over the subject or by a rightward movement of a "heavy subject", leading to two distinct constructions with different properties. Since the question of the directionality of the subject is not direclty relevant to our subsequent dicussion, we will remain neutral with respect to this issue[8]

---

8. A recent study of acquisition data (cf Pierce 1989) suggests that post verbal subjects in early stages in child language are essentially found in languages which have otherwise been argued to manifest V-raising such as French and Spanish. Characteristically, they are quasi inexistant in English with some exceptions involving, mostly, verbs of the unaccusative type. This interesting distinction can be explained if it is assumed that the post verbal position of the subject is a result of Verb movement over a leftward base generated subject and does not reflect the rightward base position of the subject assumed by Kitagawa.
There is independent evidence which support the assumption that verb movement is acquired early. This evidence is based on cross linguistic comparison of the placement of negation in early child grammar (See Deprez and Pierce (1989) forthcoming). As is well known, (Bellugi, Bellugi and Klima) English manifests a first stage in the development of negative sentences which is characterized by the (sentence) initial position of the negation. In French, on the other hand negation appears from the earliest stage in post-verbal position (Weisenborn (1988)) thus suggesting that Verb movement to I is acquired early in French. Assuming that the negation is based generated between the I projection and the VP (as argued by Pollock (1989), the correlation between the occurence of postverbal subjects and the placement of negation can be explained straightforwardly under the following scenario. Assume that both in French and in English, the subject is a left branching specifier of VP. Children fail to raise the VP subject to the IP spec in both languages. The obtained order is Neg S V in English and V Neg S in French. Such results do not follow if the subject is assumed to be generated as a right branching Spec of VP both in French and in English as would be expected under Kittagawa assumption. They do not follow either from the standard view of clause structure and thus they provide additional evidence for the VPS.

1.1.3 Empirical evidence for VPS

The empirical evidence given in this section are mostly taken from Koopman and Sportiche (1988) (henceforth K&S). Other linguists have given additional evidence based on other languages, which we do not review here. This brief review does not pertain to be exhaustive but simply to show that there are empirical evidence in favor of the VPS.

1.1.3.1 INFL as a Raising category

The first type of evidence provided by K&S supports the idea that the surface position of the subject is derived by movement. K&S argue that INFL shows properties similar to those of a canonical raising verb like *seem*. The characteristic properties of raising verbs (RV) are as follows:

(7)

1)an RV imposes no selectional restriction on its subject

2)an RV can take expletive as subject or non expletives

3)an RV allows as subject the NP licenced by the predicate of the clause embedded under it such as idiom chuncks, wheather *it* and existensial *there*

All these properties converge to determine that the hallmark characteristic of an RV is that it does not assign a θ-role to its

subject. Parallel to (7), K&S observe that the properties of INFL (visible when it contains a modal) are identical to those of RVs:

(8)

1) INFL does not assign an external θ-role,
2) INFL allows as subject an NP licenced by a predicate under it

   External argument of a predicate (John will sleep)

   wheather it (it will snow)

   idiom chunks (the cat will be out of the bag)

   Existential there (there will be a griffin on the 22 level)

Based on this identity of properties, Koopman and Sportiche conclude that INFL is a raising category[9].

1.1.3.2 Languages with S-structure VP subjects

A second empirical evidence in support of the VPS is based on the fact that there are some languages which manifest the VPS order overtly in

---

9. Pesetsky (in class 1988) speculated that INFL may in some cases assign an independent θ-role. The cases he considered involved pairs of the type: (i) A group of students has to be there./ (ii) A group of students have to be there, where the verbal agreement is either semantic or grammatical. The alternation permitted patterns with a semantic difference: (ii) cannot have the epistemic reading. This suggests that some autonomous property of INFL may be responsible for theses options. Pesetsky speculated that this may be due to some kind of particular θ-role assigned independently by INFL. If this is correct the raising properties of INFL may resemble more the raising properties of some predicates such a "threaten" or "promise" which seems to optionally assign a θ-role (see Johnson (1986) for an interesting study of the properties of these verbs). We do not pursue these interesting speculations.

the syntax. K&S argue that Irish, Welsh and Arabic are such languages[10].

Irish and Welsh are VSO languages. In sentences with complex tenses, the constituent order is AUX SVO suggesting that the VSO order is derived by verb movement. Verb movement in Irish and Welsh however, differs from verb movement in V/second Germanic languages. Typically V movement in most Germanic languages is in complementary distribution with an overt complementizer. This is not true in Irish/Welsh where the VSO order is possible in embedded sentences with overt complementizers. These two patterns of verb movement can be distinguished respectively as involving a movement of the V to COMP and a movement of the V to INFL. Since Irish/Welsh manifests only V to I, it can be concluded that in AUX S V O structure, the subject must be in its D-structure position within the VP[11]. Similar analysis have been given to various other languages such as Yddish (Diesing (1986), Icelandic (Thainsson (1986)), Old French (Adams (1986) (to account for inversion and null subjects in embedded sentences.[12]) among others.

---

10. K&S also argues that the VP subject may remain in place (although it does not have to) in Chinese, Japanese and Italian. The difference between these languages and French English and Dutch is parametric and is assumed to be due to the obligatory on non obligatory raising status of INFL.

11. If this is correct it provides additional evidence against Kitagawa's proposal. Indeed the directionality of θ marking is to the right for the object in Irish/Welsh. But in sentence where the subject remains in its D-structure position, the subject appears to the left of VP.

12. See also (Deprez (1987) for an analysis of Middle French and Dupuis (1988) for a different view.

The pattern of Arabic verbal agreement also support the VPS hypothesis. Arabic shows an alternance is word order between VSO and SVO. But while in the SVO order the verb agrees overtly with the subject, no such agreement occurs in the VSO order. This pattern of agreement receives an elegant explanation if it is assumed that the non-agreeing post verbal subject has remained in the VPS position while the preverbal subject has raised to the Spec of INFL where is triggers agreement.

1.1.3.3 Floating quantifiers

The last empirical argument for the VPS we consider is based on Sportiche's (1988) analysis of Floating Quantifers (FQ). Since this analysis is of great relevance to the rest of this thesis, we will outline it in some details. Consider the following example in French:

(9) Les enfants ont tous lu ce livre
The children have all read this book

In (9), the quantifier *tous* appears dissociated from the NP it quantifies over. Sportiche proposes that in such constructions the quantifier is adjacent to trace of the D-structure subject, i.e in the VPS position. In other words, the quantifier has been stranded in the VPS position by the movement of the subject to the Spec IP. Thus the structure of (9) is as in (10):

(10) [$_{IP}$ Les enfants$_i$ [$_{I''}$ ont [$_{V*}$ [tous t_i] lu ce livre]]]

Under this view, the position of the floating quantifier provides empirical evidence for the existence and the position of the VPS. Support for this proposal comes from restrictions on the positions in which floating quantifiers can occurs and from the previously observed fact that floating quantifiers behave in some respect like anaphors.

The anaphoric properties of the floating quantifier *tous* are shown by the following paradigm (from Kayne (1984) p 91):

(11) a.*La mère de mes amis est tous partie
The mother of my friends has all left

b.*Mes amis pensent que je suis tous partis
My friends think that I have all left

As noted by Kayne, the ungrammaticality of (11) parallels that of (12):

(12) a.*The mother of my friends likes eachother

b.*My friends think that I like eachother

Examples a. of (11) and (12) show that both anaphors and floating quantifiers need a c-commanding antecedent. Examples b. show that this antecedent must be contained in a certain domain and cannot be too far. Although FQ have anaphoric properties, it seems hardly appropriate to equate them with anaphors since anaphors are elements with referential functions. Under Sportiche's proposal, such an assumption is not needed: FQs modify the trace of an NP which is standardly assumed to have the status of an anaphor, hence the

apparent anaphoric properties[13]. Given this view, (11)a is ruled out because the NP contained in the IP subject never occupied the VP-subject position. Even if we were to assume that movement occured (11)a. would be ruled out because the trace left by the movement fails to be c-commanded by its antecedent and thus is not properly bound. (13)b is excluded for the same reason. Assuming it were possible to generate the IP subject of the main clause in the VP-subject of the embedded predicate, the movement of the NP would violate locality requirements. Sportiche's analysis predicts correctly that the distribution of floating quantifier observes the same limitations as NP movements.

Let us now turn to the possible positions of floating quantifiers in the sentence. Floating quantifiers occur generally between INFL material and VP material:

(13) Les enfants (*tous) ont (tous) vu (*tous) ce film (*tous)
The children (all) have (all) seen this movie (all)

Note first that the jugements reported on the example (13) are our own and not Sportiche's. Sportiche assumes that the VP final position is a possible position for an FQ, although he notes that this position is awkward unless the FQ is modified (presque tous) or heavily stressed. We assume that the VP final position of FQ, whenever possible, is due to an independent process (similar to heavy NP shift) and is not

---

13. The hypothesis that NP movement leaves traces with anaphoric properties has been challenged by a number of linguists (see Barss (1986) and references cited there) This however does not affect the result of Sportiche's analysis which are ameanable to a treatment under the ECP.

relevant for the present discussion. Support for this assumption comes from the fact that under the same conditions (heavy stress an FQ is possible for us in all the stared position in (13):

(14) a.Les enfants, presque tous, sont parti avant la fin
The children almost all have left before the end.

b.Les enfants ont vu, presque tous, ce film.
The children have, almost all, seen this movie

(13) shows that an FQ cannot intervene between a past participle and its object. This follows straightforwadly from Sportiche's analysis, since this position is not a possible D-structure position for the subject. (13) also shows that an FQ cannot intervene between the subject and an auxiliary. As has been argued independently by Emonds (1978) and Pollock (1989), auxiliaries and main verbs move to I in French. Consequently, an AUX always occurs above the highest VP projection at S-structure.

As has often been noted, the restrictions on the distribution of FQ are closely matched by certain types of adverbs. Consider the exemples in (15):

(15)
Les enfants (*certainement) ont (certainement) vu (*certainement) ce film.
The children (certainly) have (certainly) see (certainly) this movie.

Thus it has been suggested that FQ should be treated on a par with adverbs. In addition to the distributional similarity[14], this

---

14. Although similar, the distribution is not identical to that of any particular type of adverbs. As shown in (i) (ii) and (iii) the distribution of FQ differs from that of sentential adverbs and VP adverbs in multi-auxiliary structures:.
(i) Sentential adverbs:
*Ils ont été probablement arrêtés
There have been probably arrested

suggestion is based on two types of evidence. First, in cases such as (16), quantifiers such as *tout* can have an adverbial interpretation equivalent to the meaning of *completely*:

(16)
Ma voiture est toute cassée
My car is all broken

Second, the type of adverbial which most closely matches the behavior of FQ i.e the subject oriented adverbs also appears to have anaphoric properties. Consider (17) and (18):

(17)
Les enfants de mes amis ont intelligement répondu aux questions
The children of my friends intelligently answered the questions

(18)
Mes amis pensent que j'ai intelligement répondu aux questions
My friends think that I have intelligently anwered the questions

In (17), the quality of "intelligence" cannot be attributed to the NP *my friends* but must be attributed to the children. In (18) again, intelligence can only be attributed to the embedded subject and not to the matrix subject. (17) and (18) show that the relation that a subject oriented adverb entertains with the subject it modifies must be local, although presumably no NP trace here is at stake.

Sportiche provides several arguments which shows that FQ can in fact not be assimilated with subject oriented adverbs. First there is crosslinguistic evidence showing that the distribution of FQ and

---

(ii) VP adverbs
*Ils ont mortellement été blessés
They have fatally been wounded
(iii) FQ
Ils ont (tous) été (tous) arrêtés.

subject oriented adverbs are different. As shown by Tellier (1986) and Kinyalolo (1986) respectively, in Moore, a Gur language from Burkina Fasso, and in Kilega a Bantu language from Zaire FQs may intervene between I and VP, but no adverb can. Second, relative ordering between adverbs and FQ in French provides an other argument that the position of FQ and the position of subject oriented adverbs differ. Consider the example in (19):

(19)
Les enfants ont probablement intelligemment tous soigneusement formulé leur demande
The children have probalby intelligently all carefully worded their request

(19) shows the only possible ordering of adverbs and quantifiers when *intelligently* has a subject oriented reading. In Sportiche's view this contraint on ordering follows from a general principle governing the distribution of adverbs, which we reproduce below:

(20)

If some semantic type X "modifies" some semantic type "Y", and X and Y are syntactically realized as α and β, α is projected as adjacent either to β or to the head of β.

Subject oriented adverbs typically modify both a predicate and a subject. Following (20), they should then occur in a position in which they can be adjacent to both. Given the VPS, this requirement can be satisfied, if subject oriented adverbs are adjoined to the VP* projection which contains the subject at D-structure. Assuming that the subject occurs within the VP* projection, it is expected that subject oriented adverb should precede an FQ occuring in the position

of the D-structure subject. Thus the ordering in (19) is correctly predicted.

There are in fact additional and, I believe, more compeling evidence in support of Sportiche's conclusion that FQ and subject oriented adverbs should be treated differently. First as shown in (21), it is possible for some subject oriented adverbs to occur before negation in French:

(21)
Les enfants n'ont intelligemment pas répondu à cette question
The children have intelligently not answered this question

As shown by (22) however, this is never possible for FQs.

(22)
*Les enfants n'ont tous pas répondu à cette question
The children have all not answered this question

Second, it has been noted by Mc Connell Ginet (1982) that subject oriented adverbs can have an ambiguous reading in passive sentences: they can modify either the derived subject or the demoted agent occuring in the by-phrase, depending on the position in which they occur. Her examples are from English but similar examples can be constructed in French. Consider the following pair of sentences:

(23)
a. Ces enfants ont volontairement été instruits par leur parents
These children have willingly been instructed by their parents

b. Ces enfants ont été volontairement instruits par leur parents
These children have been willingly intructed by their parents

(23)a, with the subject oriented adverb occuring between the two auxiliaries can have a reading in which the subject oriented adverb modifies the derived subject. Under this reading, it was the children who were willing to be intructed by theirs parents. (23)b., on the

other, hand favors a reading in which "willingness" is attributed to the underlying agent, namely, the parents. The question of how these two readings come to be available is a difficult one, which we do not attempt to resolve. What is important to our present discussion is that the ambiguity of modification available to subject oriented adverbs in passive constructions is never possible with FQs. Although FQs can occur in the same positions, (i.e between the two auxiliaries and right before the past participle), they can never be interpreted as quantifying over the by-phrase. This is illustrated in (24), where the relation of quantification is encoded with indices:

(24)
a.Ces enfants$_i$ ont tous$_{\{i/*j\}}$ été intruits par leurs parents$_j$
  These children have all been instructed by their parents

b.Ces enfants$_i$ ont été tous$_{\{i/*j\}}$ instruits par leurs parents$_j$
  These children$_i$ have been all instructed by their parents

This difference shows clearly that the relation between an adverb and a subject is fundamentally different from the relation between a FQ and the subject it quantifies over. Additional evidence comes from the fact that while a subject oriented adverb can modify an understood agent this is never possible for a floating quantifier:

(25)
a.*Ces ont tous été délibérément déchirés
  These books were deleberately damaged
b.*Ces livres$_i$ ont tous$_j$ été déchirés

Note that, given some minor extension, which is independently needed, Sportiche's theory of FQ can easily account for the impossibility of (24) and (25). As (24)a. shows FQ are not limited to occurs strictly between INFL material and VP material. They can also occur between two auxiliaries. What (24) shows then, is that FQs are not limited to

occur in the VPS position; rather they seem to be able to occur in various positions along the path of a chain created by an NP movement. Confirmation for this conclusion come from examples involving regular cases of Raising such as (26) and mutiple auxiliary structures:

(26)
Les enfants ont tous semblé avoir compris le problème
The children have all seemed to have understood the problem

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

Quite clearly the positions of the FQs in (26) and (27) cannot be assumed to be that of the VPS subject. This problem was of course noted by Sportiche and he suggests that in this case, a possible assumption is that there is a landing site for NP movement in front of each of the verbs, aspectual auxiliaries included. If this is correct, and I will assume by and large that it is (with some minor modifications (see Chapter II), then the occurence of FQ in sentences like (9) given at the begining of this section does not constitute evidence for the exact position of the VPS. Nevertheless, FQs provide evidence of the existence of a chain; thus they provide evidence for the assumption that the subject is raised to the Spec IP which entails that it must start out lower than I in D-structure. Returning now to examples (24) and (25) above, the impossible relation between an FQ and the by-phrase simply follows from the hypothesis that, as opposed to the IP subject, the by-phrase is not derived by movement. No NP chain is created so that FQ cannot be adjoined to an NP trace. The impossibility of the FQ with an understood subject suggest, moreover, that FQ must be adjoined to empty categories which are <u>syntactically</u>

represented. We will return to the distribution of FQ in several places in this dissertation (Chapter II, III and IV). In Chapter IV section 4.7.1.2, we provide an account of the impossibility of FQ with operator-variable chains on the basis of this requirement.

This concludes our discussion of the Q-float theory and of the VPS hypothesis. As we have seen, the VPS hypothesis provides a principled answer to the first question raised at the begining of this chapter, namely, how the subject relates to the verb which assigns it a θ-role. The answer is simply that a local relation exists at a prior level of representation, namely D-structure. In the next section, we turn to the second question, namely the question of how the verb is related to INFL with which it fuses. As we will see the split INFL hypothesis also provides a principled and well empirically supported answer to this question.

1.2 The SPLIT INFL hypothesis

We have adopted the VPS which entails that the structure of a clause is as in (28):

(28) [$_{IP}$ [$_{I'}$ I [$_{VP}$ NP V NP]]]

Although this structure provides a solution to the first question we raised at the begining of this chapter, the locality problem between the subject and the verb, the second question, how the verb fuses with INFL, remains yet unanswered. In this section, we turn to this second question.

The first principled approach to this question was proposed by Chomsky (1955, 1976). Chomsky (1955) posited a rule called Affix Hopping which relates each verbal element to the morphological element attached to it. We refer the reader to Syntactic Structure and to The Logical Structure of Linguistic theory for a detailed and formal account of this proposal. Jackendoff (1972) (following Emonds and Klima) proposed that in some cases, the fusion between INFL and V is transformationally derived by the raising of auxiliaries (Have/be raising) to an abstract Tense node. "Have/be" raising is formulated by Jackendoff (1972) as in (29):

(29)
Have-be
X - Tense - { have/ be} - Y ===> 1 2+3 4 (obligatory)
1 2 3 4

Emonds (1978) extends the empirical domain of this tranformational rule to all verbs in French. Putting these two proposals in comparative perspective allows an interesting account of some important differences between French and English with respect to the distribution of negation, adverbs and floating quantifiers and the possibility of verb movement over the subject in interrogative sentences. The paradigm given in (30) is now classical:

(30)

a.*John understands not linguistics

b. Jean ne comprends pas la linguistique

c.*John reads often linguistics books

d. Jean lit souvent des livres de linguistique

e.*The children read all comic books

f. Les enfants lisent tous des bandes déssinées

g.*understand you linguistics?

h. Comprends-tu la linguistique?

Assuming that negation adverbs and floating quantifiers are generated in a position between the VP and INFL in both languages, the difference between the French examples and the English examples in (30) can be easily accounted for if as proposed by Emonds, main verbs raise to Tense/INFL in French but not in English. The same rule of raising restricted to Auxiliaries in English accounts for the fact that English auxiliaries manifest the same distribution as French main verbs wrt negation, adverb, FQ and the subject in interrogatives.

(31)

a. John has not understood linguistics

b. John has often read linguistics books

c. The children have all read comic books

d. Has she called?

In LGB, Chomsky (1981) argues that the fusion of INFL with V can occurs through a lowering of the features of INFL onto the verb. According to Chomsky, rule-R which lowers the INFL node onto the V is not a syntactic rule but rather a morphological rule. The major consequence of this assumption is that rule-R does not leave a trace, which would be subject to the ECP. Thus rule-R can occur either at the syntactic level of S-structure or at the PF level. In Chomsky's view, these two possibilities are parametric and account for the differences between languages which accept null subjects and languages which don't (the Pro-drop parameter).

More recently it has been argued (Chomsky (1986)b (1989), among others) that the fusion of V and I is an instance of Move α applying to X^o categories. As convincingly shown by Baker (1986), X^o movement leaves traces which obey the ECP. Travis (1984) observed that X^o movement is subject to the Head Movement Constraint. This contraint essentially prevent the movement of a head X over another head Y as in the shematic structure (32):

(32) [... X+Z ... [... Y... [t]

Under the assumption that X^o leaves traces, the HMC can be subsumed under the ECP [15]. With respect to V movement, the HMC accounts for instance for the ungrammaticality of sentences of the type (33), where an X^o auxiliary has raised over another one, presumably to C.

(33) *been John have t arrested

This that V to I movement is is an instance of X_o movement suggests suggest more generally that a large part of inflectional morphology (may be all) can be considered part of the syntax propoer. Under such a view, there is a sharp distinction between derivational morphology, which is assumed to be part of the lexicon and subject to principles operating in this component and and inflectional morphology which is governed by general syntactic principles part of UG. Following Chomsky (1989) and many others, I will assume this to be essentially correct.

---

15. See Baker (1988) for an extensive study of this hypothesis

Given this approach to inflectional morphology and the assumption that V to I is an X^0 movement, let us now return to the paradigms in (30) and (31). As noted by Pollock (1989), the assumption that this paradigms derives from the movement of main verbs in French and auxiliaries in English raises a number of interesting questions which previous account could not raise and among which we chose to focus on (34)[16]

(34)

1) Why is V movement restricted to Auxiliaries in English but not in French ?

Before we can outline Pollock's answer to this questions, we must extend the range of data taken into account and consider the properties of V movement in French infinitivals. The null assumption concerning the position of elements such as negation and adverbs is that they are identical in all types of sentences, whether finite or non-finite. Assuming this to be correct, Pollock observes that there is an important distinction between verb movement in finite clause and verb movement in non-finite clauses in French. This difference is illustrated in the paradigm (35) and (36):

(35)
a.*Ne lire pas LGB est impensable pour tout bon linguiste
  *To read not LGB is unthinkable for any good linguist

b. Ne pas lire LGB est impensable pour tout bon linguiste
   Not to read LGB is unthinkable for any good linguiste

---

16. See Pollock (1989) and Chomsky (1989) for a more thorough account of the properties of verb movement in French and in English and their consequences for the theory of UG. Throughout this thesis we will return to these analysis.

c. N'avoir pas lu LGB est impensable pour tout bon linguiste
To have not read LGB is unthinkable for any good linguist

d. Ne pas avoir lu LGB est impensable pour tout bon linguiste
To not have read LGB is unthinkable for any good linguist

(36)
a.A peine parler l'Italien après d'études est décourageant
To hardly speak Italien after 5 years of study is discouraging

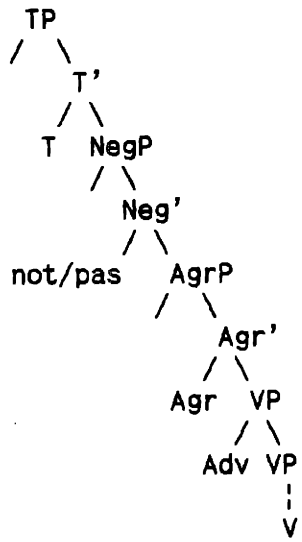
b.Parler à peine l'Italien après 5 ans d'études est décourageant
*To speak hardly Italien after 5 years of study is discouraging

The paradigm in (35) shows that only auxiliary and not main verb can move over negation in French infinitives. In other words verb movement in French infinitives shows by and large the same restriction as verb movement in English tensed sentences. This similarity between the French infinitival paradigm and the English tensed paradigm suggests strongly that the restriction of verb movement to auxiliaries in English cannot be simply due to some idiosyncratic properties of the English auxiliaries.

There are however some differences between verb movement in French infinitives and verb movement in English tensed sentences. First as shown in (35)c., auxiliary movement in French infinitives is not obligatory. Second, the paradigm in (36) shows that although main verbs cannot move over negation, they can still move over adverbs; this movement however remains impossible in English. In other words, verb movement seems to be possible in French infinitives at a short distance over adverbs but not at a longer distance over negation. To account for this complex range of phenomena, Pollock (1989) proposes to subdivide the INFL node into two separate functional projections, Tense and AGR and to concieve of verb movement to I as a succession of movements first to an AGR projection and then to a Tense projection.

Moreover, for reasons which we do not discuss here, Pollock assumes that negation both in French and in English also heads its own projection. Under Pollock's view, the structure of a clause is as in (37):

(37)



The distinctions between French and English verb movement are then accounted for in terms of the properties of the functional projections AGR and T in these two languages. It has been argued independently that AGR is stronger in French than in English[17]. Assuming this to be correct, Pollock proposes that a weak AGR blocks θ-marking but a strong AGR does not. It follows from this proposal that only non θ-marking verbs can raise to AGR in languages with a weak AGR and that all verb can raise to AGR in languages with a strong AGR. Given the

---

17. French subject verb agreement bears features of person and number in a variety of tenses. English subject verb agreement only distinguished the third person singular in the present tense. Moreover, French licences empty expletive subjects under particular circumstances (Cf Pollock (1986) and Déprez (1987)). Cf also Rizzi (1986) for an assessment of the importance of subject verb agreement across languages and its relevance to the licencing of empty subjects.

Head Movement Constraint, further raising to T cannot skip over the AGR projection. Thus unless a verb first moves to AGR, it will not be able to move to T. This proposal provides an elegant answer to the question in (34) and accounts for the paradigms (29) (29) (35) and (36): Agr is weak in English and thus prohibits movement of θ-marking verbs both in tensed and in infinitival sentences. Auxiliaries, being non-θ-marking, verbs are free to move. In French on the other hand, Agr is strong and thus permits movement of all types of verbs. In infinitival sentences, T being non finite is "weak" and thus does not permit the movement of θ-marking verbs. Nothing however prohibits the movement of θ-marking verbs to the Agr projection. This correctly predicts that although θ-marking verbs can precede adverbs in infinitives as shown in (35), they cannot precede the negation as shown in (36).

The proposal that INFL be sudivided into two separate projections appears thus to have desirable and far reaching consequences[18]. Moreover, as noted by Chomsky (1989), it eliminates the odd dual headedness of INFL which is assumed in LGB and subsequent work.

---

18. Since its first proposal, the Split INFL hypothesis has generated a profusion of novel and interesting works on a great variety of languages. Some of them will be mentioned in the following chapters of this dissertation. These works often propose further or different subdivisions of the INFL node. To quote a few, Carsten and Kinyalolo (1988) present evidence for the existence of an Aspectual functional category. Laka (1988) presents evidence for a different ordering of the functional projections in Basque. See also Fassi Fehri (1989) and Demirdache (1989) on Arabic morphology. In our own work, Déprez (1988) we have given further evidence for such a subdivision and we have suggested that a variation in the ordering of Tense and Agr may account for the different behavior of verb movement in French interrogative constructions.

Additional evidence for this hypothesis comes from the fact that in many languages verbal subject agreement and tense are realized as distinct morphemes. But Pollock's particular proposal about the ordering of the Agr and the Tense projection raises a number of questions. Since Agr is, under standard assumptions, the element responsible for Case assignment to the subject, we would expect it to dominate the Tense projection so as to be in a government relation with the subject. Moreover, as observed by Baker (1986), affix ordering appears to be of significant relevance to the syntax. Baker proposes to elevate this observation into a principle, which he calls, the Mirror Principle.

(38) The Mirror Principle:

Morphological derivations must direclty reflect syntactic derivations and (vice versa) (Baker 1988 p 13)

Although the verbal morphology of French is quite impoverished, whenever it is possible to make a distinction, the tense morpheme appears closer to the verbal stem than the subject agreement morpheme. Compare for instance the future tense paradigm of a regular French verb to the simple past

(39)	Futur	simple past
	j'achèterai	j'achetai
	tu achèteras	tu achetas
	il achètera	il acheta

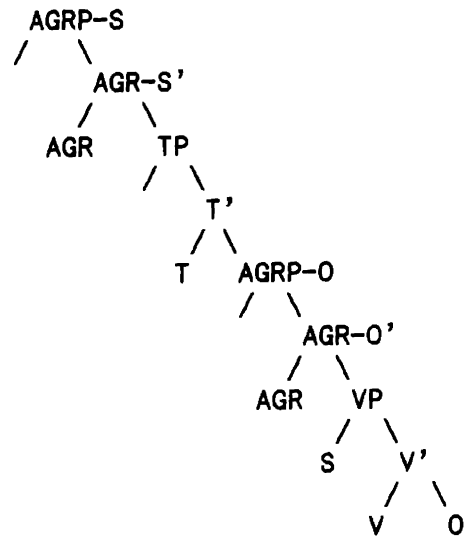
The futur is distinguished from the simple past by the presence of an "r" which is in fact the preservation of the infinitival form. The

person morphemes, however, are the same. Pollock's proposal to order Tense above AGR cannot account for this paradigm. Assuming that the verb moves first to AGR incorporating the subject agreement feature and then to Tense, we would expect the reverse ordering; in other words, Pollock's proposal is as odds with the Mirror Principle. The paradigm in (39) suggests that to obtain the right ordering of the verbal morphology the AGR projection should dominate the Tense projection. The proposal that AGR domintes INFL has been independently porposed by Belletti (1988).[19]. Adopting a structure in which the Agr dominates Tense, however, causes problems for Pollock's results. Recall indeed that Pollock's results followed from the assumption that the AGR head filters out the possible types of verb which can move to T. Chomsky (1989) suggests that the solution to this problem lies in a further sudivision of the INFL node. Noting that a number of languages manifest verbal agreement both with the subject and with the object, Chomsky proposes the existence of an other Agr node hosting the object agreement. The fully developped structure of a sentence is then as in (40)

---

19. This proposal also suggests a possible account for the pro-drop parameter. For instance, it is possible to assume that verb movement to a rich agreement licences an empty subject.(Cf Schlonsky (1989) for a proposal along these lines.

(40)



Given (40), Pollock's result can be assumed to follow from the filtering effect of the Agr-O projection, the Mirror principle is respected and the subject can receive Case in the spec of the Agr-S projection. Independent support for the existence of the Agr-O projection in French comes from the fact that French verbs can manifest overt object agreement in past participle constuctions. We return to this phenomenon in greater details in chapter II. We will also argue (Chapter III) that French manifests overt movement to the Spec of the AGR-O projection. If correct, this proposal lends further support for the existence of this projection.

1.3 Conclusions of the chapter and overview of the thesis

The success of the Split Infl hypothesis to account in a principled way for the complex facts of the distribution of negation and adverbs in French and English in turn lends support for a model in which inflectional morphology proceeds under X° movement and is thus part of

syntax proper. Pollock's approach also lends support to a conception of UG which limits possible variations to the lexicon. In the LGB model, parametric variations were assumed to be able to occur in any of the various components of UG. A considerable amount of the work in generative syntax in recent years has been directed towards an attempt to restrict the class of possible parameters (Borer (1984), (Chomsky (1986)a, b (1989), (Webelhut (1989) among others). It has been suggested, in this respect, that parametric variations do not relate to the computational system of UG but are confined to the lexicon. To quote Chomsksy (1989):

(41)

> "We might take this to mean that each parameter refers to the properties of specific elements of the lexicon or to categories of lexical items; canonical government for example. If this proposal can be maintained in a natural form, there is only one human language, apart from the lexicon, and language acquisition is in essence a matter of determining lexical idiosyncraties. " (Chomsky (1989) p 44).

Chomsky (1989) further points out that the lexicon itself is contrained by universal principles. In particular, "substantive elements (verb, nouns etc...) are drawn from an invariant universal vocabulary". Consequently, parametric variation is limited to the <u>functional elements</u> of the lexicon. Pollock's analysis is directly in the line of this conception of parametric variation since it proposes that the distinctions between French and English word order can be captured by distinctions in the properties of the AGR and Tense projection. This proposal is also in agreement with Borer's (1984) proposal that parametrisation be limited to the "inflectional" component.

In this dissertation, we pursue this line of approach to parametric variation and we provide further support for its crosslinguistic validity. Although in various places in this thesis we consider parametric variations in the properties of functional heads and their consequences (such as Case and Spec Head agreement) (see especially Chapter III), our attention focuses rather on the properties of the syntactic positions which are dependent on functional heads. The null hypothesis is that functional heads such as COMP, DET, and now AGR and Tense project categories in conformity with X' theory; like other lexical heads, they project a specifier and a complement position. Assuming this to be correct, a question arises with respect to the properties of these specifier and complement positions. In the LGB model and subsequent work, it is assumed that properties of syntactic positions determine the nature of chains or more exactly, the nature of the empty categories bound by an element in these positions. Thus empty categories bound by elements in A positions are assumed to be "anaphors" while empty categories bound by elements in A' positions are assumed to be "variables". The nature of the position of the head of a chain thus determines the properties of the chain. A'-chains have been associated with a cluster of properties which includes sensibilty to cross-over (weak and strong), parasitic gap licencing, sensibility to subjacency violations etc... A-chains on the other hand are characterized by the opposite properties. Morever they have been shown to be subject to more stringent locality properties. The central question we address in this thesis, is that of the nature of the specifier and complement position of functional heads and their consequences for the theory of chains.

Chapter II of this thesis contains a theoretical discussion of the properties of positions and chains and a proposal to adequately derive the latter from the former in a model which incorporates the VPS and the Split INFL hypothesis. We argue that in such a model, the A/A' dichotomy as defined by Chomsky (1981) and subsequent work is inadequate to achieve the necessary distinctions. We propose that the relevant distinction can be made on the basis of a dichotomy between [+HR] and [-HR] positions defined as follows:

(42) α is a Head related position ([+HR] position) iff α is a sister to X° or to X'.
α is a non Head related position ([-HR] position) otherwise.

We further argue that [+HR] positions head chains with anaphoric properties, while [-HR] positions head chains with variables properties. A consequence of this definition is that specifiers of functional projections head chains with anaphoric properties. This adequately accounts for the properties of the chain created by movement to the Spec of AGR-S in the structure (40) above. In chapter III, we turn to instances of movement to the specifier of intermediate projections such as AGR-O. Chapter III is an in depht study of the movement of "objects" in various languages. We consider scrambling in Hindi and Japanese, Object shift in Mainland Scandinavian, Icelandic, and Dutch /German Mittelfeld scrambling. We show that across these various languages, object movement creates a chain with anaphoric properties. Under our proposal this follows naturally if these object movements are analyzed as movements to the specifier of intermediate

functional projections. The proposed analysis, in turn, provides considerable empirical support for the validity of the hypothesis of a highly articulated clausal structure which comprises a number of distinct functional projections. Our analysis suggests that the exact number, the precise labelling and the ordering of these functional projections may vary across languages. An account of this particular aspect of the Split INFL hypothesis, will have to await more detailed work on the inflectional morphology of particular languages within this theoretical framework. We suggest however, that certain variations of functional heads with respect to Case assignment have consequences on the chains headed by their [+HR] positions. A consideration of German/ Dutch Mittlefeld scrambling leads us to suggest following Webelhut (1989) that a binary partition of positions and chains may not be sufficient. We identify a third type of chains which appears to combine properties of both anaphoric chains and variable chains. We speculate that this third type of chains is created by movement to a [+HR, -Case] position, supporting this proposal with evidence from bare quantifier constructions in French. Finally we propose a universal constraint on "object" movement to the specifier of intermediate functional projections : in our view, object movement to the spec of a functional projection must be parasistic on the movement of the verb. We propose that this constraint simply follows from the ECP.

Chapter IV is concerned with the properties of the specifier and complement projection of COMP. Given the definition of [+HR] position motivated in Chapter II, movement to the Spec of CP is expected to

create chains with anaphoric properties. We argue that this is indeed the case and that, contrary to current assumptions, WH-movement does not always involves a movement to or through the Spec of CP. In our view, the spec of CP is available only to local subject extractions; other instances of WH-extractions involve an adjunction to CP and thus create [-HR] chains. This restriction on movement to the Spec of CP follows from the assumption that the complement of COMP differs in certain respects from the complement of other functional projection: it is not L-marked. This, in our view, follows from the assumption that L-marking for functional projections depends on the relation that a functional heads entertains with the verbal projection. COMP differs from other functional heads in that it is not part of the inflectional system of the verbal projection. However, when COMP as a result of Spec Head agreement bears the same index as the projection it governs, it becomes similar to other functional projection and can then L-mark its complement. This proposal provides a principled account of the restriction of the que/qui effect in French (and of similar effects in other languages) to local subject extractions. It also permits an account of the "surprising subject/object asymmetries" discovered by Pesetsky (1983). We propose a number of modifications to the Barriers (Chomsky (1986)b.) theory of ECP to accomodate this view. In particular, we argue that the ban on CP adjunction which in the Barrier's framework is essential to an account of island violations can be dispensed with. Adjunction is generally free but intermediate traces as well as terminal traces must meet a dual requirement of antecedent government and head-government.

The last part of chapter IV discusses Subjacency. We propose that the weaker violations involved in argument operator variable chains derive from a recursive application of the ECP. ECP applies first at S-stucture and then again at LF after trace deletion. Chains which violate the ECP both at S-structure and at LF lead to severe violations (i.e standard ECP violations) while chains which violate the ECP only at S-structure lead to mild violations (standard Subjacency violations). This proposal provides a natural explanation for the strenght of ECP violations: the effect derives from cumulative violations. Further distinction between weak and strong Subjacency violations are accounted for in term of the calculus of what we call "the force of a barrier". As proposed by Chomsky (1986)b. maximal projections can be barriers because they are blocking categories or by inheritance. We suggest that maximal projections that are barriers under both counts lead to strong subjacency violations while maximal projections that are barrier under one count lead to mild subjacency violations.

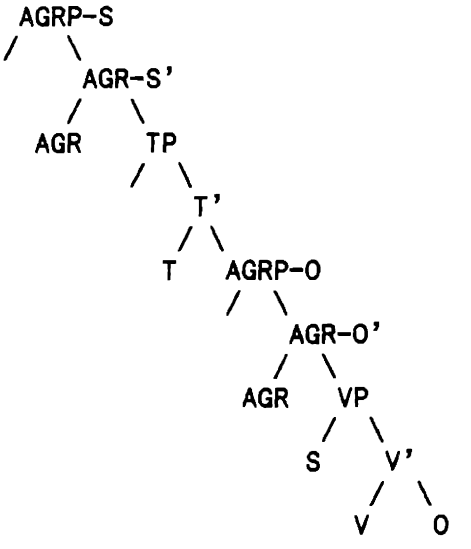
Chapter 2

Properties of positions and chains

2.1 Introduction

In the first chapter, we have presented arguments supporting two recent proposals for the structure of clauses. Following Kuroda (1985), Fukui and Speas (1986), Kitagawa (1986), Sportiche (1988), and Koopman and Sportiche (1988) among others, we have adopted the hypothesis that an external argument is generated within the verb phrase at D-structure and raises to the specifier of a functional projection where it is assigned Case. Following Pollock (1988) and Chomsky (1989), we have also adopted the hypothesis that the structure of sentences involves a number of functional projections to which verbs can move. In particular, we have adopted Chomsky (1989) proposal that the basic structure of the sentence is as follows:

(1)



As has been noted (Kitagawa (1986), Fukui & Speas (1986), Koopman & Sportiche (1988), among others), the VP-internal-subject hypothesis calls into question the traditional division between A- and A'-positions. In the LGB model, Chomsky defines A-positions as follows:

(2)

> "An A-position is one in which an argument such as a name or a variable may appear at D-structure; it is a potential θ-position. " (LGB p 47)

As entailed by the VP-internal-subject hypothesis, however, all θ-roles that are assigned by a verb are assigned internal to the VP projection. Under this view, the position to which the external argument moves at S-structure is never a potential θ-position. Thus, under the LGB definition of A-positions given in (2), the Spec of AGRP is not an A-position but an A'-position. If so, the movement of the subject to the Spec of AGRP is an instance of movement to an A'position under the LGB definition of this term, that is, it is a

movement to a position which is not a potential θ-position. This leads us to expect that raising of the subject to the Spec of AGRP will exhibit the properties standardly associated with other instances of movement to positions which are not potential θ-positions, such as topicalization or wh-movement. But as we show in the following section, this is incorrect. Movement of the subject from its VP-internal position to its surface position displays the cluster of properties which in the LGB framework is associated with cases of movement to an A-position: it involves movement from a non-Case-marked position to a Case-marked position, it does not license parasitic gaps, and it does not induce crossover violations. Coupled with the VP-internal hypothesis, the standard definition of A-positions seems thus to lead to a paradox: movement to the Spec of AGRP is movement to an A'-position under the definition (2) of A- and A'-positions, but it has the properties associated with movement to an A-position.

Given the structure (1) for a basic sentence, similar problems arise for cases of movement through or to the specifiers of intermediate functional projections. [1] Like the specifiers of the upper AGR projection, the specifiers of intermediate functional projections are neither θ-marked positions nor potentially θ-marked positions. Consequently, under the LGB definition of A- and A'-positions in (2), they are A'-positions. To see whether this is a desired consequence we need to investigate the properties of movement through or to these

---

1. See chapter III for arguments for the existence of such projections. Putting aside any empirical evidence confirming their presence, the existence of these positions is theoretically ensured since they are derivable from X' theory.

positions. In the last subsection of this chapter, we will consider cases of movement through these positions in constructions such as Raising, Passive and Clitic movement which induce past participle agreement in French. In Chapter III, we will con sider data from various languages and we will argue that movement ot the specifier of intermediate position is instanciated in a variety of language and, shows properties standardly assumed to be properties of A-movement. If this is correct, the paradox for the A/A' distinction, noted above for the movement of the subject to the Spec of AGRP, arises as well with the specifier position of intermediate functional projections.

The apparent paradox raised by the VP-internal-subject hypothesis, on the one hand, and by the hypothesis that the basic structure of the sentence contains several functional projections, on the other hand, is due in part to a terminological ambiguity inherent to the A/A' distinction defined in the LGB framework. In the LGB framework, the A/A' distinction covers two distinct concepts. First it serves to distinguish two different types of positions as in the definition (2) above. Second, it also serves to distinguish two different types of chains, namely A-chains and A'-chains. A-chains and A'-chains have been argued to manifest different clusters of properties: A-chains are subject to strict locality conditions and do not induce crossover effects or license parasitic gaps. A'-chains, on the other hand, show crossover effects, license parasitic gaps, and are non-local[2]. Thus, in the LGB framework of Chomsky (1981) the A/A' distinction refers ambiguously to the nature of the position of the head of a chain and

---

2. We review these properties in detail in section 2.3.

to a cluster of properties associated with the chains created by movement to these positions. This ambiguity is unproblematic in the LGB framework, where it is assumed that the cluster of properties of a given chain strictly correlates with the nature of the position of the head of the chain: chains which manifest the cluster of properties associated with A-chains have their heads in potential θ-positions, i.e, A-positions; chains which manifest the cluster of properties associated with A'-chains do not.[3] In the cases mentioned above, though, the strict correlation does not hold; we have chains which exhibit the cluster of properties of the LGB A-chains but whose heads do not occur in potential θ-positions. Hence the apparent paradox.

Recent literature (Cf. Koopman & Sportiche (1988), Holmberg (1986), Mahajan (1988)) on the VP-internal-subject hypothesis has added to the terminological confusion by using the term A/A' simply as a name to refer (often implicitly) to the cluster of

---

3. There is in fact a well known case where this strict correlation breaks down, namely the case of clitic chains. Clitic chains manifest the cluster of properties associated with A-chains in LGB: they are subject to strict locality restrictions and they do not license parasitic gaps. This is shown by the French examples (i) and (ii) below:

(i) *Je le souhaite rencontrer
I him hope to meet

(ii) *(Ce livre), je l'ai rangé avant d'avoir lu
(this book), I it have put away before to have read
(This book),I put it away before having read

Despite the ungrammaticality of (i) and (ii), the head of a clitic chain does not occur in a potential θ-position. This problem is acknowledged in Chomsky (1982) and has been noted by several linguists (in particular Kayne (1984) and Taraldsen (1986)).

properties associated with the two types of chains, and/or as a name to refer to a newly defined notion of what we have called "the nature of the position".

To avoid terminological confusion in the forthcoming discussion, we need to distinguish the two uses of the term A/A' which are assimilated in the LGB model. We will refer to the cluster of properties associated with the LGB A-chains and LGB A'-chains as the anaphoric-cluster (AN-cluster) and the variable-cluster (Vbl-cluster). Thus, we will say that NP-movement in Passive and Raising constructions creates instances of chains which display the AN-cluster of properties: they are "AN-chains". Crucially, the notion of AN-chain does not refer to the nature of the position of any member of the chain, but strictly to the anaphoric cluster of properties. Similarly, we will call the chains displaying the Vbl-cluster of properties "Vbl-chains". Wh-movement is, of course, the canonical instance of a Vbl-chain. As for the nature of the position of the head of a chain, we will refer to the position which is the target of movement as the T-position.

Given these distinctions, we can now adequately formulate the questions raised by the adoption of the VP-internal-subject hypothesis and the "split-INFL" hypothesis. The question is not whether movement to the specifier of a functional projection is A- or A'-movement. Given the terminological confusion, this question is in fact meaningless. Assuming that the cluster of properties is correct, and that it is the properties of the landing site of movement which determines the nature of chains, the real questions are the following:

a) what are the properties of the T-positions to which movement creates chains with AN-properties?
b) what are the properties of the T-positions to which movement creates chains with Vbl-properties?

We have observed that the distinction between potentially θ-marked positions and positions which cannot be θ-marked, which in the LGB model corresponds to the A/A' distinction, becomes irrelevant in a framework which adopts the VP-internal-subject hypothesis. We will argue that within such a framework, two other properties of positions must be taken to be relevant to the distinction among types of chains: the property of being Case-marked [± Case] and the property of being either a specifier or a complement of a head in the X' theoretic sense, which we will call the property of being Head-Related [± Head-Related] (henceforth [± HR]). Combining these properties, we obtain four possible types of positions which should in principle determine four different types of chains.

(3)
[+HR +Case] positions
[+HR -Case] positions
[-HR -Case] positions
[-HR +Case] positions

In the coming chapters, we will illustrate the first three cases. As to the fourth case, it is questionable whether it should ever be realized. It has been standardly assumed that Case is not generally assigned to an adjoined position. If this is correct then the fourth type will be ruled out be an adequated formulation of the

configurations of Case assignement. Movement to these various types of positions will be shown to exhibit the properties standardly assumed to correlate with A and A' chains. As we will show, movement to [+Case +HR] positions, has properties associated in the LGB model with the LGB A-chains; in our terminology it creates AN-chains. On the other hand, movement to [-HR,-Case] positions has properties usually associated with the LGB A'chains; it creates Vbl-chains. Interestingly, however, chains headed by [-Case +HR] positions seem to exhibit a mixture of properties. As we will see, this follows straightforwardly from our proposal.

This chapter is organized as follows. First we describe the various properties of positions that have been distinguished within the LGB framework. Second we review the cluster of properties which have been associated with standard cases of A chains and A'chains. Next, we argue for the relevance of the properties [±HR] and [±Case] on the basis of French and English data. In Chapter III, we turn to crosslinguistic data and examine in details the properties of movement to the specifier of functional projections.

2.2 Properties of positions

2.2.1 The syntactic relevance of positions

Before we can turn to an examination of the properties of positions which are relevant to recasting the A/A' dichotomy, we will discuss which position in a chain must be considered to be relevant. Taking the notion of chain here simply as a sequence of the type C = $(a_i....a_n)$ in which a_{i+1} locally binds a_i various positions could be taken into account to distinguish various types of chains. For instance, one could look at the properties of the position of the foot of a chain, ie the position of a_n in the sequence. In the LGB model, the distinction between Case marked traces and non-Case marked traces is assumed to be a factor distinguishing empty categories which are defined as variables from empty category which are defined as anaphors. Given this distinction, one could think of defining chains with respect to the properties of the foot of a chain : one could define AN-chains as chains which terminate in a non-Case marked positions and Vbl-chains as chains which terminate in a Case-marked positions. Such a definition however has obvious problems. First consider the trace of an extracted PP:

(4) To whom$_i$ do you think that Mary should speak t_i first

In (3A), under standard assumptions, t_i is not Case marked. But it is quite clear that the chain (to whom, t) does not have properties comparable to the chain created by Raising, since it is non local and

moreover, it clearly induces WCO and strong crossover, phenomena which in the standard theory have been observed to obtain with variables:

(5)

a.* To whom$_i$ does he$_i$ talk for hours

b.?? To whom$_i$ does his$_i$ mother talk for hours

Next, consider the following pair of sentences:

(6)

a.*John is likely that Mary has met t

b. John, it is likely that Mary has met t

In both cases, (6)a. and (6)b., the foot of the chain is Case-marked; but only the second chain is well formed, while the first must be excluded. The contrast in (6), thus shows that reference to the properties of the lowest position in the chain is not suffficient to appropriately distinguish among types of chains. Reference to the landing site of the movement is in fact necessary[4]. In the LGB model, the properties of the foot of a chain enter well formedness conditions on predefined type of chains; they do not distincguish among types of chains . For instance, it is a condition on the well formedness of a chain that the foot of the chain be properly governed. In other words, whether the foot of a given chain (an A-or an A'chain) is prepoerly

---

4. Moreover, it has been argued by Borer (1980) and Epstein (1987) that the requirement that a variable be Case marked is in fact too strong. If this is correct, then such a criterion is not sufficient to distinguish various types of chains. We return to a discussion of the Case requirement on variable in section 4.6.2.

governed or not neither determines nor changes the nature of a given chain: ian any even the chain will be ill-formed, whatever its nature.

To distinguish among types of chains, one could also look at the properties of the position of intermediate links in a chain. In the LGB model and its extension properties of the position of intermediate links of chains only enter well formedness conditions on chains with the notion of "improper movement" and are not taken as a definitional criterion of the type of chain itself. Let us consider again the pair of sentences in (5). It is in fact possible that all intermediate traces in each of the chains in (5)a. and (3b)b. are identical. Nonetheless, these two chains must be distinguished and again, it seems that reference to the landing site of the movement is necessary.

Finally one could look at the properties of the position containing the head of a chain, i.e, the position a_i in the sequence $(a_1 \ldots a_n)$. In the LGB model, properties of the landing site are taken to be a the necessary and sufficient criterion for the distinction among different types of chains. It is in this respect that the distinction of A vs A' positions matters: two types of chains, A-chains and A'-chains, are distinguished with respect to the status of their landing site positions, each determining the nature of the empty category which terminates the chain. Variables are defined as empty categories bound by an element in A'-position and anaphors as empty categories bound by an element in A-position. The properties of each type of chains then follows from the Binding theory. As we have seen above, reference to the landing site of a given movement seems to be necessary . Taking the properties of the landing site of a movement to be the relevant

criterion to distinguish among types of chains seems thus to be the null hypothesis. Moreover, it is assumed by Chomsky (1981 and following work) that this criterion is not only necessary but also sufficient. Once a chain is identified as belonging to a particular type, it must meet additional well-formedness conditions, but the nature of the chain is never questionned. To illustrate, consider the following example:

(7) There seems a man to be likely t' to have been arrested

In (7), the chain (a man, t',t) is ill-formed. Nevertheless, we can determine that it is an A-chain in the sense of LGB, since the landing site of the chain is an A-position. Thus, we can conclude that the chain in (7) is an A-chain which does not meet the well-formedness conditions on A-chains. Throughout this thesis, we will adopt the hypothesis that the nature of a particular chain is necessarily and sufficiently determined by the properties of the landing site of the movement which creates it. At various places in this thesis, but more specifically in Chapter III, we will provide additional support for this hypothesis and we will argue against an alternative hypothesis, proposed by Kayne (1984) and Taraldsen (1986), which regards the properties of the binder as the crucial factor determining the nature of a chain. Under the latter view, it is the properties of the element which is contained in the landing site position of a movement which determines the nature of a given chain: the properties of the position itself are regarded as irrelevant. Thus, an operator is assumed to head an operator chain (with Vbl-properties), irrespectively of the position in which it occurs, while a non-operator is assumed to head a

non-operator chain (with AN-properties). For reasons of expositional clarity, we postpone a discussion of this alternative until the end of Chapter III section 3.3.3. At this point in the thesis, we will argue on the basis of evidence accumulated throughout Chapter III, that the nature of the element heading a chain is not sufficient to determine the type of a given chain.

Within the government and binding framework several properties have been assumed to distinguish positions in a tree, for instance, θ-marking, Case-marking, and lexical government. The question of the redefinition of the dichotomy A versus A' poses in fact the question of which of these properties must be taken to be relevant for the distinction between various types of movement and/or between various types of chains. In the next subsection, we review the various position types and their relevance in the LtB framework.

2.2.1.1 θ-marked positions

As we mentionned above, in the LGB framework, a distinction is made between positions that are θ-marked and positions to which no θ-role is ever assigned, ie θ'-positions. This distinction is essentially relevant to the theta-criterion. The theta-criterion requires that every chain contains at most one θ-position, namely α_n. A consequence of the θ-criterion is that movement must always be to a θ'-position. Another distinction relevant to the classification of positions is the distinction between positions that are potentially θ-marked, that is

may be θ-marked even if they are not always θ-marked (henceforth [+Pθ] positions) and positions that are not potentially θ-marked, that is are incapable of being θ-marked (henceforth [-Pθ] positions). The notion potential (P) is important here in that it distinguishes the position of a surface subject from the position of the complement of a verb. The classical distinction obtains in a passive construction: the surface subject position is not assigned a θ-role; thus it is a θ'position. It is however not a [-Pθ] position (comparable for instance to the Specifier of CP). The set of θ'position includes the set of [-Pθ] position but not the reverse. It is this distinction between [+Pθ] positions and [-Pθ] positions which is taken to be crucial for distinguishing among types of chains in the LGB framework

.

2.2.1.2 Case-marked positions

Positions are also distinguished with respect to their Case properties: we can distinguish Case-marked positions or [+C] positions from positions to which Case is never assigned or [-C]-positions such as for instance the position of an adjunct. The notion of [+C] position[5] is relevant in the LGB framework with respect to well formedness condition on A-chains. As required by condition (170) of Chomsky (1986)a., the head of an A chain must be in a Case-marked

---

5. We assume here a standard view on Case marking as defined in LGB. Latter, in chapter III, we will modify our position on Case-marking in a way which will not affect this discussion.

position (or be PRO). To parallel the distinction made above wrt θ-marked positions, we can further distinguish between potentially Case-marked positions or [+PC] positions - (such as, for instance, the complement position of a passive verb) - from positions which are incapable of being Case-marked, or [-PC] positions, such as adjunct positions. This distinction is not relevant in the LGB framework, but as we will see, it will be relevant to our discussion of the properties of positions distinguishing AN-chains from Vbl-chains.

2.2.1.3 Base-generated and derived positions

Positions can also be distinguished with respect to X' theory: some positions are created directly by the X' schema at the level of D-structure, while others are derived through movement via adjunction. Let us clarify this distinction. Consider the X' schema as defined in Chomsky (1986):

(8)

a. X' = X Y"*

b. X" = Z"* X'

where * = 0 to n occurence

The position of Y" in (8)a. is standardly referred to as the complement of X, and the position of Z" in (8)b. as the specifier of X'. We assume with Chomsky (1986) that the number of complements of a lexical category is determined by θ-theory. Furthermore, we assume, following Kayne (1984), that the complement structure of a given head

is strictly binary. We will adopt in particular Larson's (1988) interpretation of Kayne's original idea, which distributes thematic arguments along "transitive" VP shells. Larson (1988) proposes that, besides binary branching, the projection of argument structure is subject to the following two principles:

(9) (1) if α is a predicate and β is an argument of α, then β must be realized within a projection headed by α

(2) Assuming the following Thematic Hierachy:

AGENT > THEME > GOAL > OBLIQUES (locatives, manner etc)

If a verb determines θ-roles θ_1, $\theta_2 \ldots \theta_n$, then the lowest role on the Thematic Hierachy is assigned to the lowest argument, and so on.

Larson assumes additionally, that the number of positions projected within a particular VP shell is determined by the number of theta role which a given predicate determines, and this "whether the theta-role is actually given to the the projected position or not". Note that these assumptions have a number of non-trivial consequences. First, as noted by Larson himself, principle (1) forces the adoption of the VP internal hypothesis. Second, (and this will be important for the proposal we will make in chapter 3 with respect to the passive transformation (Cf Chapter III section 3.3.5.1), Larson's interpretation of theta-theory forces the existence of a specifier position for the agent within the VP in passive constructions. Indeed, verbs which undergo passive determine an agent theta-role which under Larson's view, must be represented by the highest position projected in the VP shell, even though the agent theta-role will not be assigned

to this position. If so, this specifier position, which is presumably not occupied by the agent and not theta-marked, will be available for movement[6].

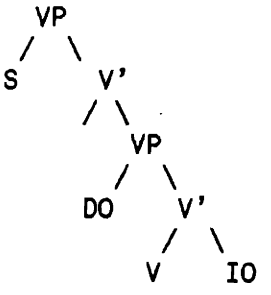
Third as noted by Larson, the proposed principles of argument projection impose the following structure on ditransitive verbs such as "give":[7]

---

6. Given Larson's proposal, a question arises with respect to the distinction between various types of verbs with a single argument such as for instance, the distinction between intransitive verbs and unnacusative verbs in the sense of Burzio (1986). If a verb has a single argument, we would expect under Larson's view that only one position is projected in the VP. Thus intransitive and unnacusative verbs will have the same structure, and we expect the differences between these two classes of verbs to be essentially thematic. It has been shown however by Burzio (among others), that differences between intransitive verbs and unnaccusative have clear syntactic consequences (wrt for instance past participle agreement). Moreover, unnaccusative verb behave, in several regards, very similarly to passive predicates. Many of Burzio's argument distinguishing the two classes lead to the conclusion that the difference in argument structure is syntactically encoded. If it is correct to assume that there is a structural difference between the two types of verbs, then this may constitute an argument in favor of the realization of the Extended Projection principle within the VP projection. The hypothesis that the EPP applies to the VP-internal subject would force the existence of a VP specifier, whether or not this specifier is actually filled with an argument. Under this view, intransitve predicates could be assumed to project their single argument as a Spec of VP, while unnaccusative predicates would project theirs as a complement and have a non-theta marked specifier, just like passive verbs. This hypothesis that the Extended projection principle applies to the VP internal subject has been proposed by Koopman and Sportiche (1988).

7. The symbols S, DO,IO should not be taken here as characterizing grammatical functions but simply as mnemonic symbols for internal arguments and the external argument.

(10)



If this hypothesis is correct, then the asterisks in Chomsky's (1986)b. schema strictly correspond to 0 or 1, and we can reformulate the X' schema as follows to express the restriction to binary branching more straighforwardly :

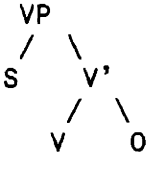
(11)

a. X' --> X (Y")

b. X" --> (Z") X'

Note that given Larson's proposal, there is no longer a correspondence between the position of an internal argument such as the direct object and the position of "complement" in the X'- theoretic sense. As we see in (10), the direct object (DO) of a ditransitive predicate is a "specifier" in the X' theoretic sense, not a complement. On the other hand, the DO of a transitive verb will be a complement and not a specifier, since the structure of the VP shell will be the following.

(12)



This suggests that we need a notion which covers both specifiers and complements. We propose, following a suggestion by Chomky (pc), to

encompass both notions under the term Head-Related positions ([+HR] positions).[8].

(13) α is a Head-Related ([+HR]) position if and only if α is a sister to X^0 or to X'α is a i.e, if and only if α is a specifier or a complement in the X'-theoretic sense

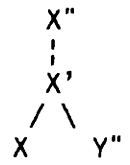
As defined in (13), the notion of [+HR] position is not limited to θ-arguments of lexical categories: specifiers and complements of functional projections are also [+HR] positions. Given the X' schemata in (11), [+HR] positions correspond strictly to positions created by the X'schemata reduced to binary branching.

Let us now return to the distinction between positions created by the X' schemata (equivalent to [+HR] positions if X' theory is as in (11)) and positions created by movement. Consider for instance, the structure of a category for which a specifier has not been generated by the X'schemata.

---

8. Fukui (1986) has argued that lexical categories do not have specifiers which close off their projections. In his view, lexical projections do not reach the level of X". They remain X' projections and allow some internal recursion. The proposal has consequences for Japanese: Fukui proposes that the Japanese VP may take several external arguments. Under Larson's proposal, Fukui's VP structure can be reinterpreted as a succession of VP shells and it is thus no longer necessary to assume that lexical categories do not have specifiers. We can assume on the contrary, that the number of specifiers or of complements is one for all types of categories. There are simply several categories internal to the VP.

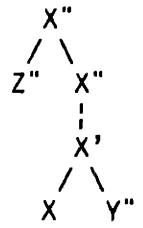
(14)



9

At S-structure, Move α may adjoin a maximal projection Z" to the category X" yielding the structure in (15):

(15)

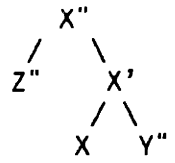


In (15) the position Z" is clearly not a position which can be generated by the X' schema as defined in (11); Z" is not a sister to X'. It is thus in our newly defined terminology a non-Head-Related position or [-HR] position. Now consider a structure in which a specifier has been created at D-structure by the X' schema. Let us assume that this specifier contains an empty category with no features, a category serving merely to mark the existence of the position. Suppose that at S-structure a constituent moves to this position, substituting for the empty maximal projection. The result will be as in (16):

---

9. As noted by Chomsky (1986)b. given the optionality of specifiers, the question arises of whether the projected structure is as in (i) or as in (ii), skipping X':
   (i) [x" [x' X]]
   (ii) [x" X]
We follow Chomsky in assuming that choice of X' is forced when there is a specifier and otherwise optional.

(16)



In (16), the position of Z" is a position conforming to the X'schema as defined in (11): Z" is the sister of X'. In our terminology, it is a [+HR] position. The difference between (15) and (16) is essentially the difference which Chomsky (1986) establishes between two types of movement, adjunction as in (15) and substitution as in (16), but it is recast here in terms of positions directly generated or not directly generated by the X'schema :[+HR] vs. [-HR] positions.

2.2.1.4 Adjunction structure

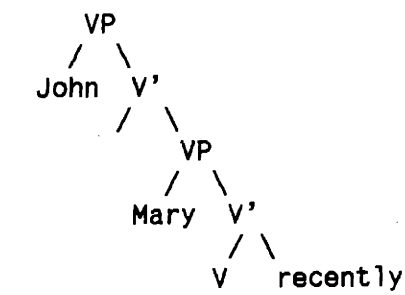
As defined in (11), the X' schema does not generate adjunction structures such as the one given in (15). If as is standardly assumed, D-structure is conform to the X' shemata this raises the question of how various types of adjuncts (such as adverbs and prepositional adverbials) will fit into the tree structure. The question whether there exist base generated adjuncts has often been raised, but remains up to now unsatisfactorily explored. It is standardly assumed that the positions of adjuncts are syntactically distinguished from that of

arguments. But the question of the level at which adjuncts are attached to the syntactic tree remains open.[10]

---

10. It has been suggested by Larson (1988) that adjuncts could be generated like arguments as members of VP shells. Thus for instance, the structure of the sentence (i) is as in (ii).
(i) John saw Mary recently

(ii)



Under this view, the position of an adjunct does not differ syntactically from that of an argument. Given this hypothesis, it becomes unclear why adjuncts should pattern differently from arguments with respect to a number of syntactic phenomena and in particular with respect to extractions and CED effects. Rizzi (1989) proposes distinguishing arguments from adjuncts on the basis of their semantic properties: arguments are referential elements but adjuncts are not. He encodes this semantic difference in the syntax by using the mechanism of referential indices. In his view, referential elements bear referential indices but adjuncts do not. This has consequences with respect to extraction properties, which we discuss in section 4.7.1.1. In this section, we will argue against this proposal. Thus, although we adopt Larson's proposal for argument structure, we maintain the hypothesis that adjuncts are structurally distinct from arguments. One problem with Larson's analysis is that it does not explain why some adverbs can occur on both sides of the VP with basically the same meaning.
(i) John recently saw Mary
(ii) John saw Mary recently.
(iii) John quickly left the room
(iv) John left the room quickly
It may be possible to derive (i) from (ii) or (iii) form (iv) by movement of the adverb (See Emonds 1978) for a similar proposal). But Jackendoff (1972), Schlyter (1973), and Ernst (1986) argue against the view that the variety of adverbial positions is due to movement. Both Jackendoff and Ernst argue that adverbs are direclty generated in the positions they occur in on the bases of numerous syntactic and semantic arguments.

2.2.1.5 Adjuncts at S-structure

The hypothesis that adjuncts are not attached at the level of D-structure has been recently defended by Lebeaux (1988). Lebeaux (1988) proposes that adjuncts are hooked up to the tree in the course of the derivation from D-structure to S-structure by an operation which he calls "Adjoin α" and describes as follows: "Adjoin α takes two tree structures and adjoins the second to the first." (p 148) Lebeaux assumes that Adjoin α creates "Chomsky-adjoined" structures. The main argument for his view relies on the lack of expected violations of principle C of the Binding theory. Consider the following paradigm:

(17)a. *[Which picture of John$_i$] did he$_i$ like

b. [Which picture that John bought] did he like

Lebeaux argues that (17)a. can be ruled out by principle C applying at D-structure, before the movement of the wh-constituent. The D-structure of (17)a. is given in (18):

(18) he likes [which picture of John]

In (18), the pronoun *he* c-commands the name *John* yielding a violation of principle C. Under this view, (17)b. should also be ruled out. To explain the grammaticality of (17)b., Lebeaux proposes that adjuncts are not hooked up to the tree at the level of D-structure: consequently, the application of principle C at this level will have no effect on adjuncts. This accounts correctly for the lack of principle C effects in (17)b.

Note that if as proposed by Lebeau, adjuncts are not hooked up to the tree at the level of D-structure, then the notion of [+HR] position which we have defined in (13) need not be defined independently. [+HR] positions strictly correspond to the positions which are generated by the X' schemata given in (11). In other words, given Lebeau's approach to adjunct structures, the distinction between [+HR] and [-HR] positions simply corresponds to the distinction between D-structure and non-D-structure positions.

There are however a number of arguments against the view that adjuncts are all attached to the tree after the level of D-structure. On the basis of contrasts similar to the ones used by Lebeaux, Speas (1989) argues that certain types of adjuncts must be present in the base. Consider the following example:

(19) *How near Dan$_i$ did he$_i$ find a snake?

Although quite acceptable when the pronoun *he* and the name *John* are not coindexed, the sentence is ungrammatical under the indexing represented in (19). (19) thus appear to violate principle C, prior to the wh-extraction. But "near John" is an adjunct, and under Lebeau's approach, we expect the sentence (19) to be able to escape condition C. Since this is not the case, the ungrammaticality of (19) clearly suggests that at leat some adjuncts must be present at the level of D-structure[11,12].

---

11. Cf Speas(1989) for a detailed analysis. Most probably, the adjucnt generated in the base will be semantically different than the one attached to the tree at a latter level. We suggest that the distinction may have something to do with the way adverbs "affect" or do not "affect" the aspectual structure of the event. We

2.2.1.6 Base-generated adjuncts

If as shown by Speas (1989), some adjuncts must be present at the level of D-structure, then the question arises as to how this can be

---

suggests that the base generated adjuncts may be those which "affect" in some ways the aspectual structure of the event described by the predicate which the adjunct modifies. Klipple (1989) argues for such a distinction among locative types of adjuncts. She shows that this difference patterns with a number of interesting syntactic contrasts. Looking at French, it seems that the cliticization of some locatives is more difficult than others. Consider the following sentences:

(i) a. Jean a planté des tomates dans le jardin
Jean has planted some tomatoes in the garden
b. Jean y a planté des tomates
Jean there has planted some tomatoes

(ii)a. Jean a reparé la television dans le jardin
Jean has fixed the television in the garden
b.?? Jean y a reparé la télévision
Jean there has fixed the television

The semantic distinction between (i) and (ii) can be described as follows: in (i) the locative is part of the make up of the event described by the predictate. The "in the garden tomatoe planting" activity is different for instance than planting tomatoes in a pot. In (ii) however, the locative does not directly affect the event. Rather it sets up the background of the event. In Klipple's terminology, (i) is an "aspectual affecting locative" while (ii) is a frame locative. Suppose that only the D-structure adverbs can move and not the adverbs attached to the tree at latter level. If so, the contrast between (i) and (ii) would receive an interesting explantation. We leave this question open for further research.

12. There are additional problems with Lebeau's analysis, centered arround the issue of the proper treatment of "reconstruction", which have been pointed out by Chomsky (1989). We will come back to some of these arguments in the course of Chapter III.

made possible. One possibility is to assume that the X' schema is augmented with the following rule:[13]

(20) X" --> X" Y"

This would allow adjuncts to be generated directly at D-structure. That this view is quite plausible is suggested by the fact that adjuncts seem to leave traces when they are extracted in the syntax. This allows their interpretation to be related to their D-structure position. As noted by several linguists (Jackendoff (1972), Ernst (1986) among others, a large number of adverbs have different interpretations depending on the position in which they occur in a sentence. Consider the following pairs of sentences:

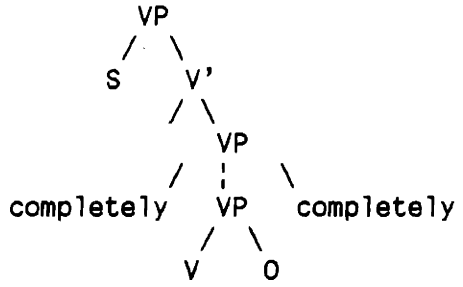
(21)

a. John intelligently has answered the question

b. John has answered the question intelligently.

---

13. (20) would generally limit adjunction structures to maximal projections. Certain adverbs, such as the adverbs of completion (completely etc..), seem to be V' adverbs. (See Pesetsky (1989) for some arguments). Possibly, however, given that we have adopted Larson's structure for the VP, they could be argued to attach to the internal VP shell; if so the structure would be as follows:



This structure is consistent with the hypothesis that adjunction only occurs to maximal projections. Moreover, it accounts for the fact that these types of adverbs generally seem to modify the object and not the subject.

(21)a. could be paraphrased as follows: "it was intelligent of John to answer the question". (21)b., on the contrary, cannot have this meaning: it can be paraphrased as follows: "John answered the question in an intelligent manner". In view of these facts, consider now the following sentence:

(22) How intelligently did John answer the question?

In (22) the adverb has been extracted. A priori, we would expect wh-extraction to be able to occur from any of the possible positions that the adverb can take. If so, (22) should be ambiguous between the "subject oriented" interpretation of (21)b and the manner interpretation of (21)b. But, contrary to expectation, in (22) the adverb can only have the meaning of (21)b, i.e., the "manner" reading: the interpretation (21)a. is unavailable. This is unexpected, if both positions for the adverb are "represented" in the same way in the structure. In other words, if as standardly assumed the adjunct leaves a trace after extraction, one would expect to be able to get a trace in both the position (21)a and the position (21)b. To explain the fact that only the "manner" interpretation is possible in questions, we suggest that while manner adverbs are present at D-structure and (thus leave traces when they are subject to movement) subject oriented adverbs (and may be more generally all adverbs which are considered to be sentential adverbs) have not yet been hooked up to the tree when wh-movement occurs. The consequence of this assumption is that subject oriented adverbs and sentential adverbs are expected to be inexctractible, an expectation which is confirmed by the facts.

(23) *How problably will John to come?

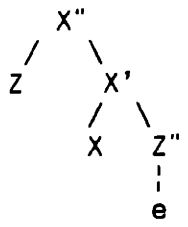
(vs How probable is it that John will come)

As proposed by Chomsky (1981), D-structure is a pure representation of θ-relations. It follows from this hypothesis that to be present at D-structure an adjunct needs to be licensed somehow by θ-theory. Higginbotham (1985), following Davidson (1970), has argued that adverbs are predicates of events. He has also proposed extending the set of possible θ-relations, positing a relation of autonymous θ-marking which applies between adverbs and predicates. If, as proposed by Higginbotham, adjuncts enter θ-relations, it is no problem for them to be directly generated at D-structure.

Note that if we augment the X'schemata with the rule (20), the structure given in (15) would now correspond to a position which is generalble by the X' schema. If so, we then need the definition in (13), to distinguish the position of Z" in (15) from the positions generated by the rules in (11). Given the defintition (13), (the position of Z" is a [-HR] position. Thus the defintition in (13) distinguishes Specifiers and complements from adjunct structure, whether these adjuncts are present at D-structure or are the result of movement.

One possible advantage to the proposal that X' theory be extended to include the rule (20) is that it permits a clear definition of the notion "Structure Preservation" which was first proposed by Emmonds (1976) in a framework which did not include the X'theory.
Consider the following structure:

(24)



In the structure (24), a head has moved to the specifier of a maximal projection. A structure such as (24) could never be produced by the rules of the X' schema. The impossibility of this structure is accounted for in Chomsky (1986) as a violation of "Structure Preservation". Assuming that the X' schema is augmented with the rule in (20), we can understand the notion of Structure Preservation in terms of the distinction between positions generable by the augmented X'schema (rules (11) + rule (20)) and positions which are not generable with the augmented X' schema. This would distinguish appropriately the structure (15) from the structure (24). Note also that this is equivalent to the proposal that the output of S-structure movement must be constrainted by X' theory, a proposal made by Chomsky (class lecture (1986)).

As we have seen, there seems to be some empirical arguments as well as some theoretical considerations which support the assunmption that adjuncts can be base generated. We will thus adopt this position, although we remain conscious that a full discussion of the question of the level at which adjuncts are attached to the tree would require an extensive study of adjunct structure which is beycnd the limits of this dissertation. The main point which is relevant to our present dicussion is the hypothesis that whatever the level at which adjuncts attach to the tree, they do not occupy positions which we have termed

[+HR] positions, i.e, they are not sisters of X' or X_0 projections. We thus assume without further discussion that adjuncts are syntactically distinct from arguments in that they occupy [-HR] positions.

2.2.1.7 Summary

The following chart summarizes the types of positions mentioned above and their particular relevance in the LGB framework:

(25)

Distinction	Relevance to GB
[+θ] positions/[-θ] positions	relevant for the θ-criterion
[+Pθ] positions/[-Pθ] positions	the LGB A/A' distinction
[+C]ppositions/[-C] positions	relevant for well-formedness condition on A-chains
[+PC] positions/[-PC] positions	irrelevant
[+HR] positions/[-HR] positions	irrelevant

In section 2.2, we discuss these distinctions and their relevance to the definition of various types of chains. But before we turn to this task, we review the various properties which are part of the cluster distinguishing between two types of chains.

2.3 Properties of chains

In the previous section, we discussed an inventory of the properties of syntactic positions in the LGB framework. In this section, we briefly review the phenomena which have been argued to distinguish AN-

chains from Vbl-chains. Recall that we have opted for this terminology in order to distinguish properties exhibited by different types of dependencies from the properties of the position of their landing sites. Thus, AN-chains are chains which have the properties of A-chains in the sense of LGB, but whose heads may not be in an A-position under the definition given to this term in LGB; likewise, Vbl-chains have the properties of A'-chains, but may not have heads in an A'-position. Our immediate purpose in this section is not to consider in detail the analysis of these phenomena. We simply intend to establish them as tests for a distinction between AN-chains and Vbl-chains. The occurrence of these syntactic phenomena represents the data that any theory of chains must be able to account for.

2.3.1 Standard distinctions

The first property which distinguishes Vbl-chains from AN-chains is that only the former type of chain is sensitive to crossover effects. As first noted by Wasow (1972), there are two types of crossover effects which, he termed strong crossover and weak crossover; these are illustrated in (26) and (27), respectively.

(26)

a.* Who$_i$ does he$_i$ think Mary loves t$_i$

b.* John$_i$, he$_i$ thinks Mary loves t$_i$

(27)

a.?? Who$_i$ does his mother love t$_i$

b.?? Mary$_i$, as you remember, John thinks her$_i$ mother likes$_i$[14]

---

14. Facts about WCO effects with topicalization seem to be quite controversial. For some speakers, the effect is weaker than with a WH-phrase (Cf. Lasnik and Stowell (1989)). For others, this is not the case. Lasnik and Stowell suggest that the absence of WCO with topicalization may be due to LF reconstruction. At LF, the topicalized NP is lowered back to its base position. There is thus no variable with which the pronoun corefers and which does not c-command the pronoun; hence WCO effects are not expected. For speakers who do get WCO effects, we can assume that either reconstruction has not occured or that the focus which is on the Topic makes it a quasi-quantifier which forces the NP to raise back under QR to take a scopal position. If so, there will be a variable which does not c-command the pronoun at LF and WCO effects will obtain. We have suggested that the difference is one among speakers. This suggestion is very odd: one does not expect a dialectal difference between speakers with respect to reconstruction possibilities. This suggests that in fact there might be two different types of topicalization, types which would be fairly difficult to distinguish. That this analysis is quite plausible is supported by the following facts pointed out by D. Pesetsky (pc): for him, WCO occurs with what we will call, following his suggestion, "non-contrastive" topicalization. Consider the following case:

(i) I am glad you mentioned Mary.
??Mary$_i$, as you may remember, John thinks her$_i$ mother likes t_i.

It does not occur, however, with "contrastive topicalization":

(ii) The old man wanted to invite many people to celebrate. That's certainly a good idea. But John$_i$, I thought that his$_i$ father should not invite t_i.

Topicalization cases such as (iii), which involve reconstruction, seem to have to be of the contrastive type:

(iii) It is John's job to promote people. Himself, John always thought he should promote better.

In the case of contrastive topicalization, Pesetsky reports that even strong crossover violations tend to disappear.

(iv) Some presidents think they do not have to abide by the law.
?Nixon, I certainly think that Nixon thought above the law.

This is consistent with the idea that this type of topicalization involves reconstruction at LF.

(26) is standardly analyzed as a violation of principle C. In the LGB model, a variable is an R-expression and must be A-free in the domain of its operator. In (26), the variable left by the WH-extraction is bound by the subject pronoun *he* in the domain of its operator *who*. The sentence is thus straightforwardly ruled out by principle C. Various analyses of the ungrammaticality of (27) have been given; we will not review them here. Characteristic structures inducing WCO violations can be described informally as follows:

(28)

WCO effects occur in a configuration where a quantifier binds both a variable and a pronoun and neither the pronoun nor the variable c-commands the other

* [Q_i [... [pronoun$_i$...]...t_i...]]

The structure of (27)a. meets this configuration, so coreference between the pronoun and the variable is prohibited.

As has often been noted, AN-chains are insensitive to crossover effects. The absence of strong crossover effects cannot be tested, due to interferences from principle B of the Binding theory. Consider the following example:

(29)

*John$_i$ is likely to seem to him$_i$ to have been arrested t_i

Superficially, the pronoun *him* seems to be in a different clause from its binder *John*. But in fact the structure of (29) is as in (30):

(30)

John$_i$ is likely [t_i to seem to him$_i$ [t'_i to have been arrested t''_i]]

Given this structure, crossover is irrelevant, since if, as represented in (30), the pronoun is construed as coreferential to *John*, there will be a straightforward violation of principle B: t_i binds the pronoun in its governing category.

The absence of weak crossover effects in AN-chains, however, can be tested. Consider (31):

(31)

Everybody$_i$ seems to his$_i$ mother [t_i to be the most intelligent person in the world.]

In (31), the pronoun can be construed as bound by the quantifier *everybody* after the raising of this quantifier to the subject position of the matrix. This is a configuration of weak crossover: the quantifier binds both a pronoun and a trace, and neither c-commands the other. Crucially for the LGB model, however, the trace is bound by an element in an A-position, and is thus not a variable. (31) thus shows that AN-chains are indeed insensitive to weak crossover.

A second property which distinguishes Vbl-chains from AN-chains is the licensing of parasitic gaps. Vbl-chains license parasitic gaps:

(32)

a.Which paper did you file before reading

b.These papers, I always file before reading

AN-chains,however, do not:

(33)

a.*The report was filed t after Bill read e

b.*The report seems to have been filed t before Bill read e[15]

---

15. The reason why A-movement does not license PGs remains somewhat obscure in Chomsky's (1986) Barriers model. In Chomsky (1982) the following conditions are given on the licensing of parasitic gaps.

(i) (i) α c-commands t and e
(ii) t does not c-command e or conversely
(iii) α does not head the chains (α,t) and (α,e)
(iv) e is governed and heads a chain with a θ-role

In this approach, a parasitic gap is licensed if it is c-commanded by the operator which licenses the real gap. The ungrammaticality of examples like (33) follows from condition (iii). To quote Chomsky: "if α is in an A-position, then ..(iii) is always violated, since α heads the chains (α,t) and (α,e)". The chain (α,e) is ruled out. There are two cases to consider: 1) both α and e have an independent θ-role, so that e must be PRO under the functional determination of empty categories. This is ruled out because PRO needs to be in an ungoverned position. 2) α does not have an independent θ-role and e is an anaphor which is unbound in its governing category, in violation of principle A of the Binding theory.
In *Barriers*, however, Chomsky proposes that parasitic gap constructions contain an empty operator and that a parasitic chain is licensed by Chain composition. Chain composition is defined as follows:

(ii) If C = $(\alpha_1,....\alpha_n)$ is the chain of the real gap, and C' = $(\beta_1,...,\beta_m)$ is the chain of the real gap, then the "composed chain" (C,C') = $(\alpha_1,...,\alpha_n, \beta_1...\beta_m)$ is the chain asssociated with the parasitic gap construction and yields its interpretation.

Chain composition must obey the following constraint:

(iii) The operator of the parasitic gap must be 0-subjacent to the head of the A-chain of the real gap.

Let us consider the structure of (33):

(iv) The report was[$_{VP}$ [$_{VP}$ filed t] [after [op Bill read t]]]

For Chomsky, VP is a barrier. However, it is usually assumed that the adjunct is adjoined to the VP; if so, the VP will not be a barrier for anything contained in the adjunct, since the VP does not dominate the adjunct; only one segment of the VP dominates the adjunct. Assuming, as Chomsky does, that the operator can adjoin to the adjunct PP, VP will not be a barrier for the operator, which will be 0-subjacent to the head of the A-chain, as follows:

(v) The report was[$_{VP}$ [$_{VP}$ filed t] [op [after [Bill read t]]]]

A third property which distinguishes AN-chains from Vbl-chains is that AN-chains create new binding possibilities, while Vbl-chains do not. Consider (34) and (35):

(34)

a.*It seems to himself$_i$ that John$_i$ is the most intelligent person in the world

b. John$_i$ seems to himself$_i$ [t to be the most intelligent person in the world]

(35)

a.*It seems to each other$_i$ that they$_i$ are happy

b. They$_i$ seem to each other$_i$ to be happy (LGB p. 45)

The a. sentences are excluded by principle A of the Binding theory because the anaphor has no antecedent in its governing category. The b. sentences, however, are perfect; raising the NP from the embedded sentence has provided the anaphor with an appropriate antecedent.

Now consider the following sentences:

(36)

a.*Pictures of himself$_i$ killed John$_i$

b.*John$_i$, pictures of himself$_i$ killed t$_i$

c.*Which boy$_i$ did pictures of himself$_i$ kill t$_i$

This framework thus has no way to exclude parasitic gaps in A-chains.

(37)

a.*Each other's$_i$ friends told John that Mary would like these men$_i$

b.*These men$_i$, each other's$_i$ friends told John that Mary would like t$_i$

c.*Which men$_i$ did each other's$_i$ friends tell John that Mary would like t$_i$

The a. cases of (36) and (37) are straightforwardly ruled out by principle A of the Binding theory: the anaphor contained in the subject NP has no antecedent. Interestingly, however, in the b. and c. cases of (36) and (37) the anaphors do have plausible antecedents, namely the topicalized NP or the WH-moved element. These constituents clearly c-command the anaphors: nevertheless, the sentences remain ungrammatical. One could try to argue that the fronted constituents are not part of the governing category of the anaphor, so principle A remains unsatisfied. Clearly, this depends on the definition of GC, but whatever the definition of GC might be, it must be able to account for the well-known fact that anaphors within subject noun phrases can be bound outside of the sentence immediately containing them. Consider the following examples:

(38)

a.John$_i$ thinks [that pictures of himself$_i$ will be on sale pretty soon]

b.These guys claim that each other's theory does not work

Given that this binding must be allowed, it is doubtful that any definition of GC can be such that it will include the matrix subject in (38) but not the fronted constituents in (37). If so, we must conclude that it is either the type of dependency involved in the examples in (37), namely what we have called Vbl-chains, or the

properties of the position containing the fronted elements which prevents the binding of the anaphor to be satisfied.

The binding of a pronoun by a quantifier in a topicalized position is also impossible, as shown by (39). However, the ungrammaticality of (39) is probably due to the intrinsic impossiblility of topicalizing quantifiers or quantified NPs, rather than solely to the properties of Vbl-chains.

(39)

a.* Everyone$_i$, his$_i$ mother loves t$_i$

b.*(??)Every boy$_i$, his$_i$ mother loves t$_i$

Consider, however, the following paradigm:

(40)

a. Each woman's$_i$ husband likes her$_i$ mother

b.*Her$_i$ mother likes each woman's$_i$ husband

c.*Each woman's$_i$ husband, her$_i$ mother likes

(40)a. shows that although the quantifier is embedded inside an NP, it can bind the pronominal in the object NP. (40)b. shows that a quantifier inside an object NP leads to a weak crossover violation. Whatever the explanation of this phenomenon[16], we can use it to determine if the binding of a pronominal is possible from a topicalized position. English speakers consulted reject (40)c. just as

---

16. One possibility is to assume that the Quantified expression contained in the NP will raise to adjoin to S at LF. This creates a configuration of WCO in (40b) but not in (40)a.. (32)c. is more complex since it probably involves reconstruction.

strongly as (40)b., which indicates that the binding of a pronoun by a quantifier from an adjoined position is not possible.

The following table summarizes the typical differences we have reviewed between AN-chains and Vbl-chains:

(41)

	AN-cluster	Vbl-cluster
WCO	-	+
Parasitic Gaps	-	+
anaphoric binding by moved XP	+	-
pronominal binding by moved NP	+	-

Additional differences which have been noted in the literature but which are of less relevance to us include the fact that certain types of Vbl-chains license resumptive pronouns. To my knowledge, similar facts have not been noted for AN-chains.[17] David Pesetsky (1982) also argues that while Vbl-chains are subject to crossing dependencies, AN-chains are immune to them. We refer the reader to Pesetsky (1982) for arguments.

---

17. Haitian Creole superficially seems to violate this generalization. Indeed, typical raising constructions require a pronominal copy which seemingly occurs in the position of the trace of the raised element.

(i) Jan sanblé li pati
Jan seems he to have left

See Déprez (1988) for an analysis of these facts which is compatible with the generalization that AN-chains do not license resumptive pronouns.

2.3.2 Floating Quantifiers

To the differences noted in the table in (41), we add another one which is revealed by taking a closer look at the distribution of floating quantifiers[18]. Sportiche (1988) has argued that the phenomenon of "quantifier floating" is better analyzed as a phenomenon of "quantifier stranding". In his view, it is not the quantifier which moves but rather the quantified NP: this NP, on its way to the "external" subject position (the SPEC of AGR-, strands a quantifier in the position of a VP internal subject.

(42)
The children$_i$ must have been cleverly [$_{VP}$ [all t_i] pretending to sleep
(Sportiche (1988))

This analysis accounts elegantly for the apparent anaphoric status of the floated qunatifiers without having to assume that the quantifier itself is an anaphor. In Sportiche's view, it is simply the trace of the moved NP which functions as an anaphor. Once we adopt the proposed analysis, it is immediately apparent that the VP-internal subject position is not the only position in which a quantifier may be stranded. Consider the examples in (43):

(43)

a.The bandits would all have been arrested before dawn

b.The bandits would have all been arrested before dawn

---

18. This phenomenon has been briefly noted in passing by Sportiche (1988).

c.The bandits are all likely to have been arrested before dawn

d.The bandits seem to have all been likely to have been arrested before dawn

As shown by the grammaticality of (43), strandec quantifiers can occur in many places along the path of NP movement.[19] In the spirit of Sportiche's analysis, we can assume that the various places in which the quantifiers can be stranded are indicative of the various specifier positions through which an NP can pass. The structure of (43)a. is then roughly as in (44):

(44)
[IP The banditsi [I' would [VP [all ti] [V' have [VPt'i [V' been [VP [VP arrested t"] before dawn]]]]]]]

In (44), [all t_i] is in the specifier of the projection headed by the auxiliary *have*. Alternatively, it could be in the specifier of any other verbal or functional projection occurring between the original site of extraction (the object position in this case) and the final landing site of the NP. As argue by Pesetsky (1989), the possibility of stranding quantifiers in various positions along the path of the moved NP provides support for the hypothesis that there are a number of functional projection between the verb and the Spec of AGR-S.

As noted by Sportiche, however, "stranding" the quantifier in the original position of the object is not possible.

(45) *The bandits could have been arrested all.

---

19. There are restrictions on a number of positions, the nature of which remains ill understood. We return to this question in a subsequent section where we will investigate in greater details the phenomenon we mention here.

Sportiche argues that the ungrammaticality of (45) is due to some construction specific properties of passive. In his view, (45) indicates that there is no postverbal NP trace in the passive construction. The claim is that, in this case, the object starts out from the Spec,VP position[20]. Although we do not have an explanation for the ungrammaticality of (45), Sportiche's solution does not seem to us to be general enough, since it is construction specific, while the phenomenon seems to be more general. As shown in (46) a quantifier cannot be stranded in the initial position of the object trace in Vbl-chains.

(46) * The bandit that you saw all...

A natural assumption would be to consider the ungrammaticality of (45) and (46) as due to the same restriction, whatever this restriction may be (see section xx for discussion). Since the fact that a quantifier cannot be stranded in the original position of an object does not distinguish between AN-chains and Vbl-chains, it is orthogonal to our present discussion. We thus leave this problem open for further research. What is however of interest to the present discussion is the observation that although quantifiers can be stranded in intermediate

---

20. For Sportiche the possibility for the object to start out from Spec, VP is related to the failure to assign a θ-role to the subject NP in passive construction. In his view, however, the subject never starts out in Spec, VP; it is thus unclear why only passive constructions and not any other construction should allow the object to start out in Spec, VP. For Sportiche, indeed, the Spec VP is never filled by the subject and should thus always be free for the object to occurs in. Since we assume, on the contrary, that the subject does start out in the VP Spec, this problem woould not arise. But in any case, we do not adopt Sportiche's suggestion.

positions of AN-chains, they cannot be associated with intermediate positions of Vbl-chains. Consider, the following sentences:

(47)
a. *the children who I will have all/both met before the end of this week

b. *the children who I will all/both have met before the end of this week

(48)
a.*These students, I will have all/both met before the end of the term

b.*These students, I will all/both have met before the end of the term

The examples in (47) and (48) are all examples of object extractions. But (49) shows that the phenomenon is not restricted to objects. Long-distance extraction of a subject does not allow the stranding of a quantifier in intermediate positions, either:

(49)
a.*These men, I would have all/both thought would be arrested.

b.*These men, I would all/both have thought would be arrested.

c.*the men that I have been all/both thinking would be arrested.

d.*the men that I have all/both been thinking would be arrested.

Quantifiers can, of course, be stranded in intermediate positions in the embedded clause, since in this case the intermediate positions are part of an AN-chain.

(50) a.?These men, I would have thought would all be arrested.

b.Which men would you have thought should all be arrested.

The contrast between (43) and (45)-(49) can be described informally as follows:

(51)

Floating quantifiers can occur in the intermediate positions of AN-chains but are incompatible with the intermediate positions of argument Vbl-chains.

We will not attempt for the moment to provide an explanation for the generalization in (51) (see Chapter 4 section 4.7.1.2 for discussion). For our present purposes, it is enough to observe that if (51) is correct, it gives us a powerful test to distinguish between AN-chains and argument Vbl-chains.

But the descriptive generalization in (51) appears to be immediatle contradicted by the distribution of floating quantifiers in French. As in English, floating quantifiers in French can be stranded along the path of an AN-chain. Examples corresponding to the English examples in (43) are common.

(52)

a. Les enfants ont tous été invités à cette soirée
   The children have all been invited to this party

b. Les enfants ont été tous invités à cette soirée
   The children have been all invited to this party

c. Les enfants ont tous semblé avoir compris les exercices
   The children have all seemed to have understood the exercises

But in French, unlike English, floating quantifiers are perfectly acceptable when associated with a Vbl-chain. Consider the following examples of restrictive and non-restrictive relatives[21] :

(53)

---

21. Kayne (1975) only gives examples with non-restrictive relatives; he does not mention any difference with restrictive relatives, and native speakers consulted did not see any, either.

a. les enfants que tu as tous grondés sont partis en pleurant
the children whom you have all scolded have left crying

b. les amis de Pierre, que j'ai tous connus à l'age de sept ans,
The friends of Pierre, whom I have all met at the age of seven,
sont très sympatiques
are very friendly

c. les fils de G, que tu peux tous voir, sont formidables
the sons of G, whom you can all see, are wonderful

d. les livres de J-P, que j'aurais tous voulu lire, sont très bons
the books of J-P, which I would have all wanted to read, are
very good (Kayne (1975), p. 19 (13)

Contrary to the generalization (51), the examples in (53) seem to show that floating quantifiers are not incompatible with Vbl-chains in French. But French differs from English (and from other Romance languages) in having a rule which permits a leftward movement of quantifiers. This phenomenon, which is known throughout the litterature as Leftward-*tous* (Tous-à-gauche) and which is described in great detail by Kayne (1975), obtains whenever an object NP is empty. Thus, it can associate a floated quantifier with a WH-extracted phrase, as in (53), or with an object clitic, as in (54)a; it can also move a bare quantifier leftward, as in (54)b:

(54)
a. Je les ai tous lus
I them have all read
I have read them all

b. Il a tout vu
He has everything seen
He saw everything

The leftward movement of a quantifier is limited: roughly, it is clause-bound.[22]

---

22. There are, however, a few exceptions to the clause-boundedness of "leftward *tous*-movement": a leftward-moved quantifier may

(55)
a.* Tu as tous cru les avoir compris
You have all believed them to have understood
You thought you understood them all

b.* Tu as tous espéré que je les rencontrent.
You have all hoped that I them meet
You hoped that I would meet them all

Given the clause-boundedness of Leftward-*tous*, consider the following examples:

(56)
a.*ces livres, que j'ai tous cru que tu avais lus
these books, which I've all thought that you had read
these books, of which I thought you had read all of them

b.*les amis de P, que Marie a tous cru que tu aimerais rencontrer
the friends of P, who M has all thought that you would like to meet
P's friends, all of whom M thought you would like to meet

c.*les amis de P, que Marie pense que Jean a tous regretté que tu
the friends of P, who M thinks that J has all regretted that you
n'aies pas pu rencontrer.
have not been able to meet.
P's friends, of whom M thinks that J regretted that you couldn't meet them all

The examples in (56) are of the same type as those in (55): that is, they involve contexts in which leftward *tous*-movement is clause-bound.

---

marginally move out of some embedded infinitivals (roughly the ones which permit restructuring and clitic climbing in other Romance languages) and out of some subjunctive sentential complements. Typical examples are given below:

(i)
a.Jean aurait tous aimé les rencontrer.
John would have all liked to meet them.

b.??Il a tous fallu que je les lise
I had all to read them

(Cf. Kayne (1975), p. 22 fnt 11 : **Ces livres-la, qu'il faut tous que votre ami lise, sont introuvables* ("Those books, which it is necessary all that your friend read, are unavailable").

They differ, however, from those in (55) in that an object has been extracted out of the lowest embedded clause. The importance of the examples in (56), resides in the fact that the long-distance extraction of the object does not license any further stranding of the quantifier. In other words, we observe that despite the long-distance extraction of the object, the leftward "floating" of *tous* remains clause-bound. This shows that the floating of quantifiers in examples such as (53) above is not strictly speaking licenced by the Vbl-chain. If, as seemingly suggested by (53), a floated quantifier could be licensed direclty by a Vbl-chain in French, the ungrammaticality of (56) would be surprising. Examples of the type shown in (56) suggest that in (53), the floated quantifier is not licensed directly by the WH-extraction, but rather that the rule of leftward-*tous* (which we assume for the moment to be an independent rule) is compatible with WH-extraction. In other words, what (56) clearly indicates is that Vbl-chains do not license any movement of quantifiers which is not independently licensed by the rule of Leftward-*tous*.

Sportiche (1988) suggests that the rule of Leftward-*tous* is of a different nature from the phenomenon of quantifier "stranding", which relates a surface subject to a quantifier. In his view, Leftward-*tous* is the overt reflection in the syntax of the rule of QR, which applies universally to every quantifier at LF[23]. We return to an analysis of

---

23. Sportiche (1988) mentions that there is a difference between the type of quantifiers which "float" rightward, "stranded quantifiers" in his view, and the type of quantifiers which can move leftward. Thus he views *chacun* (each) as only a rightward "floating" quantifier. For us and other speakers consulted, this jugement does not obtain. Kayne (1975) mentions that leftward movement of *chacun* as possible although slightly more awkward. He

Leftward-*tous* and floating quantifiers in Chapter IV section 4.7.1.2. At this point, we simply note that contrary to appearances, the distribution of floating quantifiers in French does not contradict the generalization in (51). As shown by (56), Vbl-chains do not independently license floated quantifiers the way AN-chains do: that is, "quantifier stranding", as in Sportiche's view, is incompatible with Vbl-chains. The fact that the appearant compatibility of "stranded quantifiers" with variable chain is due to an independent rule of "leftward tous" and not to Vbl-chains per-se is confirmed by subject extractions. Subject extraction patterns exactly alike in English and in French: they do not permit quantifier "stranding".

(57)
*Ces hommes que j'aurais tous cru qui auraient été arrêtés
*These men who I would have all thought would have been arrested

We conclude that the descriptive generalization given in (51) can be maintained and that the possibility of "stranding quantifiers" provides a test to distinguish AN-chains from Vbl-chains[24].

---

gives the following example: *elle les a chacun rencontrés lors d'un congrès différent* ("she they has each met at a different conference") (p. 27). I believe that the awkwardness of leftward movement of *chacun* is essentially due to semantic conditions on distributionality, which apply equally to the "rightward floating" of this quantifier. For us the following example is perfectly acceptable: *les articles, que tu as chacun envoyés dans une revue differente, ont tous été acceptés* ("the articles, which you have each sent to a different journal, have all been accepted").

24. The incompatibility of floating quantifiers with Vbl chains can be observed in a variety of languages such as Italian, Spanish, Catalan:

(i)It: *el bambini que visto tutti
Sp: *los muchachos que rencontro todos
cat:* el nens que ho vi tots

2.3.2.1 Summary

A full table recapitulating the differences between AN-chains and Vbl-chains is given below:

(58)

	AN-cluster	Vbl-cluster
WCO	-	+
Parasitic Gaps	-	+
New anaphoric binding by moved XP	+	-
pronominal binding by moved XP	+	-
"stranded Q"	+	-

With these tests in hand, we can now return to the main issue of this chapter, namely an investigation of the proper redefinition of the A/A' distinction. We know from the GB model that at least two types of dependencies must be distinguished. The first type of dependency exhibits the properties summarized in the first column; the second

---

Although in theses sentences, the quantifiers appear in post-verbal positions, there is no reason to assume that they occur in the original object position, since in all these languages, the verb clearly raises to AGR. (see Belletti (1988), Lois (1988) and Bonet(1988) for arguments). This is confirmed by the fact that stranded quantifiers with AN-chains also occur in post verbal positions:

(ii) el bambini sono arrestati tutti
los muchachos fueron arrestados todos
el nens son arrestats tots

type of dependency exhibits the properties given in the second column. It is clear that our redefinition of the A/A' distinction will have to keep this result constant and adequately predict which types of dependencies exhibit particular properties.

2.4 Toward a definition of the relevant dichotomy

2.4.1 Properties of the VP-external-subject position

As a point of departure, let us consider the properties of the movement of the subject from its VP internal position to the Spec of AGRP-S in a simple sentence. This movement has the properties of an AN-chain, as shown by the absence of weak crossover effects, the absence of parasitic gap licensing, the possibility of floating quantifiers, and locality constraints on the movement. Consider first the following pair of sentences:

(59)

a.Every student$_i$ rushed to the party because he$_i$ wanted to talk about linguistics

b.*I invited every student$_i$ to the party because he$_i$ wanted to talk about linguistics

In (59)a. the pronoun within the adjunct can be interpreted as bound by the quantifier in subject position.[25] This interpretation is not possible in (59)b. The contrast between (59)a. and b. can be explained as follows: assuming that the *because* clause is adjoined to TP,[26] the S-structure of (59)a. and (59)b. will be as in (60)a. and (60)b. respectively (irrelevant details omitted):

(60)

a.[$_{AGRP}$ Every student$_i$ ...[$_{TP}$[$_{TP}$...[$_{VP}$ t$_i$ rushed to the party]] [because he$_i$ wanted to talk about linguistics]]]

b.[$_{AGRP}$ I$_j$ ...[$_{TP}$[$_{TP}$...[$_{VP}$ t$_j$ invited every student$_i$ to the party]] [because he$_i$ wanted to talk about linguistics]]]

In (59)a. the subject quantifier binds both the pronoun *he* and the trace in VP left in the D-structure position of the subject; this is a typical configuration for WCO violations. (59)a., however, shows no WCO effects, and thus provides a first argument that the movement of the subject from the VP-internal position to the Spec of AGRP is an instance of an AN-chain. The surface structure of (60)b. is not a WCO configuration; at LF, however, the quantifier will raise to a position

---

25. Thanks to H.Lasnik for pointing this out.

26. See Rizzi (1989) for some motivation for this assumption. A potential, although weak, argument that a *because* clause is attached higher up than VP comes from condition C effects. Consider the following example:

John invited him$_i$ because he always wanted to know that guy$_i$

Assume that m-command is the relevant for condition C: from this and the grammaticality of this example, it follows that the *because* clause cannot be attached to VP.

from which it will c-command the pronoun. This will lead to a WCO violation, since QR typically creates Vbl-chains.[27]

Additional arguments that the movement of the subject from its VP-internal position to the Spec of AGRP is an instance of an AN-chain come from the absence of parasitic gaps, as shown in (61).

(61)

a. *Mary called me because Peter had met e.

b. *Mary, because Peter had met, called me.

The structure of (61) is given in (62):

(62) [AGRP-S Mary$_i$ [AGR-S [TP...[VP t_i [V called] [NP me]] [because Peter had met e_i]]]]

Here, as above, we have assumed that the *because* clause is adjoined to TP. Given this structure, the trace of *Mary* in VP does not c-command the gap inside the adjunct. If movement to the Spec of AGRP-S were an instance of a Vbl-chain, we would expect the parasitic gap to be licensed. The severe ungrammaticality of the sentence provides a strong argument that movement to the Spec of AGR creates an AN-chain. The possibility of floating quantifiers, which first motivated Sportiche's (1988) proposal for VP-internal subjects, provides

---

27. There are several possible explanations for the impossibility of the bound reading in (59)b. First, it may be due to WCO. Assuming that the quantifier raises at LF to adjoin to AGRP, it would c-command the pronoun and the trace in object position, neither of which c-commands the other. Second, if, as argued by May(1986), the quantifier is adjoined to VP, the quantifier may never c-command the pronoun; a bound reading would in this case be impossible simply because the necessary configuration for pronominal binding never occurs.

confirming evidence that movement to the Spec of AGRP-S creates an AN-chain, since, as we showed in the previous section, floating quantifiers are not licensed by Vbl-chains:

(63)
a.They all called me this morning.

b.They will all call me before tomorrow.

There is thus no doubt that whatever the criterion chosen to distinguish positions heading AN-chains from positions heading Vbl-chains, the movement of the VP-internal subject will have to be classified as an instance of an AN-chain[28]

As discussed in the previous section, given the VP-internal-subject hypothesis (henceforth the VPS hypothesis), the relevant criterion for distinguishing the position of the head of an AN-chain from the position of the head of a Vbl-chain can be neither [+θ] vs.[-θ] position nor [+Pθ] vs. [-Pθ] position.

One possibility which comes immediately to mind, but which we will argue to be insufficient, is to replace the [+Pθ] vs. [-Pθ] distinction used in the LGB framework by the distinction between [+Case] vs. [-Case] position. It is quite clear that within a framework incorporating the VPS hypothesis, the external subject position will be a [+Case] position. On the other hand, the position of an adjoined constituent or of a WH-element in a matrix clause, for

---

28. Fukui (1986) assumes that the distinction between A and A' position is not necessary. He provides a brief account for the absence of WCO effect in subject raising cases, but he neither mentions the question of parasitic gap licencing nor the issue of floating quantifiers and the other properties which we have shown in the previous section to distinguish AN-chains from Vbl-chains.

instance, are under standard assumptions [-Case] positions. Based on this distinction, let us tentatively assume the following definition:

(64)
An AN-chain is a chain headed by a [+ Case] position
A Vbl-chain is a chain headed by a [-Case] position

Given this definition, it will follow straightforwardly that VP subject raising and the standard cases of Passive will be AN-chains. But things get quickly more complicated, even in simple standard cases of Raising such as (65):

(65) John seems to have been arrested

Under standard assumptions the structure of this sentence is as follows:

(66) [John seems [t'_i to have been arrested t_i]]

Clearly, as shown by the impossibility of (67), the position of t'_i in (66) is a [-Case] position.

(67)
a.*It seems John to have been arrested
b.*There seems John to have been arrested

We must consequently wonder about the status of the chain (John, t'_i, t_i) in (66)[29]. The chain in (66) contains (minimally) the following sequence of positions : ([+Case], [-Case], [-Case]). Under the

---

29. We assume, as is natural, that the sequence ($John_i$, t'_i, t_i) cannot be broken up into two chains; if there were two chains, ($John_i$, t'_i) and (t'_i, t_i), the former could not meet the theta-criterion, since it contains no θ-position

definition given in (64) above, this chain will have a segment of a Vbl-chain followed by a segment of an AN-chain. Nothing in what we have said so far prevents this type of chain sequencing, but if this is freely allowed, cases which have been ruled out by the improper movement condition in the LGB framework will incorrectly be predicted to be grammatical. Consider for instance the example in (68):

(68) *John seems [t'$_i$ that it is important [t$_i$ to be intelligent]]

(68) has the same sequence of positions as (66), namely ([+Case] [-Case] [-Case]). There is, however, an important difference between (68) and (66). In the LGB framework, the trace t'$_i$ in (66) is considered to be in an A-position, since it is in the Spec of IP, while the trace t'$_i$ in (68) is considered to be in an A'-position, since it is in the Spec of CP. This distinction no longer holds under the definition in (67), since the Spec of IP in (66) and the Spec of CP in (68) are both [-Case] positions. The definition in (64), tentatively adopted above, replaces the A/A' distinction of the LGB framework with the [+Case]/[-Case] distinction. The contrast between (66) and (68) shows, however, that the latter distinction is not sufficient: there is no way in such a system to distinguish between [-Case] positions that were A-positions in the LGB system, such as the Spec of an infinitival clause, and [-Case] positions that were A'-positions in the LGB system, for instance adjoined positions or the Spec of CP. As a result, a condition on improper movement cannot be derived.

The condition on improper movement in LGB bars movement from an A'-position to an A-position. Quite clearly, however, a similar

formulation taking into account the distinction between [+Case] and [-Case] positions cannot be correct. We cannot bar movement from a [-Case] position to a [+Case] position, since there are, in fact, many cases in which such movement must be allowed. To mention a few, simple cases of Passive or the movement of the subject from a VP-internal position to the Spec of AGRP are typical instances of movement from a [-Case] position to a [+Case] position. So is, of course, the instance of Raising in (66). Yet it seems that some condition on improper movement is needed to ban sentences of the type in (68).

Cases of Passive and movement of the subject from the VP-internal position superficially differ from the case in (66) in that they involve movement from a [+θ] position to a [+Case] position. We could try to use this difference to reformulate the ban on improper movement as a ban on movement from a [[-Case] [-θ]] position to a [+Case] position.

(69) Condition on improper movement (to be modified)

$$*C = (\ldots a_i, \ldots . a_{i+1}, \ldots)$$
$$[+C] \qquad [-C\ -\theta]$$

(69) would adequately permit regular passives and movement of the subject to Spec AGRP; it would also correctly exclude (68), since movement to a [+Case] position in this example comes from a [[-Case] [-θ]] position, namely the Spec of CP. This would, however, also wrongly exclude (66), since in (66) the movement from the position of t'_i to the position of *John* is also an isntance of movement from a [[-Case] [-θ]] position to a [+Case] position.

A possible way of solving this problem would be to claim that the trace t_i is in fact not part of the structure of the sentence in (63). The VPS hypothesis allows us, in principle, to construe a Raising construction as direct movement from the VP-internal position to the Spec of the inflectional projection dominating *seems*.[30] Under this view, the structure of (63) would be as in (70):

(70) [John$_i$ seems [$_{IP}$ to have been [$_{VP}$ arrested t_i]]]

If we assume that movement can occur directly from a [[-Case] [+θ]] position to a [+Case] position, the sentence in (63) will not be excluded by a ban on improper movement formulated as in (69). There are, however, arguments against such a move. First, as we will show below, there exist cases which are structurally similar to (66) and for which it cannot be assumed that the position of the intermediate trace does not exist. Second, consider (71), in which the ^ symbol represents the possible positions for a floating quantifier such as *all*:

(71) The bandits^seem^to^have^been arrested at the same time.

Under Sportiche (1988) analysis, floating quantifiers provide a diagnostic for the positions through which an NP has passed. As shown in (71), there are several positions in which a floating quantifier can be stranded; this suggests that the movement of NPs in Raising constructions is not direct, but rather involves a number of steps in

---

30. This would entail that the EPP is satified by the VP-internal subject position and that nothing forces the Spec of the upper functional projection AGRP-S to be realized (see Koopman & Sportiche (1988) for a defense of this view).

[[-Case] [-θ]] positions. If so, the ban on improper movement formulated as in (69) is too strong: it incorrectly rules out sentences like (71).

Another possible way of solving the problem would be to modify the definition of chain given in (64) and to make use of the distinction between [+potential Case] position/[-potential Case] position instead of the notion of [+Case] position/[-Case] position. (64) would then be restated as in (72):

(72)
An AN-chain is a chain headed by a [+ Potential Case] position
A Vbl-chain is a chain headed by a [-Potential Case] position

As noted by Chomsky (1989 in class), the notion "potential" is rather tricky to define in a precise way. For the sake of the present argument, however, we will simply rely on a rather intuitive notion of "potential". Let us assume that the notion "potential Case-position" is informally understood as a position in which, depending on the feature value of the head to which it is related, will or will not be assigned Case, and accordingly will or will not license the occurrence of an overt NP. If this is correct, the position of t_i in (66) could arguably be regarded as a [+potential Case] position, since if the complement of *seems* is [+Tense], (presumably a change in the feature value of the sentential head) a subject NP can appear in the external subject position.

(73) It seems that John has been arrested.

Similarly, the position of the complement of a passive verb could be understood as a [+Potential Case] position. It has been standardly assumed in the LGB framework that one of the effects of Passive is to suppress the Case-marking ability of the verb. This could plausibly correspond to a change in the feature value of the verbal head.

Given (72), the ban on improper movement could be reformulated as a ban on movement from a [-potential Case] position to a [+potential Case] position.

(74) Condition on improper movement (to be modified)

$$*C = (\ldots a_i, \ldots a_{i+1}, \ldots)$$
$$[+PC] \quad [-PC]$$

Returning to the structure in (66), both t'_i, the trace in the Spec of IP, and t_i, the postverbal trace, are in [+potential Case] positions; the chain in (66) contains no [-potential Case] positions and is thus acceptable. The chain in (68), however, contains an instance of movement from a [-potential Case] position to a [+Case] position, namely the movement from the Spec of CP to the Spec of the higher AGRP, and is thus excluded. So the definition in (72) and a ban on movement from a [-potential Case] position to a [+potential Case] position apparently make the right distinctions. However, as we will argue in the next section, certain cases of AN-chains must include positions to which Case is never assigned and which never license an overt NP. Given our intuitive approach to the notion potential Case-position, these positions must be considered [-potential Case] positions. Given the existence of such constructions there are only

two possibilities: first, one could try to refine the notion of "potential Case-position" so as to include some positions to which Case is never assigned at S-structure (but which may be capable of receiving Case at LF [31]), and to exclude other positions to which Case is never assigned at all. Second, one could simply abandon the definition in (72) and adopt a different criterion to distinguish AN-chains from Vbl-chains. Since it is quite unclear what the refinement of the notion "potential Case position" could be, it is simpler to choose the second option. We thus propose that (72) be replaced by the following definition:

(75)

An AN-chain is headed by a [+HR] position

A Vbl-chain is headed by a [-HR] position

2.4.2 Properties of intermediate specifiers

We have been concerned so far with the properties of the position of the VP-external-subject. In this section, we turn tc a consideration of the properties of the specifiers of intermediate functional

---

31. Chomsky (1989) suggests that structural Case is always assigned through Spec Head agreement. Thus in his view, the Case of an object NP is assigned in the Spec of the AGR-O projection. Since an NP cannot occur in this position at S-structure in a language such as English or French, (*John has the book bought) it must be assumed that the Spec of AGR-O does not assign Case before LF. This is thus an instance of position which although a [-Case] position at S-structure is a [+Case] position at LF.

projections and we consider first cases of movement through these intermediate specifier projections.

The first construction we turn to is the French passive construction. Consider a simple case of passive in French:

(76)
Les tables seront repeintes.
The tables will be repainted(F Pl.).

The main difference between French and English passives is that French, unlike English, manifests overt agreement of the past participle with the derived subject[32]. Kayne (1989) has proposed an account of past participle agreement which establishes a parallel between the agreement of a subject with its verb and the agreement of a moved object and the past participle. In his view, the structure of auxiliary constructions in French and Italian involves a functional projection headed by Agreement: past participle agreement occurs whenever the trace of a moved NP occupies the specifier position of this AGR projection. Under Chomsky's (1989) view, this AGR-projection is the AGR-O projections. Thus, the structure of (76) is as in (77):

(77) Les tables$_i$ seront [$_{AGRP\text{-}O}$ t'$_i$ [$_{AGR'\text{-}O}$ AGR-O [$_{VP}$ repeintes t$_i$]]]

Here t'$_i$ is the trace of the moved NP *les tables* in the Spec of the AGR-O projection. As assumed by Kayne (1989), the presence of the trace is necessary to trigger the agreement of the past participle. As shown in (78), past participle agreement is obligatory in the French passive construction.

---

32. This is encoded with features in the English glosses.

(78)
*Les tables seront repeint
The tables will be repainted (M sg.)

The ungrammaticality of (78) can be taken to show that the movement of the object in the passive construction must pass through the Spec of the AGR-O projection. Thus, in this case, contrary to the standard example of raising discussed above, it could not be argued that movement occurs directly from the direct object position to the external subject position. In other words, the AN-chain created by the movement of *les tables* in (77) must necessarily include the trace t'_i in the specifier position of the intermediate AGR-O projection. The question we now need to ask is whether this specifier position is a position to which Case can ever be assigned, i.e., a [+ potential Case] position. As is well known, French has a form of impersonal passive which alternates with the regular passive construction involving the raising of the NP. An example of impersonal passive is given in (79):

(79)
Il a été repeint beaucoup de tables
There have been repainted (M sg.) many tables

As shown in (80), an NP can never occupy the intermediate position between the auxiliary and the past participle:

(80)

*Il a été beaucoup de $tables_i$ repeintes t_i

There have been many tables repainted (F Pl.)

The ungrammaticality of (80) follows straightforwardly from Case theory if it is assumed, as is natural, that the NP *beaucoup de tables*

cannot receive Case in the preverbal position either from the auxiliary in ECM fashion or through Spec-head agreement with the AGR-O head.[33] Thus, at least in the passive construction, the Spec of the AGR-O projection is not a Case marked position, and thus does not permit the occurence of an overt NP. Past participle agreement also occurs in French active transitive constructions if the object has been moved either under WH-movement or under clitic placement:

(81)
a. les tables que j'ai toutes repeintes t_i
the tables that I have all repainted (F Pl.)

b. ces tables, il les a bien repeintes t_i
these tables, he them has well repainted (F PL.)
these tables, he has definitely repainted them

In Kayne's view, these constructions involve a trace in the Spec of an AGR projection so that the structure of (81)b., for instance, is as in (82):[34]

(82) [il les$_i$ a [AGR-OP t_i [AGR-O' AGR-O [VP repeintes t_i]]]]

---

33. In French, the type of agreement which occurs on the past participle differs from the type of agreement which occurs on a tensed form of the verb in that the former shows gender and number agreement while the latter shows person and number agreement.
(i) Nous dormons. [ons] = 1 pers pl
(ii) Il nous a vues [es] = fem, pl
It is possible that only forms of agreement which include some form of person marking can be Case assigner.

34. Kayne assumes that in cases of WH-extraction, the trace triggering agreement is not in the specifier of the AGR projection, but simply adjoined to it. This assumption is important in his system to block past participle agreement with a WH-moved object in impersonal constructions. Since it is not pertinent to the line of argument we are pursuing in this section, we will not consider this alternative (see Kayne (1989) for details.

Here again, however, the movement of an object NP to the intermediate specifier of the AGR-O projection is impossible, as shown by the ungrammaticality of (83):

(83)
*Il a les tables repeintes
He has the tables repainted (F Pl.)

The same is true in an infinitival constructions, as in (84):

(84)
*Je crois avoir ces tables repeintes
I believe to have these tables repainted (F Pl.)
I think I repainted these tables

From the ungrammaticality of (80), (83), and (84), we can conclude that the specifier of the lower AGR-O projection in French is never a Case-marked position. Returning now to the case of passive in (77), we conclude that the chain (les tables, t'_i, t_i) involves an instance of movement from a [-(potential) Case] position to a [+ (potential) Case] position, namely the movement of the Spec of the lower object AGR-O projection to the external subject position, the Spec of AGR-S.

Under either formulation (69) or (74), of the condition on improper movement, this chain will be ruled out since it involves a Vbl-chain followed by an AN-chain. In other words, under (69) or (74), movement through the Spec of AGR-O is excluded. Consequently, all passive sentences in French are wrongly excluded, since as we have shown, in these cases, passage through the Spec of AGR-O is necessary. Clearly, then, the given definitions are shown to be inadequate.

The same conclusion can be arrived at from instances of Raising and Passive in English. Although there is no obligatory agreement in English Passive or Raising, the occurrence of floating quantifiers in

intermediate specifier positions indicates that the movement of the NP does not occur in one step. Floating quantifiers can occur, for instance, between two auxiliaries, a position to which Case is clearly not assigned, since it can never host an overt NP:

(85)a. The demonstrators would have all been arrested

b.*There would have many demonstrators been arrested

The ungrammaticality of (85)b. shows that there are intermediate specifiers in English which are [-potential Case] positions. On the other hand, the grammaticality of (85)b. shows that these position can be part of well formed AN-chains.

Let us summarize: we have shown that the replacement of the [+Pθ]/[-Pθ] distinction with the [+Case]/[-Case] distinction or the [+potential Case]/[-potential Case] distinction does not permit an adequate account of AN-chains. Although it is clear that the head of an AN-chain must occur in a Case-marked position, intermediate specifier positions which are part of well-formed AN-chains can be positions to which Case is never assigned. If Vbl-chains are defined as chains headed by [-Case] positions or [-potential Case] positions, this has as a consequence that some well-formed AN-chains will contain subchains which are Vbl-chains. This is undoubtedly an undesirable situation, since there are structures such as (68) in which this possibility must be ruled out.

To resolve this conflict, we propose that the property of positions relevant to distinguishing AN-chains from Vbl-chains is the distinction we established in the previous section between head-

related positions, i.e., specifiers and complements in the X'-theoretic sense, and non-head-related positions i.e., adjoined positions, either base-generated or derived.

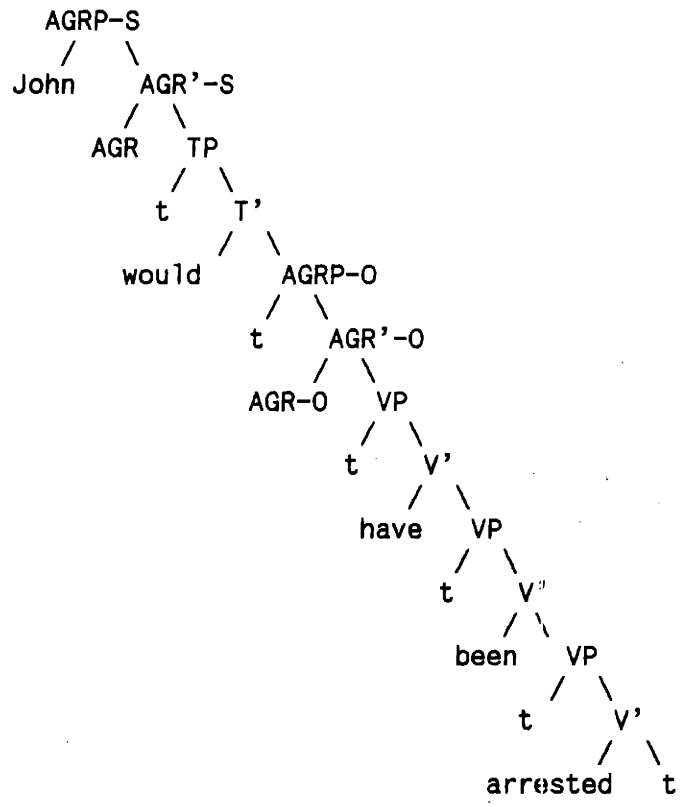
(86)

AN-chains are headed by [+HR] positions

Vbl-chains are chains headed by [-HR] positions

Given this definition, the chain of a passive, for instance, which includes a trace in the specifier of the AGR-O projection, will not involve improper movement: the moved object will be able to pass through the specifiers of various lexical and functional categories on its way to the spec of AGR-S. Thus, the derivation of a sentence such as (85)a. will be as follows:

(87)



The derivation in (87) adequately predicts the various possible positions of floating quantifiers.

French allows a variety of floating quantifiers to occur simultaneously. Consider the following example :

(88)
Les enfants ont tous pour la plupart chacun contribués deux dollars
The children have all for the most part each contributed two dollars

35

---

35. See Ruwet (1982) for arguments that *pour la plupart* has the properties of a floating quantifier. We abstract away here from the problem of deriving such a structure from the impossible subject NP **Tous pour la plupart chacun des enfants*. But if we assume simply that a floating qunatifier must be adjoined to an NP trace to be licenced, then we do not have to assume that they have been stranded under movement. This slight modification of Sportiche's view allows an account of structures such as (88).

Under the view adopted here, (88) has the following structure:

(89) [$_{AGRP-S}$ les enfants$_i$ [$_{AGR'-S}$ ont$_v$ [$_{TP}$ [tous t_i] [$_{T'}$ t_v [$_{AGRP-O}$ [pour la plupart t_i] [$_{AGR'-O}$ t_v [$_{VP}$ [chacun t_i] [$_{V'}$ t_v [$_{VP}$ t_i [$_{V'}$ contribués deux dollars]]]]]]]]][36]

Given the definition in (86), the problem of having a well-formed AN-chain containing segments which are defined as Vbl-chains no longer arises. Improper movement can be simply defined as in (90):

(90) * C = (...α_i, ...α_{i-1}...)
[+HR] [-HR]

This will unproblematically allow for derivations of the examples in (87) and (89) as well as for (66) or (77), since the Case properties of intermediate positions are not taken into consideration. Clearly, however, we must maintain the requirement that the head of an AN-chain must appear in a [+Case] position so as to prevent NP movement from stopping in any [+HR -Case] positions, as in (80), (83), (84), and (85)b. To this effect, we simply follow Chomsky (1986)a. and adopt the condition given in (91) on the well-formedness of AN-chains, defined now as chains headed by a [+HR] position, which we will from now on simply call [+HR] chains.

---

36. Whether passage through the various positions is obligatory or not will remain for now an open question.

(91)

If C = (a_1.....a_n) is a [+HR] chain, then a_n occupies its unique θ-position and a_1 its unique [+Case] position.[37]

With the definitions in (86), (90) and (91), all instances of acceptable AN-chains we have considered in this and the previous section will be appropriately accounted for. But a problem arises again with respect to sentence (67), repeated here for convenience.

(92) *[John$_i$ seems [t'$_i$ that [it is likely [t$_i$ to be intelligent]]]]

There are, of course, several possible derivations of this sentence. Suppose first that the trace t'$_i$ is not in the Spec of CP, as we assumed earlier, but that it is instead adjoined to CP. Under this derivation the definition of improper movemement in (90) will correctly rule out (92). However, it will not rule out an alternative derivation in which t'$_i$ is in the Spec of CP. Indeed, given the definition in (86), the Spec of CP is a [+HR] position, so passage through it does not violate improper movement as defined in (90). It thus appears that the problem we started with remains unsolved. Note, however, that this is not quite correct: we have made some progress, since we can at least differentiate a derivation which proceeds through the Spec of CP from one which does not. With the previous definitions ((64) and (72)), an adjoined structure and a structure where t'$_i$ is in the Spec of CP are indistinguishable. Given the

---

37. We assume with Chomsky(1986)a. that this condition should follow from other principles, but we leave open the question of how this is ultimately achieved.

definition in (86), it would be sufficient to invoke an additional condition ruling out passage through the Spec of CP in (92) to account for its ungrammaticality. In Chapter IV section 4.5.2, we will argue that the impossibility of the passage through the Spec of CP in cases such as (92) follows from a conspiracy between the ECP and the condition on improper movement as defined in (90). This will account fully for the ungrammaticality of cases such as (92).

Cases such as (92) raise the question of the status of the Spec of CP. Under the definition given in (86), the Spec of CP is indistinguishable from the Spec of other functional projections such as AGR-S, TP or AGR-O. This is in fact an important difference between the theory we propose and the LGB framework, in which the Spec of CP is distinct from the Spec of IP since it is an A'-position, on a par with adjoined positions. We will postpone a discussion of this important issue until Chapter IV where we will propose a theory of Vbl-chains compatible with the definition in (86).

2.5 Summary of Chapter II

In Chapter II, we have raised the problem of the inadequacy of the standard definition of the A/A' distinction in a framework which includes the VP-internal subject hypothesis and the Split INFL hypothesis. We have argued that the distinction which is made in the LGB model between two types of chains, which we have called AN- and Vbl-chains, should be maintained since the two types of chains manifest quite distinct properties. To this effect, we have proposed to recast the distinction between A and A' positions in terms of a

distinction between [+HR] positions and [-HR] positions defined as follows:

(93) α is a [+HR] position iff α is a sister of X' or a sister of X_0.

Under this view, every specifier and every complement of any (functional) projection may head a chain with AN-properties. On the other hand, chain with Vbl-properties must be headed by [-HR] positions. The recast of the A/A' distinction in terms of the [+HR] / [-HR] distinction has two major consequences:

-first, given the split INFL hypothesis, and the version of it we have adopted (namely the sentence structure proposed by Chomsky (1989)) there are a number of positions which we expect, in principle, to be able to head [+HR] chains with AN-properties. In particular, we expect that the Specifier of the AGR-O projection should be able to host the heads of [+HR] chains in languages where AGR-O is a Case assigning head.

-Second, since our proposed definition does not take into account the particular properties of a given head, in our view, all specifiers of functional projections are [+HR] positions. This maximally simple position seems at first to have an unwanted consequence: namely, under our view, we expect movement to the Spec of COMP to manifest the same properties as the movement to the Spec of any other functional projection.

The first of these two consequences, namely the existence of movement to the Specifier of intermediate functional projections will be the subject matter of Chapter III. We will postpone the discussion of the second major consequence until Chapter IV, where we will propose a general theory of [-HR] chains which will involve some important modification of the theory of ECP.

In Chapter III, we will pursue our discussion of [+HR]-chains. Up to now we have looked at cases of movement to the specifier of AGR-S and <u>through</u> the specifiers of intermediate functional projections. In Chapter III, we consider the properties of movements of NP <u>to</u> the specifiers of intermediate functional projections. As we have seen, French and English do not allow this type of movement.[38] For this reason in Chapter III, we turn to more "exotic" languages which allow this type of movement and consider the properties of such movements.

---

38. As we will argue in section 3.4.6 of Chapter III, although French does not permit the movement of an NP to the Spec of AGR-O, it manifests a particular instance of movement to the Spec of the AGR-O projection, namely the movement of a bare quantifier *tout* in sentences of the type: *J'ai tout envoyé par la poste* ("I sent everything by mail").

Chapter 3

Movement to intermediate specifiers

3.0.1 Introduction

In Chapter II, we proposed distinguishing AN-chains from Vbl-chains on the basis of the [+HR]/[-HR] distinction . In other words, we have proposed to recast the A/A' dichotomy of the LGB framework in terms of the [+HR]/[-HR] dichotomy and we have shown how the proposed dichotomy accounts for standard cases of AN-chains such as passive and Raising. In this chapter, we pursue our study of AN-chains but we turn to other cases of movements which have been less studied. This chapter is concerned with movement of objects (in a broad sense) to some intermediate the position within a clause, a phenomenon which is manifested in various languages--such as Hindi, Japanese, Icelandic, mainland Scandinavian, German, and Dutch (among others). As we will show, these movements of objects (which we will designate under the generic term of "object-shift") create chains with AN-properties. If so, it follows from the definitions we have proposed in Chapter II, that the chain created by "object shift" must be a chain which is headed by an [+HR] position. Under the standard view of sentence structure, (the LGB structure for instance), there is no [+HR]

positions available in intermediate position within a clause. The situation is entirely different however, with the sentence structure proposed in a model which adopts the Split INFL Hypothesis. As we have seen in what precedes, there are in fact many [+HR] positions, i.e, the specifiers of intermediate functional projections, which are potential candidates as landing sites for movement. We thus propose to analyze "object shift" as an instance of movement to the specifier of intermediate functional projections. The success of the proposed analysis to account for the various properties and constraints on "object shift" across a number of significantly different languages in turn provides evidence for the Split INFL hypothesis and the existence of various functional projections in these languages.

One of the type of "object movement" we consider is the phenomenon which is know in the literature as "scrambling". In this chapter, we propose a reanalysis of Scrambling in various languages (Hindi, Japansese,German); sentence internal scrambling is, in our view, an instance of movement to the specifier of an intermediate functional projection (ie it is an instance of what we have called "object shift"). In the literature, scambling has been most often analyzed as involving "adjunctions" to the sentence initial position or to the VP[1]. As we show, a number of the properties of the chain created by sentence internal scrambling conflict with this hypothesis. These properties are, on the other hand, accounted for naturally under our proposed analysis.

---

1. See Mahajan (1989) for a different proposal which we adopt and further develop. We review his arguments in the section to come.

Turning to languages where "object shift" is more restricted, (such as the Mainland Scandinavian languages and Icelandic), we show that the availability of "object shift" analyzed as movement to the specifier of intermediate projections is intrinsically tied with the possibility of verb movement. Under the assumption that head movement of the verb at S-structure is parametrically determined, we show that the availability of "object shift" in a given language is simply a consequence of the positive setting of this parameter and of certain properties of the functional projections such as, for instance, the possibility to assign Case at S-structure. Thus the properties and constraints of "object shift" accross various languages simply follows from general principles and parameters which are limited to the properties of the functional projections.

It has been suggested by several linguists (Borer (1984), Webehuth (1988), Chomsky (1986), (1989) among others), that parameters of Universal Grammar are restricted to properties of the lexicon and do not concern the computational system of language. Moreover, if as suggested by Chomsky (1989) " substantive elements (verbs, nouns etc) are drawn from an invariant universal vocabulary, then only functional elements will be parametrized." (Chomsky (1989) p 44). Our proposal concerning the various properties and constraints on "object shift" accross languages is directly within the lines of this view on language parametrization.

3.1 Hindi Scrambling

The first language we consider, is Hindi. In a recent paper, Mahajan (1989) notes that sentence internal-scrambling (henceforth SIS) exhibits two properties which would be very surprising if scrambling were analyzed as an instance of movement whose landing site is an adjoined position,i.e, as an instance of movement which in our view (as well as in the LGB model) is expected to have properties of Vbl-chains:

1) Scrambling suppresses weak-crossover effects[2]

2) Scrambling creates new-binding possibilities

To account for properties 1&2, Mahajan proposes to view scrambling as an instance of movement to the specifiers of functional projections, which he assumes to be "A-positions"[3].

---

2. This phenomenon was previously observed by Gurtu (1986).

3. In Mahajan's view, "A-positions" are defined as follows:

(1) An "A-position" is a potential Case-position or a θ-marked position

In this work, the term "A-position" must be simply understood as a label which does not refer to the original distinction among positions in LGB but rather to the properties of chains. Thus in our terminology we would rephrase Mahajan's definition as follows: AN-chains are headed by [+θ] positions or [+potential Case] positions.

I believe that this definition is problematic on two counts. First, it involves a disjunctive definition, which are generally taken to be symptomatic of a missed generalization. Second, Mahajan (1989) gives no clear definition of the notion "potential" as applied to Case-marking. In the chapter II, we have shown that a rather intuitive definition of "potentially Case-marked" based on feature distinction is insufficient. Thus a more complex definition of "potential Case-marked position" is needed. Pending such a

In this section, we will review Mahajan's arguments for an analysis of scrambling as creating AN-chains. Moreover, we will provide an additional argument in favor of this view based on the distribution of Floating quantifiers in Hindi.

3.1.1 Word order in Hindi

Let us first start by outlining some basic properties of the syntax of Hindi. In ordinary Hindi sentences, the order of constituents is not fixed: most constituents in a sentence are freely permutable. For instance, a simple sentence such as (2) allows all possible orders of its constituents.

(2) a.Raam-ne kellaa khaayaa
Ram (SUB) banana(DO) ate

b.raam-ne khaayaa kellaa
c.kelaa raam-ne khaayaa
d.kelaa khayaa raam-ne
e.khayaa raam-ne kelaa
f.khayaa kelaa raam-ne

However, word order in Hindi is not completely free. There are constraints, for instance, on the combination of sentential negation with other constituents in the sentence. For some description and an account of these constraints we refer the reader to Mahajan (1989)b.. One example of an impossible order is given in (3):

(3) *raam khaaegaa roTii nahiiN
Ram eat-Fut bread not

---

definition, the notion "potential Case-position" remains too vague to have predictive value.

Ram will not eat bread

Despite the great freedom in word order, many linguists (Kahkru , Bhatia, Gurtu...) have assumed that Hindi is a configurational language with a hierarchical structure. One argument in support of this hypothesis comes from the fact that the language manifests WCO effects in examples such as (4):

(4)*uskii bahin sab-koi pyar karii he
his sister everyonei loves
His sister loves everyone

In (4), the pronominal in the subject NP cannot be interpreted as bound by the object quantifier. Thus, (4) shows that the object does not c-command the subject. Consequently, Hindi cannot be assumed to have a flat structure in which every argument c-commands the others; it must be assumed to have a VP. A similar argument can be constructed to show that the indirect object asymmetrically c-commands the direct object, thus providing some argument for a hierachical structure within the VP. The relevant example is given below:

(5) *raajaa-ne unkei pitaa-ko sab daasiyaani lOTaa diiN
king(SUB) their father(IO) all maids(DO) returned
The king returned all maids to their father

The unmarked word order of Hindi is generally assumed to be S IO O V[4]

---

4. The assumed order apparently conflicts with the thematic hierachy proposed by Larson (1988) See Chapter II section 2.2.1.2 Agent<theme<goal. Here the needed hierachy should be Agent<goal<theme. Possibly one could argue that the head final nature of the language matters wrt the way the arguments are projected in the VP. This would however conflict with the Universal Theta Alignement Hypothesis proposed by Baker (1986). This problem is apparent for all the SOV languages we study and many more. A solution would necessitate an investigation of the relation between headeness and the Projection Principle, an issues which is clearly

(6) raam-ne baccoN-ko Ser dikhaayaa
Ram children tiger showed
Ram showed the tiger to the children

In addition to the evidence provided by the WCO effects, an independent argument in favor of this base order can be constructed on the basis of interesting facts about generic negative sentences.[5] In generic negative sentences, the object of a verb does not bear an overt Case marker.[6] This is shown in (7) and (8).

(7) mai Ciinii nahiiN khaataa huuN
I sugar not eat
I don't eat sugar

(8) mai gariibooN-ko paisaa nahiiN deta huuN
I poor-DAT money not give
I don't give money to the poor

When the NP does not have any overt Case-marker and has a strictly generic interpretation, the only word order which is perfectly acceptable is the one given in (7) for transitive verbs and in (8) for ditransitive verbs, namely,the word order which is generally assumed

---

outside of the scope of this dissertation. We thus leaves this problem unresolved.

5. Thanks To Utpal Lahiri for the judgments on these examples.

6. The fact that generic negative sentences behave differently in many languages, for instance Russian (Genitive of negation) and French (pas de), has led us to expect a similar difference in Hindi. It is confirmed by the facts in an interesting way. In many languages (Russian, Polish, Finnish, French...) it appears that the negative generic sentences require a special Case. In Hindi, interestingly, a language in which Case-marking is overt and seems to allow partial word order freedom, the peculiarty of negative generic sentences is expressed with rigid word order and "adjacency".

to be basic. Sentences in which the object NP occurs in initial position, before a dative object, or over an adverb are degraded.[7]

(9) a.?? Ciinii mai nahiiN khaataa huuN
Sugar I not eat

b.?? paisa mai garibo-ko nahiiN deta huuN
money I to the poor not give

c.?? mai paisa garibo-ko nahiiN deta huuN
I money to poor not give

d.?? mai garibo-ko paisa saadhaarantaya nahiiN deta huuN
I to poor money usually not give[8]

The contrast between (7) and (8) on the one hand, and (9) on the other, seems to indicate that generic NPs must occur in their base positions and are resistant to movement. If so, negative generic sentences provide independent evidence for the proposed base order S IO DO V.

---

7. What is degraded is the generic interpretation. In other words, the scrambled NP receives a contrastive interpretation.

8. Interestingly, sentence (9)d. indicates that there is an adjacency effect for a generic object. One may wonder, then, why the presence of the negation intervening between the generic object and the verb does not also cause an adjacency violation. In (7) and (8), it seems that the negation intervenes in the middle of the VP. The difference between the effect of an intervening adverb and the effect of an intervening negation may be explained if, as suggested by Pollock (1989), negation is a head while adverbs are adjuncts. As is usually assumed, Hindi is a head-final language. If negation is a head, it would be expected to be head-final like any other head. If so, the order in (7) can be assumed to be derived by rightward movement of the verb from its base position to an inflectional head above negation. Adjacency is respected in this case since the generic object is adjacent to the trace of the moved verb. In (9)d., however, the adverb is presumably adjoined to the VP. If so, the order in (9)d. must have occured by raising the object out of the VP constituent. In this case, then, adjacency of the object with the V-trace is not respected, accounting for the degraded status of the sentence.

The variation in word order is usually assumed in the literature (Bhatia (1976) Kakhru (1978) among others) to be due to a scrambling rule which is taken to be an instance of Move α. That scrambling is not simply a PF phenomenon is shown by the examples given above in (9). The marginality of these examples suggests that "scrambling" is syntactically restricted. This is confirmed by the fact that long distance scrambling obeys island constraints, for instance, the complex NP constraint:[9]

(10)
a.
araam yah baat ki mohan-ne siitaa-ko dekhaa nahiiN jaantaa
Ram this fact that Mohan Sita saw neg knows
Ram knows the fact that Mohan did not see Sita

b.
*siitaa-ko raam yah baat ki mohan-ne dekhaa nahiiN jaantaa

The contrast in (10) and the marginality of the sentences in (9) show that "scrambling" has syntactic effects and must therefore apply in the syntactic component.

3.1.2 Sentence Internal Scrambling

Given that scrambling is syntactic movement, we may now wonder about the properties of the chains created by this type of movement. Does scrambling have the properties of an AN-chain or the properties of a Vbl-chain? As pointed out by Mahajan, it is important to distinguish

---

9. As noted by Mahajan, this fact also provides an argument that the empty category left by long-distance scrambling is not pro. See Mahajan (1989) for detail on this argument.

long distance scrambling (LDS) from sentence internal scrambling (SIS), since these two instances of scrambling do not have the same properties. We will consider each type of scrambling in turn, concentrating first and foremost on SIS.

As is shown by (4) and (5), Hindi manifests WCO effects. As (11) shows however, scrambling does not induce WCO violations.

(11)
daasiyaan$_i$ raajaa-ne unke$_i$ pitaa-ko t_i lOTaa diiN
the maids$_i$ the king to their$_i$ father t_i returned

(11) is a typical s-structure configuration for WCO, since the scrambled constituent c-commands both a coindexed pronominal and a trace which do not c-command each other. But WCO does not occur. This fact shows that the movement of the NP under scrambling does not leave an empty category which has the properties of a variable. Thus, (11) provides a first argument that scrambling in Hindi has the properties of an AN-chain.

The second argument supporting this view is provided by interesting facts first noted by Gurtu (1986) and reported in Mahajan (1989): when applied to a quantifier or a WH-phrase in situ, scrambling supresses WCO violations. If an object quantifier which induces a WCO violation when it occurs in its base position is scrambled over a non c-commanding pronoun, binding of the pronoun becomes possible and WCO effects no longer obtain. Consider the pair in (12):

(12) a.*uskii$_i$ bahin sab-ko$_i$ pyar karii he
everyone$_i$ his$_i$ sister loves

b. sab-ko$_i$ uskii$_i$ bahin pyar karii he
everyone$_i$ his$_i$ sister loves

In (12)b., the quantifier *everyone* has been srambled over the subject and the bound reading of the pronoun has become possible. Similar facts can be observed with WH-in-situ constructions. When the WH-phrase remains in situ, the WCO effect obtains. When the WH-phrase is scrambled the WCO effect disappears.

(13) a.uskii$_i$ bahin kis-ko$_i$ pyar karii he
his$_i$ sister who$_i$ loves

b.kis-ko$_i$ uskii$_i$ bahin pyar karii he
who$_i$ his$_i$ sister loves

The examples given so far all involve subjects and objects. As examples (14) and (15) show, the disappearance of the WCO effects also shows up in structures with ditransitive predicates, when a direct object is scrambled over an indirect object and lands either before or after the subject:

(14)
a.*raajaa-ne uske$_i$ pitaaa-ko kOn sii daasii$_i$ lOTaa dii
king(subj) her father which maid returned
Which maid did the king return to her father.

b. kOn sii daasii$_i$ raajaa-ne uske$_i$ pitaaa-ko lOtaa dii

c. raajaa-ne uske$_i$ kOn sii daasii$_i$ pitaaa-ko lOtaa dii

(15)

a.*raajaa-ne unke$_i$ pitaa-ko sab daasiyaaN$_i$ lOTaa diiN
king their father all maids returned
the king returned all maids to their fathers

b. sab daasiyaaN$_i$ raajaa-ne unke$_i$ pitaa-ko lOTaa diiN

c. raajaa-ne sab daasiyaaN$_i$ unke$_i$ pitaa-ko lOTaa diiN

The data in (12)-(15)[10] show that sentence-internal scrambling (SIS) not only does not create WCO violations but also repairs potential WCO violations. The fact that scrambling repairs WCO violations shows that a scrambled quantifier can bind a pronominal from its landing site position. As we showed in section 2.3, this is not possible in clear cases of Vbl-chains such as WH-movement and topicalization in English. Thus (11) on the one hand and (12)-(15) on the other hand provide evidence that neither the head nor the foot of the chain created by scrambling has properties comparable to those of heads or feet of Vbl-chains; both the head and the foot of the chain created by scrambling have properties of heads and feet of AN-chains. The foot does not have the properties of a variable, as shown by the absence of WCO in (11); the landing site of the head permits new pronominal binding by the moved element, contrary to, for instance, the landing site of topicalization in English, as we showed in section 2.3.

---

10. Note that these facts immediately distinguish scrambling in Hindi from English topicalization. As we have seen above (Chapter II section 2.3), topicalization of a bare quantifier is excluded. If as we assume, following Lasnik and Saito (1989) and Baltin (1982), topicalization is an instance of adjunction, then this provides an additional argument that scrambling must be different from an adjunction structure. Possibly, bare quantifiers are generally excluded from adjunction positions at S-structure.

The hypothesis that scrambling involves an AN-chain is strongly supported by the fact that scrambling also creates new anaphoric binding possibilities. Consider the following examples:

(16)
*/??? apne$_i$ baccoN-ne mohan-ko$_i$ ghar se nikaal diyaa
self's children Mohan house from threw out
self children through Mohan out of the house

(17)
? mohan-ko$_i$ apne$_i$ baccoN-ne ghar se nikaal diyaa
Mohan self children threw out of the house

In (16), the reflexive within the subject NP is unbound, since it has no c-commanding antecedent.(16) is thus ruled out by principle A of the binding theory. In (17), the reflexive has become bound by the scrambled NP and thus satisfies binding theory. In the LGB framework, anaphoric binding requires the antecedent to be in an A-position. Since anaphoric binding is possible from the landing site of SIS, these facts provide additional evidence that the landing site of scrambling must be a T-position which heads an AN-chain.

To the above evidence given by Mahajan(1989), we add an additional argument based on the distribution of Floating quantifiers in Hindi. As shown by (19)-(22), scrambling licenses floating quantifiers. As we have argued in section 2.3.2, floating quantifiers are only compatible with AN-chains.

(18) mohan saare phal khaa jaae gaa
Mohan all fruits eat will

(19) phal mohan saare khaa jaae gaa
fruits Mohan all eat will

(20) raam-ne kitabeN mohan-ko sari lOTa di
Ram books to Mohan all returned

(21) kitabeN raam-ne mohan-ko sari lOTa di
books Ram to Mohan all returned

(22) kitabeN raam-ne sari mohan-ko lOTa di
books Ram all to Mohan returned

(19)-(22) show that floating quantifiers can appear in all the positions in which NPs can be scrambled. But these facts will only provide strong evidence for the status of SIS as AN-chains if we can shown that in Hindi, floating quantifiers cannot be associated with Vbl-chains. To this effect we will briefly consider cases of long distance scrambling.

3.1.3 Long distance scrambling

The fact that floating quantifiers occur in Hindi only with AN-chains and not with Vbl-chains is slightly more difficult to show than in languages such as English or French. Hindi indeed has no obvious instances of Vbl-chains since WH-elements usually remain in situ or are scrambled just like other NPs.[11] As argued by Mahajan, however, Hindi seems to manifest two types of Vbl-chains: rightward scrambling, which unlike leftward scrambling does not repair WCC, and long

---

11. See Mahajan (1989) for arguments that movement of a WH-element to the begining of a main clause is an instance of scrambling.

distance scrambling, which also shows WCO effects. Since the status of rightward scrambling remains rather unclear (see Mahajan fnt 18 p 22),[12] we will confine our attention to cases of long distance scrambling. An example of long-distance scrambling inducing weak crossover effects is given in (23)b.:

---

12. As mentioned by Mahajan, it could be that what appears to be rightward scrambling is in fact due to a leftward movement of the verb. With respect to floating quantifiers, however, an interesting fact may provide support to the assumption that rightward scrambling is indeed an instance of a Vbl-chain. Consider the following paradigm:

(i) raam mohan-ko saari kitabeN nahiiN degaa
Ram Mohan all book not will give
(ii) raam mohan-ko saari nahiiN degaa kitabeN
(iii)??raam saari mohan-ko nahiiN degaa kitabeN
(iv) raam mohan-ko aaj saari nahiiN degaa kitabeN
ram to mohan today all not will give books
(v)* raam mohan-ko saari aj nahiiN dega kitabeN

(i) is an ordinary sentence. In (ii), the object has been rightward-scrambled and a floating quantifier has been left in what is presumably the D-structure position of the object. In (iii), however, rightward scrambling seems to have occured after the object was first leftward-scrambled over the dative object. (v) is similar to (iii), under the assumption that the position of the adverb on the right of the negation indicates that the object was first scrambled leftward. This paradigm shows, then, that a floating quantifier cannot be stranded in a leftward-scrambled position if the object is scrambled rightward. This could be interpreted as evidence that rightward scrambling involves a Vbl-chain, which is incompatible with a quantifier floated in a non-theta position. Alternatively this may be interpreted as showing that rightward scrambling cannot occur from a leftward scrambled position, indicating that perhaps the scrambled position is a [-Case] position. I will leave the question open for further reasearch.

(23)
a.*uskii$_i$ bahin-ne sab-ko$_i$ socaa ki raam-ne t$_i$ dekhaa
His sister everyone thought that Ram saw
It's everyone that his sister thought that Ram saw

b.*sab-ko$_i$ uskii$_i$ bahin-ne socaa ki raam-ne t$_i$ dekhaa
Everyone,his sister thought that Ram saw

As shown by (23b), WCO is not repaired by long-distance scrambling of the quantifier over the subject.[13] This provides evidence that long-distance scrambling is an instance of a Vbl-chain and not of an AN-chain.[14]

Long-distance scrambling, however, does not create WCO in the embedded clause:

(24)
a. kis-ko/sab-ko$_i$ raam-ne socaa ki uskii$_i$ bahin-ne t dekhaa thaa
who/everyone Ram thought that his sister seen be past
Who did Ram think that his sister has seen

b. raam-ne kis-ko/sab-ko$_i$ socaa ki uskii$_i$ bahin-ne t dekhaa thaa
Ram who/everyone thought that his sister seen be past

The absence of WCO in embedded clauses can be easily explained if, as assumed by Mahajan, scrambling proceeds first as an AN-chain in the embedded clause and then as an Vbl-chain in the upper clause.[15]

---

13. Incidentally, (23) also provides evidence that linear precedence is not a relevant factor for WCO and more generally for the facts we have seen so far.

14. Long-distance scrambling from infinitival sentences shows properties of an AN-chains and not properties of Vbl-chains. As shown by Mahajan (1989)b., movement from an infinitival clause can trigger agreement in the matrix clause.

15. A possible explanation for this fact would go as follows: assume that there is some barrier between the highest [+HR] position in the embedded clause and the lowest [+HR] position in the upper clause. To escape the barrier, the long-distance-scrambled element would have to adjoin to it. Once it has adjoined, any further movement to an [+HR] position would be an instance of improper movement. Thus, long-distance scrambling must proceed by

Additional evidence that long-distance scrambling has properties of Vbl-chains comes from the fact that anaphoric binding is not possible with long-distance scrambling. Consider the follcwing pair of sentences:

(25)
a. *apnii bahin-ne socaa ki raam-ne mohan-ko dekhaa
self's sister thought that Ram Mohan saw
Self's sister thought that Ram saw Mohan

b. *mohan-ko$_i$ apnii$_i$ bahin-ne socaa ki raam-ne t$_i$ dekhaa
It's Mohan that self's sister thought that Ram saw

(25)a. is excluded because the anaphor within the subject NP has no available antecedent. (25)b. shows that an NP which has undergone long-distance scrambling cannot serve as an antecedent for the anaphor in the NP. This shows clearly again that long-distance scrambling has the properties of a Vbl-chain.

Let us now return to quantifier-float. As shown by the WCO in (23), long-distance scrambling has the properties of AN-chains in an embedded sentence and the properties of Vbl-chains in a main clause. If this is correct, given the generalization we have proposed in section 2.3.2 on the distribution of FQ, we would expect quantifier-float with long-distance scrambling to be possible in the embedded clause and impossible in the main clause. As shown by the paradigm in

---

adjunction to the right of the highest [+HR] position in the lower sentence and will consequently have the property of a Vbl-chain, i.e., it will induce WCO and not license Q-float. Although this is a plausible analysis for the Hindi facts, pursuing it in its details would require a detailed study of extraction facts in Hindi to determine which projection creates a barrier in this cases. This investigation is beyond the scope of this dissertation.

(26), this prediction is fulfilled. We observe that the distribution of Q-float exactly parallels the distribution of WCO: floating quantifiers can occur in all positions to which an element can be scrambled in embedded sentences but can occur in no position in main clauses:[16]

(26)

a. raam-ne socaa ki phal mohan saare khaa jaae gaa
   Ram thought that fruits mohan all eat will
   Ram thought that Mohan will eat all the fruits

b. raam-ne socaa ki phal saare mohan khaa jaae gaa

c. ?raam-ne phal socaa ki mohan saare khaa jaae gaa

d. ?raam-ne phal socaa ki saare mohan khaa jaae gaa[17]

e. phal raam-ne socaa ki saare mohan khaa jaae gaa

f. phal raam-ne socaa ki mohan saare khaa jaae gaa

g. *phal raam-ne saare socaa ki mohan khaa jaae gaa

h. *phal saare raam-ne socaa ki mohan khaa jaae gaa

These facts provide additional evidence for the difference established between long-distance scrambling and sentence internal scrambling as Vbl-chains and AN-chains, respectively. It also lends support to the generalization, proposed in section 2.3.2, that floating quantifiers are incompatible with Vbl-chains.[18]

---

16. Thanks to A. Mahajan for these data.

17. The question mark in (c) and (d) is independent of the floating quantifiers facts and is due to the fact that this position is not very natural for a long-distance scrambled NP(Mahajan(pc)).

18. Mahajan (1989) takes as additional evidence that SIS creates AN-chains the fact that (SIS) does not permit anaphoric reconstruction. Consider the following paradigm:

---

(27)

a. raam-ne$_i$ mohan-ko$_j$ apnii$_{i/j}$ kitaab lOTaaii
ram mohan self book returned

b. raam-ne$_i$ apnii$_{i/*j}$ kitaab mohan-ko$_j$ lOTaaii
ram self book mohan returned

c. apnii$_{i/*j}$ kitaab raam-ne$_i$ mohan-ko$_j$ lOTaaii
self book ra mohan returned

In (27)a. the anaphor *apnii* can be bound both by the subject and by the dative object. b. shows that reconstruction does not obtain when the NP containing the anaphor is moved to an intermediate position : as indicated by the * on the j index, the dative object can no longer bind the anaphor. c. shows that when the NP containing the anaphor is moved to the sentence initial position, reconstruction obtains with the subject but not with the dative. If as proposed by Barss (1986) reconstruction involves "chain binding", we would expect both possibilities to be allowed. To explain this asymmetry in the reconstruction possibility, Mahajan proposes (contrary to Barss (1986) and Belletti& Rizzi (1986)) that anaphoric reconstruction effects be limited to A' dependencies. The assumption that the movement of the DO (i.e., over IO) is a movement to an A position, explains why anaphoric reconstruction does not obtain with respect to the dative NP, both in (27)b. and c. To account for the reconstruction possibility with respect to the subject NP in (27c) (index i), Mahajan assumes that although scrambling can be movement to an A position, it can also involve adjunction. Assuming this to be correct, the impossibility of reconstruction with the dative object is then left unexplained. There is indeed no a priori reason why adjunction to VP or to an intermediate functional projection should be vorbidden. This problem suggest that the asymmetry in reconstruction possibility is independent of the nature of the chain created by scrambling and might be due to another factor. Evidence for this hypothesis come from Japanese scrambling. In the next section, we observe that Japanese scrambling permits anaphoric reconstruction with the dative complement. As we show, Japanese scrambling manifests otherwise properties of AN-chains. There is thus an apparent contradiction with respect to the reconstruction facts between Japanese and Hindi. We suggest that this apparent contradiction can be reconciled as follows: suppose that what prevents reconstruction in (27)b. is the fact that there is another binder available, namely the subject, and that the availability of this other binder overpowers the effects of reconstruction. Binding via reconstruction is always a marked option. In (110)b., given that there is a marked and an unmarked option to satisfy binding theory, the unmarked option is chosen. The impossibility of reconstruction with the dative in (27)b. might be due to the strong subject orientation of Hindi anaphors.

3.1.3.1 Summary

Let us summarize: The data presented above show that SIS and LDS have respectively the following properties:

(28)

	SIS	LDS
WCO effect	-	+
WCO repair	+	-
new binding	+	-
FQ	+	-

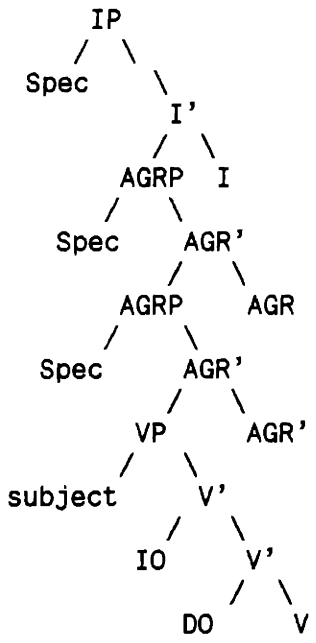
In Chapter II we gave the following definition of AN-chains and Vbl-chains:

(29)
An AN-chain must be headed by a [+HR] position

A Vbl-chain must be headed by a [-HR] position

It follows from this definition that since SIS has the properies of AN-chains it must involve movement to a [+HR] position. To account for the properties of SIS, Mahajan proposes that it involves a movement to the Spec positions of intermediate functional projections. In his view, a basic Hindi sentence has the following structure :

(30)



As Mahajan (1989)b. argues this structure receives independent support from agreement facts in Hindi. Mahajan shows that agreement facts in Hindi can receive an elegant account if we posit the existence of AGR phrases and adopt the theory of agreement proposed by Kayne(1986) for Romance past-participle agreement, which views agreement as a Spec head relation.

We refer the reader to Mahajan (1989) for details on the agreement facts in Hindi and on how they support the structure given above. Given the structure in (30), Mahajan proposes the following analysis for SIS: the subject and the direct and indirect objects can move freely to the Spec positions of any of the AGR projections, since the AGR phrases are not marked for the specific NPs they may host. This accounts for the freedom in the order of arguments to the left of a verb. To account for the properties of SIS, Mahajan proposes to redefine the LGB notions of A- and A'-positions as follows:

(31) A-positions are positions which can be assigned θ-roles

or Case.

Mahajan assumes a version of Case theory recently proposed by Chomsky(1989) (expanding ideas of Koopman & Sportiche (1988)) in which structural Case is assigned under Spec-head agreement with an AGR head. In this view, the specifiers of AGR projections are positions to which Case is (potentially) assigned. Thus, scrambling is an instance of movement from a [[-Case] [+θ]] position to a [+ P Case] position, which resembles standard cases of passive. Since scrambling is movement to A-positions, the fact that it has properties of AN-chains is expected.

Interpreting Mahajan's proposal literally, we might expect scrambling to the Spec of the projection which Mahajan has labeled IP either not to occur or to have different properties. Consider, however, the following sentence:

(32)
Sabko$_i$ uska$_i$ pitaa$_j$ apnaa$_j$ pEsaa sayad nahiiN lOTaye gaa.
To everyone($_i$ his father$_j$ self$_j$ money probably not return will
His father will probably not return his money to everyone

As we have shown above in example (9), an adverb may not intervene between a generic object and negation when the object is (arguably) in its base position. We have suggested in a foonote above that this receives a straightforward explanation under the assumption that there is a D-structure adjacency requirement between a verb and its direct object. This suggests that in (32), the object is not in its base position. This assumption is reinforced by the fact that the adverb in sentence (32) is a sentential adverb, and that under standard assumptions (Cf. Jackendoff (1972) and Travis (1987)), sentential

adverbs are generally thought not to occur in the middle of a VP constituent. If this is correct, then we may conclude that in (32), none of the arguments is in its base generated position. Adopting the structure in (30), we may assume (32) to have the structure in (33):

(33)

[IP sabko [AGRP uska pita [AGRP apnaa pEsa [adv [NEGP [VP tttt] NEG] V]]]]

Since in this sentence the dative quantifier binds the pronominal within the subject phrase, so that there is no WCO effect, we must assume that this quantifier is itself in a position which allows binding. But if, as assumed by Mahajan, the subject and the object must occupy the Spec of the two AGR projections for Case reasons, the only remaining position available for the dative NP is that of the specifier of the IP projection. In Mahajan's view, the specifier of IP is neither a [+Case] position nor a [+potential Case] position. It is thus not, in his terms, an A-position. Under such a view, the absence of WCO is surprising. Given the grammaticality of (32), we conclude that the definition of A-positions as [+potential Case] positions is insufficient.[19] Under the definition that we have proposed in (29) to replace the A/A' distinction of the LGB framework, namely the distinction between [+HR] positions and [-HR] positions, the problem does not arise: the Spec of IP is clearly a [+HR] position, so we do not expect the properties of movement to that position to differ from

---

19. One could probalby try to redefine the notion potential Case-position in such a way as to include the Spec of IP. But, again, such a definition would be rather difficult to formulate precisely.

the properties of movement to the Spec of an AGR projection. The notion of Case-marking does not come into play, so the grammaticality of (32) with the structure in (33) is predicted.

In this section we have given evidence, based on the work of Mahajan (1989) and additional data on floating quantifiers, that scrambling exhibits the properties of AN-chains. Given a framework in which clausal structure involves a number of functional projections, the properties of scrambling follow straightforwardly from an analysis which views this type of movement as a movement to the specifiers of functional projections. As we argued in the previous section, the heads of AN-chains must be in [+HR] positions. If scrambling in Hindi involves a movement to the specifiers of functional projections, this condition will be fulfilled and the properties of scrambling are expected. Under this view, scrambling is the overt manifestation of movement to the specifier positions of intermediate projections. The existence of such movement thus provides support for the hypothesis that the structure of sentences involves a rich system of functional projections and for the redefinition of the A/A' distinction in terms of the [+HR]/[-HR] distinction that we have proposed. In the next section, we review some Japanese data which suggest that Japanese scrambling may be subject to an analysis similar to the one have proposed (following Mahajan) for Hindi scrambling.[20]

---

20. Such an analysis has been independently proposed for Japanese scrambling by Yoshimura (1989).

3.2 Japanese scrambling

The phenomena which support the above analysis of scrambling in Hindi as an instance of AN-chain created by movement to the specifiers of intermediate functional projections are also found in Japanese. As in the previous section, we will discuss WCO effects, reflexive binding, and floating quantifiers. In addition, we will draw some indirect arguments from a cursory discussion of some reconstruction phenomena in Japanese.

3.2.1 WCO in Japanese

3.2.1.1 Word order in Japanese

As is well known, the word order of Japanese is quite free. It has, however, been argued extensively in the literature (Saito (1985) Hoji (1985), Fukui (1986)and reference cited there), that Japanese is a configurational language. Many arguments have been given in the literature to support this hypothesis. It has been also argued that the apparent freedom of word order found in the language is due to a syntactic rule of scrambling which is an instance of Move α. We will not review any of these arguments here, but we will simply adopt this hypothesis. Our goal is to concentrate on the properties of the rule of scrambling. Most linguists who have argued for the analysis of scrambling as an instance of Move α have proposed that scrambling

forms Vbl-chains (A'-dependencies in the LGB sense). It is this assumption which we will question in this section. We will show that scrambling appears in fact to have a number of properties typical of AN-chains.

3.2.1.2 The disappearance of WCO effects

As discussed by Saito and Hoji (1983), the study of weak crossover in Japanese is somewhat complicated by the well-known fact that the Japanese overt pronoun *kare* can never have a bound interpretation, even when it is appropriately c-commanded by a quantifier at S-structure. This is shown in (34):

(34)
** Daremo$_i$-ga [s' kare$_i$-ga Mary-ni kirawareteiru to]
Everyone he-nom Mary-by disliked Comp
omoikondeiru (koto)
be convinced (fact)
Everyone is convinced that he is disliked by Mary

However, as has been observed by Saito & Hoji (henceforth S&H), empty pronouns behave differently from overt pronouns; empty pronouns can receive a bound interpretation when c-commanded by a quantifier or an in situ WH-phrase . As expected, empty pronominals also manifest WCO effects when a quantifier or an in situ WH-phrase occurs to their right. (35) and (36) give a typical paradigm of WCO in Japanese(from Hoji (1985)).

(35)
[Dare$_j$-ga [$_{VP}$ [$_{NP}$ [$_S$ pro$_j$ hitome t$_i$ mita]
who-nom once saw
hito$_i$]-o sukininatta] no
person-acc fell in love with
Who$_j$ fell in love with a person who (he$_j$) took a glance at

(36)
*?[$_S$ [$_{NP}$ [$_S$ t$_i$ hitome pro$_j$ mita] hito$_i$]-ga
once saw person-nom
[$_{VP}$ dare$_j$-o sukininatta] no
who-acc fell in love with
Who$_j$ did the person who took a glance at (him$_j$) fell in love with

(35) shows that when c-command obtains, a QNP or a WH-word can bind an empty pronominal: e$_j$ can be a semantic variable in (35). In (36) the empty pronominal e$_j$ is not c-commanded at S-structure by the QNP. At LF, however, the QNP will raise leaving a variable to the right of the pronoun. Thus, QR creates a typical LF configuration of WCO. The schematic structures of (35) and (36) at S-structure and at LF are given respectively in (37) and (38):

(37)

S-structure [QNP$_j$ ... [...pro$_j$...]]

LF [QNP$_j$ [t$_j$...[...pro$_j$...]]]

(38)

S-structure [[...pro$_j$...] ...QNP$_j$]

LF [QNP$_j$ [[...pro$_j$...] t$_j$...]]

The same effect of WCO repair that has been observed in Hindi inthe previous section is also found in Japanese. As (39) shows, when the QNP is scrambled the WCO effect disappears, and the sentence becomes good.

(39)
$dare_j$-o [s[NP[s e_i hitome [pro_j mita] $hito_i$]-ga [VP t_j sukininatta no]]
who-acc once saw person-nom fell in love with

(from Hoji (1985)

The schematic S-structure of (39) is given in (40):

(40) [QNP$_j$[[...pro$_j$...] ..t$_j$]]

As we can see, the S-structure of (39) seems equivalent to the LF structure of (36). But while (36) is excluded as a WCO violation, (39), surprisingly, is a good sentence.[21] The disappearance of WCO effects after scrambling was first noted in S&H(1983). Both Saito and Hoji independently propose to analyze scrambling in Japanese as an instance of Move α leading to an adjunction. In their views, scrambling is an instance of A'- movement(in the LGB sense) and is thus expected to have the properties of Vbl-chains, e.g., to induce WCO effects, and not to suppress them. Thus, under Saito and Hoji's analysis of scrambling, the disappearance of WCO effects in examples like (39) is paradoxical.

Hoji(1985), who notes the paradox, proposes an elegant solution to this problem; he suggests that the gap in (39) be analyzed not as an

---

21. Note that it would be insufficient simply to assume that WCO applies only at LF, and that at LF, the scrambled constituent obligatorily reconstructs into its base position. Hoji (1985) has argued that scrambling creates quantifier ambiguities which do not exist if a quantified NP remains in its base position. This would be unexpected if scrambling were systematically and obligatorily undone at LF,that is, if the LF-structure of a scrambled NP were equivalent to the LF-structure of an NP which has remained in its base position at every level of representation.

empty pronominal but as a parasitic gap. Consider the structure of (39):

(41) QNP_i $[_S[_{NP}\ldots\ pro_i\ldots]\ [_{VP}\ldots t_i\ldots]]$

In (41) the QNP c-commands both the pronoun and its trace . But neither the pronoun nor the trace c-commands the other. As noted by Hoji, this structure shows a great similarity with that of a classical cases of parasitic gap sentences:

(42) Which book did you file without reading ?

(43) $[WH_i\ [_{IP}\ ..[_{VP}\ [_{VP}\ldots t_i\ldots]\ [_{PP}\ldots e_i\ldots]]]]$

If (39) is analyzed as involving a parasitic gap structure, the WCO effect is no longer expected. In this view, the scrambled quantifier does not bind a variable and a coreferring pronoun. It simply binds a real gap and a parasitic gap. This analysis of the disappearance of WCO is very elegant and fully in keeping with the analysis of scrambling as an instance of movement creating Vbl-chains: the licensing of parasitic gaps is indeed one of the most saliant characteristics of Vbl-chains. In the next paragraph, however, we will see that a number of phenomena cast some doubt on Hoji's analysis of the disappearance of WCO in terms of parasitic gaps. This will lead us to suggest that the effects of scrambling on WCO may be better analyzed in parallel with the Hindi cases, namely as a consequence of the hypothesis that scrambling creates AN-chain, and involves movement to the specifier of a functional projection.

3.2.1.3 Arguments against a PG analysis of WCO repair

The first argument which casts doubt on Hoji's analysis comes from the observation that contexts of WCO repair are in fact not fully parallel to contexts of parasitic gap licensing. Chomsky (1986)b. proposes analyzing parasitic gap constructions as involving empty operator chains which "compose" with overt operator chains.

(44) [Operator$_j$ ...[OP$_j$...e$_j$...]...t$_j$...]

Chain composition is defined as in (45):

(45)
If C = (α_1...,α_n) is the chain of the real gap, and
C'= (β_1...,β_m) is the chain of the parasitic gap, then
the "composed chain" (C,C') = (α_1...,α_n,β_1...β_m)
is the chain associated with the parasitic gap construction and yields interpretation. (=(130) in *Barriers*)

In the *Barriers* framework, chain composition is subject to the subjacency constraint stated in (46):

(46) Subjacency condition on Chain composition

The operator of the parasitic gap must be 0-subjacent to the head of the A-chain of the real gap. (Chomsky (1986) p 67 (154))

This condition accounts for contrasts such as the one in (47):

(47)
a.? which article should you study thoroughly before I persuade the author to publish.

b.* Which article should you study t_i thoroughly before I call the author without reviewing e_i ? (Longobardi (1983))

The schematic structure of these sentences is given below:

(48)

a.

[WH [IP you [VP [VP study t_i] [PP$_1$Op$_i$ [IP I [VP [VP persuade the author [CP ..PRO to [VP publish e_i]]]]]]]

b.

[WH [IP I [VP [VP study t_i] [PP$_1$before [IP I [VP [VP call the author] [PPwithout Op$_i$ ..[VP reviewing e_i]]]]]]]

The crucial difference between (48)a. and (48)b. is that in the former structure, the parasitic gap is within a sentential complement of the verb contained in the adjunct. We may then assume that in this case, the empty operator will be able to raise out of the lower infinitival complement to a position to which it will be subjacent to the real gap t_i. In (48)b., on the other hand, the parasitic gap is contained in two islands: it is contained in an adjunct, which is itself embedded in the adjunct clause dependent on the clause containing the real gap. Thus the empty operator is separated from the real gap by (at least) two adjunct barriers and the subjacency condition on chain composition is not met.

With this in mind, let us return to the Japanese cases of WCO repair. If, as proposed by Hoji, sentences involving a scrambled operator can license a parasitic gap construction, we expect these constructions to

be sensitive to the subjacency condition on chain composition;[22] but as noted by Tada (1987) and Yoshimura (1989), scrambling repairs WCO violations even in cases where the subjacency condition on chain composition is violated, thus disallowing a possible analysis of the gap (ie the empty pronominal) as a parasitic gap. Consider the following examples:

(49)
a. Dono ronbu$_j$-o [$_{NP}$ [e$_i$ kinoo [$_{PP}$ e$_j$ yoma-zuni] neta] hito$_i$]-ga t$_j$ hihansita-no
Which article$_j$ did the person who slept without reading (it$_j$) yesterday criticize t$_j$? (Yoshimura (1989))

b. Dono ronbun-o$_j$ [$_{NP}$ e$_h$ [$_{NP}$ [e$_i$ LI-ni e$_j$ okutta] gakusei$_i$]-ni attakoto-ga aru] sensei$_h$]-ga t$_j$ hihansita
(lit) Every article$_j$, the teacher who had met the student who sent (it$_j$ to LI criticized t$_j$. (Yoshimura (1989)

c. Dono non$_j$-o John-ga [$_{PP}$ [$_{NP}$ e$_i$ e$_j$ kaita] hito$_i$]-ga sinu mae-ni] t$_j$ yonda no
Which book$_j$ did John read t$_j$ before the man who wrote (it$_j$ died? (Tada, 1987)

In (49) a,b,and c, the empty pronominal, although embedded within two islands (an adjunct and a relative clause (a.& c.) or two relative clauses (b.)), can be understood as bound by the quantifier. Given the subjacency condition on chain composition, the gap e$_j$ in all these sentences could not be reanalyzed as a well-formed parasitic gap. The grammaticality of these examples suggests, contrary to Hoji(1985), that these structures do not involve parasitic gaps but simply base-generated gaps, i.e., pronominal empty objects.

---

22. The fact that subjacency is relevant to Japanese is shown by Saito (1986). Saito shows that long distance scrambling out of islands violates subjacency.Cf Also Fukui (1986) NLLT who shows that *naze* ("why") seems to obey subjacency at LF.

The argument constructed above is rather theory-internal. It is conceivable that a different approach to parasitic gaps could handle these facts. There is, however, stronger evidence against Hoji's (1985) analysis. As first observed by Yoshimura (1989), there is a certain kind of overt pronoun in Japanese[23] which can be interpreted as a bound variable and which induces weak crossover violations. Consider the following paradigm:

(50)
a.[Dare$_j$-ga [$_{VP}$ [$_{NP}$ [$_S$ hitome soitu$_j$ t$_i$ mita] hito$_i$]-o sukininatta] no
who-nom once saw person-acc fell in love with e$_i$
Who$_j$ fell in love with the person he$_j$ took a glance at.

b.*[[e$_i$ hitome soitu$_j$-o mita] hito$_i$]-ga dare$_j$-o sukininatta no
The person who took a glance at him$_j$ fell in love with who$_j$?

c.Dare$_j$-o [[e$_i$ hitome soitu$_j$-o mita] hito$_i$]-ga t$_j$ sukininatta no
Who$_j$ (is it that) the person who took a glance at him$_j$ fell in love with ?

This paradigm is exactly parallel to the one given above with empty pronouns. (50)a. shows that when the pronoun *soitu* is c-commanded by a WH-element in subject position, it can be interpreted as bound by it. (50)b. is a case of WCO violation. In (50)c., the WH-element has been scrambled and the sentence shows no WCO violation.

As proposed by Safir (1985), parasitic gap constructions are subject to a parallelism constraint:

(51) Variables bound by the same operator must be [a lexical]

---

23. This pronominal expression is in fact a kind of derogatory demonstrative epithet. It is considered to have pronominal status because it appears to be subject to condition B of BT, not to condition C. See Hoji & Tada Hiroaki(forthcoming) for details on this epithet. Many thanks to Tada Hiroaki and to Tajima for discussion about this topic, judgments and help in claryfying the issues.

This parallelism constraint requires that elements bound simultaneously by the same operator be either both empty or both lexically realized. Applied to sentences such as (50)c., this constraint will exclude a possible analysis of their structure as a parasitic gap constructions. Since WCO repair occurs in this case as well, Hoji's (1985) analysis cannot account for these cases. There are other cases parallel to the examples given in (50). As claimed by Saito & Hoji (1983) and Hoji (1985), WCO effects can also be observed with the anaphor *zibun* which allows a bound variable interpretation. (52) shows a configuration in which WCO occurs.

(52) ?* [$_S$ [$_{NP}$ Hanako-ga zibun$_i$-o kiratteiru kotto]-ga
Hanako-nom self-acc disliked fact-nom
[$_{VP}$ daremo$_i$-o/dare$_i$-o yuutu-ni siteiru]](no)
everyone/ who depressed made qu
The fact that Hanako dislikes him has depressed everyone/who

Scrambling of the QNP in these cases also ameliorates the WCO violation.

(53)?[daremo$_i$-o/dare$_i$-o[$_S$ [$_{NP}$ Hanako-ga zibun$_i$-o
everyone/ who Hanako-nom self-acc
kiratteiru kotto]-ga [$_{VP}$ t$_i$ yuutu-ni siteiru]](no)
disliked fact-nom depressed made[24]

Both (50) and (53) are constructions comparable to those we described in the above section for Hindi, since they involve overt elements. They both violate the parallelism constraint on parasitic gap constructions so that Hoji's proposal for the cases of empty pronouns

---

24. The sentence remains marginal since for independent and unclear reasons, scrambling is quite awkward in these contexts. Abstracting away from this awkwardness, (53) manifests no added marginality due to the binding of the anaphor by the quantified expression.

does not extend to them. This suggests that another analysis of the WCO repair facts is needed. An analysis of Japanese scrambling which would parallel the analysis proposed above for Hindi provides an easy solution to the paradox of WCO repair; if Japanese scrambling is analyzed as an instance of movement to the specifier of a functional projection,i.e a [+HR] position, it will involve an AN-chain, and the disappearance of the WCO effect will be expected.

3.2.2 Scrambling as an instance of an AN-chain

In this subsection, we will present two indirect arguments in support of the hypothesis that Japanese scrambling can be analyzed along the lines of the proposal made for Hindi scrambling. Under this view, Japanese scrambling will create AN-chains.

3.2.2.1 Vbl-chain vs AN-chain and reconstruction

The first argument we develop is based on Hoji's (1985) observation that Japanese scrambling induces what has been termed in the literature "connectivity" or "reconstruction" effects. Consider the following example taken from Hoji(1985):

(54)
(??)[$_{NP}$[$_{S}$ pro$_j$ hitome e$_i$ mita] hito$_j$]-o [$_{S}$
once saw person-acc
dare$_i$-ga [$_{VP}$ t$_j$ suki-ni natta no]]
who-nom fell in love with

[The person who took a glance at (him$_i$)]$_k$ who$_i$ fell in love with t$_k$

The schematized structure of (54) is as follows:

(55) [$_{NP}$...e$_i$...]$_j$ [$_{S}$ QNP$_i$ [$_{VP}$...t$_j$...]]

In (54), an NP containing an empty pronominal has been scrambled. As a result, the QNP does not c-command the empty pronoun. However, as indicated by the grammaticality of (54), the pronoun can still have a bound interpretation. Hoji concludes that the example in (54) involves "reconstruction", which he analyzes in terms of chain binding (Barss 1986). The empty pronoun can be bound by the quantifier because the foot of the chain created by the scrambled NP is c-commanded by the QNP.[25]

A similar conclusion can be drawn on the basis of examples containing anaphors. Consider (56):

(56)
[$_{S}$ [$_{NP}$ Zibun$_i$-no hahaoya]$_j$-o [$_{S}$ John$_i$-ga [$_{VP}$ t$_j$ aisiteiru]]] (koto)
self-mother John loves
John loves his mother

In (56), an NP containing the anaphor *zibun* has been scrambled. As a consequence of the scrambling, the anaphor *zibun* is not c-commanded at S-structure by its antecedent. The grammaticality of the sentence

---

25. We remain neutral as to whether reconstruction involves LF-lowering or an S-structure computation of the Binding possibilities. For expositional simplicity, we assume that reconstruction involves lowering, but nothing in the following argument forces this view.

shows that here too, reconstruction must have applied so as to allow the binding of the anaphor by its antecedent.

Classic cases of reconstruction in English involve anaphors contained in WH-phrases which are bound by antecedents which do not c-command them at S-structure. An example is given in (57):

(57) [Which picture of himself$_i$]$_j$ does John$_i$ like t_j

It has been frequently assumed in the literature that reconstruction effects are associated strictly with Vbl-chains. Thus the fact that pronouns and anaphors can be bound by a non c-commanding antecedent when contained in a scrambled phrase, such as in sentences of the type in (54), is interpreted by Hoji(1985) as a confirmation that Japanese scrambling has the properties of Vbl-chains. However, Belletti and Rizzi (1986) have argued that reconstruction effects can also be found with AN-chains. We will briefly review their arguments before using their result for our present discussion. The possibility of anaphoric binding in psych-verb constructions has been analyzed by Belletti and Rizzi as an instance of A-movement reconstruction. Consider the example in (58):

(58) This picture of himself$_i$ amuses John$_i$

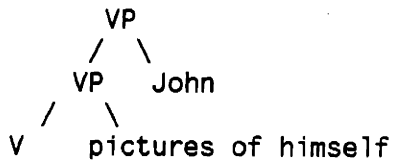
This example is problematic for the binding theory developed in LGB because it involves an anaphor which is bound by a non-c-commanding antecedent. Belletti and Rizzi (1987) show that surface subjects of psych-verbs exhibit many properties of derived subjects.[26] To account

---

26. Cf. also K.Johnson's (1985) MIT dissertation for more arguments.

for these properties, they propose (59) as the D-structure of psych verbs:

(59)



At S-structure, the direct object of psych verbs raises to the Spec of IP, leading to the surface order (58). In the D-structure (59), the anaphor *himself* is c-commanded by its antecedent *John*. The assumption that (58) involves LF reconstruction to the D-structure position of the anaphor makes it possible to interpret the binding of the anaphor as non-exceptional.[27]

If this analysis is correct, it shows that reconstruction effects are not limited to Vbl-chains. Returning to the cases of scrambling in (54), we can conclude that although reconstruction effects argue in favor of the existence of a chain, they are not conclusive as to the nature of the chain, that is, whether it is an AN or a Vbl-chain.[28]

---

27. Alternatively, Belletti & Rizzi propose that BT may apply at D-structure as well as to other levels. But see Barss (1986) for arguments supporting a reconstruction analysis.

28. An additional argument that reconstruction can occur with AN-chains is given by the following facts:

(i) These pictures of himself$_{i/j}$ seem to Bill$_i$ to have impressed John$_j$ as a poor likeness.

For a number of speakers consulted (Pesetsky, Lasnik and Higginbotham), the anaphor in (i) and (ii) can be bound by either *John* or *Bill*. If this is correct, it shows that reconstruction of anaphoric binding with AN-chains has properties similar to WH-reconstruction in that it permits any binder on its path to serve

Although reconstruction occurs both for AN-chains and for Vbl-chains, it seems to have different properties in each of these cases with respect to weak crossover phenomena. Consider (60):

(60)

* Her$_i$ mother prefers this picture of each girl$_i$

* Sa mère$_i$ prefère cette photo de chaque fille$_i$[29]

Example (60) clearly has the flavor of a WCO violation. We could attribute the WCO violation to the raising of the QNP out of its containing NP. Whatever the analysis of WCO needed to capture this fact, consider next the example in (61):

(61)* [Which picture of each girl$_i$]$_k$ does her$_i$ mother prefer t_k

For the speakers consulted, (61) has the flavor of a weak crossover violation under the relevant interpretation. Since the QNP does not directly c-command the pronoun in (61), we could simply assume that

---

as a binder for the anaphor. In other words, (i) is parallel to (ii):

(ii) Which picture of himself$_{i/j}$ does John$_j$ think that Bill$_i$ likes

Consider the following example:

(iii) Which pictures of himself$_{i/j}$ does John$_j$ think seem to Bill$_i$ to be ugly.

In (iii), again, the anaphor *himself* can be understood as bound either by *John* or by *Bill*. (iii) involves both an AN-chain and a Vbl-chain. Given that both bindings are possible, (iii) shows that reconstruction is not limited to the variable position but also occurs into the trace of the AN-chain.

29. We have added French examples to show that the phenomena under discussion are not limited to English.

the impossibility for the pronoun to have a bound reading in this case is due to the lack of c-command and not to WCO. But examples of the type given in (62) and (63) show that this view is too simple.

(62) [[Which parent of each girl$_i$] [t_k brought her to school]]

In (62), just as in (61), the QNP does not c-command the pronoun. In this case, however, a bound reading is possible. Similarly, note that when an NP which contains a QNP c-commands a pronoun and is clearly in an A-position under the LGB definition or in a [+HR] position in our view, binding of the pronominal by the QNP is possible, even though the QNP does not directly c-command the pronoun.

(63) The wedding picture of each bride was given to her mother.

We can conclude from the grammaticality of (62) and (63) that the impossibility of binding the pronoun in (61) is not due to the lack of direct c-command of the pronoun by the QNP. In other words, it is not the fact that the QNP is embedded in another constituent which prevents the bound reading of the pronominal in (61).

As we have seen in (57), reconstruction is possible for Vbl-chains. Let us suppose that reconstruction applies to (61); a plausible structure would be as in (64) (irrelevant details ommitted):

(64) [t_k [her$_i$ mother[prefers [which picture of each girl]$_k$]]][30]

---

30. Recall that we are assuming lowering here for expository reasons. In structure (64), t_k violates the ECP, but a more sophisticated analysis of reconstruction (such as "chain binding" as in (Barss 1986)) could easily get around this problem. We are thus considering this as an irrelevant factor to the present discussion.

Subsequent raising of the quantifier from inside the reconstructed constituent by QR will lead to a standard configuration for a WCO violation, as in (65):

(65) [t_k[each girl$_i$ [her$_i$ mother prefers [which picture of t_i]$_k$]

Thus, one way of accounting for the ungrammaticality of examples of the type of (61) is to assume that reconstruction has applied to Vbl-chains: if so, the intermediate structures (64) and (65) will be constructed, and QR will lead to WCO.[31]

Turning now to AN-chains, we observe that reconstruction does not have the same property wrt weak-crossover as it does for Vbl-chains. Consider again an example of psych-verbs, this time with an NP containing a QNP:

(66)[The wedding picture of each bride$_i$]$_k$ annoys her mother t_k .

If cases of AN-chain reconstruction were parallel to cases of Vbl-chain reconstruction, we would expect (66) to induce a WCO violation contrary to fact. Suppose that reconstruction applies putting the NP containing the QNP back in its D-structure position. In this case, we

---

31. Such an account does not, however, straightforwardly predict the grammaticality of (62). Reconstruction of the WH-constituent in this case will create the structure in (i), and subsequent raising of the QNP, the structure in (ii):

(i) [t_k [which parent of each girl brought her to school]

(ii)[t_k [each girl$_i$ [which parent t_i brought her to school?]

In (ii), the variable t_i left by QR does not c-command the pronoun but, as indicated in (62), the bound reading is possible. We leave this problem open.

would expect the quantifier to raise out of the reconstructed NP to take scope over it at LF.[32]

(67) [each bride$_i$ [t_k [annoys the wedding picture t_i]$_k$] her$_i$ mother]]

Raising of the quantifier by QR, however, would create the configuration for a WCO violation: consequently, we would expect examples of the type in (66) to be as marginal as examples such as (60). This expectation is incorrect: (66) is significantly better than both (60) and (61). The grammaticality of (66) shows that although AN-chains permit reconstruction (so as to allow the binding of an anaphor as in (58)), reconstruction is not obligatory, since it does not apply in cases where its application could lead to configurations of WCO violations. There are, of course, several ways of interpreting this fact; our purpose, however, is not to give an account of this distinction but simply to show that it does distinguish AN-chains from Vbl-chains. Other more classic cases of AN-chains also show the absence of WCO effects:

(68)

a. The wedding picture of each bride$_i$ seems to her$_i$ to be beautiful

b. La photo de mariage de chaque fille$_i$ lui$_i$ semble magnifique

c. The baby picture of each man$_i$ strikes him$_i$ as ridiculous

---

32. Although we have assumed here a lowering version of reconstruction for expository reasons, the problem raised here is not particular to such an approach. In a "chain-binding" view (cf. Barss (1986)) the issue would be to ensure that the chain created by a WH-movement, although permitting anaphoric binding, not permit WCO repair, while an AN-chain does.

To summarize, it appears that reconstruction with Vbl-chains has two effects: 1) it allows an S-structure violation to be overcome, as in the case of reflexive binding, but 2) in cases such as (61) it leads to a WCO violation. Roughly speaking, it appears that Vbl-chains are subject both to the "good" effects and the "bad" effects of reconstruction. AN-chains, on the other hand, enjoy only the "good" effects of reconstruction: they permit anaphoric reconstruction but do not create structures which lead to WCO violations[33]

With this in mind, let us now return to the case of Japanese scrambling. We have seen that Japanese scrambling induces reconstruction effects, since it allows a pronominal or an anaphor to be bound by a non-c-commanding NP or QP (See example (54) above).But we also know independently that scrambling does not create WCO violations, and that in fact it repairs them. Thus, scrambling shows properties of reconstruction that are parallel to cases of AN-chain

---

33. Additional evidence of the assymetry wrt reconstruction properties of AN-chains vs Vbl-chains have been given by Chomsky (1989) in class. Chomsky notes that An-chains and Vbl-chains differs in their reconstruction possibilities wrt principle C of the Binding theory. Crucial examples are given below:

(i) *Which picture of John$_i$ does he$_i$ like
(ii) These pictures of John seem to him$_i$ to be ridiculous.
(iii) These pictures of John$_i$ annoy him$_i$

He observes that Vbl-chains "reconstruct" principle C violations while AN-chain do not. These jugements seems to us less clear in French, where to our ears, coreference with AN-chains is not very good, although the severity of the ungrammaticality is not as strong as with Vbl-chain.

(iv) *Quelle photo de Jean$_i$ est ce qu'il$_i$ aime
?? Ces photos de Jean$_i$ lui semblent être ridicules
?? Ces photos de Jean$_i$ lui$_i$ deplaisent.

We have no suggestion as to why this should be the case.

reconstruction: reconstruction with scrambling only licenses the "good" effects of reconstruction, that is, reconstruction effects that permit the satisfaction of condition A of the Binding theory. Consequently, facts about reconstruction, rather than confirming the Vbl-chain status of scrambling, reinforce the hypothesis that Japanese scrambling has properites of AN-chain, allowing "good" cases of reconstruction (anaphoric reconstruction) and not forcing "bad" cases (WCO violations). These facts thus constitute an argument that scrambling in Japanese may advantageously be analyzed in parallel with Hindi scrambling as an instance of an AN-chain involving movement to the Spec of intermediate functional projections.[34]

---

34. There is an apparent problem for the hypothesis that scrambling is an instance of AN-chain:

Consider the following example:

(i)*Mary-ga kare$_i$-ni [John$_i$-no sensei]-o shookaisita
Mary introduced to him John's teacher
(i) shows a straigh principle C violation: the pronoun c-commands John. The passive version, however, shows no sensitivity to principle C.

(ii) ? [John-no$_i$ sensei]-ga kare$_i$-ni t$_i$ shookaisaneta
John's teacher to him was introduced

But at least for some speakers, the equivalent of (i) with scrambling is worse than (ii) with passive:

(iii)?? [John-no sensei]-o Mary-ga kare-ni t shookaisita
John's teacher Mary to him introduced

This shows, if the judgment holds, that there is some difference between passive Cases of AN-chains and scrambling.(As reported by Tada Hiroaki (pc), speakers disagree on the judgment. Hajime Hoji (pc) finds this sentence acceptable and so does Mamoru Saito (pc)). This could possibly be due to Case; (ii), in the traditional view, must involve movement to a [+Case] position. This is may not be the case for (iii). Plausibly this may be what comes into play here. (iii) does not seem to be a straight case of a principle C violation, since it contrasts with (i), i.e it is slightly better than (i).

3.2.2.2 Vbl-chains vs. AN-chains: reconstruction wrt Proper Binding

A second characteristic of reconstruction in AN- versus Vbl-chains leads to a similar conclusion. It has been noted by Saito (1986) that trace binding and overt anaphor or pronominal binding behave differently wrt reconstruction effects. Adopting the view on reconstruction proposed by Barss (1984) (1986) in terms of chain binding, Saito(1986) expresses the observed difference between overt anaphors and traces as follows:

(69) a.Anaphors must be chain-bound (in a domain X)

b.Traces must be bound (Fiengo (1977) (May 1977))

In other words, while it is posssible for an anaphor to be connected to a non-c-commanding antecedent through chain-binding or reconstruction, this possibility is not open for traces. Traces must meet a stricter condition: they must be c-commanded by their antecedent. Saito (1986) defines chain binding as follows:

Note that if principle C applies at D-structure, as suggested by Lebeaux (1988), then all these sentences should be excluded. Thus this paradigm constitutes an argument against Lebeaux's view that principle C applies at D-structure.
Saito (pc) also reports that there are apparently lexical differences due to the choice of predicate. Thus, the Japanese equivalent of (iv) is quite degraded but the Japanese equivalent of (v) is good.

(iv)) John's$_i$ mother he$_i$ loves
(v) John's$_i$ mother he$_i$ killed

This reinforces the idea that principle C may not be at stake, but it also casts doubts on our suggestion that Case may be involved. Thus (iv) remains a puzzle which we leave unsolved.

(70) X chain-binds Y $=_{df}$ X and Y are coindexed, and

(a) X c-commands Y, or

(b) X c-commands a trace of Z, where Z = Y or Z contains Y

Consider the sentence in (71).

(71) Which pictures of himself$_i$ does John$_i$ like

John chain-binds *himself* since *John* c-commands the trace of the moved WH-phrase which contains *himself*. But consider the following pair:

(72)

a.??Who$_i$ do you wonder [which pictures of t_i]$_k$ John likes t_k

b.**[Which picture of t_i]$_k$ do you wonder [who$_i$ John likes t_k

In (72)a., a WH-element has been extracted from inside another WH-NP. Although the sentence is not perfect,[35] it is far better than (72)b. In (72)a., *who* c-commands its trace within the WH-phrase; in (72)b, however, although *who* chain-binds its trace t_i within the WH-phrase, that is, it c-commands the trace t_k of the element containing its trace t_i, it does not directly c-command its own trace t_i; hence the severe ungrammaticality of this sentence. Abstracting away from the particular theory of reconstruction chosen by Saito, the ungrammaticality of (72) shows that although reconstruction can apply to Vbl-chains to satisfy the binding of an overt anaphor, it cannot apply to satisfy the binding of a trace.

---

35. Saito attributes the marginality of (72)a. to a subjacency violation. Cf. Saito (1986) for details.

The question now arises whether reconstruction with AN-chains shows similar effects. Although it is not possible to construct relevant cases of AN-chains which contain a gap not c-commanded by its binder in English, reconstruction effects with AN-chains can be tested in French with the construction of *en*-cliticization. Consider the paradigm given in (73)-(75):

(73)
a.(cette thèse) Jean en a révisé le premier chapitre plusieurs fois
(this thesis) John of it revised the first chapter several times

b.(cette voiture) Les voleurs en ont cassé la vitre arrière
(this car) The robbers have of it broken the rear window

c.(ce devoir) Pierre en a corrigé avec soin le dernier exercice
(this homework) Peter of it corrected carefully the last chapter

d.(son oeuvre) Marie vient d'en relire superficiellement les premières épreuves
(her work) Mary of it just proofread superficially the first draft

(74)
AN-chains:
a. Le premier chapitre en a été révisé plusieurs fois
The first chapter of it was revised several times

b. La vitre arrière en a été cassée
The rear window of it was broken

c. Le dernier exercice en a été corrigé avec soin
The last exercice of it was correct carefully

d. Les premières épreuves en ont été relues superficiellement
the first prints of it were proofread supperficially

(75)
Vbl-chains:

a. *quel chapitre crois-tu que Jean en a revisé plusieurs fois
which chapter do you believe that John of it revised several times

b. *quelle vitre as-tu dit que les voleurs en ont cassé
which window did you say that of it were broken

c. *quel exercice crois-tu que Pierre en a corrigé avec soin
which exercice do you believe that Peter of it corrected with care

d. *quelles épreuves crois-tu que Marie en a relu superficiellement

which printing do you believe that Mary of it read superficially[36]

(73) is a regular case of *en*-cliticization. The clitic *en* has been extracted from an object NP and c-commands its trace. The schematized structure of (73) is given in (76):

(76) [$_S$ NP en$_i$ [$_{VP}$ V [$_{NP}$ N t$_i$]]]

(74) is a classic case of the phenomenon which was originally termed "EN-avant" by Nicolas Ruwet (1982). The D-structure object NP which contains the trace of the extracted clitic has been raised to the external subject position through passive, an instance of an AN-chain.[37] The schematized structure of (74) is given in (77):

---

36. Incidentally, similar properties of *en*-cliticization are found in Catalan. The following paradigm illustrates these facts:
(i) En Joan n'ha corregit el primer capitol
John of it has correct the first chapter
(ii) ? el primer capitol n'ha estat corregit
The first chapter of it has been corrected
n'ha estat corregit el primer capitol
(iii) * quin capitol creus n'ha estat corregit

(vi) ??quin capitol n'ha corregit
which chapter of it have you corrected

37. The phenomenon of EN-avant seems to pose a problem for the VP-internal subject hypothesis. Indeed, one of the particularities of EN-avant is that it is only possible with derived subjects. Consider the pair in (i):

(i) a.(cette entreprise) Le patron en a été renvoyé
(this company), the boss of it was laid off
b.*(cette entreprise) the patron en a renvoyé 100 employés
(this company) the boss of it laid off 100 employees.

This is a problem for the VP-internal hypothesis because under this view all subjects are in a sense "derived". The approach to ECP that we propose in the next chapter, however, allows us to capture the distinction correctly. It has been argued by Logombardi and Giorgi (1989) and Torrego (forthcoming) that extraction out of NP must proceed through the Spec of NP (or DP, the difference being irrelevant for our purposes). In the next

(77) [IP [NP N t_i]$_k$ en$_i$ [VP V t_k]]

The examples in (74) are perfectly grammatical, although it appears that their structure violates the requirement that a trace cannot be just chain-bound but must be bound. The structure (77) is parallel to the structure of examples such as (72)b. In this case as well, the trace of an extracted NP is not c-commanded by its antecedent. But (77) is perfectly grammatical, as opposed to the ungrammatical (72). Finally, let us consider the structure of examples such as (75):

(78) [CP [NP+WH N t_i]$_k$.... [IP NP en $_i$ [VP V t_k][38]

The essential difference between the structure in (77) and the structure in (78) is the position of the constituent containing the

---

chapter we propose a theory of ECP which requires both antecedent government and head government of a trace to be satisfied. If passage through Spec NP is obligatory, we then get an alternating right branch-left branch extraction when the NP is in "object" position or in "subject" position within the VP. From our definition of head government (see Chapter IV section 4.3), it follows that a trace in the Spec of NP when an NP is in a left branch will not be able to be properly head-governed. Thus, the trace of *en* in the Spec of an NP in subject position will not be properly head-governed by the verb or by the upper head, since the trace is included in the maximal projection of NP. In our view, the Spec of NP is a [+HR] position, a position in which, we argue, traces never delete. It follows from this hypothesis that trace in Spec NP will violate the ECP. This is not the case with the trace of an NP in object position, which can be properly head-governed by the verb. This correctly predicts that *en* extraction is only possible from a D-structure object position even in a model where all subjects are "derived".

38. That the greater distance between the moved NP in (75) and the moved NP in (74) is not relevant is shown by the grammaticality of the following example:
(i) Les premières épreuves semblent en avoir été corrigées avec soin.
The first prints seems of it to have been corrected with care
See Ruwet (1982) for further examples.

trace and consequently the type of chain created by the movement. In (77) the chain between NP_k and t_k is an AN-chain while in (78) the chain (NP_k, t_k) is a Vbl-chain. The structure in (77) creates grammatical examples, while the structure (78) creates ungrammatical examples, although in both cases the trace t_i contained within the NP is not c-commanded by its antecedent. The difference between the two structures must then be a difference wrt the type of chain involved. Descriptively speaking, we could say that (77) and (78) differ with respect to reconstruction: as correctly observed by Saito, Vbl-chain reconstruction does not make the proper binding of a trace possible, hence the ungrammaticality of (78). Surprisingly, though, AN-chain reconstruction seems to pattern differently with regard to this property; it allows the binding of a trace to be satisfied. An account of this intriguing property of reconstruction with respect to the proper binding of a trace is beyond the scope of this dissertation. What is of interest for our present discussion is that this difference in the properties of reconstruction distinguishes AN-chains from Vbl-chains.

As noted by Saito (1986), Japanese scrambling seems to manifest in certain cases the reconstruction property of AN-chains. Saito (1986) argues that scrambling can be freely undone at LF, thus permitting a non-c-commanded trace contained in a constituent moved by scrambling to be properly bound at LF after reconstruction. If this is correct, this provides again an argument that scrambling has properties of AN-chains. Consider the following examples:

(79)

John-ga dono hon-o tosyokan-kara karidasita] to]
omotteiru] ka] siritagatteiru] koto
[IP Mary-nom [CP [IP all-nom [CP John-nom which
book-acc library-from checked out] Comp] think]Q] want-to-know]

([Mary wants to know [Q [everyone thinks [that John checked out which book from the library]]]])

b.??[IP **[CPJohn-ga dono hon-o tosyokan-kara karidasita] to]**$_i$ [IP Mary-ga [CP[IPminna-ga
t_i omotteiru] ka] siritagatteiru]] koto

[That John checked out which book from the library]$_i$, Mary wants to know [Q[everyone thinks t_i]

(79).b is derived from (79).a by scrambling the most deeply embedded CP to the initial position. The scrambled CP contains the WH-phrase *which book*, which finds itself out of the c-command domain of the Spec, CP, to which it moves at LF so as to take scope, namely the position marked by the interrogative morpheme *ka*. (79)b. is marginal but not excluded, according to Saito. Moreover, it is better than clear cases in which the proper binding condition is violated (see Saito (1986) for examples).

The S-structure of (79)b. is the following:

(80)

[IP [CP...wh...]$_i$[IP...[CP[IP...t_i...] Q]...]]

At LF the in-situ WH-phrase must move to the Spec, CP containing the Q morpheme. The structure will be as follows:

(81)

[IP [CP...t_k...]$_i$ [IP...[CP[IP...t_i...] WH_k]...]]

In (81), t_k is not c-commanded by its antecedent, the WH-phrase, which has moved to the position of the Q morphememe in the COMP of the

embedded sentence. Thus, t_k is not properly bound; but, as noted by Saito and contrary to expectation, the sentence is not excluded. Saito(1986) argues that (79)b. shows that scrambling can be freely undone at LF, thus allowing the unbound trace to become properly bound. If so, the LF structure of (79)b. is not (81), but rather (77):

(77)

$[_{IP}...[_{CP}[_{IP}\ [_{CP}...t_k...]\ ]\ WH_k]...]]$

In (77), the scrambled element has been put back into place, and as a result t_k is now c-commanded by its antecedent WH_k. Whatever the correct analysis of this paradigm, the point of interest for our present discussion is to note that reconstruction with scrambling manifests the property of reconstruction which we have hitherto found to be possible only with AN-chains: it allows proper binding of a trace to be satisfied. This property of scrambling provides an argument supporting the hypothesis that srambling is a type of AN-chain which involves in our terms movement to a [+HR] position.

Saito (1986) speculates that this peculiar property of scrambling is related to the fact that unlike WH-movement or topicalization in English, the position of a scrambled element is not a "scope" position. That is, the scrambled element does not have to take scope at its S-structure landing site (although it can), unlike a moved topic or a moved WH-element, which must take scope in English at its S-structure landing site or higher but cannot have a "lowered" reading. Consider for instance the following paradigm:

(82)

a. John said that this book, Mary bought

b. this book, John said Mary bought

Assuming a contrastive reading of the topicalized NP, we can tease out some differences in the interpretation of (82)a. and (82)b. In (82)a., the topic has ambiguous scope with respect to the matrix predicate. The narrow scope reading can be understood as a correction of John's utterance. John was mistaken in the report he made of Mary's actions. A possible paraphrase would be :

(83)

John said that it was this book (and not that one) that Mary bought

The wide scope reading, on the other hand, is a correction of the speaker's utterance. That is, the speaker made a mistake in reporting John's utterance, not John in reporting Mary's actions. This reading could be paraphrased as follows:

(84) It was this book that John said that Mary bought

In (82)b., however, the topic may only have a wide scope reading. The scope of the topic is determined by its S-structure position in (82)b and cannot be "lowered" to the location of the original trace.

As discussed by Hoji (1985), following Kuroda (1970), scrambling in Japanese opens up new scope possibilities. It was noted by Kuroda (1970) that sentences such as (85) below differ from sentences such as (86) with respect to their scope possibilities:

(85)
Daremo-ga dareka-o aisiteiru
everyone someone loves
Everyone loves someone

(86)
Dareka-o daremo-ga aisiteiru
someone everyone loves
Someone, everyone loves

In (85), the universal quantifier takes obligatorily wide scope over the existential quantifier. (86), on the other hand, is ambiguous, allowing both quantifiers to take either scope. Hoji (1985) formulates the following descriptive generalization:

(87) a. When two quantifiers are in their D-structure positions at S-structure, the quantified NP that c-commands the other takes wide scope with respect to the other.

b. When a quantified NP is preposed over another quantified NP, the scope interpretation is ambiguous.

Note that although scrambling allows a quantifier to have wider scope, it does not restrict it to this possibility. Descriptively speaking, it appears that a scrambled quantifier can take scope either in it surface position or in its D-structure position. This fact is very reminiscent of the scope of quantifiers in AN-chains. It has been noted by May (1977?) that in Raising constructions, a quantifier can be interpreted either in its S-structure position or in its D-structure position. This phenomenon is known in the literature as "quantifier lowering". Consider the following classic example:

(88) A man seems to be in the garden

(88) has two possible interpretations which are roughly translatable by the following paraphrases:

(89) a. There is a man that seems to be in the garden

b. It seems that a man is in the garden

It is a characteristic of AN-chains that they permit Q-lowering. As we saw above, it is not possible for a topic be interpreted with narrow scope in (82)b. Here again, we observe that scrambling and AN-chains share this characteristic.

To summarize briefly: In this section we have shown that scrambling and AN-chains are similar with regard to their reconstruction properties. Both types of chains allow an unbound trace to be properly bound at LF through reconstruction. As has been argued by Saito(1986), this possiblity is not open for Vbl-chains. In Saito's view, which I believe to be on the right track, this property of reconstruction is connected to the scopal properties of chains. Descriptively, the scope of a WH-phrase or of a topicalized NP that has undergone syntactic movement is determined by its S-structure position (It can take scope in this position or higher but not lower). This is not true, however, for either scrambling or NP-movement: both allows Q-lowering from their S-structure position. This similarity between scrambling and AN-chains provides an additional argument in support of the hypothesis that scrambling is an instance of AN-chain involving movement to a [+HR] position.

3.2.2.3 Anaphoric-binding

This hypothesis is confirmed by the fact that binding of an anaphor from a scrambled position is marginally possible. We find here cases which parallel the Hindi facts. Consider the following examples:

(90)

a.*[zibun-no$_i$[kenkoo zyootai]]-ga zetuboo-ni John-o$_i$
[self-gen physical condition] nom despair John
oiyatta
drove

b.?John-o$_i$ [zibun-no$_i$[kenkoo zyootai]]-ga t$_i$ zetuboo-ni
John [self-gen physical condition] nom despair
oiyatta
drove
John, his physical condition drove to despair

This example is somewhat exceptional, since it involves a pysch-verb[39]. Other examples with the anaphor *zibun* are usually not very good, because *zibun* is a strictly subject-oriented anaphor for most speakers. Subject orientation is not as strict with reciprocals, however, so examples of the type given in (91) show a clear constrast.

(91)

a. * [otagai-no hahaoya]-ga [John to Mary]-o kiratteiru
eachother's mother John & Mary hates
Eachother's mother hate John & Mary

b. ? [John to Mary]-o [otagai-no hahaoya]-ga kiratteiru
John & Mary, eachother's mother hates

The creation of new binding possibilities after scrambling suggests that the landing site of this movement must be a postition from which

---

39. If this case is a psych verb construction, it only partially supports our point. This example is given as good: Zibun$_i$ ga Mari-ni karakow-rare-ta koto ga Zyon$_i$-o zetubbo-e oiyatta. See p97 Masayoshi Shibatani Japanese syntax.

anaphor binding is possible, i.e., an A-position in LGB terminology, and a [+HR] position in our view.[40]

3.2.2.4 Numeral quantifiers

Additional support for the hypothesis that scrambling in Japanese creates AN-chains and involves movement to a [+HR] position comes from the distribution of numeral quantifiers in Japanese. Numeral quantifiers can be separated from the NP they modify: consider the pair in (92):

(92)
a.[3-nin-no gakusei]-ga kita
three-cl-Gen students came
Three students came

b. Gakusei-ga 3-nin kita
Students three-cl came.
Three students came

In (92)a., the numeral *three*, which is inflected with an appropriate classifier, occurs within the subject NP and is marked for genitive case. In (92)b. the numeral is syntactically independent from the subject NP (Cf. Miyagawa (1989) for supporting evidence) and appears

---

40. The sentence in (i) is ungrammatical in Japanese:

(i) *John to Mary-o otagai-ga kiratteiru
John and Mary eachother hate

This can be accounted for, however, as a violation of the LBC of Rizzi (1986). Under the assumption that Japanese scrambling is to a [+HR] position, the trace of the scrambled element will be bound by the anaphor in subject position, leading to an LBC violation. This explanation is given by Mahajan(1989) for the same ungrammatical sentence in Hindi.

without genitive Case-marking. Pairs of this type have often been analyzed (Inoue and others quoted in Miyagawa) as derivatively related by a transformation of quantifier float. We will adopt the proposal that (92)a and (92)b are transformationally related but we will assume an analysis along the lines of Sportiche (1988). We will begin by briefly surveying the distribution of these floating numerals in Japanese.

Note first that numeral quantifiers can occur in every position in which an NP can be scrambled:

(93)
a.Taroo-ga hon-o 2-satu katta
  Taroo books two bought
  Taroo bought two books

b.Hon-o Taroo-ga 2-satu katta
  Books Taroo two bought
  Taroo bought two books

c.Taroo-ga hon-o Mary-ni 2-satu atae -ta
  Taroo books to Mary two gave
  Taroo gave two books to Mary

d.Hon-o Taroo-ga 2-satu Mary-ni atae-ta
  books Taroo two Mary gave
  Taroo gave two books to Mary

e.Hon-o Taroo-ga Mary-ni 2-satu atae-ta
  books Taroo Mary two gave
  Taroo gave Mary two books

Recall that we have argued in Chapter II section 2.3.2 that floating quantifiers are compatible only with AN-chains. If this is correct, the possibility of Q-stranding in sentences such as those in (93) provides additional support for the hypothesis that scrambling in Japanese has properties of AN-chains.

Miyagawa(1989) has argued that numeral-quantifier float in Japanese provides an argument for the existence of AN-chains in passive and in

constructions with "unaccusative" verbs (in the sense of Burzio (1986)). He argues that mutual c-command between a floating numeral and the modified NP or the trace of the modified NP is required for the floating quantifier to be licensed.[41] Given this hypothesis, the contrast between the examples in (94) and those in (95) provides interesting support for the unaccusative hypothesis; recall that according to this hypothesis, the surface subjects of unaccusative verbs are derived from a VP-internal complement position by NP movement.

(94)
a.Gakusei ga [VP ofisu ni 2-ri kita]
  Students to office two came
  Two students came to the office

b.Otoko ga [VP baa ni 2-ri haitta]
  men bar two entered
  Two men entered the bar

c.Kyaku-ga [VP ryokan-ni 3-nin tuita]
  guests inn three arrived
  Three guests arrived at the inn

(95)
a.?* Gakusei ga [VP zibun-no kane de 5-nin denwa-sita]
  student self money five telephoned
  Five students called with their own money

b.?* kodomo ga [VP geragera to 3-nin waratta]
  children in a loud manner three laughed
  Three children laughed in a loud manner

c.?* Kodomo ga [VP wa ni natte 5-nin odotta]
  kids circle become five danced
  Five children danced in a circle

In Miyagawa's view, in examples of the type shown in (94) the VP contains a trace of the surface subject. Mutual c-command can obtain

---

41. See Miyagawa (1989) for additional requirements on the licensing of floated quantifiers.

between a floated numeral inside the VP and the trace of the surface subject. This accounts for the grammaticality of (94) and suggest that numeral-float is compatible with classic cases of AN-chains. In the unergative constructions in (95), however, Miyagawa claims that there is no trace of the surface subject within the VP; since in examples such as (95) the floated numeral is separated from the subject by an adjunct, mutual c-command between the NP and the floated numeral does not obtain. Moreover, since there is no trace of the subject NP within the VP which can enter into a mutual c-command relation with the floated numeral, the floated numeral is not licensed and the sentences are excluded.

Quite clearly, Miyagawa's analysis, if correct, poses problems for the VPS hypothesis: his account of the ungrammaticality of (95) relies crucially on the assumption that the subject of an unergative verb does not originate inside the VP, so it leaves no trace there which could license the floating numeral.

One possible way of reconciling Miyagawa's analysis of (95) with the VPS would be to assume that there is some additional requirement which makes the position of the specifier of VP (the position of a D-structure unergative subject in our view) ineligible for numeral float in Japanese. Miyagawa argues that a trace with which a floating numeral can be associated must be governed by an "affecting" verb. If the government relation required to meet this additional condition is defined in terms of strict c-command, then an NP in the specifier position of the VP will not be able to meet the condition.

Consequently, a floating numeral in the Spec of an unergative or transitive VP will be ruled out.

A number of facts suggest, however, that contrary to Miyagawa's conclusion, a floating numeral may sometimes be associated with the subject of unergative or a transitive verbs. These facts cast some doubt on the precise formulation of the analysis given by Miyagawa, even though they leave untouched the central idea that the relation between a floating numeral and the modified NP or its trace must be local. Consider the following examples:

(96)
kodoma-ga 3-nin waratta.
children three laughed
Three children laughed.

(97)
Gakusei-ga kyoo 3-nin hon-o katta
Students today three book bought
Three students bought the book today

(96) and (97) show that a floated numeral can occur separated from an unergative surface subject. In (96), the relation of mutual c-command required by Miyagawa may be met directly by the NP, since the floated numeral and the NP are adjacent. This is not the case in (97), however: this shows that strict adjacency between the subject and the numeral is not required.[42] Moreover, since *today* is not a VP adverb

---

42. Miyagawa proposes to get around this problem by positing a more abstract level which does not take into account adverbs of the type found in (97). At that level the NP and the floating numeral mutually c-command each other. It seems to us, however, that if such adverbs can be disregarded with respect to the mutual c-command requirement, it is unclear why the type of adjunct in sentences like (95) should act any differently. Possibly the degraded status of sentences like (95) is essentially due to stylistic factors since, the intervening adjuncts used in these

under standard assumptions (see Jackendoff (1972)), (97) suggests that the subject is no longer in its D-structure position and may have moved above the adverb from the position in which the floating numeral occurs. Assuming, along lines suggested by Sportiche (1988), that a floating numeral is derived by NP movement, these examples indicate that there is some position from which the subject of an unergative verb has moved.[43] Further support for this hypothesis comes from the following paradigm ((98)a. and (98)b. are taken from Miyagawa(1989)).

(98)
a.Tomodati-ga 2ri Sinzyuku-de Tanaka sensei-ni atta
  Friends two Shinjuku Tanaka professor met
  Two friends met professor Tanaka in Shinjuku

b.*Tomodati-ga Sinzyuku-de Tanaka sensei-ni 2ri atta

c.(?)Tomodati-ga Sinzyuku-de 2ri Tanaka sensei-ni atta

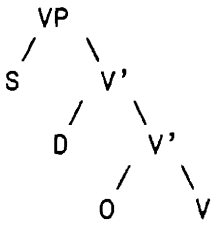
The ungrammaticality of (98)b. shows that if a quantifier related to a subject is stranded after an argument, the sentence is excluded. But as shown by (98)c., the occurrence of a floating numeral between an non-subcategorized adjunct and an argument is considerably better. This is unexpected under Miyagawa's analysis, but if, as argued by Fukui(1986)and Hoji (1986), the structure of the Japanese VP is as in (99), then the contrast between (98)b. and (98)c. is expected.

---

examples are rather "heavy". We will leave this question unresolved.

43. Other adverbs lead to rather marginal sentences when they occur between a subject and a floated numeral. We will assume that this marginality is due to independent factors which for the moment remain mysterious.

(99)



Given the structure in (99), we can assume that the occurrence of a floating numeral after the dative argument D in (98)b. is impossible because this is not a position in which the subject of an unergative verb ever occurs at D-structure. On the other hand, assuming, as is standard, that the non-subcategorized locative adjunct is adjoined to the VP, we expect a floating numeral to be able to occur after it in the D-structure position of the subject, as in (98)c. Confirmation of this hypothesis would require further investigation of the distribution of floating numerals with respect to the argument structure of various predicates, which is beyond our concern at the moment.

Summarizing, it appears that Japanese floated numerals generally occur with clear cases of NP-movement as well as with scrambling. For this fact to be of real significance for the hypothesis that scrambling forms an AN-chain, it remains to be shown that floated numerals are not compatible with Vbl-chains in Japanese. As in Hindi, Japanese long-distance scrambling seems to have properties which differ from those of short-distance scrambling. Although the judgments are somewhat tricky, it appears that long-distance scrambling does not permit WCO repair in matrix clauses. The relevant example is given in (100):

(100)
*Daremo-o/dare-o$_j$ [e$_i$ e$_j$ /soitsu-o$_j$ hitome mita]
hito$_i$]-ga] [musume-ga t$_j$ sukininaru]-to ommota no

Everyone$_j$/who$_j$ the person who took a glance at him$_j$
thought his daughter will fall in love with t$_j$

(100) is excluded under the relevant reading.[44] That is, a pronominal contained in the matrix clause cannot be bound by the long-distance scrambled quantifier or WH-phrase: long-distance scrambling thus appears to have some of the properties of Vbl-chains. If this is correct, we can use long-distance scrambling as a test case for compatibility with floating numerals. Interestingly, we observe the same phenomenon in Japanese that we previously observed in Hindi: although floating numerals can occur within an embedded sentence in all the positions to which a NP can be scrambled, they cannot occur in any position within the matrix sentence. In the example below, we use the symbol ^ to mark the positions in which a floating numeral can occur and * to mark the positions in which it cannot:

---

44. Speakers seem to disagree about this judgment: Saito (pc) judges this sentence as being quite good and Yoshimura (1989) gives it as perfect. As suggested to us by Saito and as argued for by Yoshimura, it may be that Japanese long-distance scrambling in fact forms an AN-chain. Saito (pc) suggests that the apparent difference between Hindi and Japanese may be a consequence of the fact that Japanese does not clearly distinguish tensed sentences from infinitival sentences. Assuming that all sentences in Japanese are in some sense tenseless, then Japanese long-distance scrambling can be like raising in English, a case of a long-distance AN-chain. Under the theory of ECP which we develop in Chapter IV, passage through the spec of CP is possible if IP is not a barrier; it may be that because of the lack of tense, Japanese IP is not a barrier, thus allowing long distance scrambling to form AN-chains.

(101)
Hon-o$_i$ John-ga * Mary-ni * kossori-to [^(?) Peter-ga ^
t$_i$ katta] to itta.
books John to Mary secretly Peter bought that told
These books, John told Mary secretly that Peter had bought

The distribution of floating numerals in Japanese exactly parallels the distribution of floating quantifiers in Hindi. We conclude from this fact that, as in Hindi and other languages, floating numerals in Japanese are only compatible with AN-chains. Thus, the fact that they are compatible with scrambling provides additional support for the hypothesis that Japanese scrambling, like Hindi scrambling, is an instance of movement to a [+HR] position and thus creates an AN-chain.

Conclusion:

In this section we have considered the nature of the chain created by scrambling in Japanese. Based on WCO effects, reconstruction effects, binding effects, and the distribution of floating numerals, we have argued that the chain created by scrambling has the properties of an AN-chain. Given the definition of AN-chains we proposed in Chapter II, this entails that scrambling can be an instance of movement to a [+HR] position.[45] This raises the question of what these [+HR] positions

---

45. Although we have given arguments that scrambling can be movement to a [+HR] position, we have given no arguments showing that it cannot also involve adjunction, i.e., movement to a [-HR] position. Suppose that scrambling can involve either movement to a [-HR] position or movement to a [+HR] position: we would expect that if no violation results from adjoining, movement to either a [+HR] position or to a [-HR] position should be possible. If some violation would result, the syntax dictates that movement to the position which leads to the lesser violation is the only one possible. Let us clarify this view : assume that the scrambled constituent contains a quantifier. Then movement to a [-HR] position would induce WCO effects. In this case, then, adjunction

might be. Mahajan (1989) has given independent evidence based on agreement phenomena for the existence of multiple functional nodes in Hindi; these provide several possible [+HR] landing sites for scrambling. It is not immediately obvious however, that these same nodes exist in Japanese: although Japanese verbs have tense marker, it is a well-known fact that they show no overt marking for person, number, or gender. However, the absence of agreement does not necessarily entail the absence of functional projections in Japanese;[46] although it has been pointed out by various linguists that Japanese lacks most of the properties which indicate the existence of a functional category INFL, such as subject-verb agreement, structural nominative Case, and the effects of the nominative island condition, many linguists assume its presence (in particular Miyagawa (1989), Kuroda (1986), Hoji (1985), and Fukui (1986)).

---

is excluded, but movement to a [+HR] position remains a possibility. Thus the latter possibility is chosen. If, contrary to what we have argued above, anaphoric reconstruction is strictly a property of Vbl-chains, then when reconstruction is needed, adjunction is permitted. Possibly, when nothing is at stake, the two types of scrambling are distinguished by intonation. It has been argued by Webelhuth that a scrambled NP in German is necessarily unfocused (see section 3.4). This may not be correct in Japanese, and it may be the case that focused scrambling is adjunction with properities similar to English topicalization, but unfocused scrambling is a movement to a [+HR] position. This would explain the fluctuation of speaker judgments on some of the facts we have mentioned, since it might be quite difficult to isolate the right intonational pattern.

46. But cf. Fukui (1986) for the proposal that Japanese has (almost) no functional projections. Fukui claim is that Japanese has neither a D nor a C projection. He maintains, however, that Japanese has a (rather defective) INFL and suggests in a footnote that possibly the Case-markers (-ga ..-o..-ni) head their own functional projections.

3.2.2.5 Functional projections in Japanese

As noted by Fukui (1986), positing the existence of an INFL node has several descriptive advantages. It allows a simple explanation of why the tense morpheme always occurs at the end of a verbal complex: as shown in (102), tense morphemes can never intervene between a causative morpheme or a passive morpheme and the predicate root. They must always occur at the end of the verbal complex.

(102)
a.tabe - sase - rare -ta
  eat cause pass past
  Was caused to eat

b.*tabe -ta -sase -rare
c.*tabe -sase -ta -rare

Positing an INFL node also permits a straightforward account of the fact that the tense morphemes can freely appear outside of the *soo su-* construction, the Japanese equivalent of the English *do so* construction. Consider the following examples from Fukui (1986):

(103)
a.John-wa kinoo tosyokan-de benkyoo-si-ta
  John studied in the library yesterday

b.Bill-wa ototoi soo si-ta
  Bill the day before did so
  Bill did so the day before

c.Bill-wa asita soo-su-ru-daroo
Bill tomorrow do so Pres seems
Bill seems to do so tommorrow

d.Bill-mo asita soo si-tagatte-i-ru
Bill also tomorrow do so want to Pres
Bill wants to do so tommorrow.

As can be observed in (103), the tense morphemes (underlined) occur outside the constituent replaced by the form *soo-su*. As proposed by Fukui, if it is assumed that the tense morphemes occupy the INFL position ouside the V projection, these data can be accounted for simply by saying that *soo su-* replaces a projection of V, excluding the element in INFL. (See Fukui(1986), p. 211). Additional arguments have been given by Tateishi (1989) for the existence of an INFL projection.[47] Asssuming that Japanese has an INFL node, whether this INFL node is further subdivided, as proposed by Pollock (1989) and Chomsky (1989) for French and English, remains an open question. As suggested by Pesetsky (pc), it may be possible to consider the Japanese honorific markers of the "respectful language"(*sonkei-go*) as heading functional projections, possibly comparable to the AGR projections we have assumed to exist in French and in English.[48] There are two "subject-oriented" honorific markers in Japanese: *rare*, which is homophonous with the passive morpheme but is distinct from it in

---

47. See Tateishi (1989), *Proceedings* of NELS 19, edited by J. Carter & R-M Déchaine.

48. See Kitagawa for an analysis of one variety of these honorific affixes. Kitagawa proposes to analyze honorific markers as lexical heads which can be affixed to a verbal root through "affix Raising". In Kitagawa's view, affixes are attached to a verbal head in the lexicon and undergo affix raising at LF. This analysis can be easily recast in terms of head movement within the framework of derivational morphology elaborated by M.Baker (1986).

that it cooccurs with an accusative object, and a discontinuous morpheme which surrounds a verb: *o-* V *ni-nar*. The two forms differ essentially with respect to the context of their use, the former being slightly less formal than the latter. With respect to the empirical argument given above for the existence of an INFL projection, the honorific markers behave exactly the same. For instance, as shown in (104) and (105), honorific markers are strictly ordered with respect to other verbal morphemes.

(104) Honorific marker *o-V-ni-nar*

With passive

a. Tanaka sensei-ga seito-ni o-nagu-rare-ni-nat-ta
   Tanaka professor student by H-hit- PASS-HON -PAST
   Professor Tanaka was hit by a student

b. *... o-nagu-ni-nar-rare-ta
   H- V - HON -PASS-PAST

Schematically, the only acceptable order is:

o-Verb-Passive-Honorific-tense

with negation

a. Tanaka sensei-ga seito-o o-naguri-ni-nar-ana-kat-ta
   Prof Tanaka student H-hit -HON NEG PAST
   Professor Tanaka did not hit a student

b. *... o-naguri na- ni-nat-ta
   H- V - NEG-HON- PAST

c. *... o-naguri na-kat-ta -ni nar
   H- V NEG-PAST -HON

(104B) Honorific Marker *rare*

With passive

a. Tanaka sensei-ga seito-ni nagu-rare-rare-ta
   hit -PASS-HON- PAST
   *hit -HON -PASS-PAST
   Professor Tanaka was hit by a student

With negation
a. Tanaka sensei-ga seito-o nagu-rare-na-kat-ta
hit -HON -NEG -PAST
Professor Tanaka did not hit a student

b. *... nagur-ana-rare-ta
hit- NEG-HON -PAST
*... nagur-ana-kat-ta -rare
hit- NEG -PAST -HON

Moreover, as shown in (105) and (106), the honorific markers share with the tense markers the possibility of occurring outside the constituent replaced by the pro-form *soo-su*.

(105)
John-wa kinoo tosyokan-de benkyoo-si-ta
John studied in the library yesterday

o-V-ni-nar

Tanaka sensei wa ototoi soo o-shi-ni-nat-ta
Prof Tanaka the day before so do HON PAST

Tanaka sensei wa asita soo-o-shi-ni-nar-ru-daroo

(106) *rare*
Tanaka sensei wa ototoi soo-s-are-ta

Tanaka sensei wa soo-s-are-ru daroo

The fact that the honorifc markers and the tense marker behave alike with respect to these two properties suggests that honorific markers may head their own separate functional projections. We will not pursue this hypothesis in any detail. A careful study of Japanese morphology would be needed to substantiate this proposal: see Kubo (1989) (forthcoming MIT generals paper) for such an analysis. If this is on the right track, we may assume that the existence of functional projections in Japanese will receive some independent evidence. Japanese scrambling can then be analyzed in parrallel with Hindi

scrambling as involving movement to the specifier of various functional projections.

If the analysis we have proposed for Japanese scrambling is tenable, it provides support for the hypothesis that the functional structure of sentences may be complex even in languages whose morphology shows no overt making for agreement.[49] In other words, it provides support for the idea that there is no direct correspondence between overt morphological marking and the presence or absence of functional projections. As we have shown, Japanese scrambling shows many similarities to Hindi scrambling, although the two languages differ radically with respect to the richness of their overt morphological marking. Given the striking similarities between Japanese and Hindi scrambling, the hypothesis that Japanese has a highly articulated clause structure with a number of functional projections (possibly invisble) appears in any case to be the null hypothesis.

3.3 Object shift

In the previous sections we considered instances of scrambling in in Hindi and Japanese and argued that they instantiate movement to the specifiers of functional projections. In this section, we turn to languages with a more rigid word order which allow restricted movement of objects under certain circumstances. This phenomenon, first

---

49. Incidentally, this is assumed generally for English. Chomsky (1989) suggests that English has an AGR-O projection, even though object agreement is never manifested in English.

described and analyzed in Holmberg (1984) and (1986), has come to be known in the literature as object shift. The term is used to describe the movement of objects to some intermediate position in the sentence, the status of which remains yet to be determined. Object shift is relevant to our present discussion since, as argued by Holmberg (1986), it is an instance of syntactic movement which creates AN-chains. This phenomenon is common to all Scandinavian languages, although it obeys different constraints in mainland Scandinavian and in Icelandic. We will first briefly review the properties of Object shift in mainland Scandinavian and then turn to the properties of object shift in Icelandic. We will then propose an analysis of object shift as involving movement to the specifier of functional projections.

3.3.1 Object shift in Mainland Scandinavian

Consider the following sentences from Swedish and Norwegian:

(107) Swedish
a. Vafor laste studenterna inte alla den
   Why read the students not all them
   Why didn't all the students read them

b. Vafor laste studenterna den inte alla

(108) Norwegian
a.(?)vi leste ikke dem
   we read not them
   We didn't read them

b. vi leste dem ikke[50]

In (107)b. and (108)b., an object pronoun has moved to the left of the negation. In the mainland Scandinavian (MS) languages Swedish, Danish, and Norwegian, the leftward movement of an object is stricty restricted to unstressed pronouns. As shown in (109) for Swedish, a full NP cannot undergo this type of movement:

(109)
a. Vafor laste studenterna inte alla artikeln
   Why read the students not all the article
   Why didn't all the students read the article

b.*Vafor laste studenterna artikeln inte alla

This is not the case in Icelandic, where the movement of a full NP is perfectly acceptable. We turn to Icelandic object shift in the next subsection.

As is well known, Swedish and other MS languages are typical verb-second languages. Holmberg argues extensively that V-second in these languages involves the movement of the verb to C in main clauses. As observed by Homlberg, object shift (OS) is in some sense "parasitic" on the movement of the main verb to C. Thus, it is ruled out either when an auxiliary verb moves to C and the main verb remains in its D-structure position or in embedded sentences where no verb moves to C. Examples showing this restriction are given below for Swedish:

(110)
a. Vafor har studenterna inte alla last <u>den</u>
   Why have the students not all read it

---

50. Many thanks to Arild Hestvik for the Norwegian data. The question mark in (108) indicates that the pronoun must bear stress to be acceptable in post-negative position

b.*Vafor har studenterna den inte alla last
Why have the students it not all read

c. att studenterna inte alla laste den
that the students not all read it
d.*att studenterna den inte alla laste den
that the students it not all read (from Holmberg (1986))

As argued by Holmberg (1986), OS is an instance of syntactic movement, and the chain created by it has the properties of an AN-chain. Independent evidence that OS is syntactic comes from the fact that it "feeds" the syntactic rule of topicalization. Consider the example in (111), again from Swedish:

(111)
Tillgivna var dom honom alla
devoted were they him all
Devoted to him, they all were (from Holmberg (1986))

Transitive adjectives in Swedish have a particular property which allows their NP complement to occur on either side of the adjectival head. Since Swedish is a left-headed language, Holmberg argues that the preadjectival position of the complement of AP is derived by movement. Thus, in Holmberg's view, (111) has the following structure:

(112) $[_{AP}\ e_i\ [\text{tillgivna } e_i]]_j\ [\text{var dom honom}_i\ \text{alla } e_j]$

In (112), the pronoun *honom* has first been moved to the preadjectival position, then extracted out of the AP constituent by object shift. Finally, the AP constituent has been topicalized to the front of the sentence. (113) retraces this derivation. Since topicalization is standardly assumed to be a syntactic rule, the fact that OS occurs

prior to it provides an argument that OS must also be a syntactic rule(See Holmberg pp. 168 -169 for details of the argumentation)[51]

(113)

a. NP preposing in AP
[Hundarna var alla [$_{AP}$ honom$_i$ [tillgivna e$_i$]]]
The dogs were all to him devoted

b. Object shift
[Hundarna var honom$_i$ alla [$_{AP}$ t$_i$ [tillgivna e$_i$]]]

c. Topicalization
[$_{AP}$ t$_i$ [Tillgivna e$_i$]]$_k$[hundarna var honom$_i$ alla t$_k$]

The hypothesis that the position of the pronoun in (113)b. is due to OS is confirmed by the impossibility of a similar construction with a full NP. As shown in (114), a full NP cannot occur to the left of a floated quantifier related to the subject.

---

51. Incidentally, Holmberg argues for the syntactic status of topicalization on the basis of reconstruction. Observing that adjectival topicalization can leave behind it an object-shifted pronoun, he concludes that reconstruction must apply to satisfy Binding. The argument seems to us rather weak, since if adjectival topicalization were a PF rule, we would expect apparent violations of the Binding Theory to have no effect on the grammaticality of the sentence. Recall, moreover, that we showed above that A'-movement does not allow reconstruction to satisfy trace binding. If our reasoning is correct, then Holmberg's argument can in fact be turned against itself to show precisely that either adjectival topicalization or OS in Swedish may be a PF rule, since the trace of the shifted object does not need to be properly bound. Although this particular argument is rather weak, it seems to us that the conclusion reached by Holmberg is correct. It is perhaps better shown by the interaction of OS with Quantifier float or with ECM constructions.

(114)
a.*Hundarna var Erick alla tillgivna.
the dogs were to Erik all devoted

b. Hundarna var honom alla tillgivna.

Holmberg shows convincingly that the properties of the chain created by OS are properties which are typical of AN-chains. Object shift has the following properties:

1) it is clause-bound

2) it does not license parasitic gaps

3) it is insensitive to crossover violations

The clause-boundedness of OS is shown by the following pair of sentences:

(115)
a. Eva anser honom mojligen vara inkompetent
Eva considers him possibly to be incompetent

b. Eva anser han mojligen ar inkompetent
Eva considers him possibly (that) is incompetent
Eva considers that possibly he is incompetent (H 1986)

(115)a. is ambiguous: the sentential adverb can be interpreted either as part of the matrix clause or as part of the embedded infinitival clause. As argued by Holmberg, this follows from the fact that OS, like NP movement, is possible across an infinitival in ECM contexts. (115)b., however, is unambiguous: the sentential adverb can only have scope within the embedded tensed sentence. This shows that when a embedded sentential complement is tensed, a pronoun cannot move to the higher clause.

The fact that OS does not license parasitic gaps is shown by (116):

(116)
a. vilken artikel kastade dom t_i innan du hade last e?
   Which article threw they before you had read
   Which article did they throw away before you had read?

b.*Artikeln kastades t_i innan jag hade last e
   The article was thrown away before I had read

c.*Jag kastade den inte t_i innan jag hade last e
   I threw away it not before I had read (H 1986)

As can be seen in (116), OS in (116)c. patterns like NP-movement in (116)b., and not like WH-movement in (116)a., in that it does not license a parasitic gap.[52]

The insensitivity of OS to crossover violations is shown in (117):

---

52. It might be argued that in (116)c. the object pronoun has not moved high enough to c-command the gap and thus cannot license the parasitic gap for independent reasons. But some adjuncts can attach to a position higher than the VP, as shown by their position in embedded sentences (in section 3.3.4.1 we argue that the verb moves to T in embedded sentences):

(i)
att Olof [sa nart han var fardig], kastade artikel i papperskorgen.
that O. as soon as he was finished threw the article in the wastebasket

(ii) *Olof kastade den, sa nart han hade lat e, t i..
O. threw it, as soon as he had read e, t i .. (H 1986)

(ii) shows the main clause verb-second word order, where the verb has moved into C and the subject has been fronted. The adjunct in (ii) is adjoined to a projection of TP and the pronoun has moved over it. Thus, it is quite clear that in (ii) the pronoun c-commands the gap. The sentence, however, remains impossible, showing that it is indeed the property of the chain created by OS which rules out the parasitic gap and not simply a lack of c-command.

(117)
a. ?Vem$_i$ tilldelade dom i hans$_i$ franvaro t$_i$ priset
who awarded they in his absence the prize

b. Dom tilldelade honom$_i$ i hans$_i$ franvaro priset
They awarded him in his absence the prize (H 1986)

In (117), the adjunct constituent [i hans franvaro] is not inside the VP, but is adjoined to some projection of INFL. This can be deduced from the position of this adjunct in embedded sentences. As shown by the paradigm in (118), this adjunct can only precede the verb in embedded sentence, never follow it:

(118)
a. om dom i min franvaro tilldelade homom priset
if they in my absence awwarded him the prize

b.*om dom tilldelade i min franvaro homom priset

c.*om dom tilldelade homom i min franvaro priset (H 1986)

Since embedded sentences represent the base order of constituents, Holmberg concludes that the position of the adjunct in (117)b. results from two distinct instances of movement over the adverb: the movement of the verb to C and the movement of the object pronoun under object shift. If so, the structure of (117)b. is is as in (119):

(119) [$_{CP}$ Dom [$_C$ tilldelade [honom$_i$ [i hans$_i$ franvaro] v t$_i$ priset]]]

(119) is a typical S-structure WCO configuration: after movement, *honom* c-commands both its trace t$_i$ and a pronominalwithin the adjunct phrase with which it is coindexed; but neither the bound pronoun nor the trace c-commands the other. The fact that no WCO violation occurs thus provides additional evidence that the chain (*honom*, t$_i$) has properties of an AN-chain.

To the evidence provided by Holmberg, we add the following facts about floating quantifiers[53] As (120) shows, there is a sharp contrast between a construction in which a quantifier has been floated after OS and a construction in which a quantifier has been floated from an object after WH-extraction:

(120)
a.Jeg leste dem ikke alla
  I read them not all

b.*Boeken att jeg ikke leste alla
   the books that I not read all

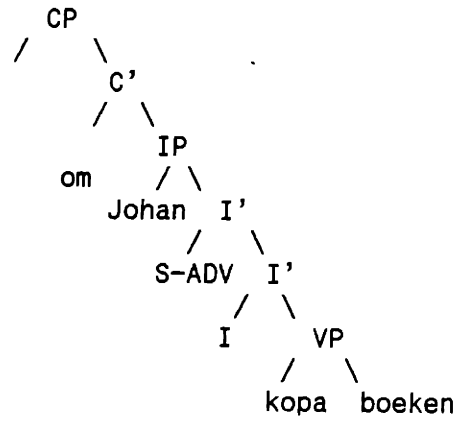
Here again, we observe that floating quantifiers are impossible with clear cases of Vbl-chains, an observation which reinforces the hypothesis that the chain formed by OS must be an AN-chain.

The various arguments presented above classisfy the chain created by OS as an AN-chain. In Holmberg's view, however, object shift in Swedish and other MS languages involves the movement of the pronoun to a position adjoined to I'. The base structure assumed by Holmberg for Swedish is the following:

---

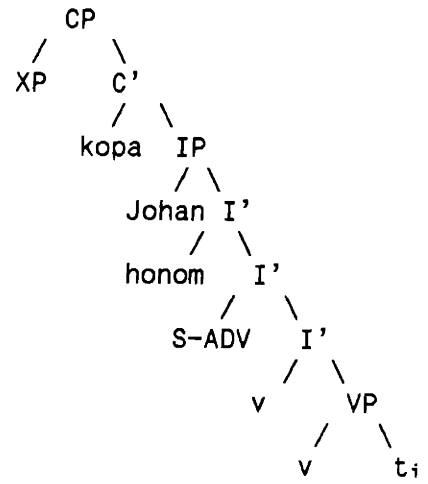
53. The data we use are from Norwegian rather than Swedish since we only had access to Norwegian informants.

(121)



In Holmberg's view, OS moves a object pronoun from its base position and adjoins it to I' above the sentential adverb. As mentioned above, OS is parasitic on the movement of the Verb to C. Thus a typical structure for object shift is the following:

(122)



Note, however, that under the LGB definition of A- vs. A'-movement, object shift as analyzed by Holmberg is an instance of A'-movement, since it is an adjunction. Similarly, in our view, adjunction to I' is a movement to a [-HR] position and is thus expected to have the properties of a Vbl-chain, contrary to fact. Thus, if Holmberg's

proposal that the landing site of OS is an adjoined position is correct, the properties of the chain created by OS would remain unexplained both in the standard LGB model and in our system. In other words, Holmberg's proposal conflicts with both the LGB definition and our definition of AN-chains. In Holmberg's view, however, there is no conflict between the nature of the landing site of OS and the properties manifested by the chain created by OS. Following Taraldsen (1986b),[54] Holmberg adopts a view in which it is not the landing site of a movement, but the nature of the element which heads a chain which determines the nature and the properties of that chain. Holmberg proposes the following definitions:

(123)

a.A trace is a variable iff it is locally operator bound

b.A trace is an anaphor iff it is locally non-operator bound

Since the trace of mainland Scandinavian object shift is locally bound by an unstressed pronoun,i.e., an element which has no operator status, it follows from the definition in (123) that the trace is an anaphor and that the chain created by OS will have the properties of an AN-chain.

Holmberg's proposed analysis of object shift raises an issue which up to now we have not discussed. We have assumed, following Chomsky (1981), that the nature and properties of a given chain are determined by the nature and properties of the <u>position</u> of the head of the chain.

---

54. See also Kayne (1984)," On Complex Inversion".

A possible alternative view, and the one which Holmberg has adopted, is to assume that the nature and properties of a given chain are determined by the nature and properties of the element heading the chain, disregarding the position in which it occurs. For expositional clarity, let us call the hypothesis which views the properties of the landing site of the head of a chain as the determining factor the Landing Site Determination Hypothesis, or LSDH, and the hypothesis which views properties of the head of the chain as the determining factor the Head Determination Hypothesis, or HDH.

As we will show in the next subsection, the HDH adopted by Holmberg to account for the properties of OS in MS runs into serious problems once we extend our range of data to Icelandic OS. We have seen that MS object shift is restricted to unstressed pronominals; it is thus not possible to examine the nature of the chain created by OS when the chain is headed by an operator. As we will see in the next subsection, however, such a verification is possible in Icelandic, since OS in this language applies to a broader set of constituents. We will show that Icelandic object shift provides strong support for the more standard determination of chains, namely the LSDH, over the HDH.

3.3.2 Object shift in Icelandic

Icelandic object shift differs from MS object shift in two main respects:

First, unlike mainland Scandinavian OS, Icelandic OS occurs both in main and in embedded clauses. We will assume, following Holmberg, that

this difference follows from the properties of the movement of the verb in Icelandic[55]. As is well known, Icelandic differs from mainland Scandinavian in exhibiting the verb-second phenomenon both in main and in embedded clauses. As has been argued by Holmberg, there is a direct relation between verb movement and OS. Thus, the fact that OS is possible in embedded clauses in Icelandic can be seen as a consequence of the verb-second property of Icelandic embedded clauses.

Second, and more important for our present concerns, Icelandic OS is not limited to pronouns but also occurs with full NPs. Examples of Icelandic OS are given below[56]:

(124)
Main clause:
a. Jon keypti ekki baekunar
   Jon bought not the books

b. Jon keypti baekunar ekki

Embedded clause:
c. pað var gatt að hann keypti ekki baekunar
   it was good that he bought not the books

d. pað var gatt að hann keypti baekunar ekki

(125)
a. Sigga setur aldrei hlutina a rettan stað
   Sigga puts never the things in the right place

b. Sigga setur hlutina aldrei a rettan stað

Holmberg (1986) assumes that Icelandic OS, like mainland Scandinavian OS, has properties of AN-chains. A number of facts support this

---

55. See the following section for an analysis of the relation between OS and verb movement

56. We thanks Gudrun Gunarsdottir and for the data in this section

assumption. First, as shown in (126), Icelandic OS, like MS OS, is compatible with floated quantifiers.

(126)
Eg las baekunar alltaf/ekki allar
I read books always/not all

As shown in (127), however, floating quantifiers are incompatible with clear cases of Vbl-chains such as relative clauses.

(127)
*baekunar sem Jon keypti ekki allar
the books that John bought all

This provides a first argument that Icelandic OS has the properties of an AN-chain. Second, Icelandic OS does not induce WCO violations. To be able to show this we need to make a brief digression to describe the ECM construction in Icelandic. Consider the following sentence:

(128)
Jon$_i$ telur mig [i karnaskap sinum$_i$] hafa etidh hakarlim
John believes me in foolishness his to have eaten the shark
In his foolishness, John believes me to have eaten the shark

(128), due to Thrainsson (1979), is a somewhat peculiar and very interesting ECM construction. The peculiarity of this ECM construction resides in the position of the adjunct phrase [i karnaskap sinum]. This adjunct phrase contains an anaphor which can refer to the subject of the matrix clause, even though it appears to occur within the boundaries of the embedded ECM complement (since it occurs after the pronoun *mig*, the subject of the embedded ECM complement). As shown in (129), if the complement is tensed, the anaphor contained in the adjunct can no longer refer to the subject of the matrix sentence.

(129)
*Jon$_i$ telur [$_{CP}$ adh eg [i barnaskap sinum$_i$] hafa etidh hakarlim
John believes that I in his foolishness have eaten the shark

Examples such as the one in (128) have been used by Thrainsson to argue for the existence of a transformation of Raising to Object. In his view, the object pronoun moves to the object position of the matrix verb. In the LGB model, however, the existence of a transformation of Raising to Object conflicts with the projection principle and the theta-criterion, which preclude any movement to a θ-marked position. Preserving the spirit of Thrainsson's analysis, Holmberg (1986) proposes to reanalyze sentences such as (128) as instances of object shift. In Holmberg's view, the structure of (128) is the following:

(130)

[$_S$ Jon telur [$_{VP}$ mig$_j$ [$_{VP}$ [i barnaskap sinum]
[$_{VP}$ tv [$_S$ t$_j$ hafa etidh harlakim]]]]]

The pronoun *mig* has moved by OS from its base position to adjoin to the matrix VP. Thus, the fact that the anaphor inside the adjunct phrase can refer to the matrix subject is no longer surprising, since it is part of the matrix sentence. Holmberg provides many arguments in support of this hypothesis, which we will adopt without any further discussion. Let us return now to the question of WCO. Consider the following example:

(131)
Maria$_j$ telur Jon$_i$ i adhaun sinni$_j$ a modhur hans$_i$ hafa etidh harlakim.
Maria believes Jon in her admiration for his mother to have eaten the shark
In her$_j$ admiration for his$_i$ mother, Maria$_j$ believes John to have eaten the shark

Since in (131) the adjunct phrase contains an anaphor bound to the matrix subject *Maria*, we conclude that the position of the NP *Jon* has been derived by OS. If so, the structure of (131) is as in (132):

(132)

[Maria$_j$ telur$_v$ [Jon$_i$ [[i adhaun sinni$_j$ a modhur hans$_i$] t$_v$ [t$_i$ hafa...

(132) is an S-structure configuration of WCO, since the NP *Jon* binds both the pronoun *hans* and its own trace t$_i$ in the subject position of the ECM complement. But the grammaticality of (132) under the desired reading shows that the movement of the object *Jon* does not induce WCO. This confirms the hypothesis that OS creates an AN-chain.

A third argument in support of the hypothesis that object shift creates AN-chains is the fact that it does not license parasitic gaps. As shown in (134), Icelandic allows regular parasitic gap constructions:

(134)
a.(?) Hvadha bok las tu an thess ath kaupa?
which book did you read without buying?

b. Hvadha granmeti bortar thu an thess ath sjotha?
Which vegetable do you eat without cooking

As shown in (133), however, parasitic gaps are not possible with OS.

(133)
a.*Eg las baekurnar alltaf/ekki an thess adh kaupa
I read books always/not without buying

b.*Eg bortadhi petta granmeti ekki/aldrei an thess adh sjodha
I eat this vegetable not /never without cooking

One could argue that in (133) the parasitic gap is ruled out not because of the nature of the chain, but because the shifted object does not raise high enough to c-command the adjunct clause and the parasitic gap. As shown in (134), however, an object-shifted pronoun induces a condition C violation with a lexical object contained in the complement clause. This shows that c-command must obtain between the raised object and the adjunct clause.

(134)
*Eg hitti hani ekki/aldrei an thess adh hitta systur Petersi
I met him not/never before meeting Peter's sister.
I did not meet him before I met Peter's sister

As we have seen, Icelandic OS can move a full NP constituent. It is standardly assumed that when they are not moved to the initial position of the sentence, WH-phrases can generally occur in positions available to NP. If so, we expect a WH-phrase in situ to be able to undergo object shift. That this is indeed possible is shown in (135):

(135)
a.Hver las ekki hvadha baekur
who read not which book

b.Hver las hvadha baekur ekki

c. Hver keypti ekki/aldrei hvadha bok
who bought not which book

d. Hver keypti hvadha bok ekki/aldrei

e.Hver helt adh eg laesi ekki/aldrei hvadha baekur
Who thinks that I read not/never which book

f.Hver helt adh eg laesi hvadha baekur ekki/aldrei

g.Hver helt adh eg laesi ekki/aldrei hvern
Who thinks that I read who not/never

h.Hver helt adh eg laesi hvern ekki/aldrei

This fact is particularly interesting since it provides us with a possible test of the validity of the Head Determination Hypothesis of the nature of chains. Since the HDH takes into account the nature of the head of the chain to determine the nature of a given chain, there should be a crucial difference between the cases of OS given in (124) and (125), which involve an NP, and those given in (135), which involve a WH-phrase. The HDH predicts that a chain headed by a non-operator should have the properties of an AN-chain and that a chain headed by an operator should have the properties of a Vbl-chain. Thus, OS with a WH-phrase should create a Vbl-chain and not an AN-chain. As a consequence, we expect OS with a WH-phrase to manifest the typical properties of Vbl-chains, such as the licensing of parasitic gap constructions and WCO effects. As shown in (136) and (137), however, this prediction is not supported by the facts. Parasitic gap constructions with shifted WH-phrases have the same status as the parasitic gap constructions with shifted NPs as in (133) above: Both are completely ungrammatical.

(136)
a. *Hver helt adh eg laesi hvadha baekurnar alltaf/ekki an thess adh kaupa
Who thinks that I read which books always /not before buying

b.* Hver etur hvadha granmeti ekki/aldrei an thess adh sjodha
Who eats which vegetable not/never without cooking

Moreover, WCO effects are not induced by an object-shifted WH-phrase, as would be expected under the HDH. Consider the example in (137):

(137)
Hver telur hvern [i adhdaun sinni a modhur hans] hafa etid hakarlim
Who$_i$ believes who$_j$ in his$_i$ admiration for his$_j$ mother to have eaten the shark.

(137) has basically the same S-structure representation as (131), except that in (137) the element which has been moved by OS out of the subject of the ECM construction is a WH-element and not an NP. We observe, however, that in this case, just as in (131), the binding of the second pronominal by the shifted element is possible: no WCO effects are induced.

The fact that the chain created by an object-shifted operator does not manifest properties of Vbl-chains raises a serious problem for the HDH and consequently for Holmberg's proposed analysis of object shift. Recall that under Holmberg's proposal, the landing site of OS is an adjoined position. Consequently, under the LSDH we expect the chain created by OS to have properties of Vbl-chains. As we have seen above, this prediction is incorrect. The HDH correctly predicts that a chain created under the object shift of a non-operator, that is, a pronoun or a full NP, has properties of an AN-chain; but it incorrectly predicts that the chain created under the object shift of an operator should have the properties of a Vbl-chain. This suggests that an analysis of OS which views the landing site of this movement as an adjoined position is not adequate under either hypothesis. In following section, we will suggest a modification of Holmberg's analysis which will capture the properties of OS without running into the problems raised by the adoption of the HDH. We propose that the landing site of OS is not an adjoined position, but the specifier of a

functional projection. Recall that in our view the specifier of a functional category is a [+HR] position. We have argued further that the chain created by movement to a [+HR] position has the properties of an AN-chain. Under such a view, the LSDH can be maintained without undesirable consequences, and the absence of any distinction between an OS chain headed by an operator and an OS chain headed by a non-operator is expected. Before we turn to such an analysis, we briefly discuss some additional evidence against the HDH.

3.3.3 Evidence against the Head Determination Hypothesis

In the previous subsection we saw that the properties of OS with a WH-phrase in Icelandic provide an argument against the Head Determination Hypothesis (HDH) of the nature of chains. Evidence against the HDH can also be found in English. Consider the following example:

(138) Who thinks that who seems to be happy.

The S-structure of (138) is given in (139):

(139) [$_{CP}$ who [$_{IP}$ thinks [$_{CP}$ that [$_{IP}$ who$_i$ seems [t_i to be happy]]]]]

In (139), the second WH-phrase is in the subject position of *seems* after raising. In a theory in which it is the nature of the head of a chain which determines the nature of the chain, the chain (who$_i$, t_i) is expected to have the properties of a Vbl-chain. Consequently, we expect such a chain to induce crossover effects and to license

parastic gaps. As we show, however, neither of these predictions is fulfilled.

Consider first the example in (140):

(140) who thinks that who$_i$ seems to his$_i$ mother to be happy.

A partial S-structure for (140) is given in (141):

(141)... [$_{CP}$ that [$_{IP}$ who$_i$ seems to his$_i$ mother [t$_i$ to be happy]]]

(141) is an S-structure configuration for WCO: the WH-phrase binds both its trace and a coindexed pronoun, neither of which c-commnds the other. But as shown by the grammaticality of (140), WCO effects are not induced. Consequently, we can conclude that despite the fact that the chain (who$_i$, t$_i$) is headed by an operator, the chain has the properties of an AN-chain. The absence of WCO effects in (140) is correctly predicted by a theory which takes the nature of the position of the chain to be the relevant factor, i.e., the LSDH, but not by the HDH, as represented by Holmberg's definitions given in (123) above.

One could argue that the grammaticality of (140) is due not to a failure of the HDH but to the fact that WCO effects only obtain at LF. The schematic LF structure of (140) is the following:

(142) [who$_i$ [who$_j$ [t$_j$ thinks [(that) [t'$_i$ seems to his mother [t$_i$ to be happy]]]]]]

At LF, t$_i$ is locally bound by t'$_i$, which is not an operator. Thus under a literal interpretation of the HDH, t$_i$ is not a variable, since it is not locally bound by an operator, and WCO effects are not

expected. Note, however, that if the non-operator status of the intermediate trace is taken to be the relevant factor which explains the absence of WCO in (140), and this quite independently of any consideration of the position in which it occurs, we expect any intermediate trace to have similar effects and thus to suppress WCO. This cannot be correct: consider for instance the following example of long-distance extraction:

(143) *Who$_i$ do you think that his$_i$ mother loves

(143)) shows WCO effects, even though it is generally assumed that the Spec of the CP of the embedded sentence contains an intermediate trace left by the movement of *who* accross that CP.

(144) [$_{CP}$ who$_i$[you think [$_{CP}$ t'$_i$ that [$_{IP}$ his$_i$ mother loves t$_i$]]]]

If so, t$_i$ in (144) is locally bound by t'$_i$. Since t'$_i$ is not an operator, t$_i$ will not be defined as a variable and we do not expect WCO effects to obtain in the embedded clause, contrary to fact. This shows that even if WCO does not obtain until LF, the HDH cannot account properly for the facts. The contrast between (140) and (143) shows that reference to the nature of the landing site of a movement is needed to determine the properties of a chain.[57] That the chain (who$_i$, t$_i$) in structures such as (138) and (140) is an AN-chain, as

---

57. Pushing things a little further, one might argue that in (143) the intermediate trace gets deleted. Note, however, that one would have to find a reason for why the trace in the subject position of *seems* in (140) does not also get deleted. Here again, it seems quite difficult to avoid reference to the nature of the position in which the trace occurs in making the relevant distinction.

predicted by the LSDH but not by the HDH, is confirmed by the fact that it cannot license a parasitic gap. Consider the example in (145):

(145)* Who thinks that who was killed before Mary could meet

The S-structure of the embedded clause in (145) is given in (146):

(146) ...[$_{CP}$ that [$_{IP}$ who$_i$ was [$_{VP}$ killed t_i] [before Mary could meet e_i]]]

In (146), *who*$_i$ c-commands both its trace within the VP and the gap within the adjunct. Thus, this structure is comparable to that of good cases of parasitic gap licensing such as (147), except for the position of the WH-element.

(147) ? I wonder who he killed before Mary could meet

Under the HDH, this difference does not matter; the chain (who$_i$, t_i),since it is headed by an operator, is a Vbl-chain in both cases. We would thus expect a parasitic gap to be licensed in structure such as (146), contrary to fact. The ungrammaticality of (145) thus provides further support for the LSDH over the HDH.

Strong crosslinguistic evidence for the LSDH also comes from the cases of scrambling we have studied in sections 3.1 and 3.2. Recall that in both Hindi and Japanese, the scrambling of a WH-operator has the properties of an AN-chain, since it does not induce WCO effects. Again, this is unexpected under the HDH but expected under the LSDH if, as we propose, scrambling involves movement to [+HR] positions and movement to [+HR] postions always creates AN-chains.

As we have argued in this subsection, there are a number of reasons to choose the LSDH over the HDH. In light of these arguments, it appears that the phenomenon of object shift should not be analyzed as movement which creates an adjunction: an alternative analysis is required. We turn to this task in the next subsection.

3.3.4 Object shift as movement to an [+HR] position

3.3.4.1 Verb movement in Scandinavian

In this section, we outline a proposal for an analysis of object shift in Mainland Scandinavian (MS) and Icelandic (IC) as involving a movement of the object to the specifier of an intermediate functional projection. Recall that in our view, the specifier of a functional projection is a [+HR] position. Since movement to a [+HR] position creates an AN-chain, the fact that object shift has AN-chain properties will follow straightforwardly under such an analysis. As we mentioned earlier, Object Shift appears to be "parasitic" on the movement of main verbs both in MS and in IC. It is clear that any analysis of object shift in these languages must acccunt for this restriction. Thus, as a first step to our account of object shift, we must turn to an analysis of the properties of verb movement in MS and IC respectively. We will proceed as follows: first we summarize the main aspects of the link between verb movement and object shift through a review of Holmberg's analysis. Then we turn to the

exposition of our proposal. Our account will cover both the main properties of verb movement in MS and IC. and the restrictions on object shift. As we will show, the restriction on object shift will follow in part from our analysis of verb movement.

The following paradigm summarizes the main differences in word order between MS and Icelandic:

(148) Swedish:

Main clause

a. Johan kopte inte den
   Johan bought not it

b. Johan kopte den inte

c. *Johan inte kopte den

Embedded clause

d. *att Johan kopte inte den
   that Johan bought not it

e. *att Johan kopte den inte

f. att Johan inte kopte den

(149) Icelandic:

Main clause

a. Jon keypti ekki hann
   Jon bought not it

b. Jon keypti hann ekki

c. *Jon ekki keypti hann

Embedded clause

d.að Jon keypti ekki hann
  that Jon bought not it

e.að Jon keypti hann ekki

f.*að Jon ekki keypti hann

As is apparent from the paradigm in (148) and (149), the possible word orders of MS and IC are exactly parallel in main clauses, and exactly reversed in embedded clauses. In Holmberg's view, these similarities and differences betwen the word order of MS and IC follow from the set of hypotheses given in (150):

(150)

1) In main clauses, the verb raises obligatorily to C in MS but not in IC.

2) The verb moves obligatorily to I in both languages. VP is pruned after V to I in MS but not in IC.

3) The base position of sentence adverbs differs in MS and in IC:

In MS, S-adverbs are attached to I':
I'---> S-adv I'

In IC, S-adverbs are attached to VP:
VP ---> S-adv VP

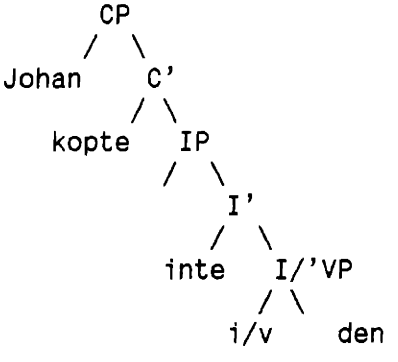
4) In both languages, OS is only possible when the object is governed by a verbal trace and not by the verb itself.

In Holmberg's view this latter assumption derives from the Case marking property of a verb vs the Case marking properties of a verb trace. Holbmberg assumes that a verb must obligatorily assign Case but that a verb trace can assign Case optionally. He assumes additionally that the head on an A-chain must satisfy the Case filter but that the

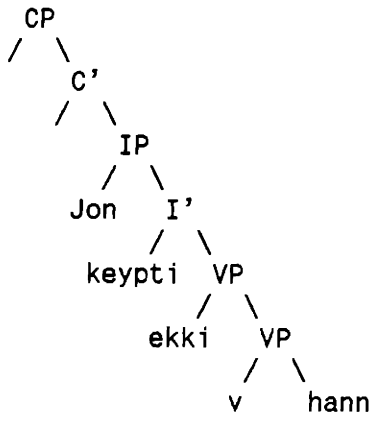
foot of an A-chain must not be Case marked. Given these assumptions, if OS occurs when the object is governed by a Case assigning verb, the resulting chain will be excluded because Case is assigned to the foot of the chain. On the other hand, if OS occurs when the object is governed by a verb trace, the foot of the chain will not be marked for Case. Holmberg moreover proposes that the Case requirement on the head of an A-chain can be satisfied by overt morphological Case. We return to a discussion of this latter assumption at the end of this section.

For the moment, let us see how this set of hypothesis captures the facts in (148) and (149). First, it follows from (150) 1) that main clause word order, although superficially identical, reflects a different structure in MS and in IC. The structures of MS and IC main clauses are given in (151) and (152):

(151) **MS:**



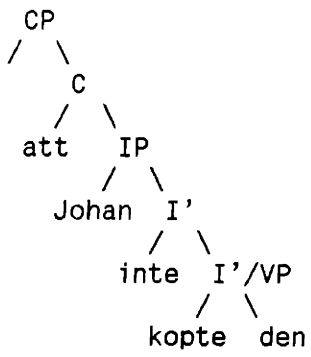
(152) **IC:**



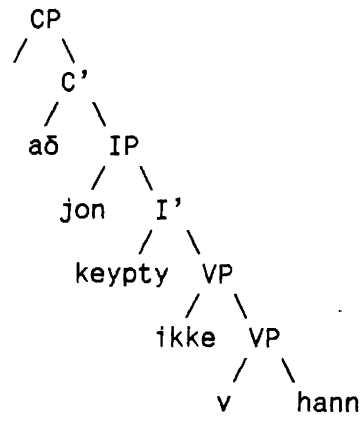
In (151), VP pruning has taken place after the movement of the verb to I. But since the main verb has further moved to C, the object is governed by a verbal trace; consequently object shift is possible. In (152), the verb has raised to I, but VP pruning has not occurred. Thus, the object is also governed by a verbal trace, and object shift is possible. This accounts straightforwardly for examples a. and b. in (148) and (149). The c. examples are accounted for because the movement of the verb to C or to I is assumed to be obligatory in both cases.

Let us now turn to embedded clauses. The structure of embedded clauses in MS and IC are given in (153) and (154):

(153) **MS:**



(154) IC:



Movement of the verb to C is impossible because C is occupied by an overt complementizer. Movement of the V to I has occurred in both languages, but VP pruning has taken place in MS and not in IC. Consequently, the object is governed by a verbal trace only in IC, so object shift is possible only in IC.

Although the set of hypotheses made by Holmberg accounts correctly for the paradigm in (148) and (149), it is conceptually problematic in two respects. First, it is unclear why there should be a difference in the placement of sentential adverbs in MS and IC. Second, it is unclear why the movement of the verb to I should lead to VP pruning in MS but not in IC. Note that each of these hypotheses is crucial to account respectively for 1) the fact that in MS embedded clauses, sentential adverbs must precede the verbs (Cf. (148)d.) and 2) the fact that object shift is impossible in MS embedded clauses (Cf. (148)e. and f.).

As mentioned by Holmberg, there is no empirical evidence supporting movement of the V to I in MS. Thus, it is never possible to find the word order V Adverb NP, even if the adverb in question is a VP

adverb[58] Given this state of affairs, one may wonder what the motivation is for the claim that V raises to I in MS.

The motivation appears in fact to be cross-linguistic. It involves a comparison of the possibilties of verb movement in English and in MS. As assumed by a number of linguists (Emonds, Pollock, Chomksy, among many others), verb movement distinguishes between auxiliary verbs and main verbs in English: auxiliary verbs can move to I, and to C, but main verbs never do. Roughly speaking, this assumption is supported essentially by two facts:

1) negation and sentential adverb can follow auxiliaries, but never main verbs.

2) auxiliaries, but not main verbs, can occur in C in interrogative sentences in English.

Combining 1) and 2), we obtain the following generalization about verb movement in English:

(155)

a.If a V moves to I then it can move to C

b.If a V does not move to I then it cannot move to C

---

58. An intriguing fact is that VP adverbs can apparently not occur (at least in Norwegian (Arild Hestvik (pc)) and Danish (Maria Bitner (pc)) between the AUX and the past participle in embedded sentences; nor can they occur before the Aux; they must always follow the object. Assuming as is natural that the past participle is in the D-structure position of the verb, this fact suggests that VP adverbs can never be left adjoined to the VP. If so the absence of the sequence V Adv NP might just reflect this fact and not the absence of movement of the V to I.

The question now arises whether this generalization holds cross-linguistically, that is, whether it is a universal property of verb movement (presumably due to the Head-to-Head constraint) or whether it is simply a specific property of verb movement in English.
We will digress somewhat from the main concern of this section to argue that the generalization in (155) should indeed be taken as having the status of a universal.

3.3.4.2 Universal property of verb movement to C

As is well known, Romance languages allow main verbs to move both to I and to C, movement to C occuring for instance in direct questions. We can thus conclude that generalization (155)a. clearly holds for Romance languages as well. What about generalization (155)b.? As argued by Pollock (1989), there is at least one case in which a main verb cannot move to I in Romance: this is the case of main verbs in French infinitival sentences. Evidence for this comes from the fact that french infinitivals cannot precede the negation:

(156)a. Ne pas finir sa thèse à temps est très dangereux
To not finish your thesis on time is very dangerous

b.*Ne finir pas sa thèse à temps est très dangereux

To check the validity of (155)b. for French, we must find an infinitival context in which the movement of the verb to C should in principle be possible. Non-embedded interrogatives are generally assumed to be contexts in which V can (or must) move to C. As shown in

(157), French allows infinitival interrogatives to occur quite freely as non-embedded interrogatives. Consider the following examples:

(157)
a. Qui voir ?
   Whom to see?

b. Que faire?
   What to do?

c. Où aller?
   Where to go?
d. Comment resister à la tentation?[59]
   How to resist temptation?

Since the interrogative sentences in (157) are not embedded, we might expect them to allow the movement of some verbal element to C. If so there are in fact two possibilities. Either an empty I moves to C, or else the main verb itself moves to C. If the generalization in (155)b. were only a specific property of English verb movement, we may expect the movement of main verbs to C to be possible in these French infinitival questions. Recall that, as assumed by Pollock, negation occurs lower than I/T in French. If a main verb could move to C in infinitival interrogatives, bypassing the I/T projection, we would expect the following pattern to be possible: a main verb should be able to precede the negation in infinitival interrogatives even though

---

59. That these sentences are not just embedded sentences whose embedding verb has been contextually deleted is shown by the fact that these interrogatives can have the value of yes/no questions: *Etre ou ne pas être? Dormir? Partir? Rester?* ("To be or not to be? To sleep? To leave? To stay?"). These are always impossible in embedded contexts: **Je me demande partir / *Je me demande si partir* ("I wonder (whether) to go"). Note that in English, an infinitival question can never have the form of an embedded yes/no question: *What to do? Where to go? *Whether to leave or to stay?* This suggests that English infinitival questions also have the properties of matrix questions.

it cannot precede it in regular infinitival complements. The schematic structure would be as in (158):

(158) [CP quii [C' V [TP..T..[NEGP..pas [VP. tv..ti]]]]]

This prediction, however, is not fulfilled. As shown in (159), main verbs can never precede the negation in infinitival interrogatives:

(159)
a. Qui ne pas croire ?

b. *Qui ne croire pas

c. *Qui croire ne pas
   Whom not to believe?

d. Où ne pas aller?

e. *Où n' aller pas

f. *Où aller ne pas
   Where not to go?

As shown by Pollock (1989), auxiliary verbs, unlike main verbs, can move to I/T in infinitivals. As predicted, auxiliaries can also precede negation in infinitival interrogatives:

(160)
a. Où inviter sa fiancée, où ne pas inviter sa belle-mère?

b. *Où inviter sa fiancée, où n'inviter pas sa belle mère?
   Where to invite one's fiancee, where not to invite one's mother-in-law?

c. Où être invité, où ne pas être invité?

d. Où être invité, où n'être pas invité?
   Where to be invited, where not to be invited?

The paradigm given in (159) and (160) shows that generalization (155)b. also holds in French: main verbs do not move to I/T in infinitivals, and therefore they cannot move to C, either. These facts

support the hypothesis that the generalization (155) is valid accross languages and is not simply a language-particular restriction.

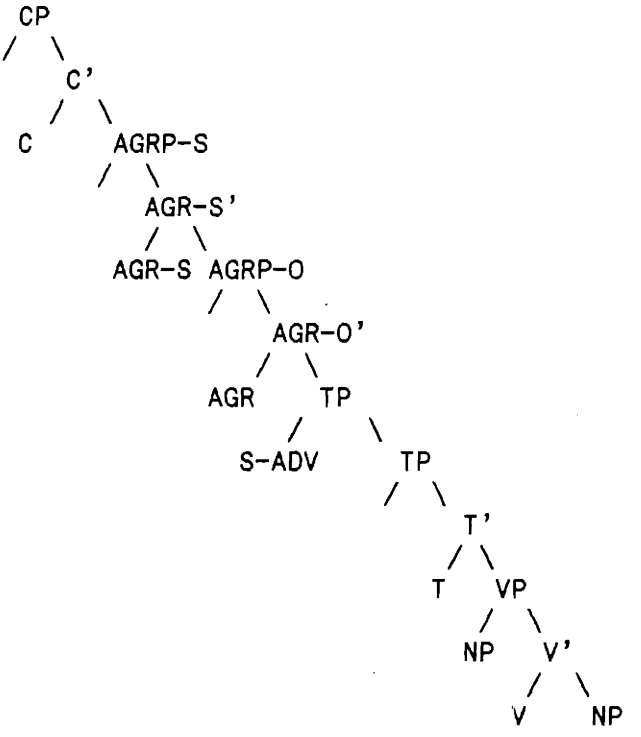
3.3.4.3 Clausal structure in Scandinavian

With this in mind, let us now return to MS. Assuming generalization (155) to be universal, we conclude that main verbs must move to I in MS, since there is little doubt that they can move to C in main clauses and in interrogative sentences such as (161).

(161) Vafor kopte Johan den?
Why bought Johan it?

The challenge is to allow for this possibility and to account as well for the distinctions between IC and MS regarding the position of the sentential adverbs and the possibility of object shift. To meet this challenge, we propose to make use of Pollock and Chomsky's assumption that the Inflectional node I is sudivided into several functional projections. We propose the following base structure for both Icelandic and MS:

(162)



We assume, in accordance with Holmberg and Platzack (1989), that negation in Icelandic and MS is simply a sentential adverb and does not head its own projection.[60] Nothing in what follows hinges on this particular assumption, however.Given the structure in (162)[61] a single

---

60. Pollock has proposed that negation must c-command tense at S-structure. Laka (1988) proposes to extend this view and assumes that the S-structure c-command requirement is a universal. Note that if we were to assume that MS has a NEGP we would have to assume that this NEGP occurs above TP to account for the position of the negation in Swedish embedded sentences. If so the S-structure c-command requirement is not met. Suppose that the S-structure relation of the negation and T can be expressed in terms of m-command or non-exclusion (in the sense of Chomsky (1986)) from a category. In the representation we have given, since NEG is adjoined to TP it is not excluded by it and thus the requirement is satisfied.

61. The structure we propose for IC and MS differ from the structure proposed by Chomsky (1989) for English and French with respect to the ordering of the functional projections. Note first that the assumption that functional projections may have a different

well-motivated assumption is sufficient to distinguish Verb movement in Icelandic from Verb movement in Mainland Scandinavian. We propose to make crucial use of the distinction between MS and IC with regard to the overt manifestation of subject-verb agreement.[62] As is well known, Mainland Scandinavian languages show no overt subject-verb agreement. Thus we will assume that in these languages the movement of

---

hierachical order across languages is consistent with Chomsky's (1989) suggestion that parametric variations concern essentially functional heads; variation in the order of functional projections can be understood in terms of variation in the selectional properties of functional heads. Moreover, the assumption that AGR-O is ordered before T in languages such as Swedish and Icelandic (and may be across Germanic) but not in French allows a unified account of the fact that sentential adverbs can occur after the direct object in Icelandic and German but not in French or in English (cf Travis 1987 and the next section on German). Let us assume that cross-linguistically sentential adverbs are adjoined to the projection of TP. If AGR-O dominates TP in Germanic, then we expect an object to be able to precede sentential adverbs when they move to AGR-O. On the contrary, in French even if an object moves to AGR-O, we expect it to follow a sentential adverb. There are evidence that this is indeed correct. In section 3.4.6, we argue that French allows a restricted version of object shift, in the form of bare quantifier movement to the Spec of AGR-O. As shown in (i), when OS has occured, a bare quantifier can still not precede a sentential adverb:

(i) Il a probablement tout mangé
*Il a tout probablement mangé

Il n'a pas tout mangé
* Il n'a tout pas mangé

As we will argue later in this section, the impossibility of object movement in English follows from the impossibility of verb movement. Thus the fact that object cannot precede sentential adverbs in English is predicted.

62. A similar proposal has been made by Holmberg and Platzack (1989). They provide additional evidence for this proposal and show that a number of interesting facts can be made to follow from this distinction. The same range of facts can be captured under the analysis we propose. For reasons of space and time, it is not possible to show how the various properties follow. We will take up this task in forthcomming work.

the verb is limited to T in embedded clauses. Icelandic, on the other hand, does show overt subject-verb agreement. Consequently, we will assume that the verb must move as high as the upper AGR-S projection in all Icelandic sentences so as to fuse with the AGR-S morpheme. The different course of the verb movement in Icelandic and MS is represented in structure (162) as follows: the dotted lines represent the movement of the V to T in MS. The solid line represents the movement of the V to AGR-S in Icelandic. Given this assumption, the apparent difference in the position of sentential adverbials in embedded clauses in IC and MS, i.e., the contrast between examples d. and f. in (148) and (149), is straightforwardly accounted for. The position of the sentential adverbs remain constant but the verb occurs higher in Icelandic, where it moves to AGR-S, than in MS, where it only moves up to the T projection. This view allows us to eliminate Holmberg's stipulated difference in the phrase structure component of MS and IC. Next, we further reformulate constraint (159) in terms of T rather than I:

(163) If a V does not move to T then it cannot move to C.

Given this reformulation, the difference between English main verbs and MS main verbs follows. Main verbs do not move to T in English; therefore they cannot move to C. Main verbs do move to T in MS, and therefore they can move to C.[63]

---

63. We leave open the reason why the movement of the V to C is obligatory in MS and other V/2 languages. Most proposals in the literature concerning the feature composition of C, as in Holmberg (1986) or Holmberg and Platzack (1989), are compatible with our proposal.

One technical question arises with respect to the movement of the V to C in MS. Does the movement of the V to C proceed directly from T to C, bypassing the head of AGR-O and the head of AGR-S, or does the movement proceed successively through these two heads? Direct movement from T to C would violate the Head movement constraint in most of its current formulations. Possibly some adjustment could be made so that a head which is totally devoid of content does not "count" for the Head movement constraint.[64] We will simply assume that the verb does move through the position of AGR-S and AGR-O on its way to C. In our view, nothing in principle forbids the movement of the verb to these projections. The assumption that the verb does not move to these heads in embedded MS sentences simply follows from a view of movement as a last resort: since nothing forces the movement of the Verb to AGR-S or AGR-O in MS embedded sentences (because AGR-S and AGR-O are completely empty), the movement is in fact impossible.

Having provided an account for the movement of the verb in MS and IC, we can now return to the main topic of this section, namely object shift. Given the structure in (162), we propose that object shift involves the movement of an object to the Specifier of the AGR-O projection. Both the position of the object with respect to the sentential adverbs and the nature of the chain as an AN-chain will follow straightforwardly from this proposal. Recall that in our view, the specifier of the AGR-O projection is a [+HR] position and, as we have argued, movement to [+HR] positions creates AN-chains. What now

---

64. See Platzack and Holmberg (1989) for such a proposal.

remains to be accounted for is the restrictions on object shift with respect to verb movement and with respect to the type of constituants which can be moved.

3.3.5 Restrictions on Object shift

3.3.5.1 Dynamic Minimality

We begin with a discussion of the restrictions related to the movement of the verb. Recall that in both MS and in IC, object shift is possible only when a main verb has moved over the position of the sentential adverbs. Given the structure in (162), this means that object shift is only possible when the main verb has moved higher up than TP.

In our view, movement of the object is an instance of movement to a [+HR] position. Suppose that when the verb does not move higher than TP, some barrier intervenes between the extraction site of the object and the Spec of AGR-O in (162). Following Chomsky (1986)b., we assume that a barrier can be crossed via adjunction. But as we have proposed in section 2.4, an adjoined position is a [-HR] position and movement from a [-HR] position to a [+HR] position is an instance of improper movement. Since object shift involves a movement to a [+HR] position, it cannot proceed via adjunction without violating the improper movement condition. Consequently, if any barrier intervenes between

the Spec of AGR-O and the object trace, object shift will be ruled out by the ECP and by the condition on improper movement as formulated in Chapter II. The effect of the movement of a main verb will be to eliminate the barrier in question, so that when the barrier is eliminated, object shift will be possible. This is in essence the proposal that we develop.

As is clear from this short preview, our proposal will rely on precise assumptions about the theory of movement. A deep exploration of the theory of movement is not our primary concern in this section: this will in fact be the topic of Chapte IV. In what follows we will restrict our discussion to the type of movement under consideration here. Before we develop our proposal, let us clarify the results we need to obtain.

1) We want to preclude the movement of the object to the spec of AGR-O, whenever the main verb does not raise above TP in (162). Recall however that, throughout this thesis, we assume that the subject originates at D-structure in the Spec of VP and that it raises at S-structure to satisfy the Case Filter. We also argued in section 2.4.1, that the movement of the subject creates an AN-chain. Given these assumptions, we must be careful that in our attempt to exclude the movement of the object to the SPEC of AGR-O in the required cases, we do not also exclude the movement of the subject. In other words, there is an asymmetry between the raising of the subject and the raising of the object which must be accounted for. Raising the subject out of the VP must be possible in all cases; raising the object, however, must be limited.

2) As we have mentioned above, object shift is not possible in either IC or MS when an auxiliary verb has raised higher than TP and the main verb has remained in place. To put it more clearly, we must account for the fact that the movement of the object can never cross over an overt verbal head which occurs in its D-structure position.

(164)

a. Vafor har Johan inte kopte den
   Why has Johan not bought it

b. *Vafor har Johan den kopten inte

For expository clarity, let us call these two issues respectively the "subject/object asymmetry" and the "aux/main-verb asymmetry".

It is apparent that neither VP nor TP if taken as a relevant barrier to prevent the movement of the object will provide a satisfactory account of either the subject/object asymmetry or the aux/main-verb asymmetry. To see why, let us briefly consider each case in turn.

If TP is taken to be the relevant barrier the following problems will arise:

1) In MS embedded sentences, where the verb does not move higher than T, the subject will not be able to move out of TP. Consider the schematic structure in (165):

(165) $[_{CP}$ att $[...S_i ...[_{TP}\ t_i\ V_j...[_{VP}\ t_i\ t_j\ ]]]]$

In (165), if TP is a barrier the trace of the subject in the Spec of TP will not be properly antecedent-governed and will cause an ECP violation. That the subject must be able to occur higher than TP is shown by the following sentence of Swedish:

(166) att Johan inte kopte den
that Johan not bought them

Thus, the hypothesis that TP is a barrier does not account properly for the subject/object asymmetry.

2) The choice of TP as a barrier does not permit a satisfactory account of the aux/main-verb asymmetry, either. As shown by their positions in embedded sentences in MS, auxiliary verbs must be assumed to be base-generated no higher than the projection of T in (162). This is shown by the fact that they cannot be followed by sentential adverbials.

(167)
a. att Johan inte har kopte den
that Johan not has bought it

b. *att John har inte kopte den

Consider a structure in which an auxiliary has raised out of TP to move either to C in MS or to AGR-S in IC:

(168) [AUX_i....[$_{TP}$...t_i... [$_{VP}$ V O]]]

If we assume that the movement of the auxiliary verb cannot eliminate the barrierhood of TP, the trace of the auxiliary verb in T will not be antecedent-governed, thus leading to an ECP violation. If we assume, on the other hand, that the auxiliary verb can eliminate the barrierhood of TP, object shift will be incorrectly allowed.

If VP is taken to be the relevant barrier, the following problems arise:

1) the subject/object asymmetry will not be accounted for since neither the subject nor the object will be able to raise out of the VP.

2) Recall that we have assumed that in MS the verb raises to TP. In this case, we must assume that the barrierhood of the VP projection has been eliminated or else the trace of the V will fail the ECP since it will not be antecedent-governed. If the barrierhood of VP is eliminated, however, we wrongly predict that object shift should be possible in MS embedded sentences, contrary to fact.

We conlude from the above discussion that to obtain the correct result, namely allowing object shift only when a main verb has moved over TP and accounting for the subjec/object asymmetry, the appropriate barrier cannot be set once and for all but must be computed dynamically. The definition of barrier we propose will have exactly this effect. This definition bears a resemblance to the notion of Minimality barrier of Chomksy (1986) but it in fact has a rather different effect. The crucial intuition behind the notion of Minimality is to prevent government from any governor into the domain of a head. Chomsky's (1986) notion of Minimality is static. A Minimality barrier τ for a trace β will always remain a Minimality barrier for β no matter what happens to the head δ which defines the Minimality barrier. Our proposal is to allow for a <u>dynamic</u> notion of Minimality. In our view, a Minimality barrier for β will not remain a Minimality barrier if the head of the projection has been extracted. Likewise, a projection which was not a Minimality barrier for β may become a Minimality barrier for β by virtue of head movement. Since

the notion of barrier we propose here differs significantly from the notion of Minimality barrier proposed by Chomsky (1986), we will use a different terminology and call them Dynamic Minimality barriers or D-barriers:

Consider the following configuration:

(169) $[...\alpha...[_{\tau}...\delta...\beta]]$

In (169) α is a governor, either a head or an antecedent-governor which is excluded by τ.[65]

In a configuration like (169):

(170)

τ is an D-barrier for β with respect to a governor α iff

τ is a maximal projection which

(i) excludes α and

(ii) contains δ, δ a LEXICAL head which immediately c-commands β.

(171)

α is a LEXICAL head iff

(i) α is not a functional head

(ii) α is overt[66]

---

65. The notion of "exclusion" is here borrowed from Chomsky (1986).

66. The proposal we make is in fact very similar in its effects to what M.Baker (1986) call the Government Transparency Corollary (GTC):

(i) The GTC:
" A lexical category which has an item incorporated into it governs everything which the incorporated item governed in its original structural position" p 64.

We now turn to a discussion of the consequences of these definitions. Note first that since the notion of D-barrier is defined in terms of exclusion, a governor α which is adjoined to τ will be able to govern ß, since it will not be excluded by τ. The consequence is that this

---

In Baker's theory the GTC follows from the definition of barriers which we repeat below:

(ii) Let D be the smallest maximal projection containing A. Then C is a barrier between A and B if and only if C is a maximal projection that contains B and excludes A, and either:
(a) C is not selected, or
(b) the head of C is distinct from the head of D and selects some WP equal to or containing B.

There is however an important difference between Baker's GTC and our proposal with respect to the VP-internal subject hypothesis. "Selection" is defined by Baker as follows:

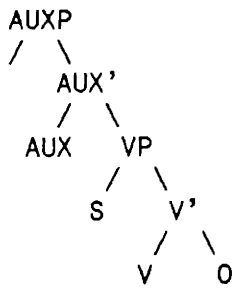
A selects B if and only if:
(i) A assigns a theta-role to B
(ii) A is of category C and B is its IP
(iii) A is category I and B is its VP

The definition of selection requires some obvious adjustements to be made compatible with the Spit INFL hypothesis. Assuming these adjustements, let us consider the category VP containing (under the VP internal subject hypothesis) both the subject and the object. Assume a structure in which the verb has not moved to any functional projection. Consider the case in which the object has moved to the Spec of the functional projection immediately dominating VP. VP (= C) will be a barrier for the trace of the object under (ii)b. since its head selects the object and it is distinct from the head dominating functional projection. As a consequence "object shift" will be correclty excluded. If the verb moves, VP will no longer be a barrier and object shift will be permitted. But a problem arises with respect to the subject. Consider the standard case where the subject is theta-marked by V. In this case, VP will also be a barrier for the subject under (ii)b. As a consequence, the subject will be prevented from raising to the Spec of any functional projection if the Verb does not raise (i.e in a language like English or in any language when there is an auxiliary structure). We conclude that Baker's formulation of barriers is in fact incompatible with the VP internal subject hypothesis.

definition will have no effect on movement which proceeds by adjunction; it will not affect, for instance, WH-extraction. In subsequent discussion the term "movement" should be understood as refering only to [+HR] movement.

The first consequence of the definitions in (170) and (171) is that a VP which contains its verbal head will be a D-barrier for its object but not for its subject. This follows from the requirement of immediate c-command. Although the verb immediately c-commands its object, it does not immediately c-command its subject. Thus when a verb occurs in its D-structure position, movement of the subject will be allowed but movement of the object will not be allowed. Consider next the case in which a verb is dominated by an auxiliary.

(172)

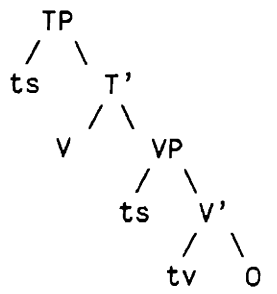


Here the AUX phrase is a D-barrier for the subject, since it contains a LEXICAL head. But the subject can move at least as far up as the specifier of the AUXP. There it will no longer be c-commanded by the head of AUX and it will be free to move further. In this configuration, however, since VP is a D-barrier for the object, the object cannot move out of the VP unless it adjoins to it. Thus, although WH-extraction of an object will be possible, it will be impossible to move the object to an [+HR]-position, since in the latter case this will lead to a violation of improper movement. Note

that the situation remains unchanged even if the AUX moves to a higher functional position. This thus predicts correctly that the object will not be able to move out of the VP when an auxiliary moves to C in MS or to I in IC.

Let us now turn to cases where a main verb raises to a functional projection. The first case we need to consider is the case where the V moves to T in embedded sentences in MS. The structure will be as follows:

(173)



Here VP is no longer a D-barrier for the object, since the trace of V is not a LEXICAL head. But given the definition above, TP will now be a D-barrier for the object since it is a LEXICAL head which strictly c-commands the object. As a consequence, the object will not be able to move out of TP unless it adjoins to it. Thus, object shift is correctly predicted to be ungrammatical in these cases as well.

Suppose now that the verb has moved further up than TP, either up to C as in MS, or up to AGR-S as in IC. Given the structure we have proposed in (162), the result will be that there are now no D-barriers between the D-structure position of the object and the spec of AGR-O. Consequently, "object shift" will be possible.

This predicts correctly that object shift will be possible in matrix clauses in MS when a main verb moves all the way to C. It also predicts correctly that object shift will be possible both in matrix and in embedded clauses in IC when the main verb moves to AGR-S. This completes our account of the link between verb movement and object shift in MS and IC.

3.3.5.2 Pronouns vs NPs

Finally, we need to discuss one last constraint on object shift, namely the constraint on the type of constituents which can undergo the movement. More precisely, we must now ask the question: why is object shift limited to pronouns in MS but not in Icelandic? In Holmberg's view, the limitation to pronominal object shift in MS is a consequence of the Case filter. Holmberg proposes that the Case filter can be satisfied essentially in two ways: first under structural Case-marking and second under overt morphological Case-marking. Holmberg redefines the Case filter as follows:

(174) All and only heads of chains have CASE
(except if they are PRO)[67]

In Holmberg's view, to "have CASE" is to satisfy the Case filter either under structural Case assignment or because the head of a AN-chain bears overt morphological Case. Recall that in his analysis, the

---

67. In Holmberg's terminological system, the term "chain " is restricted to AN-chains.

landing site of object shift is not a position to which structural Case is ever assigned since it is an adjoined position. As a consequence, only elements which can satisfy the Case filter under morphological Case can undergo object shift: the difference between Icelandic Object shift and MS object shift follows straightforwardly. While in Icelandic all NPs bear overt morphological Case features, in MS, only pronouns bear morphological Case features. Thus it is expected that in Icelandic all NP will be able to undergo object shift while in MS object shift will be restrictied to pronouns.

Although Holmberg's proposal is entirely compatible with the proposal we have made and seems at first rather attractive, we will not adopt it. It seems to us that there is strong evidence which argues against considering overt morphological Case-marking as a relevant factor to distinguish among various types of object movement to functional specifiers. As we will see in the next section, German and Dutch are languages which also manifest object movement to intermediate position within a clause. These movements shares several properties with Icelandic object shift but also differs from it in some crucial respects. As is well known, German NPs share with Icelandic NPs the property of bearing overt morphological Case, while Dutch NPs do not. However, the properties of object movement are similar in German and in Dutch, and both differ from object shift in Icelandic. TO sum up, we have the following situation: German and Icelandic share the property of overt Case-marking of their NPs, but the properties of object movement in both languages are not identical. On the other hand, German and Dutch, although they do not share the property of

overt Case-marking of their NPs, have essentially identical properties of object movement, a fact that Holmberg's approach cannot account for. In Holmberg's view, the essential reason why the properties of Icelandic object shift and Dutch object shift differ is because Dutch NPs do not have overt Case-marking while Icelandic NPs do. This view, however, predicts wrongly that German should pattern like Icelandic and not like Dutch, contrary to fact. This clearly suggests that overt morphological Case-marking is not a factor which distinguishes among various types of object movement cross-linguistically. We will thus propose an alternative to Holmberg's proposal.

As a starting point, we will assume that Icelandic and MS differ with regard to the Case-assigning properties of their AGR-O projection. In MS, both AGR projections are devoid of features: in other words, the AGR positions in MS are "abstract" AGR-positions similar to the abstract AGR projection posited by Kayne (1989) for certain dialects of English.[68] We will thus assume that in MS the AGR-O projection does not assign Case. In Icelandic, however, the AGR-S projection is active, and although to my knowledge there is no overt object agreement in this language, we will assume that AGR-O can assign

---

68. This position may pose problems for certain cases of past-participle agreement in Norwegian discussed by Taraldsen and Kristi Koch Kristensen (1989). For these cases however, I would adopt the position taken by the authors that this type of agreement involves Spec-Head agreement in the VP projection. One of the characteristics of this kind of agreement is that it occurs only in constructions where the Spec of VP contains no thematic subjects.

Case.[69] The non-Case-assigning property of the AGR-O projection in MS will prevent correctly NPs from moving to the spec of AGR-O. But the question arises now as to why the pronominal can move to this position. There is a possible alternative to the assumption that pronouns in MS move to the specifier position of the AGR-O projection, and that is that they move instead to the head of the AGR-O projection. Note that in our view, this would still be a [+HR] position, since it is clear that a position adjoined to a head will be a Head related position, i.e, a [+HR] position.[70] Thus the AN-properties of the movement of pronouns in MS are still correctly predicted.

The proposal that the MS pronoun moves to the head of AGR-O seemingly assimilates the pronominal movement in MS to pronominal cliticization in the Romance languages.[71] There is, however, a major distinction between Romance cliticization and MS pronominal movement: while the former is always obligatory, the latter seems optional and is sometimes (e.g. when the main verb does not raise) impossible. To

---

69. In our view, Case and agreeemnt are not direclty related. It has been proposed by Chomsky (1989) that Case and Agreement may in fact be the same form of licensing. Although we believe that there undoubtly is a relation between Case and Agreement, we do not believe that these two relations should identified. Cf Deprez (1988) " On case and Agreement" ms MIT for some argument for this position.

70. It has been suggested by Chomsky (1989) that Head movement has properties very similar to the properties of A-movement (in the LGB sense). In our view, this follows straightforwardly, since it it clear that head movement is a type of movement which involves [+HR] positions.

71. See Kayne (1988) for arguments that clitic movement in Romance involves head movement.

maintain a distinction between Romance cliticization and MS pronominal movement, we propose to make use of Muysken's (1983) view on X' theory in terms of feature distinctions. Muysken (1983) proposes that the levels of categorial projection are determined not in terms of bar-level but in terms of the features [± maximal] [±projected]. The correspondence between Muysken's feature system and the standard X' theory is given in (175):

(175)

[+maximal] [+projected]	------->	X"
[-maximal] [+projected]	------->	X'
[-maximal] [-projected]	------->	X°

Note however that the correspondence is not complete. There remains one combination of features which does not correspond to any distinction in the standard X'theory, namely the combination [+maximal/ -projected]. Keeping in mind the distinctions of the standard X' system of categorial projection, we expect categories of the type [+maximal, -projected] to have the mixed properties of X° and X" projections. If we are correct in assuming that MS pronouns move to the head of the AGR-O projection and not to the specifier, these are in fact precisely the type of properties manifested by MS pronouns. MS pronouns are pronouns which may occur in the positions of maximal projections, i.e., in the same positions as regular object NPs, but they are also pronouns which may occur in X° projections. This is correctly predicted if we assume that MS pronouns are categories of the type [+maximal -projected].

To distinguish Romance clitics from MS pronouns we can assume that Romance clitics are [-maximal/-projected] categories. If so, we predict correctly that the clitics cannot occur in X^0 positions, i.e., head positions at S-structure. This accounts in a very simple way for the fact that Romance cliticization is obligatory, but MS prononominal movement is not. Support for the hypothesis that pronominal movement is to the head of a projection and not to the Spec comes from the behavior of pronouns in Icelandic. Holmberg notes that while object shift of an NP object in Icelandic is purely optional, object shift of a pronoun both in Icelandic and in MS is almost always obligatory when the pronoun is unstressed. If we assume that pronoun movement is to the head of the projection of AGR-O while NP movement is to the Spec, we can account for this difference. It is sufficient to assume that the movement of a [+maximal -projected] category is always preferable, whenever it is possible.

Given the proposed analysis, the difference between Icelandic and MS can be reduced to a difference in the Case-assigning property of the AGR-O projections. AGR-O is a Case-assigner in Icelandic but not in MS. This may plausibly be related to the fact that generally speaking the agreement system is rather robust in Icelandic but inexistant in MS. Pronominal movement, being head movement, is not subject to the Case filter, although it is subject to the Dynamic Minimality system.[72]

---

72. There are also cliticization effects in German. It is possible that these cliticization effects can receive the same analysis as the movement of MS pronouns. Note that an interesting consequence of this proposal is that we expect pronouns in German, unlike NPs,

3.3.6 Summary

We have proposed the following analysis of object shift for mainland Scandinavian languages and Icelandic: object shift involves the movement of an object either to the specifier or to the head of a functional projection. This movement is restricted, occurring only in relation to the movement of a main verb by the system of Dynamic Minimality that we have proposed. Moreover, we have suggested that Icelandic and MS may differ with regard to the Case properties of their functional projections. We suggest that in Icelandic, the projection AGR-O allows Case-marking of an NP under Spec-head agreement. This is what permits the movement of all types of object NPs, pronominal and others, in Icelandic. In MS, on the other hand, we suggest that the functional AGR-O projection is not a Case-assigner. As a consequence, non-pronominal NPs cannot move to the Specifier of AGR-O.

Our analysis posits no structural difference between Icelandic and MS. In particular, we assume that in both languages, sentential adverbials have the same basic position: they are adjoined to TP. The difference we find in the distribution of tensed verbs and sentential adverbs in embedded sentences in IC and MS follows from a well-motivated distinction in the two languages. Since Icelandic manifests overt

---

not to be able to license parasitic gaps. In the literature, no case of parasitic gaps that has been noted mentions a pronoun.

subject-verb agreement, we have proposed that the verb moves as high up as the AGR-S projection in the structure (162). Since MS, on the other hand, manifests no overt subject-verb agreement, we have proposed that in embedded sentences the verb does not move to AGR-S but simply stops at the level of the TP projection.

Finally, it is clear that the system of Dynamic Minimality we have proposed has far-reaching consequences elsewhere in the grammar. We will discuss some of its consequences in Chapter IV.
Note that an important consequence of this proposal is the prediction that S-structure object movement to the specifiers of functional projections is only possible in a language which manifests S-structure Verb movement. This prediction appears to be confirmed in the languages we have discussed so far. Mahajan (1989) has given arguments for the movement of the verb in Hindi. With regard to Japanese, we may assume, following M.Baker's (1986) theory of derivational morphology that affixation of the tense, negation, and honorific markers is done through movement of the verb to various functional projections. The relationship between the movement of the verb and the movement of the object, as we have just seen is very clear in Scandinavian languages.

Moreover, as argued by many linguists (Emonds, Pollock and Chomsky), English is a language which manifests no S-structure movement of main verbs and thus, as predicted by our principle of Dynamic Minimality, does not permit any surface movement of its objects. Although English pronouns just like Swedish pronouns manifest overt morphological Case-marking, they cannot undergo any movement comparable to object shift.

In our view, this follows from the fact that in English, main verbs never move to functional projections, so VP always remains a D-barrier[73] [74]

3.4 German and Dutch Mittelfeld Scrambling

The last languages we consider in this chapter are German and Dutch; the emphasis of this section will be on German. The phenomenon of object movement we study seem to pattern in essentially the same way in Dutch and in German, with some differences that we will not discuss.[75] For some discussion of the differences see Den Besten (1984) and Van Wyngaert (1989). For discussion of the similarities see Koopman and Sportiche (1988).

---

73. This may seem too strong, since we must of course allow object movement outside of VP in passive constructions and unaccusative constructions. Note, however, that in both of these cases, Spec VP is not a θ-position. Consequently, objects can move outside the VP by first moving into the Spec of VP. Similar considerations account straightforwardly for the fact that in English an object appears to be allowed to move (out of VP or inside VP) in passive constructions such as (i), but not in active contructions such as (ii):

(i) There was a man arrested
(ii *The police has a man arrested.

74. Note that Old English contrary to Modern English allowed Verb movement. Our prediction is that Old English must have manifested at least movement of pronominals over negation. This seems to be confirmed. As mentioned to us by Wayne O'Neil (pc), Old (Middle) English allowed constructions such as (i):

(i) I love thee not.

75. One important difference is that at least for some speakers the movement of an object over a dative is quite restricted in Dutch. We leave open the question of why this should be the case.

3.4.0.1 Scrambling vs Base-generation

It has been claimed by Bennis and Hoekstra (1985) and Webelhuth (1989) that the freedom of constituent order which is characteristic of the German and Dutch *Mittelfeld* is due to a syntactic rule of scrambling. Under this view, the word-order alternation between objects and prepositional phrases in Dutch examples such as (176) and in German examples such as (177) manifests movement of the object from its D-structure VP internal position to some intermediate position within the sentence.

(176)
a. dat Jan in Amsterdam zijn vriendin ontmoette
   that Jan in Amsterdam his girlfriend met

b. dat Jan zijn vriendin in Amsterdam ontmoette

(177)
a. Daß Hans im Berlin seine Freundin gekennt hat
   that Hans in Berlin his girlfriend met has

b. Daß Hans seine Freundin in Berlin gekennt hat

In both (176) and (177), the a. examples are taken to show the base order of constituents; the b. examples are assumed to be derived from a. by movement of the object over the prepositional phrase. Under this view examples in b. have the following structure:

(178) [daß [IP Hans [VP seine Freundin$_i$ [VP im Berlin [VP t_i gekennt hat]]]]]

A similar hypothesis accounts for argument alternations in ditransitive predicate constructions. Although accusative NPs can

occur on either side of the dative complement, as shown in (179), Den Besten has proposed that the D-structure of a ditransitive VP for both German and Dutch is as in (180):

(179)
a. Ich habe meiner Muter dieses Buch gegeben
   I have my(Dat) mother(Dat) this book given

b. Ich habe dieses Buch meiner Mutter gegeben

(180) [VP DAT ACC V]

The order in (179)b. is assumed to be derived by movement of the accusative NP over the dative NP.

As shown by Webelhuth (1989), there are a number of asymmetries among the various possible positions which object NP can occupy in German. The existence of such asymmetries supports the hypothesis that the variation in constituent order is derived by movemert, rather than by free base-generation. Webelhuth observes that there is a distributional difference between semantically definite and indefinite NPs. As shown in (181), definite NPs (a-b) may appear on either side of a sentential adverb, while indefinite NPs (c-d) must appear to the right.

(181)
a.weil er wohl das Buch gelesen hat
because he probably the book read has
because he has probably read the book

b.weil er das Buch wohl gelesen hat

c.weil er wohl ein Buch gelesen hat

d.*weil er ein Buch wohl gelesen hat

The same distributional constraint obtains in ditransitive constructions. Examples of the type shown in (182)a., where an indefinite NP occurs to the left of a dative, are unacceptable:

(182)
a.* weil er Blumen der Frau gab (Webelhuth p. 394)
because he flowers the(Dat) woman(Dat) gave

b. weil er der Frau Blumen gab

The ungrammaticality of (181)d. and (182)a. shows that the different positions in which an accusative NP may occur have different properties. This would be unexpected if both positions were base-generated. Some differences are expected, on the other hand, if it is assumed that some positions are base-generated while others are derived by movement.

As shown by Jackendoff (1972), sentential adverbs must occur outside the VP projection in English. Assuming, as is natural, that this is also true in German,[76] we conclude that in (181)a. the definite NP occurs outside the VP constituent.[77] The ungrammaticality of (181)d.

---

76. See Renate Bartsch (1984) for arguments to this efffect.

77. Travis(1988) claims that sentential adverbs can occur inside VP in languages which allow movement of the verb to INFL. In her view, when V movement to I occurs the features of I can percolate down into the VP and license a sentential adverbial in situ. In previous work, we have given a number of arguments against

suggests that an indefinite NP, on the other hand, remains inside the VP projection. The fact that both the position of an NP to the left of a sentential adverb and the position of an NP to the left of a dative are subject to restrictions, and further that they must obey the same semantic restriction--namely, they require definite or specific NPs--suggests that the position of the NP in both cases is a position external to the VP projection. If this is correct, it provides some support for the hypothesis that (180) is indeed the base representation of distransitive verbs and that (182)a. has been derived by movement of the accusative NP over the dative. Additional arguments supporting this hypothesis can be constructed on the basis of weak crossover facts. Consider the following sentences:

(183)
a. *Ich habe seinem$_i$ Professor jeden Schuler$_i$ vorgestellt
I have to his professor every student introduced

b. *Ich habe in seine$_i$ Schachtel jedes Buch$_i$ getan
I have in its box every book put

(183) shows WCO effects. This suggests that at S-structure, the accusative complement does not c-command the dative NP, since the

---

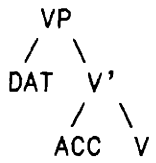
Travis's view. One of the most serious arguments against this view comes from the Romance languages; as argued by Pollock and others, Romance languages in general and French in particular show V-to-I movement. Travis's approach thus predicts that French should license a sentential adverb within the VP projection. Consider, however, the following examples:

(i) Jean donnera probablement de l'argent aux pauvres.
Jean will give probably (prt) money to the poor.
(ii) Jean donnera de l'argent probalement aux pauvres.

(ii) is grammatical, but in this sentence the adverb *probablement* has scope only over the dative complement. It cannot have scope over the whole sentence, as it does in (i). This indicates that contrary to Travis's prediction, a sentential adverb is not licensed within the VP in French, even though V movement to I has occurred.

quantifier cannot bind the pronominal in the dative position. If so, at LF, raising of the quantifier will create a WCO configuration, which explains the judgments given for (183). The WCO facts of (183) show that there is an asymmetry between the indirect object and the direct object, and suggest that the base structure of the German VP is as in (184):

(184)



In (184), the accusative NP does not c-command the dative NP, so we do not expect a quantifier in the object position to be able to bind a pronominal in the dative argument position. This conclusion is also supported by binding facts. Consider (185):

(185)

a. weil ich diesen Autoren Bucher von einander geschenkt habe
because I these authors books by each other given have
because I gave these authors books by each other

b. *weil ich Freunden von einander diese Autoren vorgestellt habe
because I friends ov eachother these Authors introduced have
because I introduces these authors to friends of eachother

(185) shows that binding relations between the direct and indirect object are asymmetric; in (185)a., the indirect object can bind the anaphor contained in the direct object, but this is not the case in (185)b: its ungrammaticality indicates that in this case the indirect object cannot c-command the direct object. The binding possibities follow straightforwardly from the hypothesis that (184) is the base structure of both (185)a. and (185)b.; (185)a.retains the base order, while (185)b. is the result of movement of the direct object over the

indirect object. Thus it is only in (185)a. that the proper c-command relation obtains between the indirect and direct object. The contrast in (185) would be unexpected if the direct object could be freely base-generated on either side of the indirect object.On the basis of the evidence provided in (183) and (185), we will adopt the hypothesis that (184) is indeed the correct D-structure for German ditransitive VPs,[78] and that the order in which the direct object precedes the indirect object is derived by movement.

Assuming that objects do move, what is the nature of the chain created by the movement of the object in sentences like (176),(177) and (181). Does this movement manifest AN-chain properties or Vbl-chain properties? As we will show in this section, the properties of the chain created by this movement appear to be mixed; German sentence-internal scrambling appears to manifest properties of both Vbl-chains and AN-chains. To account for these mixed properties, we will suggest that the distinction we have made so far between chains headed by [+HR] positions and chains headed by [-HR] positions may be insufficient. We will suggest that within the system we have proposed, a further distinction can be made between a [+HR, + Case] position and a [+HR, -Case] position. We will tentatively propose that while movement to a [+HR, +Case] position creates chains which have "pure AN-chain properties", movement to a [+HR,-Case] position creates chains which have "mixed properties".

---

78. See Hoji (1985) for a similar argument for the base structure of the Japanese VP.

The rest of this section is organized as follows. First, we give evidence of the Vbl-chain properties of German scrambling. The arguments are essentially taken from Webelhuth (1989). Next we turn to its AN-chain properties. Finally we argue, following Webelhuth (1989), that the mixture of AN-chain properties and Vbl-chain properties cannot be accounted for independently . If correct, this suggests that the binary classification of chains into AN-chains and Vbl-chains is too coarse and that a third type of chain may be needed. We will conclude this section by showing that the division we propose between chains headed by [+HR, +Case] positions and chains headed by [+HR, -Case] positions, although invisible in German, is supported by some interesting facts in French.

3.4.1 Vbl-chain properties of German scrambling

Webelhuth (1989) argues that German scrambling has Vbl-chain properties by showing that the chain created by scrambling manifests striking similarities to the chain created by WH-movement in German. The similarities between these two types of movement are summarized in the table below and illustrated in the paradigm in (187):

(186)

Properties	WH-movement	scrambling
1) Moves DP	+	+
2) Moves PP	+	+
3) Strands prepositions	+	+
4) Licenses PG	+	+

(187)
Moves DP
a. weil Hans das Buch wahrscheinlich t gelesen hat
because Hans the book probably read has
because Hans has probably read the book

Moves PP
b. weil er mit mihr nicht tanzen wollte
because he with me not to dance wanted
because he didn't want to dance with me

Strands prepositions
c.weil da niemand [PP t gegen] war
because noone it against was
because noone was against it

Licenses PG
d.?weil er den Artikel [ohne PRO vorher e zu lesen] t ablegte
because he the article without before reading filed
because he filed the article without previously reading

Despite these similariies between WH-movement and scrambling in German, it does not seem to us that German scrambling must be taken as an instance of Vbl-chain. Most of the properties which Webelhuth assumes to be characteristic properties of WH-movement do not seem to us to be so clearly distinctive. As noted by Webelhuth, property (1) is common to both WH-movement and NP-movement and can thus not serve as a test to distinguish AN-chains from Vbl-chains. Webelhuth argues that property (2), on the other hand, is a property which distinguishes WH-movement from NP-movement. Clearly, standard cases of

NP-movement such as passive do not usually permit the movement of a PP. It has been argued, however, that some cases of movement to a "surface subject position", i.e., to the Spec of TP or AGRP-S, can sometimes involve PP movement in constructions such as locative inversion (see Bresnan 1988) or in examples such as (188):

(188)

a.Under the table is a safe place to be during an earthquake.

b.Under the table seems to be a safe place to hide during an earthquake.

If this is correct, PP movement may not be a distinctive characteristic of WH-movement. In addition, the scrambling of a PP may have different properties from the scrambling of an NP.[79] We will

---

79. See Mahajan (1989) for an argument that this is indeed the case in Hindi. Mahajan's argument relies on binding facts. As we saw in section 3.1, leftward scrambling of arguments creates new binding possibilities in Hindi. We will show in the next paragraph that this is also the case in German. In both Hindi and in German, however, leftward scrambling of a PP does not allow new binding possibilities. Consider the following example:

(i)* weil ich gegen Paul und Maria$_i$ wahrscheinlich die Kinder von einander$_i$ vertidigen mußte
because I probably must defend each other's children against Paul and Maria

In (i) the PP has been scrambled outside of the VP, as can be seen from the fact that it precedes a sentential adverb. It cannot, however, serve as an antecedent for the anaphor contained in the object position. (i) is thus ruled out by the Binding theory, since the anaphor remains unbound. This contrasts with the example in (ii), in which an NP has been scrambled over a dative argument, allowing the anaphor to be bound:

(ii)weil ich Paul und Maria Freunden von einander vorgestellt habe
because I introduced Paul and Maria to friends of each other.

The ungrammaticality of (i) may be explained in two different ways. First, it could be argued that because of the presence of

suggest below that the scrambling movement of an NP in German may involve the movement of an NP to a [+HR] position. This of course does not rule out the possibility of movement to adjoined positions. It may be that both are allowed for NPs and that PP movement is confined to adjunction. Although we do not take up this question in more detail, we will assume that the fact that PP scrambling is possible says nothing about the type of chain which is created by NP scrambling. There is indeed no contradiction in assuming that the two types of scrambling may involve different processes. Property (3), which, according to Webelhuth distinguishes WH-movement from NP movement in German, is, as is well-known, quite possible with both in English.

(189)

a.This evidence was tampered with by the judges

b.Which evidence did the judge tamper with[80]

Given (189), it is apparent that preposition stranding cannot be taken in all cases as a property distinguishing AN-chains from Vbl-chains cross-linguistically. It is possible that preposition stranding in German has quite different properties from English preposition stranding. In any event, although we will not offer an account of the differences, nor of the reason why NP movement does not allow

---

the preposition , c-command does not obtain between the potential antecedent and the anaphor. (Incidentally, this is also true for the Hindi case.) Second, and more interestingly for our discussion, it could be that scrambling of a PP is to an adjoined position and disallows binding.

80. Although it is unclear to what extent this is correct, it is generally claimed that preposition stranding with NP movement is somewhat more restricted than preposition stranding with WH-movement.

preposition stranding, it seems to us that given (189), the preposition-stranding property cannot be taken as a reliable test to distinguish AN-chains from Vbl-chains. As we have assumed all along, however, property (4) is a distinctive property of Vbl-chains. The fact that German scrambling appears to license parasitic gaps thus clearly distinguishes this type of movement from the object shift movement that we analyzed in the previous sections and from the previous cases of scrambling we studied in sections 3.1 and 3.2.[81]

---

81. Some caution is required with the type of examples given in (187)d. Let us first note that for a number of speakers, these examples are essentially unacceptable if they contain a gap instead of an overt pronoun. Even for these speakers, however, a constrast remains between sentences in which the NP precedes the adjunct clause and sentences in which the NP follows it. Thus, (ii) is even worse than (i):

(i) *weil er den Artikel [ohne PRO vohrer zu lesen] ablegte
because he filed the article without previously reading it

(ii)**weil er [ohne PRO vohrer zu lesen] den Artikel ablegte

No such contrast occurs, however, if the adjunct contains a resumptive pronoun.
Interestingly, for speakers who do not accept these examples, regular parasitic gaps with a WH-extraction are equally unacceptable. We observe the following distribution: speakers who do not accept parasitic gaps with WH-constructions will not accept parastic gaps with scrambled constituents. On the other hand, speakers who accept parasitic gaps with WH-construction fairly well will accept parasitic gaps with scrambled constituents. It is not the case for either set of speakers that there is a significant contrast between parasitic gaps with scrambled constituents and parasitic gaps with WH-constituents. Given the system we propose, this difference could plausibly be explained with respect to the Case-marking possibilities of the functional projections (see below). For the speakers who do not accept parasitic gaps with WH-movement and scrambling, we would suggest that all specifier positions are [+HR, +Case] positions and thus license only pure AN-chain properties. For speakers who accept parastic gaps in both constructions, we would suggest that in their grammar functional projections do not allow Case-marking of their specifiers. Thus, the difference between the two dialects can be explained simply as the consequence of a difference in the properties of functional projections. Such a parameter is in keeping with the hypothesis that all parametric variations occur

We conclude from this brief review that German scrambling shows at least one property characteristic of Vbl-chains, namely the ability to license parasitic gaps. Bennis and Hoekstra (1985) have argued that Dutch scrambling manifests a similar property. According to Bennis (1986), (see also Koopman and Sportiche (1988)) there is a clear contrast between the two sentences in (190):

(190)
a. Jan heeft die boeken [zonder ze/e te bekijken] t weggelegt
   Jan has the book without it/e reading put away
   Jan put the book away without reading it/e

b. Jan heeft [zonder ze/*e te bekijken] t weggelegt

In his view, (190)a. is possible with a gap because the gap is licensed by the movement of the NP from its VP-internal position to a

---

in the lexicon and may in fact be restricted to functional projections, as suggested by Chomsky (1989).
This explanation, however, suffers two problems: first, if the difference between the [+Case] or [-Case] properties of some functional projection is a parametric difference within German, it is unclear how that parametric difference could be set, since it does not seem obvious that this distinction correlates with any other distinction. Possibly, however, such correlations may be discovered with further research. Second, on cursory inspection, the distinction among speakers does not seem to correspond to any geographical region, for instance Northern and Southern dialects, which manifest other known differences with respect to long distance extraction.

It is quite possible that further research will determine that the constructions in question are not parasitic gap constructions, but correspond rather to some object drop phenomenon, the conditions of which remain to be investigated.(See Authier for an interesting proposal which does not, however, address the cases in question explicitly). Note that if it were to turn out that German scrambling cannot license parastic gaps, this would not cause any problem for our view. German scrambling would simply be closer to Hindi scrambling and would thus receive a similar analysis. In any event, it appears that for the speakers who accept these constructions, the judgments given by Webelhuth for these constructions do hold. The purpose of this section is to show how our model can accommodate these constructions if the data are indeed as reported by Webelhuth.

position in which it c-commands both its own trace and the gap within the adjunct. (190)b. is ungrammatical with a gap, however, because the NP has remained in its base-structure position, and consequently there is no path or chain for the gap in the adjunct clause to be parasitic on.

As has been noted by both Haider (1984) and Webelhuth (1989), German NP scrambling also manifests a number of properties which are characteristic of AN-chains. We turn to these properties in the next subsection.

3.4.2 AN-chain properties of German scrambling

The first AN-chain property manifested by German NP scrambling is that it does not create WCO violations, as (191) shows.

(191)
weil Maria den Hans$_i$ seinem$_i$ Professor vorgestellt hat
because Maria Hans to his professor introduced has
because Maria introduced Hans to his professor

Consider the structure of (191):

(192) [weil [Maria den Hans$_i$ [$_{VP}$ seinem$_i$ Professor t_i vorgestellt] hat]]

In (192), the NP *den Hans* binds both a pronoun and its trace, neither of which c-commands the other. Thus, (192) is a typical S-struture configuration for WCO violations. But as the grammaticality of (191)

shows, no WCO effect obtains, a property which we have argued to be typical of AN-chains.

Moreover, as shown in (193), German scrambling also permits WCO repair. As noted above, a quantified accusative NP induces WCO effects with respect to a c-commanding indirect object containing a pronoun, but if the accusative NP is scrambled to a position which precedes the indirect object containing the pronoun, binding of the pronoun by the quantifier becomes possible and the WCO effect no longer obtains.

(193)
a. *weil ich seinem$_i$ Professor jeden Schuler$_i$ vorgestellt habe.
because I his professor to every student introduced have

b. weil ich jeden Schuler$_i$ seinem$_i$ Professor vorgestellt habe.

Interestingly, the same contrast obtains whith scrambled WH-phrases. Consider the following examples:

(194)
a.*Ich frage mich wer seiner$_i$ Frau wen$_i$ vorgestellt hat
I ask myself who his$_i$ wife(Dat) whom$_i$(Acc) introduced has
I wonder who introduced whom$_i$ to his$_i$ wife

b. Ich frage mich wer wen$_i$ seiner$_i$ Frau vorgestellt hat

In (194)a., an accusative WH-phrase in situ induces a WCO violation with respect to the pronoun contained in the c-commanding indirect object. But as shown by the grammaticality of (193)b., the WCO violation disappears when the WH-phrase is scrambled over the indirect object; the fact that the scrambled element is an operator is irrelevant. The grammaticality of (194)b. thus provides more evidence that it is the landing site of movement which determines the properties of a chain, not the nature of the element which heads it.

WH-phrases are clearly operator phrases, yet the properties of scrambling remain unchanged when they are involved.

As noted by Webelhuth, NP scrambling shows another property typical of AN-chains: movement of an NP creates new binding possibilities.

(195)
a. *weil ich einander diese Frauen vorgestellt habe
because I each other these women introduced have
because I introduced these women to each other

b. weil ich diese Frauen einander vorgestellt habe

(195)a. is ruled out by Principle A of the binding theory: the anaphor *einander* has no appropriate antecedent, since the object NP does not c-command it. In (195)b., however, the anaphor is appropriately bound by the scrambled NP. This shows that the landing site of scrambling must be a position from which binding is possible, ie. an [+HR] position in our view. Other examples which lead to the same conclusion are the following:

(196)
a.*weil ich Freunden von einander Paul und Maria vorgestellt habe
because I friends of each other Paul and Maria introduced have
because I introduced Paul and Maria to friends of each other

b.weil ich Paul und Maria Freunden von einander vorgestellt habe

German NP scrambling manifests yet another property which we have argued to be typical of AN-chains: it is compatible with quantifier float.

(197)
weil ich diese Bucher einem Kind alle geben wollte
because I these books a child all to give wanted

becaue I wanted to give all these books to a child

The AN-chain properties of German NP scrambling are summarized in the table below:

---

82. Properties of German reconstruction remain somewhat unclear, essemntially because of the difficulty of the jugements. We summarize briefly teh properties we have been able to distinguish in this footnote. But the jugements are essentially those of one speaker and may thus not be representative.
German permits reconstruction with WH-elements and topicalization:

(i)
(?) Welches Buch uber sich$_i$ hat Hans$_i$ gekauft
Which book about himself has Hans bought
(ii)
(?) Dieses Bucher uber sich hat Hans gekauft
his book about himself, Hans has bought

It also permits reconstruction with scrambling:

(iii)
(?) weil dieses Buch uber sich$_i$ Hans kaufen wollte.
because this book about himself, Hans wanted to buy

But consider the following example:

(iv)
weil seine Eltern$_i$ dieses Buch uber sich$_{i/??j}$ dem Hans$_j$ gegeben hat.
because his parents this book about them/??himself to Hans given have
because his parents gave Hans this book about them/??himself

The anaphor *sich* in German is neutral with respect to number. This means that in sentences (i) and (ii) above, both NPs can logically be the antecedent of *sich*. In sentence (iv) above however, it appears that the anaphor can only refer to the subject c-commanding it. In other words, it appears that in this case the anaphor cannot be reconstructed in its base position to take *Hans* as its subject. As mentioned by Mahajan (1989), the same fact holds in Hindi (although not, apparently, in Japanese). Mahajan interprets this impossibility as follows. Crucially, he assumes that anaphoric reconstruction effects are limited to Vbl-chains. In his view, movement to an intermediate position within the sentence is movement to the Spec of a functional projection and creates AN-chains. Since reconstruction is limited to Vbl-chains, the fact that binding by the the dative is impossible in (iv) is taken to follow from the fact that the chain created by scrambling is an AN-chain. On the other hand, movement to right edge of the

---

sentential constituent can be, in Mahajan's view, either movement to a functional projection or adjunction. Since adjunctions create Vbl-chains, reconstruction from this position is possible. This explains the grammaticality of (iii).
A number of questions arise for this analysis. First, it is unclear why adjunction to an intermediate position in the sentence (VP or any functional projection) shoud be ruled out. Second, consider the following example:

(v)
weil dieses Buch uber sich$_{i/??j}$ seinen Eltern$_i$ dem Hans$_j$ gegeben hat
because this book about them/??himself his parents to Hans given have

In (v) the favored binding of the anaphor remains the subject rather than the dative, even though the constituent has now been fronted to the right edge of the sentence and can thus be in an adjoined position. Again, the same fact holds in Hindi. Mahajan speculates that the reason that the binding by the dative element remains impossible is that the fronted constituent is reconstructed only to the intermediate position and not to the base position of the object. Under this view, the LF reconstructed structure of (v) is equivalent to the S-strucutre of (iv). In other words, it is assumed that movement to the rightward position of the sentence proceeds in two steps: first movement to the Spec of an intermediate position, then movement to an adjoined position. The first step forms an AN-chain and the second a Vbl-chain. Since only Vbl-chain allow reconstruction, reconstruction will only go as far as the intermediate position. In footnote (xx) of section 3.1, we speculated that the obligatory binding of the anaphor by the subject in Hindi was essentially the result of a difference in markedness. We suggested that binding through reconstruction was more marked than binding through c-command and that this explained the impossibility of binding the anaphor by the dative in (iv) and (v). If this is correct, the prediction is that if the subject is not a possible binder for the anaphor, reconstruction should be possible. This is confirmed by the following sentences:

(vii)
(?) weil wir Lehrer dieses Buch uber sich$_i$ dem Hans$_i$ empfohlen haben
because we teacher this book about himself to Hans recommended have

(viii)
(?) weil dieses Buch uber sich$_i$ wir Lehrer dem Hans empfollen haben

(ix)
weil Maria diese Bucher uber einander seinen Freunden gegeben hat

(198)

Properties	Scrambling
WCO	-
WCO repair	+
New binding	+
Q-Float	+

Let us recapitulate. We have seen that German sentence-internal scrambling manifests at least one (and maybe more) characteristic properties of Vbl-chains, as well as several properties of AN-chains. We have so far assumed, following Chomsky (1981), that there are essentially two types of chains, Vbl-chains and AN-chains, which are in complementary distribution. In our view, as in Chomsky (1981 and following), the distinction between the two types of chains corresponds to a distinction between two types of positions, [+HR] positions and [-HR] positions. There is strict complementary distribution between the two types of positions, and thus movement to either type of position is expected to manifest one given set of

---

because Maria this book about each other her friends given has
because Maria has given this book about each other to her friends

These sentences have the same status as the examples of reconstruction above in (i), (ii), and (iii). The grammaticality of these sentences thus confirms our hypothesis. Moreover, they support the hypothesis, argued for in section 3.2.2.1, that both AN-chains and Vbl-chains allow for anaphoric reconstruction.

(?) [Dieses Buch uber sich$_i$]$_j$ hat seine Eltern dem Hans t$_j$ gegeben

(?) [Welches Buch uber sich$_j$]$_i$ hat seine Eltern dem Hans$_j$ t$_i$ gegeben]

properties. The mixed properties of the chain created by German scrambling, however, seem to conflict with this simple binary distinction.

3.4.3 Toward a ternary partition of positions

Note that the mixed properties of the chain created by German scrambling could be compatible with a binary distinction in the typology of positions if it turned out to be the case that each set of properties (Vbl-properties and AN-properties) never occurred simultaneously. This would entail that scrambling can "choose" to move either to a [+HR] position or to a [-HR] position, creating a different type of chain in each situation. If so, the fact that scrambling can show either AN- or Vbl-properties is expected. But we expect properties of both sets never to cooccur. As shown by Webelhuth, however, this prediction is incorrect. The following example simultaneously displays AN-chain properties and Vbl-chain properties:

(199)
? Peter hat jeden Gast$_i$ [ohne e$_i$ anzushauen] seinem Nachbarn t$_i$ vorgestellt. (W 1989)
Peter introduced each guest without looking at (him) to his neighbor

(200)
? Peter hat die Gaste$_i$ [ohne e$_i$ anzushauen] einander t$_i$ vorgestellt. (W 1989)
Peter introduced the guests [without looking at (them)] to each other.

(199) is an example which displays both WCO repair (or pronominal binding by a scrambled quantifier), an AN-chain property, and

parasitic gap licensing, a Vbl-chain property. (200) displays anaphor binding by the scrambled element and parasitic gap licensing. These examples show clearly that the Vbl-chain properties and the AN-chain properties of scrambling can cooccur. Consequently, the suggestion that scrambling can "choose" its landing site and thereby determine the unique status of its chain is not tenable. Suppose that the landing site of scrambling is an A'-position in (200), for instance; then the binding facts would remain unexplained. If, on the other hand, the landing site is an A-position, the occurrence of the parasitic gap would remain unexplained.

Based on this observation, Webelhuth proposes to modify the binary distinction of positions of the LGB model. In his view, the landing site of scrambling is an adjoined position, an A'-position in the LGB classification. To account for the mixed properties of German scrambling, he proposes that positions be classified in a ternary system which comprises, first, "pure" A-positions (potential θ-marked positions) from which only AN-chains can be formed, second, "pure" operator positions from which only Vbl-chains can be formed, such as the Spec of the CP projection, and third, adjoined positions, which may head chains with ambiguous properties. In Webelhuth's view, chains headed from adjoined positions are unrestricted with respect to their binding-theoretic properties.

Webelhuth's proposal is articulated within a model which assumes neither the VP-internal hypothesis nor the "split INFl" hypothesis, with its multiple functional projections. In a model which assumes

both hypotheses, such as the model we have adopted throughout this thesis, the conclusion drawn by Webelhuth for examples of the type of (199) and (200) is not necessary. Although (199) and (200) show that an approach in which scrambling discriminately "chooses" its landing site is untenable, this does not entail that the binary distinction of positions needs to be modified. It is possible to construct an alternative derivation for examples such as (199) and (200) which satisfies both the AN-chain properties and the Vbl-chain properties of German scrambling and at the same time maintains the binary distinction. This alternative derivation is one in which the scrambled element first moves to a position from which AN-chains can be headed, i.e., a [+HR] position, and subsequently moves to a position which can head Vbl-chains, a [-HR] position. Under this view, the structure of examples such as (199) and (200) would be as follows:

(201)

? [Peter hat [jeden Gast$_i$ [ohne e$_i$ anzushauen] [t'$_i$ [seinem Nachbarn t$_i$ vorgestellt]]]

In (201), we assume that the trace t'$_i$ is in the specifier of some functional projection, a [+HR] position in our terms. Support for this assumption comes from the fact that, as shown in (202), a floating quantifier(underlined) can occur in this position.

(202)

? Peter hat den Gasten$_i$ [ohne e$_i$ anzushauen] alle$_i$ t'$_i$ ihren Nachbarn t$_i$ vorgestellt.

Recall that under Sportiche's (1988) view of floating quantifiers, which we have adopted, floating quantifiers may be taken to indicate the intermediate landing sites of NP movement. In (202), the position of the floating quantifier suggests that there is a possible [+HR] landing site for the scrambled NP which is outside the VP constituent,since it is to the left of the indirect object and occurs before the adjunct. If so, the AN-chain property which is manifested in (199) by pronominal binding will be straightforwardly accounted for. Recall that in our view, AN-chains are headed by [+HR] positions: thus the chain (t'_i, t_i) will be an AN-chain. We can now assume that the NP *den Gasten* has subsequently moved over the adjunct to an adjoined position, from which it heads a Vbl-chain which licenses the parasitic gap.[83] Given this derivation, the apparent simultaneous mixed properties of German scrambling shown in examples such as (199) and (200) can receive an explanation compatible with a binary distinction between AN-chains and Vbl-chains with no modification. In other words, (199) and (200) no longer constitute counterexamples for a binary distinction of positions.

Note, however, that Webelhuth's intuition that a ternary distinction in the typology of position may be needed may still be correct. Although examples (199) and (200) do not force a further subdividsion of the positions relevant for the AN/Vbl-chain distinction, it is possible to construct examples where the AN-chain and Vbl-chain

---

83. A similar analysis has been proposed by Wyngaerd (1989).

properties must truly hold simultaneously and cannot hold successively. Consider the following example:

(203)
a. ?weil Maria jeden$_i$ Gast [ohne seinem$_i$ Partner e$_i$ vorzustellen] t$_i$ allein laßt.
because Maria every guest without his partner introducing left alone
because Maria left every guest alone without introducing to his partner

b. ?weil Peter jede$_i$ Frau ohne ihrem Partner vorzustellen allein laßt.
because Peter every woman without her partner introducing left alone
because Peter left every woman alone without introducing to her partner

The sentences in (203) show the following phenomena. First, there is a parasitic gap in their adjunct clause. Second, the indirect object in their adjunct clause contains a pronoun which is bound by the quantifier in the main clause.

Note that the acceptability of sentences in (203) is on a par with Webelhuth's examples in (199) and (200). That is, they are far from perfect. What is crucial, however, is that there is no difference in acceptability between a reading in which the pronoun in the adjunct clause is bound by the quantifier and a reading in which it is not. In other words, the binding of the pronoun does not reduce the acceptability of the sentence.

As we showed above, a pronoun contained within an indirect object cannot be bound by a quantifier in the direct object position. Thus, the bound reading of the pronoun in (203) clearly cannot be licensed by the parastic gap in object position. If this is right, we conclude that it is the quantifier in the matrix clause which serves directly

as the antecedent for the pronoun in the adjunct clause. The structure of (203) is given below:

(204) [weil [Peter jede Frau$_i$ [ohne ihrem$_i$ Partner e$_i$ vorzustellen] [$_{VP}$ t$_i$ allein laßt]]]

The NP *jede Frau* has been scrambled over the adjunct phrase and binds both its trace inside the VP and the pronoun inside the adjunct, neither of which c-commands the other. This is an S-structure configuration for WCO, but WCO effects do not obtain. Thus, scrambling in (203) manifests AN-chain properties. Since a parasitic gap is licensed, scrambling also manifests Vbl-chain properties. (203) then truly shows the simultaneity of the AN-properties and Vbl-properties of German scrambling. Note that an analysis involving two consecutive movements creating an AN-chain and a Vbl-chain successively is of no help in this case. To allow pronominal binding, either the quantified NP itself or its trace must c-command the pronoun. But if either the trace or the quantified NP c-commands the pronoun, it will also c-command the parasitic gap. We conclude that it must be the same chain which licenses both the binding of the pronoun and the parasitic gap, whether that chain is headed directly by the quantified NP or by its trace, in cases where the quantified NP itself has moved further. Given examples such as those in (203), we are driven back to the conclusion that some positions seem to create chains with ambiguous properties. Thus, we are driven back to Webelhuth's proposal that a ternary rather than a binary distinction in positions may be needed. To account for examples of the type in (203), we will indeed suggest that a further distinction between the types of positions we have so

far distinguished can be made. We will, however, make a proposal which differs from that of Webelhuth.

3.4.4 On the nature of positions which head chains with mixed properties

Recall that in Webelhuth's view adjoined positions are positions which can head chains with mixed properties. It seems to us, however, that such a proposal runs into some theoretical and empirical problems.

First, it is standardly assumed that quantifier raising creates adjoined structures at LF. Clearly QR, unlike scrambling, has Vbl-chain properties, since it induces crossover violations. Recall the contrasts given above between quantifiers which have remained in situ and quantifiers which have been scrambled; only the former induce WCO violations. If it is assumed, however, as suggested by Webelhuth, that adjoined positions have ambiguous properties, then either the hypothesis that QR is movement creating adjunctions must be revised, or else it must be assumed that S-structure adjunctions and LF adjunctions have different properties: in particular it must be assumed that only S-structure adjunctions create chains which are mixed with respect to their binding-theoretic properties.

Second, if adjoined positions are positions from which chains with mixed properties are licensed, it becomes quite unclear how a condition on improper movement could be derived. Consider a case of super-raising:

(205)*John$_i$ seems [that it is considered [t_i to be intelligent]]

To exclude this sentence, Chomsky (1986) crucially assumes that adjunction to VP is excluded by the condition on improper movement.[84] If adjunction to the VP projection of *consider* were allowed, the trace in the subject of the infinitival complement would be properly antecedent-governed, and (205) would be at worst a mild subjacency violation. Recall that the LGB condition on improper movement rules out structures which involve movement from an A'-position to an A-position. If, as proposed by Webelhuth, however, adjoined positions are ambiguous with respect to their binding properties, then they should be compatible with both A-movement and A'-movement. Consequently, movement from an adjoined position to an A-position should be possible, which among other problems would wrongly predict that (205) should be only slightly marginal.

Finally, it is standardly assumed that Heavy NP-shift constructions in English create adjoined structures (see Baltin (1972) Johnson (1986) among others).[85] As argued by Webelhuth, among others, HNPS creates chains which have Vbl-properties. Thus, HNPS licenses parasitic gap constructions, as shown in (206), and it is incompatible with quantifier float, as shown in (207), two properties which we have argued to be characteristic of Vbl-chains.

(206)
a.John offended by not recognizing, his favorite uncle from Cleveland.

---

84. See *Barriers*, p. 21 for details.

85. Cf. Larson (1988) for a different view.

b. I filed without reading thoroughly all the papers you gave me yesterday

(207)
*I bought all yesterday the books that were required for the course.

If adjoined positions were positions which could head chains with ambiguous properties, we would expect a HNP-shifted NP to permit WCO repair and to license new binding from its landing position. As we show in (208) and (209), however, this expectation is not fulfilled.

(208)

a.* John showed his$_i$ mother everybody$_i$

b.* John showed his$_i$ mother without showing his father, everybody who finished on time.

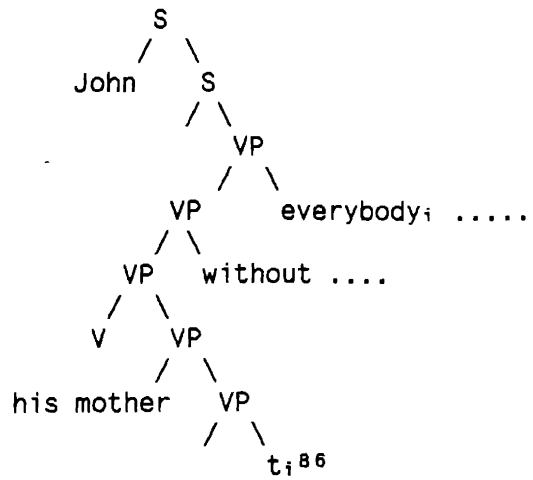
(209)

a.* John showed each other's friends Bill and Mary

b.* John showed each other's friends early in the morning, Bill the happy fellow and Mary the beauty

(208)a. is a double object structure. As shown by Barss and Lasnik (198), in double object structures the first NP c-commands the second but not conversely. Thus (208)a. is a straight WCO violation. In (208)b., however, the second NP has undergone HNPS over the adjunct phrase. Thus the structure of (208)b is roughly as in (210):

(210)



In (210), the quantifier c-commands the pronoun *his* after Heavy NP shift. But the binding of the pronoun by the quantifier remains impossible. We thus conclude that the adjoined position which is the landing site of HNPS does not head chains with mixed properties. This conclusion is also supported by the binding facts given in (209). We observe that the landing site position of HNPS does not permit new binding. In (209), the NP *Bill and Mary* cannot serve as an antecedent for binding. It is thus clear from the examples in (208) and (209) that adjoined positions cannot always be taken to head chains with mixed properties.

For these reasons, we do not adopt Webelhuth's proposal. We will make the speculative proposal that the types of positions which can head chains with mixed properties are [+HR] positions and not adjoined

---

86. In (210) we adopted the structure proposed by Larson to represent the double object construction merely for convenience. Any structure which correctly accounts for the fact that the first NP does not c-command the second prior to HNPS would make no change to our immediate point.

positions. As we argued in section 3.3.5, some [+HR] positions are positions to which Case is assigned, either by a head or through a process of Spec head agreement. Other [+HR] positions are positions to which Case is not assigned. Thus, [+HR] positions subdivide into two types: [+ HR, +Case] positions and [+ HR, -Case] positions. Given this distinction, we suggest the following typology of positions:

(211)

[+HR + Case] positions are positions which head strictly AN-chains

[+HR - Case] positions are positions which head chains with mixed properties.

[-HR] positions are positions which head strictly Vbl-chains

Since we propose that the [± Case] distinction is relevant to subdivide the class of [+HR] positions, we need to ask whether it is also necessary to partition the class of [-HR] positions. In principle, the classification we propose allows for four types of positions: [+HR/-Case], [+HR/+Case], [-HR/-Case], and [-HR/+Case]. It is standardly assumed, however, that adjoined positions are not positions to which Case is directly assigned. Since in our view [-HR] positions are adjoined positions, we would not expect the [± Case] distinction to further partition this class of positions. Thus, we will tentatively suggest that no [-HR] position can receive Case directly. The following definitions will ensure this result:

(212)

α can assign Case to β iff:

(i) α is a Case-assigning head and either:

(ii) β is narrowly included in the maximal projection of α or

(iii) β is narrowly included in τ, τ the maximal projection of a head δ, which is a complement of α and not a barrier for β, and δ does not assign Case to β

(213) β is narrowly included in τ, τ a maximal projection, iff β is in a [+HR] position included in the projection of τ

Note that (212)(iii) is only needed to ensure that Exceptional Case marking is possible. If we were to assume that ECM proceeds differently, (iii) could be suppressed and Case-marking would reduce to (ii): that is, Case would be assigned by a given head strictly to its specifier and/or to its complement, a very simple and quite natural definition.

It has been proposed by Reuland (1983) that in ECM constructions Case is in fact assigned to the subject of the ECM clause by the head of the ECM complement clause and not by the verb. Assuming, as proposed by Kayne (1986) (in class) among others,[87] that ECM infinitival clauses and ECM small clauses are headed by functional heads, we suggest, following Reuland, that the functional head assigns Case to an ECM subject only when it is governed by an appropriate verb. (See

---

87. See Uriageraka (1987) for some interesting arguments for such a position

Reuland (1983) for further details).[88] If such a view is tenable, then Case-marking can be reduced to (ii) in (212). The issue of the modality of Case assignement in a model such as the one we explore in this thesis is no doubt an important question but one which would require another dissertation to investigate properly. We thus leave this important question open and we simply adopt some hypotheses which we will not take the time to motivate. In particular, we will assume that the configuration for Case assignment reduces to (212) (ii) and adopt the conjecture that Case assignment in ECM constructions proceeds roughly along the lines suggested by Reuland (1983). Given these hypotheses, Case will never be assigned to a [-HR] position, and consequently the [± Case] distinction will not further subdivide the class of [-HR] positions.

Let us turn now to a discussion of some of the consequences of our proposal for German scrambling. Given the above typology of positions, we can now suggest that German scrambling is similar to Hindi scrambling, Japanese scrambling, and object shift in Icelandic in that it involves movement of an NP to the specifier of a functional projection.[89] German scrambling differs from Object shift in

---

88. Other views on Case-marking may be compatible with our approach, such as the one proposed in Chomsky (1989) under which only functional projections can assign Case.

89. We do not mean to claim that under our proposal, we expect scrambling in all these languages to have exactly the same properties. Quite clearly, it is entirely possible that there are additional properties which will distinguish scrambling in these different languages. We expect this, maybe as a reflection of the various properties of functional projections in each of these languages. Our purpose has been to show that across such wildly

Icelandic, however, in that while the former may involve movement to a [+HR -Case] position, the latter may only involve movement to a [+HR +Case] position. This distinction accounts straightforwardly for the fact that Icelandic object shift does not license parasitic gaps, since the chain headed by the moved object will have strict AN-chain properties, given the typology in (212).[90] German scrambling, on the other hand, will be able to license parastic gaps because the chain created by scrambling may have mixed properties, since it can be headed by a [+HR, -Case] position.[91]

We do not mean to propose that German functional projections never assign Case to an NP which occurs in their specifier. Our proposal is rather that when a functional projection does not assign Case, movement of an NP to its specifier is still possible in German. Let us clarify this statement somewhat. Let us assume that Case assignment is an S-structure procedure which is generally optional (the null hypothesis): under this view a head which is a Case-assigner may or may not assign Case to its specifier and/or to its complement. With this in mind, let us briefly consider object shift in Icelandic by way of contrast. As we have assumed, the verb optionally assigns Case to

---

different languages, there are striking similarities in scrambling which need to be accounted for.

90. It is to be hoped that if this distinction is indeed significant, it will correlate with other properties yet to be discovered.

91. Since both Hindi and Japanese are languages which allow free object drop, it is not possible to determine, on the basis of the phenomena we have analyzed in section 3.1 and 3.2, whether scrambling in either of these languages involves movement to a [+HR -Case] position or to a [+HR +Case] position. We thus leave the question open for further research.

its object. Suppose now that the verb takes the option of not assigning Case. We assume that both the AGR-S and the AGR-O functional projections are potential Case assigners in Icelandic. If so, an object which does not receive Case from its verb may move to the Spec of a functional projection and receive Case in this position under Spec-head agreement. Under this view, object shift in Icelandic is parallel to standard Cases of A-movement such as passive in the LGB framework: movement of the object is forced by the Case filter. If, on the other hand, the verb chooses the option of assigning Case to the object, then movement of the object is impossible.

Turning to German, we suggest that in this language, even when Case is assigned by the verb inside the VP projection movement of the object to the Spec of a functional projection is still possible. Under this view, German differs from Icelandic in that it licenses an NP in the specifier of a functional projection even if the NP does not receive Case in this position.

Consider the two types of chains that we are positing:

(214)

a. (NP.....t)
[+HR +Case] [+HR -Case] (Icelandic)

b. (NP.....t)
[+HR -Case] [+HR +Case] (German)

We assume that the type of chain in (214)a. is well-formed because it obeys Chomsky's (1986) well-formedness condition on AN-chains, a condition which requires that the head of an A-chain (in the LGB sense) be Case-marked. The type of chain in (214)b., however, does not obey this well-formedness condition. We need to determine what the

well-formedness condition on chains of the type (214)b is. Clearly, chains of this type are not simply freely generated. For instance they are impossible in French, a language which we argued in section 2.4.2 to have an AGR-O functional projection which does not assign Case. Recall that in French, a sentence such as (215) is impossible:

(215)
* Pierre a [AGRP-O la pommei [VP mangée ti]]
  Pierre has the apple eaten

We argued earlier in section 2.4.2 that (215) is ruled out by the Case condition on AN-chains because the NP *la pomme* is not assigned Case by the AGR-O head.[92] But this is not sufficient if chains of the type in (214)b. exist. We must additionally rule out these types of chains. Recall, as we discussed in section 3.3.5.2, that Holmberg (1986) proposed a modification of Chomsky's (1986)a. Case condition on A-chains and argued that overt morphological Case can license the head of an A-chain in a non-Case-marked position. As is well known, German manifests overt Case morphology on its argument NPs. We could thus adopt Holmberg's suggestion and assume that it is the overt morphological Case-marking on German NPs that allows them to move to [+HR, -Case] positions. This suggestion, however, is not compatible with the facts of a closely related language, namely Dutch. Recall

---

92. Contrary to appearances, this sentence cannot be excluded by the D-Minimality principle we proposed in section 3.3.5.1. It is argued by Pollock on the basis of examples like (i) that the past participle may move to the AGR-O projection.

    (i) Jean a mangé souvent des pommes.
    Jean has eaten often apples

    In (i) an adverb intervenes between the verb and the object NP, indicating that the verb must have moved.

that, as argued by Bennis and Hoekstra (1985) and by Bennis (1986), Dutch scrambling, like German scrambling, licenses parasitic gaps. Moreover, as argued by Wyngaert (1989), Dutch scrambling also exhibits the same AN-chain properties as German scrambling. In short, it appears that Dutch allows the same type of chains with mixed properites as German. If this is correct, then we must assume that Dutch, like German, allows the formation of chains of the type in (214)b. But as is well known, Dutch differs minimally from German in that it shows no overt morphological Case-marking on its NPs. Thus, the morphological Case-marking hypothesis cannot account for the fact that Dutch seems to license chains of the type in (214)b. There is, however, a property which is common to Dutch and German and which distinguishes them from Icelandic, French, and English, namely the head-initial/ head-final parameter. Both German and Dutch are head-final languages, while Icelandic, French, and English are head-initial languages. Let us speculate, then, that chains of the type in (214)b. can only be licensed in languages which are head-final, or maybe more generally in languages in which specifiers and complements are in the same direction with respect to the head. This restriction will allow the formation of chains of the type in (214)b. in German and in Dutch, but will exclude such chains in Icelandic, French, and English. Under this view, the possibility of NP object movement in Icelandic comes from the conjunction of verb movement with the hypothesis that the functional projection of AGR-O can assign Case to its specifier. Its impossibility in French will follow from the lack of Case-marking ability for the AGR-O projection in French. Finally, in English the impossibility of the movement of any object to the specifier of

functional projections follows from the absence of verb movement, as we argued in the previous section.

Let us summarize briefly. The data we have presented suggest that the chain created by German scrambling can manifest mixed properties, both AN-chain properties and Vbl-chain properties. To account for this possibility, we have speculatively proposed to replace the binary division of positions we have presented earlier in Chapter II, by the ternary division of positions outlined in (211). Moreover, we have suggested that German and Dutch allow a Case-marked NP to move to a [+HR -Case] position, a position from which it can head a chain with mixed properties. Finally, we have suggested that the movement of an NP to a [+HR -Case] position is licensed only in languages where specifiers and complements occur in the same direction with respect to a head.

This condition on the formation of chains of the type in (214)b., although descriptively correct, appears somewhat suspicious. We take this to indicate that further research is needed to determine more clearly the status of the mixed chains we have posited. Following Webelhuth (1989) and Bennis and Hoekstra (1988), we have analyzed the gaps which are licensed by object movement in German and Dutch as "parasitic gaps". Note, however, that unlike the more robustly attested parasitic gaps such as the parasitic gaps licensed in English by WH-movement, we have speculated that these gaps are licensed from a [+ HR -Case] position. As we will argue in the next chapter, WH-movement involves movement which has as a landing site a [-HR] position. Given this distinction, we might expect the parasitic gaps

which are licensed by the movement of an object to have different properties than the parasistic gaps licensed by WH-movement, i.e., chains licensed by [-HR] positions. If this turns out to be correct then movement to [+HR -Case] positions might not really create chains with mixed AN and Vbl properties, but rather AN-chains which have the property to license a particular type of "dependent gap", the properties of which remain to be investigated.

3.4.5 Functional projections in German

Since we have proposed that German scrambling involves movement to specifiers of functional projections, the question arises whether there is independent evidence for the existence of multiple functional projections in German. Arguments which have been so far given in the literature for the existence of functional projections are basically of two types: either they are based on morphological evidence or, as in Pollock (1989), they are based on the positions of adverbs with respect to other constituents. On the morphological side, let us first note that German has overt marking of subject-verb agreement. We can thus assume that German has at least one functional projection in which subject-verb agreement occurs. Turning now to the adverbial positions, as we mentioned above, German appears to allow sentential adverbials to occur in the middle of the VP:

(216)
weil der Hans dieses Buch warsheinlich dem Maria geben wird
because Hans this book probably to Maria give will
because Hans will probably give this book to Maria

Jackendoff (1972) has convincingly argued on the basis of English facts that the base position of sentential adverbials is neither a position within the VP nor adjoined to the VP, but rather a position adjoined to some projection of the AUX constituent. Transposing this analysis to the split INFL framework we have adopted in this thesis, we have suggested that sentential adverbials are base-generated as adjuncts to the maximal projection T, perhaps universally.[93] If so, the position of the subject and the position of the object in (212) suggests that there must be at least two functional projections above the T level. Since in German it is also possible to move an object above the subject in sentences such as (212), this suggests that German functional projections, like Hindi projections, may host any argument.[94]

We propose that German has a structure similar to the one we posited in the previous section for Scandinavian languages, with the

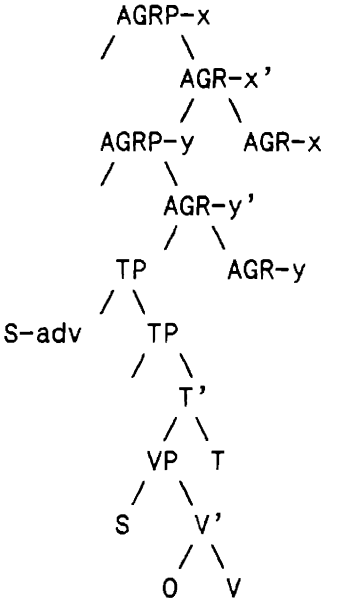
---

93. See Rizzi (1989) for a similar proposal.

94. Possibly this may be a consequence of the fact that Case-marking by a functional projection is not necessary for NP movement to this projection in German. This entails that Case can be assigned inside the VP projection to all arguments, including the subject. See Den Besten (1985) for a proposal that Case can be assigned to the Subject inside the VP in German. See Platzack and Holmberg for a similar proposal for Icelandic and Scandinavian languages. May be object over subject movement reflects a configuration in which the subject has remained inside the VP. This appears to be true in Dutuch (Den Besten) where in such contructions (OSV) the subject must be indefinite. CF also Diesing (1988) for some semantic arguments and Kratzer (1988)

difference that German functional projections are assumed to be head-final, in conformity with the VP constituent. Under this view, the basic structure of a sentence in German is as in (217):

(217)



Given (217), we can now view scrambling as involving a movement to the specifier of any of the functional projections posited above.

One remaining condition on the movement of objects to the specifier of functional projections in German needs to be discussed. Recall that in the preceding section, we argued that the movement of objects to the specifier of functional positions is conditioned by the movement of the verb. Since there is no asymmetry in object movement between main and embedded sentences in German, we must assume that the verb in German may move to the highest functional projection under C. Recall that in the previous section we argued that movement of the verb to T is a necessary condition for allowing the movement of the V to C.

Since German is a typical verb-second language which has been argued to involve a movement of the verb to C in main clauses, then according to the generalization proposed in section 3.3.4.2, we conclude that the verb must move at least to the projection of T. Moreover, since German, like Icelandic, shows overt subject-verb agreement, we can assume that it moves to the highest projection containing the subject agreement morpheme. Unlike Mainland Scandinvian and Icelandic, though, the movement of the object in German is insensitive to the presence of a complex tense involving an auxiliary verb and a past participle. We speculate that the reason for this insensitivity is that in German the past participle independently moves to the functional head immediately following the head containing the auxiliary verbs. Some support for this speculation comes from the fact that in embedded sentences, nothing may intervene between the auxiliary and the past participle. In particular, no VP adverb may occur between these two verbal heads.

(218)
*weil Maria dieses Buch gelesen grundlich hat.
because Maria this book read thoroughly has

An alternative way of ruling out (218) would be to suggest that VP adverbs in German may never adjoin to the right of VP. It is unclear, however, what would motivate such a condition. Possibly one could argue that this restriction follows from the head-final parameter, but there is no correspondence to such a condition in head-initial languages. As the distrubution of manner adverbs in English shows, VP adverbs can occur on either side of the VP.

(219) a.He thoroughly read this book

b.He read this book thoroughly.

Moreover, recall that, as argued by Webehuth (1989),indefinite NPs, unlike definite NPs, appear to remain inside the VP. But as shown in (220), certain VP adverbs can occur on either side of an indefinite NP.

(220)
a.weil deine Mutter immer sorgfeldig ein Kleid geputzt hat
because your mother always carefully a dress cleaned has

b.weil deine Mutter immer ein Kleid sorgfeldig geputzt hat

Assuming, as has been argued by Stowell (1981) and Kayne (1983), that no adverb can intervene between the D-structure postion of a verb and its complement NP, we must conclude that in (220) either the object has moved to the left, or the verb has moved to the right. If we assume that the object has moved out of the VP, the fact that an indefinite NP cannot occur to the left of a sentential adverb becomes rather mysterious. Indeed, if an indefinite object can move a certain distance out of the VP, one would expect it to be able to move further. On the other hand, under the assumption that the verb and the past participle can move to the right, we can continue to assume that (for a reason yet to be discovered), an indefinite NP must remain in its D-structure position. We will assume the latter alternative and take sentences of the type in (218) and (220)b. as evidence that both the verb and the past participle move in German. Under this view, the insensitivity of the movement of the object to auxiliary structure in German follows straightforwardly.

As a final argument that German scrambling appears to be conditioned by the movement of the verb, comes from studies in German language acquisition. As noted by Clahsen (1982), in the child's grammar, prior

to the acquisition of verb movement, the object occurs adjacent to the verb:
(221)

"Before the acquisition of V-restriction II [the verb-second phenomenon VD]the children prefer an order of the complements in which the object appears adjacent to the verb. After the acquisition of V-restriction II, this constraint is given up, i.e., now adverbials can also stand between the object and the verb".
Clahsen 1982, 70
Cited in Webelhuth (1989), p.359

This suggests that the mobility of objects is directly related to the mobility of the verb. In our view, this correlation simply follows from the ECP, given our proposed definition of Dynamic Minimality. This concludes our analysis of scrambling in German.

3.4.6 Bare quantifier movement in French

In the previous paragraph, we presented a speculative typology of positions which allows us to distinguish three types of chains:

1) chains headed by [+HR +Case] positions, which manifest strictly AN-chain properties
2) chains headed by [+HR -Case] positions, which manifest mixed properties
3) chains headed by [-HR] positions, which manifest strictly Vbl-chain properties

We have suggested that the [±Case] distinction subdivides the set of [+HR] positions. Although this proposal is compatible with the theory

we have elaborated so far, it receives no direct support from the phenomena we have studied. As we have shown, object movement in Icelandic is distinguished from object movement in German by the fact that the former but not the latter licenses parastic gaps: there is, however, no clear evidence that this phenomenon is related to the [± Case] property of the landing site, as we suggest. In this section, we will briefly consider a particular instance of object movement, namely the movement of bare quantifiers of the class including *tout*("everything") and *rien* ("nothing") in French. The analysis we propose for this phenomenon will provide some supporting evidence for the hypothesis that the [±Case] distinction may be relevant for a further dictinction among [+HR] positions.

We showed earlier (Chapter II, section 2.4.2) that French shows no overt NP object movement to the specifier of a functional projection. We argued that the absence of overt NP movement in this language is due to the inability of the functional projection AGR-O to assign Case. Suppose, however, that French had some type of objects which did not need to be directly Case-marked; nothing in principle would then forbid their movement to the Spec of AGR-O in French. Fullfilling this expectation, there is indeed one type of object in French which is able to undergo movement to an intermediate position in a sentence, namely, bare object quantifiers like *tout* and *rien*. Consider the following examples:

(222)
a.Marie a tout cassé
  Mary has everything broken
  Mary broke everything

b.Pierre n'a rien cassé
  Peter has nothing understood
  Peter broke nothing

A question immediately arises with respect to examples like (222) where a quantifier which is the object of the verb appears in a preverbal position: what is the position to which these bare quantifiers move: is it a [+HR] position or a [-HR] position? To begin answering this question, let us first consider the nature of the other positions in which these bare quantifiers can occur. As is shown by the examples in (223), *tout* and *rien* behave in other respects like regular objects: they can undergo passive.

(223)
a. Tout a été cassé
   Everything was broken

b. Rien n'a été changé
   Nothing was changed

(223) also shows clearly that French bare quantifiers can occur in [+HR] positions; in this particular case, they occupy the SPEC of AGRPs. When they occur in this position, they behave like regular NP subjects: they can be the antecedent of anaphors as shown in (224), they can undergo clitic inversion as shown in (225), and they can undergo subject raising as shown in (226):

(224)
a.Tout se transforme mais rien ne se modifie
  Everything transforms itself but nothing modifies itself

b.Tout s'autodétruira sur commande
  Everything will autodestruct on command

(225)
Tout-a-t-il été bien rangé?
Was everything well tidied up

(226)
a.Tout semble avoir bien changé
  Everything seems to have changed a lot

b.Rien ne semble l'affecter
  Nothing seems to affect him

95

From (223), (224), and (225), we conclude that bare quantifiers of the type of *tout* and *rien* can occur in [+HR] positions. (223) and (226) also show that bare quantifiers can undergo NP movement. As shown in (227), moreover, the raising of a bare quantifier is subject to the same restriction as the raising of a regular NP: it is local and does not admit super-raising.

(227)
*Tout semble qu'il est susceptible d'avoir changé
 Everything seems that it is likely to have changed

Given that bare quantifiers can occur in [+HR] positions and can undergo NP-movement (movement which creates typical AN-chains), it seems plausible to suggest that in examples of the type in (222), bare quantifier also occurs in a [+ HR] position, namely in the Spec of the AGR-O projection, which (following Kayne (1989)) we have assumed to

---

95. Although an overt dative NP as a complement of *sembler* ("seem") is not very felicitous in French, if we disregard its marginality we can observe that in examples of the type given in (i), the binding of the pronoun by the raised quantifier is the most natural reading. This shows that the raising of a quantifier does not induce WCO. From this fact, we deduce that the chain created by the raising of a bare quantifier has characteristic properties of AN-chains:

    (i)
    ??[Tout$_i$ a semblé à son$_i$ propriétaire [t$_i$ avoir beaucoup changé]]
    Everything seemed to his boss to have changed much

dominate a past participle in French (see section 2.4.2). Under this view, the structure of sentences like (222) will be as follows:

(228) Marie a [$_{AGRP}$-o tout$_i$ [AGR-o [cassé t_i]]]

Note that under this view, we do not need to assume that a bare quantifier can occur both in adjoined positions ([-HR] positions[96]) and in [+HR] positions. We can generally assume that bare quantifiers are restricted to [+HR] positions at S-structure, just like floating quantifiers.[97]

An obvious alternative analysis to the one proposed in (228) would be to assume that in (222) bare quantifiers are adjoined to VP. There are, however, two indirect arguments which suggest that, although plausible, this view may not be the best hypothesis.

---

96. Note that bare quantifiers cannot occur in a position adjoined to S, as shown by (i):
(i)
a. *Tout, elle l'a appris
   Everything, she learned it
b. *Je crois que tout elle veut savoir
   I believe that everything she wants to know

   This impossibility could be straightforwardly explained if some principle disallowed S-structure adjunction for quantifiers in general.

97. The reason for this restricion remains unclear and we do not have any suggestion to make for the moment. Note, however, that if this view is adopted, then the existence of the "leftward *tous*" phenomenon in French and its absence in English simply follows from properties of verb movement in both languages. Leftward *tous* is possible in French because S-structure movement through the Spec Agr-O is made possible by the movement of the verb. In English, on the other hand, movement through spec Agr-O is not possible because since the verb does not raised VP reamins a D-barrier.

First, note that when a bare quantifier undergoes passive or NP movement in general, we must assume that it does not adjoin to VP on its way to the Spec of AGRP-S. Adjunction to VP would in fact preclude further movement to the Spec of AGRP-S, since movement from an adjoined position (a [-HR] position) to a [+HR] position is generally ruled out by the condition on improper movement. Recall that in section 2.4.2, we argued that since past participle agreement is obligatory in French passive, passage through the Spec of AGR-O is obligatory. If so, it must be the case that a bare quantifier can at least land in the Spec of AGR-O on its way to the Spec of AGRP-S. If bare quantifiers must be assumed to be able to pass through the Spec of AGR-O, there is no clear reason why they should not be permitted to stay there, assuming, as is plausible, that they do not need to occur in a Case position, maybe because of their operator status.[98] Thus, we have at least two cases which suggest that bare quantifiers can occur in [+HR] postions. There is thus no reason to assume a priori

---

98. Note that this does not mean that we expect quantifierss to be able to occur in all non-Case-marked positions. Although the quantifier itself does not seem to need to be in a Case-marked position, it must be associated with a Case-position. In other words, bare quantifiers seem to be able to head two types of chains: 1) chains in which the head is Case-marked but the foot is not, and 2) chains in which the foot is Case-marked but the head is not. Chains which do not get any Case at all are generally excluded, so a case of passive in infinitivals like (i) is impossible:

(i)
Il semble tout avoir été changé
It seems everything to have been changed

In (i), the Case requirement is satisfied neither by the head of the chain nor by the foot, so the structure is ill-formed, presumably as a θ-criterion violation. The object θ-role is never visible.

that bare quantifiers can also occur in adjoined positions, i.e., [-HR] positions.

Second, recall that it has been argued by Pollock (1989), on the basis of examples of the type in (229), that the French past participle can move to the functional projection AGR-O.

(229)
a.Pierre a presque mis fin au conflit
  Pierre has almost put an end to the conflict

b.Pierre a mis presque fin au conflit
  Pierre has put almost an end to the conflict

c.Pierre a à peine vu Marie
  Pierre has hardly seen Marie

d.Pierre a vu à peine Marie
  Pierre has seen hardly Marie

Pollock (1989) argues that adverbs of the type *à peine* ("hardly") and *presque* ("almost") are strictly VP-initial adverbs in French. The fact that these adverbs can occur on both sides of the past participle shows that the past participle must have moved over the adverb in examples like (229)b. and d.

Now, assuming that bare quantifiers can adjoin to VP, we would expect them to occur, like the adverbs in (229), on either side of a past participle. They should occur on the left of the past participle when it remains in its D-structure position inside the VP and on the right of the participle when it moves to the functional head which dominates it. The two strucutres are given in (230):

(230)

a. $[_{AGRP\text{-}O}$ .....$[_{VP}$ tout$_i$ $[_{VP}$ Past participle t$_i$]]]

b. $[_{AGRP\text{-}O}$ (Past participle)$_V$ $[_{VP}$ tout$_i$ $[_{VP}$ t$_V$ t$_i$]]]

This expectation, however, is not supported by the facts. There is a clear contrast between examples in which the bare quantifier occurs to the right of the participle and examples in which it occurs to the left. The latter are at best marginal and at worst unacceptable:

(230)
a. J'ai tout vu, tout entendu
??J'ai vu tout, entendu tout
I have seen everything, heard everything

b. Je n'ai rien vu, rien entendu
* Je n'ai vu rien, entendu rien
I did not see anything, hear anything

99

The fact that the bare quantifiers are always better when they occur to the left of the past-participle is somewhat problematic for the VP-adjoined hypothesis. It can, however, be straightforwardly accounted for if it is assumed that bare quantifiers occur in the Spec of AGR-O; under this view, even if the past participle raises to AGR-O the bare quantifier will always precede the past participle. The question why bare quantifiers do not appear in post-participial positions can now be reduced to the more general question of why quantifiers which are not attached to overt NP heads cannot remain in deep-structure object

---

99. There is a clear difference between the case with *rien*, which is clearly ungrammatical, and the case with *tout*, which, although pretty bad when unstressed, is not completely unacceptable (see Pollock (1989) for an analysis of *rien*). Note that both types of examples get better when the bare quantifier is made "heavy":
(i)
a. Il a dit absolument tout ce que tu voulais
He said absolutely everything you wanted
b. Elle n'a mangé absolument rien de ce que vous lui avez preparé
She ate absolutely nothing of that you has prepared for her.
We will not be concerned with these questions, which only a more thorough analysis of bare quantifiers could begin to answer.

position.[100] Although we do not address this question, it seems to us highly desirable that the marginality of all the examples in the previous footnote be reduced to a single principle, whatever the principle. Under a VP-adjoined analysis, we need to rule out at least two structures: one in which the Q is in base position, and a second in which the Q is adjoined to the VP and the past participle has moved over it to the head of the AGR-O projection.

Although none of the arguments we have given above for an analysis of sentences like (222) in terms of the structure in (228) is compelling, the facts we have discussed seem to favor the hypothesis that a bare quantifier like *tout* occurs in the spec of AGR-O. Note that this

---

100. Cf.the following examples:
(i)
a. Tous ont été arrêtés
All were arrested
b.*Ils ont arrêté tous
They arrested all
(ii)
a. Je les ai tous appelés
I called them all
b. ??Je les ai appelés tous
(iii)
a. Ils ont tous dormi
The have all slept
b. ??Ils ont dormi tous
To account for this descriptive generalization, Pesetsky (1988 class notes) has proposed that a floated quantifier cannot occur in a θ-marked position. Note that this idea also prohibits an FQ from ever occuring in the VP internal subject position. This proposal in fact causes no problem wrt Sportiche's analysis since, the position of an FQ can always be that of the Spec of an intermediate functional projection. Under this view, the fact that English quantifiers such as *everything* occur in object position may be due to the fact that they are not "bare", but have the noun *thing* incorporated.

hypothesis also makes bare quantifiers more similar to floating quantifiers, with which they share a number a properties.[101,102]

---

101. See Kayne (1975), (1983) for details.

102. Note also that the hypothesis that bare quantifiers occur in [+HR] positions gets us a step closer to the plausible hypothesis that all floating and displaced quantifiers in French occur in [+HR] positions at S-structure (with the possible exception of *beaucoup* ("many")). The assumption that even leftward-*tous*-floated quantifiers occur in [+HR] positions (at least in non-marginal cases) may provide the beginning of an explanation for why something like leftward-*tous* does not occur in English. French and English have been argued by Pollock (1989) to differ with respect to both the movement of the main verb and the movement of the past participle. If leftward-*tous*-floated quantifiers, like rightward *tous*, involved movement to a [+HR] position, it could be ruled out in English because of the lack of verb movement. Recall that we have assumed that VP is generally a barrier for the object when the V does not move (see section 3.3.5.1). Since the verb never moves out of the VP in English, the VP always remains a barrier for the object. Escaping from this barrier is only possible through adjunction, and further movement to a [+HR] position will be ruled out by the condition on improper movement. Under this view, Q movement in English is barred for the same reason that object shift is barred. In French, on the other hand, since object shift is ruled out not because of the barrierhood of VP but because of the lack of Case assignment by AGR-O, Q movement is expected. This view, although plausible for English and French, seems to run into trouble with other Romance languages such as Spanish, Catalan, and Italian: leftward-*tous* is apparently prohibited in these Romance languages, even though they have been argued to manifest verb movement. As a plausible solution to this problem, we would speculate that leftward-*tous* is in fact possible in other Romance languages but that its effects are not overtly visible due to the movement of the past participle along with the auxiliary verb. The resulting structure is comparable to examples like (i) in French, where the effects of leftward-*tous* have been overridden by the effects of the movement of the verb to a higher functional projection:

(i) [AGRP-S Je [AGR-S les verrai$_i$[TP [AGRP-O tous$_j$ [VP t_i t_j]]]]]

We develop these ideas in forthcoming work.

Now recall that we have argued on the basis of examples like (231) that the Spec of AGR-O in French is not a position to which Case is assigned, that is, it is a [+HR -Case] position.

(231)
a. *Jean a la pomme mangée
Jean has the apple eaten

b. *Il a été un homme arrêté
There was a men arrested

Thus, according to the typology of positions we have outlined above, we would expect the chain created by the movement of bare quantifiers to exhibit both AN-chain properties and Vbl-chain properties. As shown in (232), a pronominal can have a bound reading when it is c-commanded by *tout* in the Spec of AGR-O:

(232)
J'ai tout rangé dans sa boite
I put everything back in its box

Note, however, that a pronominal in a locative PP can also be bound by a quantifier that has remained in its D-structure object position, as shown in (233).

(233)
J'ai rangé chaque instrument dans sa boite.
I put each instrument back in its box.

(233) shows that WCO effects do not obtain even when the NP occurs in its base position. This suggests that the bound reading of the pronominal in (232) may be due to the position of the trace of *tout*, and not to the surface position of *tout*. Similar effects obtain with various other kinds of PPs, so it is unfortunately not possible to construct examples which would show whether the chain created by the movement of *tout* can manifest AN-chain properties.

Turning now to Vbl-chain properties, we observe a contrast between a moved bare quantifier and a quantified NP which has remained in object position: as shown in (234), a moved quantifier can license a gap in an adjunct clause, but an in-situ quantified NP cannot.

(234)
a. *J'ai noté tous les exercices sans relire
I graded all the exercises without re-reading

b. J'ai tout noté sans relire
I graded everything without re-reading

c. *J'ai vérifié tous ces colis avant d'expédier
I checked all these parcels before sending

d. J'ai tout vérifié avant d'expédier
I checked everything before sending

103

Examples (234)b. and (234)c. are similar to cases of gaps, such as those in (235), which are licensed by WH-movement:

(235)
a. Quels exercices as-tu noté sans relire
Which exercises did you grade without re-reading

b. Qu'as-tu acheté sans bien inspecter
What did you buy without checking well

The examples in (235) have the structure of classic parasitic gaps.The examples in (234)b.and c., on the other hand, have a structure parallel to the cases of German parasitic gaps, which are induced by

---

103. As shown in (i), a gap cannot be licensed by a cliticized object, even when the clitic is paired with a floated quantifier:

(i)
a. J'ai tout acheté sans bien inspecter
I bought everything without inspecting well
b. *Je les ai achetés sans inspecter
I bought them without inspecting
c. *Je les ai tous achetés sans inspecter
I bought them all without inspecting

the scrambling of a definite object. Thus, (234)b. and c. show that the chain created by the movement of bare quantifiers in French has Vbl-properties in that it licenses parasitic gaps. Given our typology of positions, this is expected if, as we suggest, bare quantifiers move to the Spec of AGR-O, a [+HR -Case] position. Under this view, the structure of examples like (234) b. is as follows:

(236) [J'ai [AGRP-O $tout_i$ [AGR-O $notée_v$ [[VP t_v t_i] [PP sans PRO relire e_i]]]]]

We assume that in (236) the adjunct PP is adjoined to the VP. The quantifier *tout* c-commands both its own trace t_i and the parasitic gap e, which is licensed by the chain (*tout*, t_i), while neither the parasitic gap nor the real gap c-commands the other. Consider now the sentences in (237):

(237)
a. *Tout a été rangé par Marie pour bien protéger
Everything was put away by Marie to protect well

b. *Tout a été noté par le professeur sans relire
Everything was graded by the professor without re-reading

c. *Tout a été classifié pour bien conserver
Everything was classified to keep well

In (237) the bare quantifier has moved under Passive to the Spec of AGRP-S. The structure of (237) is thus as follows:

(238) [AGRP-S $Tout_i$ [AGR-S a [TP [AGRP-O t''_i été [[VP t'_i rangé t_i par Marie] [PP pour [PRO bien proteger e_i]]]]]]]

In (238), the bare quantifier c-commands its own trace within the VP and the gap in the adjunct, neither of which c-commands the other.

Both the PG and the original trace are c-commanded by the intermediate trace of *tout* in the Spec of AGR-O, but we expect this trace not to rule out the PG, since it is in a mixed position.[104] Thus, we expect that a parasitic gap be licensed just as in the previous cases in (236). As shown by the ungrammaticality of (237), however, parasitic gaps cannot be licensed. Note that we cannot simply assume that the parasitic gap is excluded because it is c-commanded by an element in a [+HR] position. If we are correct in our analysis of bare quantifier movement in French, this is also the case in (236), where the parasitic gap is licensed. The main difference between the structure in (238) and the structure in (236) is that in the latter, the bare quantifier has moved to a [+HR +Case] position, a position which, in our view, heads chains with strict AN-chain properties. Thus a parasitic gap is not licenced.

Let us summarize: we have argued that the S-structure movement of French bare quantifiers is to a [+HR] position. If this is correct, we further observe that when bare quantifiers occur in [+HR, -Case] positions, they can license parastic gaps. When, on the other hand, they occur in [+HR +Case] positions, parasitic gaps are impossible. If this analysis is tenable, it provides some support for the typology of positions we have posited in the previous section, and in particular for the idea that [± Case] properties of positions are relevant to a further partition of [+HR] positions.

---

104. Note that this trace does not license the PG, either, which indicates that the properties of a chain (except for improper movement) must be assessed on the maximal chain, and not on segments of chains.

3.4.7 Conclusion of Chapter III

This chapter was concerned with the movement of objects (direct and indirect)[105] to intermediate positions in the sentence. We have shown that throughout a number of significantly different languages--Hindi, Japanese, Mainland Scandinavian, Icelandic, German, and Dutch--this type of movement presents striking similarities: it does not induce WCO effects; it permits WCO repair;, it allows new anaphoric binding; it is compatible with quantifier float; and it manifest reconstruction properties which are typical of more classic cases of AN-chains. To account for these similarities, we have proposed that in all cases these instances of movement be analyzed as movement to specifiers of functional projections. As we argued in the previous chapter, movement to the specifier of functional projections is, in our terms, movement to a [+HR] position, and chains headed by [+HR] positions have AN-chain properties. Thus, the analysis we propose accounts straightforwardly for the properties of chains created by object movement. Given that our analysis is based on the assumption that the basic structure of sentences contains a number of functional projections, it success provides strong support for a model of UG which includes the "split-INFL hypothesis".

---

105. The notion "object" must be extended somehow so as to include the subject of a small clause or of an ECM construction.

In section 3.3, we showed that in the Scandinavian languages the possiblity for object movement is constrained by the possibility for verb movement. We proposed that this restriction is valid cross-linguistically, and that it follows from the ECP, given the principle of Dynamic Minimality which we defined in section 3.3.5.1. The cross-linguistic validity of this hypothesis is amply supported in the languages we have discussed: evidence for verb movement in Hindi has been given by Mahajan (1989), and in Japanese, morphological affixation of various elements strongly suggests that verb movement occurs.[106] We gave evidence for the relationship between the movement of the verb and the movement of objects in Mainland Scandinavian and Icelandic. Finally, following Den Besten (1986) and Schwartz and Vikner (1989) (among others), we argued for the movement of verbs to

---

106. There is more potential evidence for the movement of the verb in Japanese: it is a well-known fact that no adverb can occur to the right of the verb in Japanese. One could suggest that the absence of rightward adjunction follows from the SVO status of the language, relating it to the head-final parameter. Although this may be correct for adjunction derived by movement (SVO languages do not exhibit rightward adjunction of moved constituents), this seems incorrect for base-generated adjuncts such as VP adverbs. As we noted before, in head-initial languages VP adverbs (but not sentential adverbs) may adjoin either to the left or to the right of VP. Note, however, that the absence of apparent rightward adjunction to VP in Japanese would follow straightforwardly from the assumption that the verb always moves to a functional projection higher than VP. Under this view, an adverb adjoined to the right of VP at D-structure would always occur to the left of the verb at S-structure. A third potential argument for V-movement in Japanese comes from the behavior of adjectives vs. verbs. In certain negative copular sentences, an adverb can occur between the adjective and the negation morpheme and the tense marker morpheme, the latter two being morphologically attached together. The structure is ADJ Adv NEG+Tense . Such a structure is never possible with a verb, suggesting that verb raising is necessary while in copular sentences there may be raising of an empty copula (thanks to Akiro for pointing this out to me).

inflectional projections in German and Dutch. Thus, the dependence of the movement of objects on the movement of main verbs is exemplified in all the languages we discussed. The principle of Dynamic Minimality we proposed also makes a clear prediction with respect to languages which do not exhibit any movement of main verbs: these languages should not allow any movement of objects to the specifier of a functional projection. This prediction is straighforwardly confirmed by English.

As we have seen, there are variations among languages with respect to the types of objects which can undergo movement to the Spec of functional projections. While Hindi, Japanese, Icelandic, German, and Dutch allow all types of object movement, object movement is limited to pronouns in Mainland Scandinavian and to bare quantifiers in French. We have argued that these restrictions follow from the Case-assigning properties of functional projections in various languages. For instance, we have assumed that in Icelandic the projection of AGR-O is a Case-assigner: consequently, any type of NP can move to the Spec of AGR-O. In French, on the other hand, AGR-O is not a Case-assigner; thus, only operator-like elements may move to the Spec of AGR-O. Similarly, we have proposed that in mainland Scandinavian the AGR-O projection is not a Case-assigner; the movement of pronouns in this language is, in our view, an instance of movement to the head of AGR-O. Movement to the head of a functional projection is, under our

definition, also an instance of movement to a [+HR] position and is thus expected to manifest AN-chain properties.[107]

Finally, we discussed some cases of movement to the Spec of functional projections which appear to create chains manifesting both AN-chain properties and Vbl-chain properties. If this is correct, it suggests that a binary distinction among positions is insufficient and that a ternary distinction of positions may be necessary. We speculated that these movements involve cases of NP movement to a [+HR -Case] position and that chains headed by [+HR -Case] positions can have mixed properties.

The table in ($1) summarizes the results of this chapter. The asterisks in the table represent chains with mixed properties. The column labeled "others" represents pronoun movement (in MS and other Scandinavian and Germanic languages) and bare quantifier movement (in French), i.e., morphologically Case-marked elements and operator-like elements.

---

107. That movement to a head essentially creates chains with AN-properties is shown by the properties of the Romance clitic chains which, following Kayne (1987), we assume to involve movement to a head.

($N1)

	Object movement to [+HR]		Main Verb Movement
	NP	others	
Hindi	Yes	Yes	Yes
Japanese	Yes	Yes	Yes
MS	No	Yes	Yes
Icelandic	Yes	Yes	Yes
German	Yes*	Yes*	Yes
Dutch	Yes*	Yes*	Yes
French	No	Yes*	Yes
English	No	No	No

To recapitulate, we have arrived at the following conditions on object movement to the specifier of a functional projection:

In a language L, the movement of an object α to a [+HR] position is possible iff:

(i) main verbs in L move to some functional projection (cf. Dynamic Minimality) and:

if α = NP then either

(ii) the [+HR] position is a [+Case] position

or (iii) [+HR] positions are unidirectionally oriented wrt the head in L

otherwise

α = [±maximal -projected]

or

α = [+operator]

Chapter 4

The Spec of COMP

4.1 The problem

4.1.0.1 Introduction

In chapter 2, we gave arguments for recasting the LGB A/A' dichotomy in terms of the more general opposition between [+Head Related] ([-HR]) and [-Head Related] ([-HR]) positions. Recall that as we defined them, [+HR] positions are the specifiers and complements of heads, while [-HR] positions are positions adjoined to maximal projections. This leads us to a very simple distinction among positions and types of chains: we have suggested that [+HR] positions can head chains with AN-properties, while [-HR] positions can head chains with Vbl-properties. In the last section of Chapter 3, we were led to question this simple binary partition of positions and chains and suggested a further subdivision of [+HR] positions into [+HR +Case] positions and [+HR -Case] positions. The resulting types of positions and the type of chains each licenses are summarized in (1):

(1)

[+HR +Case] positions head chains with AN-properties

[+HR -Case] positions head chains with mixed AN- and Vbl-properties

[-HR] positions head chains with Vbl-properties

The discussion in chapters 2 and 3 concentrated on the status of positions related to functional projections which can be informally described as part of the inflectional system of the verb, namely the projections of Tense and Agreement. In these chapters we were concerned essentially with NP movement through and to these positions. One question which has so far been intentionally left aside is that of the status of the Complementizer projection, more precisely the status of the specifier of the CP projection and the properties of movement to it. This is the topic of the present chapter.

Whether the complementizer position is a position endocentric or exocentric to the head of the sentential projection has been debated since the origin of X' theory. Until quite recently in the generative literature, the use of the labels S' and S reflected a certain hesitation over the issue. Stowell (1981) argued that both the category INFL and the complementizer should be assumed to head their own projections. More recently, Chomsky (1986) proposed that both the INFL projection and the complementizer should be integrated to the X' system and assumed to head their own functional projections, namely IP and CP. As noted by Chomsky, this proposal is in fact the null hypothesis, and we will therefore assume it without any further discussion. Note, however, that given the typology of positions we have proposed, if C is assumed to head its own functional projection, the Spec of CP will be non-distinct from the specifier of any other lexical or functional projection; it will be a [+HR] position. This

implies that movement to the Spec of CP, like movement to the Spec of any other functional projection, is expected to manifest properties of AN-chains.

Since Chomsky's (1986) proposal to integrate the complementizer projection into the X' system, it has been standardly assumed that WH-movement is an instance of movement to the Spec of CP.[1] Now, the Spec of CP of a matrix sentence is a position to which Case is not assigned, under standard assumptions. Given the approach to Case assignment we adopted in Chapter 3 section 3.3.4, the Spec of an embedded CP is also not a position to which Case is assigned.[2] Given the typology in (1), if WH-movement is movement to the Spec of CP it is an instance of movement to a [+HR -Case] position. If so, we predict correctly that WH-movement has Vbl-chain properties. But this is not sufficient, since, as we proposed in Chapter III section 3.4.4, [+HR -Case] positions can license chains which simultaneously exhibit AN-chain properties and Vbl-chain properties. As we showed in Chapter II section 2.3, however, WH-movement in English manifests no properties of AN-chains: it induces WCO effects, it does not permit new binding, and it cannot be associated with a floating quantifier. We repeat the paradigm showing these properties for WH-movement in (2):

---

1. But cf. Lasnik and Saito (forthcoming) for a theory in which WH-movement is movement to COMP and not to the Spec of CP.

2. It has been argued by Kayne (1984) that the Spec, CP of some infinitivals can receive Case. See section 4.6.2 a discussion of these cases.

(2)

a. *Who$_i$ does his$_i$ mother love

b. *Which man did friends of himself congratulate

c. *Which women did each other's friends invite

d. *Which books$_i$ did you all$_i$ read

Thus, the hypothesis that WH-movement is an instance of movement to the Spec of CP, a [+ HR,-Case] position, makes incorrect predictions for the properties of the chain created. The assumption that WH-phrases move to the Spec of CP is therefore incompatible with the theoretical model we have developed so far.

Chomsky (1986)b. distinguishes two types of movement: substitution and adjunction. In this framework, NP-movement and WH-movement are distinguished as involving respectively movement to an A-position and movement to an A'-position (in the LGB sense), but they are also assumed to be similar, since they are both considered to be substitutions and not adjunctions.

(3) "Substitution will always move a maximal projection to the specifier position. The two major cases will be NP-movement to the subject of IP or NP and WH-movement to the specifier of CP".(Chomsky (1986b) p 4)

In the *Barriers* framework, this similarity between NP-movement and WH-movement follows in part from Chomsky's definition of A'-position, which intersects with the substitution/adjunction dichotomy at one particular point, namely the Spec of CP, and in part from the

assumption that adjunction to CP is prohibited.[3] The assumption that WH-movement is to the Spec of CP and is therefore a substitution introduces a curious imbalance in the way WH-movement is viewed: it is assumed that WH-movement may proceed through adjunction most of the way, while some intermediate steps (intermediate traces in the Spec of CP) and the final step are substitutions.[4] Other types of A'-movement, such as topicalization, rightward extraposition or Heavy NP shift, do not show this imbalance: they proceed strictly by adjunction.[5] NP-

---

3. Under the standard definition of the A/A' dichotomy, all A'-positions are adjoined positions except for the Spec of CP. In the *Barriers* framework, long-distance WH-movement must proceed through the Spec of CP. This means that some intermediate steps in the movement will be substitutions into the Spec of an intermediate CP, while others will be adjunctions (to VP, for instance). This lack of uniformity would not be forced if adjunction to CP were permitted; but in *Barriers*, adjunction to CP is explicitly excluded.

4. In LGB, WH-movment was considered an adjunction to COMP:

   "Consider WH-movement, which we may assume to be adjunction to COMP in the sense of OB (Chomsky (1980a)). The grammatical function chain (GF_1, GF_2) produced by a single application of WH-movement contains GF_2, necessarily an A-GF, and GF_1, a non-A-GF (A'-GF) that we may denote 'adjunct to COMP'. Assume that there are two types of movement rules: substitution and adjunction, the latter always forming a structure of the form $[_\beta\ \alpha\ \beta]$ or $[_\beta\ \beta\ \alpha]$, where α is adjoined to β by Move α. Then the only GFs are heads, complements, adjuncts and subjects. A principled approach to the theory of GFs, which I will not undertake here, will begin by defining such general notions as 'heads' etc., then defining particular GFs in terms of them". (LGB, p.47). Our view in that sense is a conservative one.

5. We have assumed, following Lasnik and Saito (1989), among others, that topicalization involves adjunction. It has been proposed by Chomsky (1973) that topicalization is movement to a TOP phrase and by Chomsky (in class (1989)) that topicalization is movement to the Spec of CP. If this is correct, then topicalization may also show this imbalance, involving adjunction in intermediate steps and substitution in the final step. There are however reasons to doubt the proposal that topicalization is movement to the Spec CP. Topicalization, contrary to wh-movement does not require AUX-inversion. If both topicalization and wh-movement are movements to

movement does not show an imbalance, either: it is viewed as pure substitution even in its intermediate steps. For NP-movement, indeed, adjunction is explicitly disallowed, by the condition on improper movement[6]

According to our definitions, the Spec of CP is a [+HR] position. The curious imbalance which we noted in the way WH-movement is conceived in the *Barriers* framework becomes an instance of improper movement in our terms: recall that in Chapter 2 section 2.4, we defined improper movement as follows:

(4) *(XP.....t...)
[+HR] [-HR]

If, as assumed in *Barriers*, WH-movement proceeds by successive adjunction to intermediate projections (an assumption which we have tacitly adopted), then movement to the Spec of CP, a [+HR] position, will be an instance of improper movement and will be ruled out. This is a second reason why the assumption that WH-movement involves movement to the Spec of CP is incompatible with the theoretical model we have developed so far.

In this chapter, we address this incompatibility. We will discuss the problems raised for the theory of movement by the hypothesis that the

---

Spec CP it is quite unclear how this fact can be accounted for. In our view, topicalization is distinguished form wh-movement by its landing site. Topicalization is adjucntion to IP and Wh-movement adjunction to CP.

6. Cf. the analysis of Super Raising in *Barriers*

Spec of CP is a [+HR] position, and thus a position a priori non-distinct from the specifiers of other functional projections. The solution we propose to this problem is very simple: WH-movement does not involve substitution for the Spec of CP, but rather adjunction to CP. The discussion in this chapter will lead us to develop a view of WH-movement which involves no mixtures of substitution and adjunction, either in intermediate steps (landing in the Spec of CP in a long-distance extraction) or in the final step. Under this view, adjunction to CP will be allowed (contrary to the *Barriers* framework), and any mixture of substitution and adjunction in a single chain will be ruled out as improper movement. The general picture of movement theory is thus greately simplified: there are essentially two types of chains, [+HR] chains and [-HR] chains, and any mixture of the two types of movement which does not form an independently well formed chain[7] chain and the [-HR] chain) is independently well formed is ruled out as improper movement.

This chapter is organized as follows. First, we discuss the problems posed by the assumption that WH-movement is movement to the Spec of CP for the theory we have proposed. Next, we justify the assumption that WH-movement involves adjunction to CP. We then consider the ECP and propose a revision of the *Barriers* system compatible with the rest of our assumptions. Finally, we consider a number of consequences of the theory we propose and suggest a new approach to Subjacency.

---

7. In cases of wh-extraction preceded by raising (as in "Who do you think seems to be on top of everything") each chains (the [+HR

4.1.1 Short-distance WH and NP movement

As mentioned in the introduction, the Spec of CP is in our terms a [+HR] position. Given this approach, if WH-movement is movement to the Spec of CP, it is wrongly predicted that WH-movement should have some AN-properties. Restricting ourselves for the moment to main clauses, several solutions to this problem come immediately to mind.

First, as has been proposed by several linguists(and most recently by David Pesetsky), we could assume that root clauses do not have a CP projection. Under this view, WH-movement could be simply conceived of as an adjunction to the projection of highest functional projection in the sentence, AGRP-S in our view. One serious problem with this assumption is that it wrongly predicts that adjoining a WH-element to AGRP-S is an embedded sentence should be possible, contrary to fact:[8]

(5) *Mary said that what John bought?

We will thus not adopt it. Another possible alternative is to assume that main clauses do have a CP projection, but that C, unlike other functional projections, does not project a specifier. As assumed by Chomsky (1986), specifiers are generally optional.[9] If C does not project a specifier, WH-movement will involve adjunction to CP, so

---

8. Pesetsky (1989) proposes that the WH-element moves to the Spec of IP. This option is clearly not available to us, since the Spec of IP is in our view also a [+HR] position.

9. "In general specifiers are optional but the choice of complements is determined by the projection principle" Chomsky (1986) p. 4.

that the structure of a simple interrogative is as follows (we omit details of the internal structure of the clause):

(6) [$_{CP}$ what$_i$ [$_{CP}$ did [$_{AGR\text{-}S}$ John buy t_i]]]][10]

Assuming this to be correct, the next question which arises is whether it is ever possible for the C projection to have a specifier or whether we should assume that the absence of a specifier is simply a lexical property of C. If we confine ourselves to English, evidence apparently favors the second option, since it is clear that a sentence in which both an NP and a WH-element occur at the begining of a sentence is strictly ungrammatical and must be excluded.

(7) *[$_{CP}$ How$_k$ [$_{CP}$ Mary$_i$ [$_C$ did [$_{AGRP\text{-}S}$ t_i understand that t_k]]]]

But he ungrammaticality of such sentences is clearly not universal: interrogative sentences where both a WH element and an NP precede a verb raised into C are common in French. Consider the following example:

(8) Comment Marie-a-t-elle compris cela?
How Marie has-she understood that
How did Marie understand that

Kayne (1984, Chapter 10) argues convincingly that in sentences of the type in (8), which he calls Complex-Inversion structures (henceforth CI), both the subject NP and the auxiliary with the clitic have undergone fronting to a category which Kayne leaves unspecified, but which it is natural to assume to be CP. In other words, in CI

---

10. We ignore for the moment the motivation for movement of the AUX, which we assume moves to the head of C, as is standard. See footnote (xx) for a discussion of this question.

constructions, neither the subject nor the AUX occurs in the position in which it occurs in ordinary declarative sentences. Evidence for this hypothesis are based on the observation that the CI contruction has the same distribution as the AUX-inversion construction of English. Like English AUX-inversion, CI does not occur in embedded sentences and is in complementary distribution with an overt complementizer:

(9) *Je me demande comment Marie a-t-elle compris cela
I wonder how Marie has-she understood that
*I wonder how has Mary understood that

(10) a. *Si Marie eut-elle compris cela...
If Marie had-she understood that...
*If had Mary understood that...

b. Marie eut-elle compris cela...
Marie had-she understood that...
Had Mary understood that....

A root vs embedded dinstinction and a complementary distribution with an overt C are standard arguments for the movement of V to C. We thus assume that the V occurs in C in (8). If so it is clear that the preceding NP does not occupy the spec of AGR-S. A natural assumption is that this NP occupies the Spec of CP. If so[11], the structure of (8) is the following:

(11) [$_{CP}$ Comment$_k$ [$_{CP}$ Marie$_i$ [$_{C'}$ a-t-elle [$_{AGRP-S}$ t_i compris cela t_k]]]]

Such a structure is expected to be possible in our theoretical model: the NP is in the Spec of CP, a [+HR] position, and the WH-phrase is

---

11. See Deprez (1989) for a detailed analysis of this constuction within a model adopting the Split INFL hypothesis and the VP-internal-subject hypothesis

adjoined to CP. A plausible alternative would be to assume that the NP in (8) is adjoied to CP. But there is evidence agaisnt this alternative. As we argued in chapter 3 section 3.4.6, the bare quantifier *tout* in French occurs only in [+HR] positions. Notably, it can never occur in a position adjoined to S (or S'), as shown in (12):

(12)
*Tout, Marie a emporté
Everything, Marie took

But as shown by (13), the bare quantifier can occur in CI constructions .

(13) Comment tout est-il parvenu à sa destination?
How everything did-it arrive at its destination?

From the contrast between (12) and (13), we conclude that the bare quantifier is not in an adjoined position in the CI construction. Rather it is in a [+HR] position, namely, the Spec of CP. This shows that the position between the WH-element and the V raised to C is a [+HR] position, as predicted in our theoretical model. We take the grammaticality of (8) and (13) to indicate that in some languages, the co-occurrence of a WH-element adjoined to a CP projection and an NP in the specifier of CP is allowed. Thus CP adjunction of a WH-element in matrix clauses must be assumed to be possible.

Just as there is evidence supporting the idea that the subject NP in CI must be in a [+HR] position, there is also evidence supporting the assumption that the WH-phrase in CI is in an adjoined position. Consider the following pair of sentences:

(14)
a. Que fais-tu?
   What do-you
   What do you do?

b.??Que Marie fait-elle
   What Marie does-she
   What does Marie do?

As argued by Hirshbuhler(1980), the French WH-element *que* ("what") is a clitic form which alternates with a non-clitic form, namely *quoi*. As observed by Hirshbuhler, among others, the two alternative forms *quoi* and *que* bear a striking resemblance to the clitic and non-clitic forms of French pronouns. The French pronouns present the following paradigm:

(15)	Clitic form	Non-clitic form
1prs. sg.	me	moi
2prs. sg.	te	toi
3prs. sg.(refl.)	se	soi

If *que* is a clitic, it is expected to occur obligatorily attached to a verbal head at least at the level of PF. Given this expectation, the marginality of (14)b. can be explained straightforwardly. In both (14)a. and (14)b. we can assume that the WH-word *que* is adjoined to CP at S-structure. In (14)a., it can cliticize at PF to the verb in C, so the sentence is acceptable. In (14)b., however, *que* is not adjacent to the V in C, and assuming, as is standard, that PF-cliticization requires adjacency, *que* cannot cliticize to the V. It thus remains in

an adjoined position, an impossible position for a clitic, since it is associated with sentence-initial stress.[12]

On the basis of the evidence provided by the French CI construction, we propose that WH-movement generally involves movement to a position adjoined to CP, at least in main clauses. Given this proposal, the question arises how (7) should be excluded. We could assume that it is a lexical property of the complementizer in English that it generally precludes the projection of a specifier. As we will show, however, such a stipulation is not necessary, and (7) can in fact be excluded by general principles. To exclude (7) in English, we suggest that an AUX or a verb in C is generally incapable of assigning Case to the NP in its Spec. Consequently, the NP in the Spec of CP in (7) violates the Case filter, more precisely the principle which requires the head of a [+HR] chain to be Case-marked (see Chomsky (1986)a.). In French, however, the overt appearance of the postverbal clitic, which we assumed to be a "reinforced" agreement marker, will confer on the V in C the capacity to Case-mark the NP in the Spec of C. This proposal could be understood as follows: assume that the movement of AUX or V to C is an adjunction to an empty C, which, as is standardly assumed,

---

12. The question arises why the alternative form *quoi* cannot occur in this position. Note that there are contexts in which this alternation is possible:

(i) Que faire? / Quoi faire?

The main distinction between (i) and (14)a is the tense properties of the verb. The absence of alternation in sentences like (14)a. is thus due to an independent factor , namely the fact that "quoi" is incompatible with tensed sentences. (See Koopman (1980) for an account of the distribution of quoi).

is not a Case-marking head. Under this view, even if AUX or V contains an agreement marker, it will not be able to assign Case to an NP in its Spec. In French, we propose that the adjunction of the clitic to the verb "restores" the Case-marking capacity of the verb. Consequently, the subject can be directly Case-marked in the Spec of CP.

Since a structure such as (7) is now excluded on a principled basis, we can assume that in English, just as in French, the Spec of CP is optionally present and WH-movement involves adjunction to CP.[13]

---

13. One of the arguments given by Chomsky (1986) for the movement of WH-phrases to the Spec of CP is that it gives a reason why WH-phrases do not adjoin to any other projection. In his view, a WH-complementizer contains the feature [+WH] which is then transmitted to the WH-phrase under Spec-head agreement. Given an approach to WH-movement which assumes adjunction rather than substitution, one may wonder why it is that WH-phrases cannot adjoin just anywhere. The first reason that can be given is that WH-phrases need to have scope over the whole sentence. Although this is certainly correct, one may wonder why adjunction to IP, for instance, does not satisfy this scope requirement and why it is that WH-phrases must occur adjoined specifically to CP. A possible answer lies in the observation that adjuncts seem to be generally restricted with respect to the category they can be adjoined to at S-structure. Thus, for instance, as clearly shown by Jackendoff (1972), VP adverbs cannot be adjoined just anywhere, but must be adjoined to some projection of the VP. The same holds true essentially for sentential adverbs, which, we have proposed to be adjoined to the TP projection, following Rizzi (1989). These restrictions suggest that there must be some relation which holds between the head of a given projection and elements which can be adjoined at S-structure to the maximal projection of this head. Whatever the formal status of this relation, whether it is conceived of as feature sharing, as suggested by Travis (1988), or as some form of predication or theta-marking (Higginbotham), we may assume that the same relation must obtain between a WH-Comp and a WH-phrase, and perhaps also between a Topic and the head of AGRP-S. This would straightforwardly account for the fact that neither a topic nor a WH-phrase can adjoin to VP.

A plausible alternative analysis of CI which we have not yet explored would be to assume that the structure of CI involves a recursive CP structure. There is evidence against this hypothesis, however, based on the obligatoriness of V movement to C. As shown in (16), V-to-C is obligatory in matrix interrogatives in standard French.

(16)
a. *Comment Marie est partie?
How Marie has left?

b. *Que Marie a fait?
What Marie has done?

Whatever the motivation for this fact, we observe that if the CI construction involves a recursive CP structure, we expect the verb to be able to move to the higher C position. Characteristically, however, this is not possible, as shown by (17).

(17)
a. **Comment est-elle Marie partie?
How has-she Marie left?

b. [$_{CP1}$ Comment$_k$ [$_{C'1}$ est-elle$_j$ [$_{CP2}$ Marie$_i$ [$_{C'2}$ t_j [$_{IP}$ t_i [$_{I'}$ t_j [partie t_k]]]]]]] (i=j)

(17)a., with the structure in (17)b., will be trivially excluded if we assume, following Kayne (1984), that CP recursion is impossible. [14]

---

14. Although (17)a. is accounted for under the assumption that CP recursion is impossible with the structure in (17)b., the question arises why (17)a. would be impossible with a structure in which the subject NP has remained in place, that is, with a structure such as the one in (18):

(18) [$_{CP}$ Comment [$_{CP}$ [$_{C'}$ est -elle [$_{IP}$ Marie partie]]]]

The ungrammaticality of sentences with a structure as in (18) remains unexplained in Kayne's analysis. He speculates that for some reason the NP in the Spec of IP cannot be assigned Case; we believe this speculation to be on the right track. It can be made

As we have shown, an at first seemingly plausible analysis of CI in terms of CP recursion is in fact rather problematic. This provides an additional argument that CI in fact involves the adjunction of a WH-phrase to CP. Based on this and previous evidence, we propose that WH-movement in matrix clauses always involves an adjunction to CP. Under this view, WH-movement in main clauses is movement to a [-HR] position, and the properties of the chain created by WH-movement are correctly predicted to be strictly Vbl-chain properties. Moreover, additional NP movement to the Spec of CP is permitted in French and ruled out in English on principled grounds. Assuming that WH-movement involves adjunction to the projection of CP in its final step, the proposal that the Spec of CP is in fact indistinguishable from other specifiers, which follows from our definition of [+HR] positions, causes no particular problem for short-distance extraction in main clauses. In fact, it permits an elegant account of the Complex Inversion structures of French.

Note that the hypothesis that WH-movement involves adjunction to CP in main clauses is compatible with the *Barriers* theory of movement, which we have by and large adopted so far (apart from some modifications of the definition of Dynamic Minimality (see section 3.3.5.2)). In *Barriers* Chomsky proposes that adjunction to CP is banned when CP is a

---

more precise along the following lines: assume that prior to V movement, the clitic is in fact occupying the head of AGR-S and absorbs its Case-assigning properties. If so, a subject NP will not be able to occur in the Spec of AGR-S without violating the Case filter. It will, in fact, have to move to the Spec of CP, the only position in which it can be assigned Case. If this is correct, then (18) is simply excluded by Case theory and nothing more needs to be added.

complement and (following a suggestion by Kyle Johnson) that this ban follows from theta-theory. Given the latter assumption, adjunction to CP is not banned in main clauses, since main clauses are clearly not arguments, and thus cannot be complements. Although the assumption that WH-movement involves adjunction to CP does not raise any problems for simple WH-extractions in main clauses, the situation is quite different in the *Barriers* framework for long-distance extraction out of various embedded clauses. In the *Barriers* framework, the ban on CP adjunction is crucial for an account of CED effects and WH-island violations. We will address this issue in section 4.2.2 and following, and we will propose further modifications to the *Barriers* theory of the ECP. But before we turn to this issue, we will first outline the problems posed by the assumption that the Spec of CP is a [+HR] position for long-distance extraction. There are essentially three cases in which this hypothesis will create problems for the *Barriers* framework:

1) For intermediate traces of long-distance WH-movement
2) For embedded WCO effects
3) For super-raising

We will consider each case in turn.

4.1.2 Long distance WH and NP movement

4.1.2.1 Long-distance WH-movement and intermediate traces

Given our definition of [+HR]/[-HR] positions, intermediate traces of WH-movement in the Spec of CP should be in [+HR] positions. In the *Barriers* framework, it is assumed that WH-elements adjoin successive-cyclically to maximal projections before moving through the Spec of CP. If the Spec of CP is a [+HR] position, a simple case of object extraction from an embedded question will create chains of the type [-HR] [+HR] [-HR]: as we argued in section 2.4, this is a typical case of improper movement. A variable created by the movement of a WH-operator will be [+HR]-bound in the domain of its operator, in violation of Principle C of the Binding theory. Variables are R-expressions, comparable to names, and consequently must remain [+HR]-free. [15] Consider a simple case of object extraction out of an interrogative:

(19) What does Simone think that Jean-Paul will write

Under the *Barriers* theory of WH-movement, the structure of (19) will be as in (20):[16]

---

15. The term and concept of improper movement are due to May (1979).

16. So as to simplify the structures, I will use IP throughout this section to avoid having to spell out every intermediate functional

(20) [CP What does [IP Simone [VP t_1[VP think [CP t_2[C that [IP Jean-Paul [I' will [VP t_3 [VP write t_4]]]]]]]]]]

t_3 is adjoined to VP and is thus in a [-HR] position. t_2, however, is in the Spec of CP, an [+HR] position under our definition. This is thus a case of [-HR] [+HR] [-HR] movement. In this particular type of example, the problem could be avoided by assuming (following a proposal by Lasnik and Saito (1984)) that all intermediate traces of argument WH-extraction delete prior to LF. If, as has been argued by Safir (1982) and Chomsky (1986), Binding theory applies only at LF, (20) will not violate Principle C.[17] But the same improper movement violation will occur with WH-chains and long-distance adjunct extraction for which trace deletion is crucially assumed to be non-applicable. In the latter cases, improper movement cannot be avoided.[18] Given the assumption that the Spec of CP is a [+HR]

---

projections, except when necessary. The label IP is used as shorthand for the sequence AGR-S, TP, AGR-O.

17. Safir argues that Principle C does not apply at S-structure to explain the well-formedness of expletive chains of the type: *There$_i$ is a man$_i$ in the garden.* In his view, the expletive and the postverbal subject are coindexed. This coindexation would induce a principle C violation if Principle C applied at S-structure. In Safir's view, expletive constructions are saved from a principle C violation by the raising of the indefinite NP at LF under QR. But, as argued in Chomsky and Lasnik (1973), other constructions require that Principle C apply before QR. Consider the following example: **He$_i$ likes everyone that John$_i$ knows.* The impossibility of coreference betweeen *he* and *John* follows straightforwardly from Principle C. At LF, however, QR will apply to *everyone* and pied-pipe the RC containing *John*: *everyone that John$_i$ knows he$_i$ likes.* No violation of principle C occurs at LF. Thus, examples of this type provide arguments that principle C must apply before QR.

18. This presupposes that adjunct traces are subject to improper movement violations. See section 4.7.1.1 for arguments that adjuncts are generally subject to principle C. Moreover, in our

position, the grammaticality of long-distance extractions is thus in need of explanation.

4.1.2.2 Weak Crossover

Another problem arises with cases of embedded weak crossover. We showed in Chapter II and III that movement to a [+HR] position does not induce WCO violations: in fact, it permits what we have called WCO repair. Since in our terms movement to the Spec of CP is movement to a [+HR] position, we would expect no WCO violation in embedded sentences. However, as shown by (21), this prediction is incorrect.

(21) * Who$_i$ do you think that his$_i$ mother loves t$_i$

The structure of (21) is given in (22):

(22) [$_{CP}$ Who$_i$ do [$_{IP}$ you [$_{VP}$ think [$_{CP}$ t'$_i$ [that [$_{IP}$ his$_i$ mother loves t$_i$]]]]]]][19]

In (22), t' binds the pronoun *his* from a [+HR] position and is thus expected to license a bound variable interpretation. The fact that WCO violations occur in embedded sentences indicates that the movement through the Spec of CP must not count for WCO repair. This fact will have to be explained.

---

view, improper movement is a requirement on chain homogeneity and thus applies to all chains, independently of the argument/adjunct distinction.

19. We have ommitted the trace of VP adjunction in this example, since it plays no role in this argument.

4.1.2.3 Super-raising

Finally, let us turn to cases of long-distance NP movement. Ungrammatical cases of raising have often been analyzed as involving either an ECP violation or a violation of improper movement.[20] A number of these cases crucially involve the Spec of CP. Consider (23)a., with the derivation given in (23)b.

(23)

a. *John is possible to sleep.

b.[$_{IP}$ John$_i$ is possible [$_{CP}$ t_i [$_{IP}$ t_i to sleep]]]

c. It is possible for John to sleep

As shown in (23)c., the adjectival predicate *possible* subcategorizes for a CP complement and does not assign a θ-role to its subject.[21] In a framework where the Spec of CP is an A'-position, such as in LGB, (14)a. can be ruled out as a case of improper movement. Consider the derivation in (14)b. Chomsky (1986)b. suggests that the trace t'_i in the Spec of CP will force the subject trace t_i to be an R-expression;

---

20. Cf. *Barriers*, pp. 76-77.

21. That the subactegorization for CP is maintained with the choice of an infinitival complement (unlike predicates such as *likely* or *seems*) is particularly clear in French. The infinitival complement of this predicate always has an overt complementizer:

    (i) a. Il est possible de réussir
    b. *Il est possible réussir
    It is possible to succeed

    Cf. Kayne (1984), "On some Differences between French and English" for arguments that *de* is a complementizer.

this will then violate Condition C of the Binding theory, since the R-expression t_i will be A-bound by John$_i$.[22] This explanation is no longer available in a framework where the Spec of CP is a [+HR] position non-distinct from the specifier of any other functional projection. Although NP movement <u>to</u> the Spec of CP is ruled out by the lack of Case-marking ability for C, NP movement <u>through</u> the Spec of CP violates none of the conditions we have discussed so far. Since improper movement cannot be used in our theoretical model to rule out cases such as (14)a., another explanation must be found to rule out cases such as (14) and other cases of super-raising.

We have outlined the three major problems which occur if the Spec of CP is considered to be a [+HR] position, i.e., a position from which AN-chains are licensed. In the brief discussion of each of these problems it is apparent that many solutions are possible and that the choice of a particular solution will be highly dependent on particular choices made for a theory of movement. For instance, we have tacitly

---

22. As mentioned by Epstein (1987), certain sentences have an alternative derivation which does not violate Condition C. Consider the following example: **Who is possible to sleep?* This sentence can have the following derivation: First, the WH-element is moved to the Spec of IP, violating the ECP. Second, the WH-element moves to Spec of the higher CP. Third, from this position it lowers into the Spec of the intermediate CP. Finally, it moves back to the Spec of the higher CP. Under such a derivation, movement always proceeds from an A-position to an A-position and from an A'-position to an A'-position. Improper movement is not violated at any time, and the resulting representation is equivalent to (14)b. Consequently, the ECP is satified. This suggests that the condition on improper movement may be better viewed as a condition on representation. As we suggested in the introduction to this chapter, improper movement is a condition on chain homogeneity. If so, the representation in (14)b. will violate this condition on chain homogeneity on any derivation. See Epstein (1987) for a discussion of this case.

assumed above (Cf. (20)) that an extracted WH-element first adjoins to VP on its way to the Spec of CP. It could be assumed instead that VP is not a barrier.[23] As has often been noted, no empirical facts crucially depend on the assumption that VP is a barrier in Chomsky's (1986)b. framework. Suppose, then, for the sake of the argument, that VP is not a barrier. If so, adjunction to VP would not be necessary and we would have the following derivation for (19):

(24) [$_{CP}$What does [$_{IP}$ Simone think [$_{CP}$ t_1 [$_{C'}$ that [$_{IP}$....t_2]]]]]

In (24), t_2 is directly bound by t_1 in the Spec of CP. But since there are no intermediate adjoined traces, that is, no traces in [-HR] positions, this derivation instantiates a case of [+HR] to [+HR] movement. No improper movement violation arises. The assumption that VP is not a barrier seemingly offers an immediate solution to our first problem. But this apparent solution makes our second problem, the absence of WCO repair, much harder to solve. If long-distance WH-extraction of an object is in its first steps a case of [+HR] to [+HR] movement, WCO should never occur in embedded sentences, contrary to fact.

These brief considerations illustrate that any solution to the three problems we have identified will be crucially dependent on the hypothesis made with respect to the theory of movement and the ECP. Consequently, before we can discuss solutions to our three problems, we will outline the particular theory of ECP that we propose. We turn

---

23. This is assumed by several linguists: Lasnik and Saito (forthcoming) and Kayne (1989), for example.

to this topic in the next section. Although the theory we propose remains close in spirit to the theory outlined by Chomsky (1986)b., it will incorporate some recent proposals by Rizzi(1989); the result will be a simpler approach to the ECP and will extend empirical coverage to some cases which have remained up to now unexplained. In particular, it will allow a straightforward account of the *que/qui* alternation in French.

Our discussion of the ECP will proceed as follows: in section 4.2 we address the question of what categories should be assumed to be potential barriers. In section 4.2.2 we turn to the question of which categories can be adjoined to. Regarding these two questions, we will argue for a simplification of the *Barriers* model; we will remove the "defectiveness" of the IP category and will eliminate the stipulated ban on adjunction to CP. Throughout this discussion, we will use the label IP as shorthand for the succession of functional projections which we have assumed throughout this thesis to be part of the structure of a basic sentence. Following Pollock (1988) and Chomsky (1989), we will assume that the functional projections which occur between AGR-S, (the higher AGR projection) and VP, namely AGR-O and TP, are never inherent barriers. This assumption in fact derives from the definition of L-marking we give in section 4.4. The discussion of the barrierhood status of the IP projection should then be understood as posing the question whether the functional projection which is the complement of C, namely AGR-S, should or not be considered a barrier.

4.2 ECP: a first approach

4.2.1 What counts as a barrier?

We will begin our discussion of the ECP by proposing some revisions of the notion of barrier as it is conceived in the *Barriers* framework of Chomsky (1986). In this section we will argue that IP is not a "defective category" and that VP, although a D-barrier, as we proposed in section 4.3.2 is not an inherent barrier.

4.2.1.1 The barrierhood of IP

Recent revisions of the *Barriers* framework have proposed removing the stipulation that IP is a defective category.[24] For Chomsky (1986), IP is never an inherent barrier or a barrier for Minimality. Certain cases of Head-to-Head movement, however, strongly suggest that IP can sometimes act as a Minimality Barrier. The definition of Minimality barrier is given in (25):

Minimality Barrier

(25) τ is a barrier for β if τ is (a projection, the immediate projection) of δ, a zero-level category distinct from β.

---

24. See Frampton(1989), Fukui and Speas, Uriagereka, and Ambar, among others, for various interesting arguments.

Keeping this definition in mind, let us consider the following example:

(26) [$_{C}$ Have$_{i}$ [$_{IP}$ John$_{j}$[$_{I'}$ will [$_{VP}$ t_{i} been arrested t_{i}]]]]

Under Chomsky's (1986) assumptions, none of the categories dominating any of the intermediate auxiliaries will be barriers for the movement of the NP *John*. The NP must be permitted to cross all these projections on its way to the Spec of IP without ever adjoining, since adjunction would induce a violation of the condition on improper movement.[25] Quite clearly, however, some projection must count as a barrier to prevent the extraction of the lower auxiliary in (26). The ungrammaticality of (26) will naturally be accounted for by the ECP if it is assumed that the projection of the X^{o} which contains *will*, namely INFL under current assumptions (see Emonds, Pollock, Chomsky, and chapter 1 for justification), is a Minimality barrier for the movement of *have* into C. This suggests that INFL is not always transparent for Minimality.

If, as shown by (26), the projection of INFL is not defective for Minimality, the null hypothesis is to consider that IP is not defective in any respect. Such an assumption has been made by a number of linguists, including Fukui and Speas, Frampton (1989), Lasnik and Saito (to appear), and Ambar (1987), among others, and we will in turn

---

25. Any adjunction would be an instance of improper movement; claiming that (26) is ruled out because the NP has adjoined would also incorrectly exclude a sentence like (i), where the NP has presumably followed the same path as in (26):
(i) Mary will have been arrested

adopt it. We will further assume that IP is not L-marked by C, and is thus in most cases a barrier.

In Chomsky (1986), IP, although never an inherent barrier, is a blocking category, and thus crucially cannot be adjoined to. A consequence of the ban on IP-adjunction is that any element extracted from within an embedded IP is forced to land in the Spec of CP. This is forced by the mechanism of inheritance:

(27) τ is a barrier iff (a) or (b):

a. τ immediately dominates δ, δ a BC for β

b. τ is a BC for β

Consider the following structure:

(28) $[_{CP1}$ WH [....V ..$[_{CP2}$ t_1 $[_{C'}$ that $[_{IP}$ NP V $t_2]]]]]$[26]

By assumption, IP is a BC. Since CP2 immediately dominates it, CP2 is a barrier by inheritance for t_2, despite the fact that it is L-marked. Thus, unless the WH-element lands in the Spec of CP2, a barrier will be crossed, leading to either a (weak) Subjacency or ECP violation. We have assumed, contrary to Chomsky, that IP is an ordinary barrier which can be adjoined to. If adjunction to IP is permitted, the following question arises: will CP still inherit barrierhood from IP after adjunction? If not, movement to the Spec of CP will become unnecessary, since CP will not be a barrier. The answer to this question depends on the status of the upper segment of an IP

---

26. We will ignore adjunction to VP, since it is irrelevant to the present point.

projection created by adjunction for the definition of blocking categories. Let us consider Chomsky's definition of a BC.

(29) τ is a BC for β iff τ is not L-Marked and τ dominates β.

Domination is defined as in (30):

(30) α is dominated by β if it is dominated by every segment of β.

Condider an adjunction structure:

(31) $[_{\tau}...\alpha...[_{\tau}..\beta]]$

Under Chomsky's definition, β is dominated by τ but α is not, since there is at least one segment of τ which does not dominate α. Assuming that τ is a BC, it will be a BC for β but not for α. Returning to the case under consideration, we conclude that after adjunction, the upper segment of IP will no longer be a BC. Consequently, if IP is a regular barrier allowing adjunction, landing in the Spec of an intermediate complement CP will never be forced: CP is not a barrier, since it is L-marked by the matrix VP and it does not inherit barrierhood from IP. Movement will always be able to proceed up to the next barrier (VP or IP). Chomsky (1989) has argued for the relevance to syntactic theory of a principle of Economy, which rules that the shortest possible derivation which satifies all UG principles is always the only one possible. If this principle is correct, we are led to the conclusion that landing in the Spec of an embedded complement CP is in fact always impossible: it represents an additional unnecessary step in the derivation, and since there is a shorter derivation which violates no UG principle, landing in the Spec of an intermediate CP will be

impossible. Hence, unless some principle forces landing in the Spec of an embedded complement CP, we never expect it to occur.

Is this a desirable conclusion? It appears to provide an immediate solution to the first two problems we mentioned at the beginning of this section. Clearly, if movement through the Spec of an embedded CP never occurs, the nature of this position with respect to the [+HR]/[-HR] dichotomy will be entirely irrelevant. Therefore, neither the problem of improper movement for intermediate traces nor the problem of WCO repair would ever arise. This conclusion, however, raises other problems. It is well known that many languages show overt effects of successive cyclic extraction on their complementizers. Some languages show obligatory complementizer agreement, for instance Irish, Bantu, Palauan, etc. Other languages have constructions which seem to depend on the presence of intermediate traces in CP. This is the case for French stylistic inversion, as shown by Kayne and Pollock (1978), and to some extent for Spanish inversion, as shown by Torrego.[27] To account for these various phenomena, it appears necessary to assume that long-distance extraction may affect intermediate complementizers.

For this reason, we will assume that even if a WH-element has adjoined to IP, the upper segment of IP will still count as a BC. If so, CP will inherit barrierhood from IP, and either landing in the Spec of CP or adjunction to CP will be necessary.

---

27. See also Kayne, "ECP extensions", for an argument that *combien* extraction must go through the Spec of CP.

Concretely, this assumption involves a slight modification in the definition of BC. We can simplify the definition of BC as follows:

(32) τ is a BC for β iff τ does not exclude β and τ is not L-Marked.

Given (32), any segment of a non-L-marked maximal projection will be a BC. If any BC is automatically a barrier when it is not L-marked, it will no longer be possible to escape a barrier through adjunction. In order to maintain the mechanism of adjunction as a means to escape from a barrier, we must accordingly redefine the notion of barrier as follows:

(33) τ is a barrier for β iff (a) or (b):

a. τ immediately dominates δ,δ a BC for β;

b. τ is a BC for β and τ dominates β

This obtains the desired result. Under the definitions in (32) and (33), CP will be a barrier by inheritance from any segment of IP; adjunction to IP will have no effect on inheritance. Consequently, movement into the Spec of CP or adjunction to CP will be forced even for a trace adjoined to IP, so as to avoid crossing the inherited CP barrier. Adjunction to CP will satisfy antecedent government of a trace adjoined to IP, since the notion of government as defined by Chomsky (1986) refers to the notion of exclusion:

(34) α governs β iff α m-commands β and there is no τ, τ a barrier for β, such that τ <u>excludes</u> α.

Exclusion is defined as follows:

(35) α excludes β if no segment of α dominates β.

Let us consider the structure in question:

(36)[$_{CP1}$...α...[$_{CP}$...σ...[$_{C'}$...[$_{IP1}$...β...[$_{IP}$...]]]]]

In (36), α is adjoined to CP, σ is in the Spec of CP, and β is adjoined to IP. CP is a barrier by inheritance from IP. However, since CP does not exclude σ, σ can antecedent-govern β. Similarly, CP does not exclude α, since α is dominated by CP1, a segment of CP. Consequently, α can antecedent-govern β.

4.2.1.2 The barrierhood of VP

In recent discussions of the *Barriers* framework, the question of the barrierhood of VP has also been controversial. For instance, Fukui and Speas (1986) have proposed that only functional projections can be barriers. Lasnik and Saito (forthcoming) have assumed that VP is not a barrier. Kayne (1987) has proposed that VP is L-marked by INFL. Finally, Rizzi's (1989) system of Relativized Minimality has as a consequence that VP does not count as a Minimality barrier for movement of any XP.

In Chapter III section 3.3.5.1, we proposed a definition of Dynamic Minimality which entails that unless the V moves, VP is a Minimality barrier, thus preventing extraction of an object, but crucially not the movement of the subject under [+HR] movement. Although in our view VP is a Minimality barrier, we will assume that it is not otherwise a barrier: this will eventually derive from our definition of L-marking (see section 4.4.1). Given the VP-internal hypothesis, this in fact a

necessary assumption. Indeed, if VP were a barrier, the subject NP could not raise out of it to move to the Spec of AGR-S without first adjoining to VP. But adjunction would lead to improper movement and preclude the movement of any VP-internal subject to the Spec of AGR-S, a most undesirable consequence. We therefore assume that VP is not an inherent barrier.

4.2.2 What can be adjoined to? Toward the elimination of the ban on CP adjunction

Another stipulation of the *Barriers* framework is that adjunction to CP is impossible. Chomsky (following a suggestion by Kyle Johnson) suggests that the ban on CP-adjunction follows from the theta-criterion. Under this view, adjunction to an argument violates the strict sisterhood requirement on theta-role assignment. Notably, however, in the *Barriers* system it is essentially in cases where the CP is <u>not</u> a direct argument, namely in cases of adjunct and subject islands, that the ban on CP-adjunction is crucially needed. Adjunctions to adverbial CPs or to subject CPs must be forbidden to prevent extraction out of islands. Consider a case of extraction out of an adjunct for illustration, assuming the structure of (37)a. is as in (37)b:[28]

(37)

---

28. We discuss the structure of adjunct clauses in more detail in a later section.

a. *How did you buy this book before leaving Paris t

b. [CP1 How did [IP1 you buy this book [CP2 before [IP2 PRO leaving Paris t]]]]

This sentence is an ECP violation, so there must be at least one barrier interrupting the chain of antecedent government. We have assumed that VP is not a barrier. Under our view IP2 is a barrier, but it can be escaped by adjunction. The adverbial CP2 is a barrier by inheritance from IP2. Now, if CP2 could be adjoined to, then clearly there would be no barriers between a trace adjoined to CP2 and the antecedent in the higher CP1. The ban on adverbial CP-adjunction is thus crucial to create at least one barrier in these cases.

There is no particular reason, however, why adjunction to a regular complement CP should be forbidden. Consider a case of adjunct extraction from an embedded interrogative sentence:

(38) How did you think that Mary would fix the car

As we saw above, although CP is L-marked by the matrix V, it will be a barrier by inheritance from IP. In principle, there are two ways that the barrierhood of CP can be eliminated: the extracted adjunct either lands in the Spec of CP or simply adjoins to CP. In cases such as (38), if the ban on CP-adjunction were to be eliminated, no problem would arise, as it it simply redundant. Note that if adjunction to CP were generally permitted, we would have a more elegant theory, allowing adjunction in all cases, and the problems which occur in our theory if long-distance WH-extraction involves movement through the Spec of CP, namely NP movement and WCO repair (cf section 4.1.2.2),

would be eliminated. In the rest of this section and in the following section we will propose further modifications to the *Barriers* framework which will allow us both to simplify the theory of the ECP and to eliminate the ban on CP-adjunction.

4.2.2.1 WH-islands

The first case we discuss is the that of WH-islands. Consider a simple example:

(39)* How do you wonder what to fix

In the *Barriers* framework, the assumed structure of (39) is the following:

(40) [$_{CP1}$ How$_k$ do [you wonder [$_{CP2}$ What$_i$ [PRO to fix t_i t_k]]]]

Two assumptions are crucial to an account of WH-islands: 1) that the Spec of CP is occupied by the first WH-phrase *what* so that the Spec of CP is unavailable for the successive cyclic movement of the adjunct, and 2) that adjunction to CP is banned. These assumptions have as a consequence that the movement of the adjunct will have to cross CP_2, which is a barrier by inheritance from IP. There is evidence, however, which suggests that the first assumption is not clearly relevant to an account of WH-islands. Consider the following example from French:

(41)
*Comment te demandes-tu si Jean va réparer la voiture.
How do you wonder if Jean will fix the car.

Like (40), (41) is an ECP violation, but in this case there is clear evidence that *si* does not occupy the Spec of CP, but rather the head of C. Evidence for the status of *si* as the head of a complementizer phrase comes from the fact that *si*, like English *if*, prevents the movement of the verb to C. If *si*, like other WH-elements, occurred in the Spec of CP, its complementary distribution with verbs in C would remain unexplained. (42) gives the paradigm showing the complementary distribution between *si* and a verb in C and the absence of complementary distribution between a verb in C and other WH-phrases:

(42)
a.Si tu avais fait ce que l'on te demandais
If you had done what you were asked

b.Aurais-tu fait ce que l'on te demandais
Had you done what you were asked

c.* Si aurais-tu fait ce que l'on te demandais
* If had you done what you were asked

d. Pourquoi as-tu fait ce que l'on te demandais
Why have you done what you were asked

Let us now return to example (41), the structure of which is given in (43):

(43) [$_{CP1}$ Comment$_k$ te demandes-tu [$_{CP2}$ t'$_k$ [$_{C'}$ si [Jean a réparé la voiture t$_k$]]]]

Note that in this case, the Spec of CP2 is free for the adjunct to move into. Consequently, the movement of the adjunct does not have to cross CP, and no barriers are crossed. Examples like (41) suggest

that assumption 1 above is insufficient to account for certain cases of WH-islands.

An alternative account of the ungrammaticality of extraction out of WH-islands has been proposed by Obenauer (1984 GLOW) and recently reformulated and expanded by Rizzi (1989). The basic idea of the Obenauer-Rizzi proposal is that in a given chain C = (a_1.....a_n), antecedent government cannot obtain if an element similar to the head of the chain intervenes in some position between the head and the foot of chain C. Consider the structure in (44):

(44) [A.....B....C]

In (44), A cannot govern C if there is a closer potential governor B for C. Rizzi (1989) proposes generalizing this constraint to account for a number of violations standardly regarded as ECP violations. He calls this generalized constraint Relativized Minimality, which is defined as follows:

(45) X α-governs Y iff there is no Z such that

(i) Z is a potential α-governor for Y and

(ii) Z c-commands Y and does not c-command X

The notion of potential governor is defined as follows:

(46)

Z is a potential antecedent-governor for Y if either:

1)- in an A-chain: Z is an A-specifier c-commanding Y

2)- in an A'-chain: Z is an A'-specifier c-commanding Y

3)- in an X^o-chain: Z is a head c-commanding Y

Rizzi conceives of Relativized Minimality as a constraint on antecedent government in a general theory of ECP which requires empty categories to be both antecedent-governed and head-governed. For the sake of completeness, we repeat his definitions here. The ECP is defined as follows:

(47) a [- pronominal] ec must be
(i) properly head-governed
(ii) antecedent-governed

Antecedent government is defined as in (48) and proper head government as in (49):

(48) α antecedent-governs β iff
(i) α and β are non-distinct
(ii) α c-commands β
(iii) no barriers intervene
(iv) Relativized Minimality is satisfied

(49) α properly head governs β iff
α is an X^0 and β is within the first projection immediately dominating α.

The case which concerns the present discussion is condition 3 in definition (46). Under the standard view that the Spec of CP is an A'-specifier, this condition accounts straightforwardly for the classic cases of WH-islands. To see this, let us consider (50):

(50) a.*How do you wonder which problem Mary could solve? b.[$_{CP1}$ How$_k$ do [$_{IP1}$ you wonder [$_{CP2}$ which problem$_i$ [$_{IP2}$ Mary could solve t_i t_k]]]]

Under Rizzi's definition of Relativized Minimality given in (45) and (46) above, the WH-constituent *which problem* in the Spec of the intermediate CP acts as a potential governor for the trace of the extracted adjunct. Consequently, antecedent government will fail. The number of barriers intervening between the extracted argument and its antecedent is irrelevant to the failure of antecedent government in this case. Note, moreover, that under this view the ban on CP-adjunction, assumption 2 above, becomes entirely irrelevant. Even if we were to assume that the adjunct WH-element has adjoined to the intermediate CP on its way up, the WH-element in the Spec of CP would still be a closer potential governor for the trace adjoined to IP. As a result, antecedent government would not obtain. Clearly, in this case the ban on CP-adjunction is superfluous.

Although Rizzi's proposal accounts in an elegant fashion for the classic cases of WH-islands, it fails to account for (41) because *si* is not an A'-specifier, but a head. Despite this technical difficulty, there is an intuitive sense in which the Obenauer-Rizzi hypothesis seems correct. To account for the ungrammaticality of (39) and (41), and to make the condition compatible with our model, in which the A/A' distinction has been abandoned, we propose to reformulate the RM condition pertinent to A'-chains so as to avoid the problem of Rizzi's definition:

(51) OPERATOR Opacity: (O-O)
In a [-HR] chain, α antecedent-governs β iff there is no δ, δ an OPERATOR, such that δ c-commands β.

We define OPERATOR as follows:

(52)

α is an OPERATOR iff

(i) α is a lexical operator and

(ii) α is in a [-HR] position

A precision is necessary to clarify the effects of our formulation of the constraint. In Chapter II we defined [+HR] positions as follows:

(53)

α is a Head-Related ([+HR]) position iff α is a sister to X^0 or to X', i.e., iff α is a specifier or a complement in the X'-theoretic sense.

[-HR] positions are any other positions. Note that according to our definition in (53), a Head is not defined as a [+HR] position. It is thus by default a [-HR] position.[29] From this it follows that the ungrammaticality of (41) will be correctly accounted for: *si* is a lexical operator which is the head of CP, and thus in a [-HR] position. Consequently, OPERATOR Opacity will be invoked, and the trace of the adjunct will fail to be antecedent-governed.

To summarize: The *Barriers* approach to WH-islands relies on two hypotheses: 1) that the Spec of CP is filled and therefore cannot serve as an escape hatch for the movement of a second WH-element, and 2) that adjunction to CP is forbidden. We have argued that cases of the type in (41) suggest that these hypotheses are insufficient. Following Obenauer (1984) and Rizzi (1989), we have opted for an

---

29. The idea that the position of heads is distinct from that of specifiers and complements was suggested by Chomsky (1989, class notes).

alternative account of WH-islands involving the condition on Operator opacity. Finally, we have noted that given this independently motivated approach,[30] the ban on CP-adjunction becomes superfluous.

Before we leave the topic of WH-islands, we need to raise a final question. If, as we suggest, WH-extraction involves adjunction to CP both in matrix sentences and in long-distance extractions, it follows that in an embedded interrogative the WH-phrase will also be adjoined to CP. Thus, the structure of an embedded interrogative is the following:

(54)
a.Je me demande comment Jean va réparer la voiture
I wonder how Jean will fix the car

b.Je me demande [$_{CP}$ comment [$_{CP}$ Φ [$_{IP}$ Jean va réparer la voiture]]]

We assume that the head of CP contains a Q morpheme which satisfies the requirement for the semantic selection of a question by the matrix verb, and that the WH-phrase must be adjoined to the category containing the Q morpheme. Note that this assumption creates no problems for our proposed account of WH-islands. Consider again a standard example of a WH-island: in our view (55)a. has the structure in (55)b.:

(55)

a. Comment te demandes-tu quel problème résoudre

b. [$_{CP}$ Comment$_k$ te demandes tu [$_{CP}$ t_{1k} [$_{CP}$ quel problème[$_{CP}$ Φ [t_{2k} [$_{IP}$ PRO résoudre t_i t_{3k}]]]]]]]

---

30. See Obenauer (1984 and following) and Rizzi(1989) for additional arguments in support of this approach.

The assumption that the first WH-phrase is adjoined to CP instead of occurring in the Spec of CP does not matter for the Operator Opacity condition. Since the OPERATOR *what* intervenes between the trace of the adjunct t_{1k} adjoined to CP and the trace of the adjunct t_{2k} adjoined to IP, t_{2k} will not be properly antecedent-governed, which leads to an ECP violation.

One question which remains is why it is never possible for two WH-phrases to occur simultaneously adjoined to CP. Consider the following cases:

(56)
a.*Je me demande qui comment Jean a rencontré.
I wonder who how Jean has met

b.?(?) Qui te demandes-tu comment Jean a rencontré
Who do you wonder how Jean met

In both (56)a. and (56)b., the Operator Opacity condition is violated. But given that the second extraction is that of an argument, we expect (56)a. to lead to a mere Subjacency violation, as does (56)b.;in fact, however, it is much worse. To account for this added ungrammaticality, we borrow an idea from Kayne (1984). Kayne proposes that in multiple questions such as (57), the first WH-position must be in the position of an operator, while the second must be in the position of a variable. In his view, this is a necessary condition for operator absorption (in the sense of Higginbotham and May (1981)) to be possible. This condition is met in (57):

(57) I know what you put where

In a sentence such as (56), however, although the first WH-position is indeed in an operator position, the second WH-phrase is not in the

position of a variable, if, as is standardly assumed, variables must occur in A-positions, [+HR] positions in our view. Thus, a complex operator cannot be formed at the relevant level of interpretation (presumably LF), and the sentence is ruled out. (56)b. differs from (56)a. in that the WH-phrases do not "compete" for the same scope position. *Qui* has matrix scope, *comment* embbedded scope. This is thus not a case of multiple questions such as (57), where both WH-phrases have scope over the matrix sentence. This completes our account of WH-island extractions. In all cases we have shown that given the Operator Opacity condition, WH-island violations can be accounted for without assuming either that WH-phrases must occur in or pass through the Spec of CP or that CP-adjunction is ruled out.

Given that the ban on CP-adjunction is now irrelevant to an account of WH-island violations, we are led to an interesting conclusion: we observe that the ban on CP-adjunction appears to be redundant in exactly the cases where CP is an argument, namely WH-islands and extractions out of ordinary sentential complements. In other words, the ban on CP-adjunction appears to be redundant in the cases where it is motivated under the theta-criterion in the *Barriers* framework. To eliminate the redundancy, a plausible alternative, which we will not, however, adopt, would be to reformulate the ban on CP-adjunction as follows:

(58) *[$_{CP}$ t [$_{CP}$ iff CP is not a complement

This would obtain the right results and adequately ban extractions out of subject islands and adjunct islands. (58) is, however, a purely

descriptive statement which lacks intuitive motivation and which does not derive from other principles.[31] Further consideration suggests that the ban on CP-adjunction might also be eliminated in cases of adjunct islands and subject islands. We will consider each of these cases in turn in the next section. Finally, we turn to a last case of islands which we have so far not mentioned: Complex NP islands. We will argue that in each of these cases, the ban on CP-adjunction can be eliminated. As a consequence, it will be possible to conceive of WH-movement as generally involving an free adjunction to any category including CP.

4.2.2.2 Adjunct islands

Following previous work by Obenauer, Rizzi(1989) argues that certain types of adverbs induce island effects.

(59)
*Comment as-tu beaucoup resolu de problemes?
How have you many solved of problems
How have you solved many problems?

---

31. The ban on CP-adjunction for adjunct islands could possibly be motivated as part of a general ban on "double adjunction". It is known that in English overt double adjunction to IP is generally impossible.

(i) *I think that to John, this book I will give.

We could exploit the adjoined status of adjunct clauses to ban further adjunction to their CP. This motivation, however plausible, does not extend to subject islands. We will thus not adopt it.

(60)
*Combien as-tu beaucoup consulté de livres?
How many have you much consulted of books
How many books have you consulted a lot?

(61)
[$_{CP}$ Combien [$_{IP}$ t_1[$_{IP}$ as-tu [$_{VP}$ beaucoup [$_{VP}$ consulté
[$_{NP}$ t_2 de livres]]]]]]

In Rizzi's view, the quantificational adverb *beaucoup* occurs in the Spec of VP and is thus a potential A'-binder for the trace of the extracted WH-element *combien*. Although no barriers are crossed by the WH-movement, Relativized Minimality or, in our view, Operator Opacity, is violated. Consequently, t_2 is not antecedent-governed by t_1 and an ECP violation occurs.

Extending this hypothesis slightly, we propose that the adverbial in the CP of an adjunct clause (or alternatively an operator in CP linked to this adverbial[32]) counts as an OPERATOR and thus as a potential antecedent-governor for WH-extractions out of adverbial adjunct clauses. If so, even if adjunction to CP is allowed, antecedent government will not be met in these cases either. Consider the following case:

(62) *How did John leave Paris before Peter fixed the car

(63) [$_{CP}$ How [did [$_{IP}$ t_1 [$_{IP}$ John [$_{VP}$ [$_{VP}$ leave Paris] [$_{CP}$ t_2 [$_{CP}$ before [$_{IP}$ t_3 [$_{IP}$ Peter fixed the car t_4]]]]]]]]]

---

32. Cf. Larson (1987) MITWPL for the proposal that certain averbial adjuncts have an operator in the specifier of their CP projection. Cf. also Hegarty's (1989) generals paper.

The adverb *before* intervenes between t_2 and t_3 and is thus a closer governor for t_3 Consequently, Operator Opacitity will prevent t_3 from being properly antecedent-governed; hence a violation of the ECP results. Here again, as in the case of WH-island violations, the presence of t_2, the trace adjoined to CP, does not matter. This extension of the Obenauer-Rizzi analysis, which is directly in keeping with the intuitive idea of Relativized Minimality or Operator Opacity, allows us to eliminate almost completely the stipulated ban on CP-adjunction without losing the core of the intended effects, namely an account of island violations.

There is, however, a residual case, namely the case of extractions out of subject islands, to which we now turn.

4.2.2.3 Subject Islands

To account for cases of ECP violations induced by extractions out of subject islands without making use of the ban on CP-adjunction, we could adopt a suggestion made by Frampton (1989). Frampton proposes that adjunction to any category is generally constrained by the following principle:

(64) Head Government Condition on Adjunction:
A WH-element can only be adjoined to a maximal projection XP from a position which is canonically governed by the head of XP. (Frampton (1989), p. 6)

This condition crucially bars adjunction to IP when the extracted element comes from a left branch. It thus bars adjunction to IP in cases of extraction from a subject clause. Consider a typical example of extraction out of a subject clause:

(65)

a.*How did that John fixed the car annoy Mary

b.[$_{CP1}$ How did [$_{IP1}$ [$_{CP2}$ t_1 [$_{CP2}$ that [$_{IP2}$ John fixed the car t_2]].....]]]

(65)b. gives a derivation in which adjunction to CP has occurred. Assuming (64), t_1, the trace adjoined to CP2, will be separated from its antecedent *how* by one barrier, namely IP1, even if adjunction to CP is allowed. This is sufficient to induce an ECP violation and to account for the ungrammaticality of (65). We have shown that the stipulated ban on CP-adjunction can be eliminated in the case of subject islands, provided we adopt Frampton's condition on adjunction.

Note, however, that Frampton's condition on adjunction bears a striking resemblance to another assumption of Rizzi's (1989) proposed theory of the ECP. Recall that Rizzi assumes a conjunctive statement of the ECP requiring both antecedent government and proper head government to be satisfied. We repeat Rizzi's formulation of the ECP here for convenience, as well as his definition of proper head government:

(66)ECP:

A [-pronominal] ec must be

(i) antecedent-governed

(ii) properly head-governed

(67) Proper head government:

α properly head-governs β iff α is an X^0 and β is within the first projection immediately dominating α.

Frampton's condition on adjunction requires that intermediate traces be governed by the head of the projection to which they adjoin. This requirement holds most of the time except, crucially, in the case of a trace moving from a subject position, since the subject position is not canonically governed. Frampton's condition strongly recalls Rizzi's statement of proper head government, which allows adjunct and complement traces, but not subject traces in the Spec of IP, to be properly head-governed.[33] Suppose we were to extend the conjunctive statement of the ECP to all intermediate traces.[34] It is quite clear, given Rizzi's formulation of proper head government, that a trace which is within a category in the Spec of IP (and which is itself not properly head-governed) will never be properly head-governed by I,

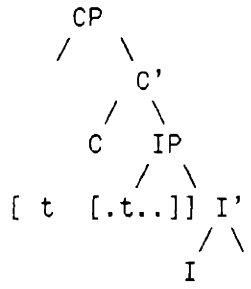
---

33. Complement traces will be properly head-governed by V. Adjunct traces will be properly head-governed by any functional head dominating VP, namely AGR-O, T, or AGR-S, depending on their adjunction site.

34. In fact, if we interpret Rizzi's formulation of the ECP literally, it should naturally apply to intermediate traces, since these are clearly non-pronominal ecs. But Rizzi never exploits the failure of head government for any other trace than the trace of a subject to account for the *that*-trace effect.

since it is not within the first projection containing I, namely I'. Consider the structure under discussion:

(68)



A trace either included in τ or adjoined to τ is not within the first projection containing I, namely I', and consequently it will not be properly head-governed by I. Moreover, such a trace will not be properly head-governed by C, since IP is a barrier for anything included in its projection.

This conclusion suffices to obtain the effect of Frampton's condition on adjunction at no additional cost and without the assumption that adjunction to IP is restricted. Suppose that in the structure in (68) there is an additional trace adjoined to IP. This will have no effect on the status of extractions of the type in (68): although antecedent government of a trace within the Spec of IP will obtain, head government will not. Thus, the ECP will not be satisfied.

We thus propose that head government as well as antecedent government must be met by all intermediate traces for the ECP to be satisfied. Let us call this proposal the Generalized Head Government Requirement:

(69) Generalized Head Government Requirement (GHGR)

All [-pronominal] ecs, including intermediate traces, must be properly head-governed.

Let us consider more closely the notion of proper head government as defined by Rizzi (1989):

(70) α properly head-governs β, α an X^0, iff

(i) α head-governs β and
(ii) β is contained in the first projection dominating α

Rizzi defines head government as follows:

(71) α head-governs β iff

(i) α ∈ { A,N,P,V,Agr,T}
(ii) α *m-commands* β
(iii) no barriers intervene
(iv) Relativized Minimality is respected

In Rizzi's framework, C^0 is excluded from the set of head-governors essentially to prevent a subject from being head-governed by a complementizer. This is at the core of his proposal to account for the *that*-trace effect. In Rizzi's view, the *that*-t effect results not from the failure of antecedent government, but from the failure of head government. Consider the classic example in (72):

(72) Who do you think [t'_i that [t_i left?]]

In Rizzi's view, *that* is not a potential head-governor for t_i, by assumption. Moreover, since t_i is in the Spec of IP, it is not within the first projection of I. It can thus not be head-governed by I. Consequently, t_i is not head-governed at all, and the ECP is violated. We return to a more extensive discussion of the *that*-trace effect in section 4.4.1. For the moment, let us consider the effects of Rizzi's definition within the system we have adopted so far. In section 4.2.1.1, we argued that IP is not a defective category and that it is

a barrier because it is not L-marked by C°.[35] Given this hypothesis, we do not need to exclude C° from the range of head-governors. Given Rizzi's definition, no X° is able to properly head-govern across a barrier. Consequently, in our system, even if C° is a head-governor, the subject trace in the Spec of IP will not be able to be properly head-governed across the IP barrier.

Given the GHGR, it is in fact crucial to our view to assume that C° is a head-governor. Under our proposal, every intermediate trace must be properly head-governed. Consider the case of a trace adjoined to IP. The structure is the following:

(73) [$_{CP}$ [$_{C'}$ C° [$_{IP}$ t [$_{IP}$.....

If C° is not a head-governor, an IP-adjoined trace will not be properly head-governed and will thus lead to an ECP (or Subjacency) violation. If, on the other hand, C° is a head-governor, an IP-adjoined trace will be properly head-governed, since 1) it is contained in the first projection containing C°, namely C', and 2) no barrier separates this trace from the head-governor C°, since the upper segment of IP is not a barrier. In our view, C° is not a deficient head-governor; it simply differs from other functional heads in that it does not L-mark the functional projection it dominates, namely IP.[36] We will argue in section 4.4.1 that agreement turns C°

---

35. In other words, in our view an X° can be a head-governor without being an L-marker. We will discuss the notion of L-marking in section 4.4.1

36. See section 4.4 for a discussion of L-marking. We assume the following definition: a functional projection L-marks only under

into an L-marker. Since in our view C° is a head-governor, we can simplify the definition of head government in the following way:

(74) α head-governs β iff

(i) α is an X°

(ii) α m-commands β

(iii) no barrier intervenes[37]

As a consequence of this simplification, any trace adjoined to IP will be properly head-governed by C°. Other traces adjoined to any projection of XP within a clause will be properly head-governed by the head of the projection YP which dominates the projections XP they are adjoined to. The abstract structure is the following:

(75)... [YP..Y...[XP t [XP ...]]]

In other words, head government will generally hold for intermediate adjoined traces within a clause. Note that head government of a trace in the specifier of a projection XP by a head Y will also be met if the the projection XP is not a barrier. For instance, the intermediate trace of a passive NP in the Spec of AGR-O will be properly head-governed by T, since, as we have assumed, the projection of AGR-O is not a barrier. Consider the structure in question:

---

head agreement. A lexical projection L-marks under head government. L-marking only occurs under strict c-command.

37. We have eliminated the Relativized Minimality requirement on head government in the definition, first because it is irrelevant to the present discussion, and second because in section 4.3.2 we will argue that Relativized Minimality is not needed.

(76)
a. Les tables seront repeintes
   The tables will be repainted

c. [AGRP-S les tables$_i$ seront $_v$ [TP t"$_i$ t$_v$ [AGR-O t'$_i$ [repeintes t$_i$ ...]]]]

In (76), t"$_i$ in the Spec of TP is properly head-governed by the V in AGR-S, t'$_i$ in the Spec of AGR-O is properly head-governed by the trace of T, and t$_i$ is properly head-governed by the past participle. Since no barrier intervenes, antecedent government is also met, so the ECP is satisfied and the sentence is correctly predicted to be grammatical.[38]

Given definitions (70) and (74) and the assumption that intermediate traces need to meet both clauses of the ECP to be fully licensed, we can now return to the central issue of this section, which is to account for ECP violations in extractions out of subject sentences. Let us consider the derivation of (65), given in (78):

---

38. Note that the GHGR suggest an explanation for why *do*-support is needed in main clause interrogatives. To see this, let us consider a simple case of object extraction.

(77) What did [IP t'$_i$ [IP John buy t$_i$]]

Since in our view IP is a barrier, adjunction to IP is necessary, leaving the trace t'$_i$ adjoined to IP. Under the GHGR, t'$_i$ needs to be head-governed. Assume that a C$_o$ which is not head-governed cannot itself be a head-governor. This will force the movement of a lexical head to C to satisfy the head government of the trace adjoined to IP. In embedded sentences, C$_o$ is head-governed and selected by the matrix verb, so we can assume that it acquires enough strength to head-govern the t adjoined to an embedded IP. Under Economy, because AUX to C is not needed, it will be impossible. In our view, the distribution of AUX to C follows directly from the GHGR.

(78)

[$_{CP1}$ How [$_{C'}$ did [$_{IP1}$ t_1 [$_{IP1}$ [$_{CP2}$ t_2 [$_{CP2}$ that [$_{IP2}$ John fixed the car t_3]]] .....]]]]

As we have assumed, the WH-extracted adjunct is free to adjoin anywhere on its way to CP1. In particular, it is free to adjoin to CP2 and to IP1. If it does not adjoin to either CP2 or IP1, then clearly a barrier will be crossed, since neither CP2 nor IP1 is L-marked, and the sentence will be ruled out. If the WH-element adjoins to both CP2 and IP1, as represented above, there are apparently no barriers crossed. Thus, an ECP violation will arise, not through a failure of antecedent government, but rather through a failure of proper head government. Consider the trace t_2 adjoined to CP2. This trace cannot be properly head-governed by any head and consequently the ECP is violated. Being adjoined to CP2, it is not properly governed by the head of CP2 because it is not contained in the first projection dominating this head. Next, it is not properly head-governed by I2 because it is not contained in I'2,the first projection immediately dominating I2. Finally, it is not properly governed by C1, either, because IP1 is a barrier for t_2. Thus, t_2 violates the ECP through a failure of head government. Adjunction to IP_1 does not help, because in this case IP1 is not a barrier for antecedent government, but a barrier for the head government of t_2 by CP1.[39]

---

39. This analysis may seem problematic for contexts in which the subject island is apparently head-governed, e.g. in the subject position of a small clause.
(i) ??Je trouve [que Jean soit venu] révoltant
I find [that Jean has come] revolting
*Comment trouves-tu [t [que Jean soit venu t]] révoltant
How do you find [that Jean has come] revolting

As we have shown, the requirement that intermediate traces be both antecedent-governed and head-governed suffices to derive the effects of Frampton's condition on adjunction for extraction from subjects: adjunction to CP does not help to prevent an ECP violation. We thus conclude that with the conjunctive definition of the ECP and the GHGR, extraction from subject islands can be accounted for without any recourse to the ban on CP-adjunction. Adjunction to CP can thus be freely allowed without any unwanted consequences.

4.2.2.4 CNP islands

Before concluding this section, we need to consider a final case of islands, namely complex NP islands. There are two basic cases of these islands: relative clause constructions and noun-complement constructions. Relative clause constructions are parallel to extractions out of WH-islands. Under standard assumptions, the complementizer of a relative clause construction contains an operator; this operator will induce an Operator Opacity violation, just as it

---

Here the adjunct trace adjoined to CP is seemingly head-governed by the upper V. We would argue, however, that this is not the case. Kayne() has proposed that SCs are in fact projections of an AGR phrase. Although L-marked, the maximal projection of AGR is a barrier for the trace adjoined to CP by inheritance from CP. Independent arguments that SCs are projections of a functional category can be derived from Ruwet's paper on absolutive phrases of the type *Avec Jean pour guide*("with Jean as a guide") vs. *Avec Jean malade*("with Jean sick"). For reasons of time and space we do not develop these arguments in the present thesis, but will do so in forthcoming work.

does in cases of extraction out of WH-islands. As we argued above, the ban on adjunction to CP in this case is superfluous. Since this case does not introduce any new problems, we will not consider it in detail.

Let us turn to noun-complement constructions. Grimshaw (forthcoming) has argued that in these constructions the sentential complement is in fact not really a complement of the head noun. Although she argues for this conclusion on a semantic level, such a conclusion can be reinterpreted syntactically as involving an adjunction structure rather than a complementation structure: the embedded sentence is adjoined to the NP projection containing the head noun. We will adopt this proposal.[40] The structure of sentences such as (79) is given in (80).

(79) Mary confirmed the rumor that John was going to move

(80) Mary confirmed [NP [NP the rumor][CP that [IP John was going to move]]]

Let us now examine an extraction out of this construction:

(81)*How did you examine the rumor that John was going to move

---

40. Cf. Stowell (1981) for a comparable proposal. Alternatively, we could assume that although the embedded sentence is a complement of the head noun, the head noun is not a "proper" head-governor; this alternative is similar to a proposal by Cinque. If N is not a proper head-governor, we obtain the right results, as the reader can easily check: traces adjoined to CP will not be properly head-governed.

(82) [$_{CP1}$ How [did [$_{IP1}$ t_1 [$_{IP1}$ you examine [$_{NP}$ t_2 [$_{NP}$ [$_{NP}$ the rumor] [$_{CP2}$ t_3 [$_{CP2}$ that [$_{IP2}$ t_4 [$_{IP2}$ John was going to move t_5]]]]]]]]]]

In the structure given in (82), we have assumed free adjunction to all categories. Given such a derivation, the extraction of an adjunct does not cross any barriers, and antecedent government is satisfied. t_4 antecedent-governs t_5, and is itself antecedent-governed by the t_2 adjoined to the NP, which in turn is antecedent-governed by t_1 adjoined to IP. Head government of each trace also holds. As we have assumed, the original trace t_5 is head-governed by I in the lower clause, t_4 is head-governed by C° in CP2, t_3 is head-governed by the matrix verb, and t_1 is head-governed by the matrix C°. We thus wrongly predict that the structure (81) should be grammatical. To account for this last case without giving up the idea that adjunction to CP is freely allowed, we will adopt the idea, proposed by Belletti and Rizzi (1986), that a segment of a projection can inherit barrierhood from a barrier that it dominates. This proposal is formalized as in (83):

(83)

τ, a segment of an XP, is a barrier for β iff τ dominates δ, δ a non-L-marked BC which does not exclude β.

With this in mind, let us return to the derivation in (82). Under (83), the upper segment of the NP will be a barrier for t_3 adjoined to CP2. Indeed, it dominates CP2, which is a barrier, since it is not L-marked by the NP *rumor* and it does not exclude t_3. We could assume, contrary to the derivation represented in (82), that adjunction to NP is forbidden. This would be sufficient to exclude (82), since the trace t_3 would fail to be antecedent-governed. But the stipulation

that NPs cannot be adjoined to goes against the spirit of the theory we are developing, in which adjunction is generally free. Fortunately, we do not need such a stipulation. Our extension of the head-government requirement to all intermediate traces, the GHGR, will be sufficient to provide an account for this case. Consider the trace t_3. Since the upper segment of the NP is a barrier, head-government of t_3 by the matrix verb will not obtain. Although t_3 is within the first projection of the V^o *examine*, it is separated from it by the NP segment, which is a barrier under (83). Thus, proper head-government of t_3 will fail, leading to an ECP violation. This accounts for the ungrammaticality of (81).

We conclude again that the ban on CP-adjunction is not needed in this case, provided we adopt the proposal in (83).[41,42]

---

41. The impossibility of adjunct extraction out of CNPs is accounted for under Minimality in *Barriers*. Chomsky assumes, contrary to us, that CP is a complement of the nominal head *rumors*. As a consequence, the nominal head will induce a Minimality barrier and adjunct extraction will be ruled out. This account is crucially not available in an RM framework, and it is in fact unclear in Rizzi's theory what accounts for ECP violations in extraction out of CNPC islands.

42. The account we offer of CNPC violations can be extended straightforwardly to sentential adjuncts not introduced by an adverb, such as the following example cited in *Barriers* p. 33: "it's time to fix the car".

4.2.2.5 Conclusion

In this section we have argued that the stipulation that adjunction to CP is banned can be eliminated. Four assumptions have been crucial to this result:

1) the Operator Opacity condition
2) the assumption that the ECP involves a conjunctive formulation, i.e., an ec must meet both proper head government and antecedent government and that this conjunctive requirement maut be met by all intermediate traces (GHGR)
3) the hypothesis that the upper segment of an adjunction can inherit barrierhood form a non-L-marked BC.

We have shown that hypothesis 1 can account for cases of WH-island violation which are not straightforwardly accounted for either in the *Barriers* framework or in the Relativized Minimality framework of Rizzi (1989). Hypothesis 43has been argued for independently by Rizzi and Belletti (1986) to capture data not accounted for in the *Barriers* framework. Thus, each of the hypotheses we have made is independently motivated.

To conclude this section, we will review some empirical evidence which shows that a complete ban on CP-adjunction is too strong.
To begin with, as is well known, there are a number of languages, for instance Bulgarian, Romanian, Serbo-Croatian, Czech, Polish, Hindi, and Russian, which allow a succession of overt WH-elements.

(84)

Koj kogo vizda (Bulgarian)

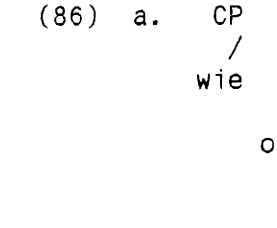
who whom sees

See Rudin (1988) and references cited there for arguments that in at least some of these languages multiple WH-structures involve adjunction to CP. Moreover, an interesting example cited by Koster (1987) also suggests that adjunction to CP should be possible in Dutch (a non-multiple-question language) in complement sentences.[43] Koster mentions that in certain dialects of Dutch up to three elements are allowed to cooccur in CP:

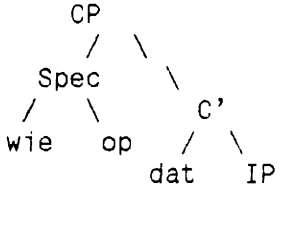
(85)
ik wet niet wie$_i$ of (dat) t$_i$ het gedaan heeft
I know not who if that it done has
I do not know who did it

Further questions arise as to the structure of these triply-filled complementizers; several possible structures come to mind. Adjunction could be either to the Spec of CP or to the maximal projection of CP. These two options are represented in (86):[44]

(86) a.



b.



---

43. Many thanks to David Pesetsky for pointing out this fact and its relevance.

44. An alternative to (86)a. and b. is a double CP structure.

An exploration of the diverse possible structures of multiply-filled complementizers and of their consequences on the islandhood of various constituents is beyond the scope of this dissertation. The main purpose of this section was to show that Chomsky's ban on adjunction to CP may be eliminated from the *Barriers* framework. Given this possibility, WH-movement can now be analyzed as generally involving adjunction to CP and not movement to the Spec of CP. As we will argue in the next section, however, there is at least one case in which, in our view, movement to the Spec of CP is possible and in fact required. This case involves the extraction of a subject. Before we turn to the subject extraction cases, however, we will discuss in detail the theory of the ECP we propose.

4.3 Toward a definition of the ECP

In the previous sections, we proposed a number of modifications to the *Barriers* framework. The modifications we argued for are summarized below:

1) WH movement is not movement to the Spec of CP, but involves an adjunction to CP. Thus, WH-movement forms [-HR] chains.
2) There are no defective categories and adjunction is generally free.
   a. We have argued that IP is a BC and a barrier when it is not L-marked. Contrary to Rizzi, we have argued that C° is a head-governor but that it does not L-mark IP. We have assumed additionally that CP always inherits barrierhood from IP, even after adjunction.
   b. We have assumed that VP, although a D-barrier, is not an

inherent barrier.

c.We have abandoned the ban on CP-adjunction and argued that adjunction is generally free.

3) We have adopted a conjunctive definition of the ECP:

ECP: a [-pronominal] ec must be

(i) properly head-governed

(ii) antecedent-governed

Antecedent government is defined as follows:

(87) α antecedent-governs β iff

(i) α c-commands β and

(ii) no barriers intervene

(iii) Operator Opacity is satisfied

We have so far, by and large, adopted Rizzi's definition of proper head government which we repeat here for convienience:

(88) α head-governs β iff β is contained in the first projection of α

4) In Chapter 2 we proposed a definition of Dynamic Minimality which, although it has no direct effect on [-HR] chains, is in our view part of the ECP.

Although the system as it stands has a relative elegance and covers a wide range of empirical facts, there are a number of conceptual issues which still need to be discussed.

-First, we need to provide some motivation for the assumptions we have made for the status of various categories as barriers. In other words, we need to answer the question why VP is not an inherent barrier and why IP is not L-marked by C and is thus a barrier. The answer to this

question involves a discussion of the "proper" definition of L-marking. We will postpone this discussion until section 4.4., where we consider the analysis of subject extractions.

-Second, the definitions of head government and antecedent government in Rizzi's approach involve a striking parallelism, but remain crucially distinct. We will discuss the reasons for this distinction in section 4.3.2 and propose a revision of the definition of ECP in which we will consider the requirement of head government as a condition on antecedent government.

-Third, the theory we have so far adopted resembles in some aspects the theory proposed by Rizzi (1989). The Operator Opacity condition we have proposed is similar to Relativized Minimality for A'-chains. The question then arises why we do not simply adopt this approach for other types of dependencies such as [+HR] chains and Head-to-Head movement. One obvious reason is the incompatibility of Relativized Minimality with the notion of Dynamic Minimality we proposed in Chapter III section 3.3.5.1, but there are additional reasons why the RM theory appears in fact too strong. We discuss these issues in section 4.3.2.

4.3.1 C-command and M-command

Rizzi (1989) defines the notion of proper head government in terms of the notion of head government. The main distinction between the two involves 1) the restriction on the set of head-governors and 2) the

notion of c-command vs. m-command. Concerning the first difference, we have argued that given the GHGR (a natural and logical extension of the conjunctive formulation of the ECP) and the assumption that IP is a barrier, the set of head-governors can be assumed to include all X^o. With respect to the second difference, we observe that in Rizzi's theory only the notion of c-command is directly relevant for the head government relation involved in the formulation of the ECP. The notion of m-command, on the other hand, concerns the head government relation involved in other types of licensing such as Theta-marking and Case-marking. (It may also be relevant to the distribution of PRO, but we will leave the issue of Control as an open questior.[45]) The c-command/m-command difference between the notion of head government and proper head government in fact reflects the LGB distinction between the notion of government and the notion of proper government. It thus poses the question of which if these two relations, m-command or c-command, is relevant for the ECP. This issue is pervasive throughout the ECP literature. In what follows, we will show that the notion of m-command raises problems for most approaches to the ECP. We thus propose that the notion relevant for the theory of ECP is the notion of c-command.

---

45. Note first that even under a standard view, if the notion of m-command is relevant to Control, it is unclear why PRO should be considered "ungoverned", since it is m-commanded, at least in English, by the prepositon *to*, standardly assumed to occur in the Infl node of infinitival sentences. If, as argued by Pesetsky (1989), PRO must in fact be assigned dative Case, then the distribution of PRO falls under Case theory, as originally proposed by Rouveret and Vergnaud (1980). If this is correct, it is natural that the notion of m-command is the one which is relevant to the licensing of PRO.

In the LGB model, both Theta-marking and Case-marking are assumed to occur under government by a head. Since the subject is assumed to be both Case-marked and Theta-marked in the Spec of INFL, the relation of head government must include the relation between a head and its specifier as well as the relation between a head and its complement.[46] However, Chomsky noted in LGB that the relation of head government relevant for Case- and Theta-licensing is too permissive for the proper formulation of the ECP. This is why he proposes to define a more restrictive relation of government relevant only to the ECP, namely the relation of proper government. Broadly speaking, we can see the LGB definition of proper government as effecting a split between specifiers and complements: specifiers must be antecedent-governed, while complements must be head-governed. On the other hand, the relation of head government necessary for Case- and Theta-licensing essentially collapses specifiers and complements.

Thus, there seems to be an important difference between the relation needed for Theta-marking and Case-marking and the relation needed for the licensing of a trace under the ECP. In other words, there seems to be a tension between the notion of government relevant to Case- and Theta-licensing and the notion of government relevant to the ECP. In Rizzi's theory this tension stricIty corresponds to the distinction between c-command and m-command in the definition of head government. Note that crucially for Rizzi's view, it is the notion of c-command

---

46. Note that this remains true under the VP-internal subject hypothesis

which is relevant for antecedent government. Thus, it is apparent that only the relation of c-command is relevant in his theory of the ECP.

Another fundamental relation of UG has been argued by Chomsky to involve solely the notion of c-command, namely the relation of anaphoric-binding. To quote Chomsky (1986), p. 8:

(89)

" α c-commands β iff α does not dominate β and every τ that dominates α dominates β

Where τ is restricted to maximal projections (following Aoun and Sportiche (1983)), we will say that α m-commands β. It seems that for the binding theory, τ should be taken to be any branching category along the lines of Reinhardt (1976)" (our emphasis)

It is quite clear that one does not expect to find a situation in which, for instance, a head can serve as an antecedent for an anaphor in the Spec of its projection, (concretely, allowing a clitic which is in a relation of m-command with the specifier of a projection to license an anaphor in this specifier position).[47]

In recent years, the basic conceptualization of the ECP has evolved toward putting more emphasis on the relation of antecedent government, thus making the ECP more similar to anaphoric binding. It has been shown that the split between specifiers and complements with respect to the relation of "proper government" is in fact too strong. Kayne

---

47. Cf. the notion of accessible SUBJECT in the LGB definition of governing category.

(1981) showed that the relation of antecedent government must also obtain for complements.[48] Pushing similar ideas further, Aoun (1985) developed a theory which subsumes the ECP under the broader notion of binding understood as a general theory of anaphora. In the same vein, Chomsky (1986) suggested that the essence of the ECP was antecedent government. He developed a theory of anaphoric binding which involves movement of anaphors at LF and attempts to subsume anaphoric binding under the ECP. The prevalence of the notion of antecedent government is also present in Rizzi's (1989) theory of Relativized Minimality. The core notion of Relativized Minimality is that antecedent government cannot obtain if a potential binder intervenes. RM can be seen in a sense as an attempt to extend the Specified Subject Condition (Chomsky (198..)). Note that if the assumption that the core of the ECP is essentially antecedent government is correct, and we believe that it is, one would expect that, as with anaphoric binding, the relevant relation for the ECP be that of c-command, not m-command.

Despite the conceptual move of placing emphasis on antecedent government, the notion of head government has not been eliminated from the various theories of ECP. Thus, Kayne (1981) argues that head government is also needed for the ECP. Head government is also present in Aoun's theory in the definition of Governing Category. Broadly speaking, the notion of head government is at the core of the definition of Minimality in *Barriers*.

---

48. See "ECP extensions". In this paper Kayne in fact shows that the relation of antecedent government and the relation of lexical government must hold simultaneously for complements.

In *Barriers*, while the relation considered to be pertinent for government, and thus also for antecedent government, is m-command, (α governs β if α m-commands β and no barrier for β excludes α (*Barriers*. p. 88)), the relation pertinent for Minimality appears to be rather the relation of c-command (although Chomsky also mentions an alternative possibility, the Minimality condition: " a category τ is a barrier for β if it is the immediate projection (alternatively, a projection) of a zero-level category δ not equal to β." (*Barriers* p. 88)). Here again, we can detect the dichotomy which is explicit in LGB between two different notions of government.

Although Chomsky suggests that the relevant distinction for anaphoric binding is c-command, he assumes that the relevant relation for antecedent government is m-command. It is clear, however, that as in the case of anaphoric binding, we do not expect a head coindexed (through agreement) with its specifier to be able to serve as an antecedent-governor for a trace in the specifier. Thus, for instance, the head of AGR-S cannot serve as an antecedent-governor for the subject trace, even though it m-commands it and it is coindexed with it. To account for NP-movement in passive, Chomsky proposes that a coindexed head may in fact serve as an antecedent-governor for a trace (see *Barriers* p. 77). Given this hypothesis and the definition of antecedent government in terms of m-command, we expect AGR-S to be able to antecedent-govern a subject trace. This clearly makes the wrong prediction with respect to the *that*-t effect. It thus seems clear that even in the *Barriers* framework, the relevant relation for

antecedent-government should be assumed to be c-command and not m-command.

If this is correct, then, a clear picture emerges from the above discussion. There are essentially two types of formal relations which are fundamental in UG: c-command and m-command. Each of these relations seems to enter into fundamentally different types of licensing. M-command is relevant for Case- and Theta-licensing and c-command is relevant for BINDING (understood here as a general notion of anaphora which includes anaphoric binding and antecedent government).

Although it seems clear that the relation relevant for antecedent government should be c-command, there remains a question about head government. As is apparent from Rizzi's definition of proper head government and from Chomsky's definition of Minimality, the notion of head government relevant for the ECP involves c-command and the notion of head government relevant for Theta-marking and Case-marking involves m-command. It is standardly assumed that head government is but one condition on the licensing relations of Case-marking and Theta-marking. Theta-marking, for instance, must additionally meet the condition of Visibility (Chomsky (1981)). Along these lines, we propose that the head-government requirement involved in the ECP be considered as a condition on antecedent government, itself conceived in terms of c-command. With this in mind we can now simplify the definition of ECP.

(90)

ECP: (final version)

a [-pronominal] ec must be ANTECEDENT-governed.

ANTECEDENT government:

α ANTECEDENT-governs β iff

(i) α c-commands β

(ii) β is c-commanded by δ, δ = X^0

(iii) no barrier intervenes between α and β and between δ and β.

(iv) Operator Opacity is satisfied.

For expository convenience I will continue to refer to the notion of antecedent government and head government separately in the sections to come, so as to detail which part of the definition (90) is at stake. To avoid terminological confusion, I will use the terms a-government and h-government.

4.3.2 Relativized Minimality vs Dynamic Minimality

As is apparent, the definition of the ECP we propose bears some similarity to Rizzi's Relativized Minimality framework. In particular, the Operator Opacity condition resembles Rizzi's condition on A'-chains. The question arises, then, whether we should use Relativized Minimality for other types of chains, such as [+HR] chains and Head-to-Head movement. In this subsection we will briefly discuss [+HR] chains and the notion of Relativized Minimality applied to it. Then we turn to a discussion of Head-to-Head movemement. We will conclude that

Relativized Minimality is of no use for these two types of chains, given our approach to the ECP.

Let us first consider [+HR] chains. Rizzi formulates the condition in the terms given in (91) (=(48) in the previous chapter):

(91)

Z is a potential antecedent for Y if:

- in an A-chain, Z is an A-specifier c-commanding Y

Transposed into our model, the condition would be reformulated as follows:

(92)

- in a [+HR] chain, Z is a [+HR] specifier c-commanding Y

Let us first note that the notion of [+HR] specifier is not really expressible in our terms. Recall that in our view, the notion of [+HR] positions covers both specifiers and complements in an X'-theoretic sense. In other words, we would need to distinguish a special kind of position to be able to integrate Rizzi's proposal into our theoretical model.

Second, let us consider the cases in which Rizzi invokes the RM condition on A-chains. The major cases are cases of super-raising such as the following:

(93) *John$_i$ seems that it is important t_i to leave

In Rizzi's view, the ungrammaticality of this sentence is due to the fact that the expletive pronoun *it*, which is in an A-specifier, c-commands the trace t_i of *John*. But, as aknowledged by Rizzi, this

analysis does not in fact suffice to account for the severity of the ungrammaticality of (93). To account for the distinction between ECP violations and Subjacency violations, Rizzi proposes to distinguish adjuncts and arguments in terms of their referential properties. In his view, adjuncts are essentially non-referential while arguments are referential:[49] arguments bear referential indices, but adjuncts do not. This implies that adjunct chains are subject to strict locality, i.e., no barriers can intervene to separate any link from another. In argument chains, on the other hand, if antecedent government does not obtain, there is an alternative way of establishing the relation, namely binding. With this in mind, let us return to the example in (93). (93) involves the movement of an argument. Thus, even though antecedent government does not obtain because of the intervention of the A-specifier containing the expletive *it*, binding of the trace should be possible. As a result (93) should only have the force of a Subjacency violation. This is clearly an incorrect prediction. To account for (93), Rizzi proposes an additional condition on theta-chains, but the necessity of this proposal is in a sense vitiating the force of the argument that A-chains are sensitive to Relativized Minimality. Indirectly, A-chains are thus shown to be insensitive to RM. For this reason we do not adopt Rizzi's proposal and will account for super-raising cases in a different way (see section 4.5.2).[50]

---

49. We discuss this distinction and its consequences in more detail in section 4.7.1.1.

50. There are additional arguments against Rizzi's definition of Relativized Minimality for [+HR] chains which come from our analysis of scrambling and object movement. If we are correct in analyzing the movement of objects as movement to a [+HR] position, under an RM analysis, we would expect the subject to be incapable

The next case to consider is the case of Head-to-Head movement. Although it seems that Rizzi's Relativized Minimality accounts correctly for the Head-to-Head movement constraint, it is quite clear that this constraint is somewhat redundant with the definition of Dynamic Minimality we have proposed in Chapter III section 3.3.5.1. Let us consider this definition.

(94) τ is a Minimality Barrier for β with respect to α iff

τ is a maximal projection which

(i) excludes α and

(ii) contains δ, δ a LEXICAL head which c-commands β.

α is LEXICAL iff α contains an overt lexical head.

The definition in (94) has two main results: 1) the subject can move to any [+HR] position independently of the movement of the verb, and 2) object movement (except passive (see section 3.3.5.1)) is dependent on verb movement. Moreover, since functional projections which do contain a LEXICAL head are not D-barriers, the movement of the subject can skip the specifier of some functional projections on its way to its final position. In our view, this is important, because it allows the subsequent movement of the object into the Spec of AGR-O. Suppose, on the other hand, that the subject NP were forced to land in the specifier of every functional projection on its way to AGR-O. This would block further movement of the object to the Spec of the AGR-O projection since, as we will argue in section 4.7.1, traces in [+HR]

---

of moving over an object in the Spec of a functional projection or vice-versa. Clearly, this is an incorrect prediction, as we saw in Chapter III.

positions do not erase. Thus, even in cases where the verb has moved, object movement is impossible for independent reasons. Note that this would be the case if we were to assume that functional projections not containing lexical heads can induce Minimality barriers.

Let us now consider the consequences of our definitions for the Head Movement Constraint. Note first that our definition correctly blocks an example like (95), where a verb has moved over an auxiliary to a higher functional projection, namely C.

(95) *been$_i$ Mary has t_i arrested

In (95), the auxiliary *been* has moved over another auxiliary. But the projection containing the auxiliary *have* will be a Minimality barrier, since it contains an overt lexical element. Thus, t_i will not be antecedent-governed and the sentence is excluded by the ECP. However, our definition of Minimality is more permissive than the Head Movement Constraint. In principle it allows a verb to move to C directly without stopping in any of the intermediate functional projections, since these do not contain any lexical material. Whether such a definition is too permissive is not easy to assess. Suppose that the verbs moves to C, stranding an affix in AGR-S. Such a hypothetical case is represented in (96):

(96) [$_{CP}$ V_v [$_{AGRP\text{-}S}$ affix [....t_v..]]]

A priori, it seems desirable to exclude such a structure. But the question arises whether such a structure must be excluded by the ECP or can be independently excluded. It is clear that any theory that

assumes verb movement (or affix lowering) must include a principle which excludes unattached affixes, and this independently of any consideration of the ECP. Let us call this principle principle P. Depending on its ultimate formulation, principle P will plausibly exclude structures such as (96) independently of the ECP. Now, of course the affix could move independently to escape principle P and attach to the verb after the raising of the verb. Although such a derivation would violate a "rigid" formulation of the ECP (in which all X° count for the definition in (94)), it would not violate principle P. Whether such a derivation should be allowed or precluded is not immediately obvious and would require a thorough investigation of head movement. Since this is orthogonal to the main topic of our section and more generally to this dissertation, we will leave the question open. Note that if it turned out to be necessary to assume that every functional projection must count as inducing a Minimality Barrier, there would still be a derivation which would permit both the movement of the object and the movement of the subject. Consider a derivation in which the verb moves first to the highest functional projection, namely AGR-S. In such a derivation all the intermediate projections would contain only a trace. Assuming that the trace of an X° does not induce a Minimality barrier[51], the movement of the subject could proceed directly to the Spec of AGR-S without stopping in the Spec of AGR-O. In fact, since stopping in the Spec of AGR-O would represent one more step in a derivation, this intermediate landing

---

51. See Baker (1986) for a similar proposal. WE have discussed his proposal in a footnote of section 3.3.5.1 and concluded that his definition encounters problems with the VPS hypothesis.

would be precluded by Economy considerations. Consequently, the Spec of AGR-O will be free for the object to move into. Given this possible derivation, where the movement of the V precedes any movement of NP, it is quite clear that the requirement for a LEXICAL head in the definition in (94)) is not crucial to any of the facts we have accounted for so far, although the assumption that the trace of an X^o does not induce a Minimality barrier is. We could thus simplify the definition in (94) as in (97), keeping in mind, however, that such a strengthening is not obviously necessary.[52]

(97) τ is a Minimality Barrier for β with respect to a governor α
iff τ is a maximal projection which
(i) excludes α and
(ii) contains δ, δ an X_o which immediately c-commands β.

To conclude, let us note that in this definition of Minimality, c-command is the crucial notion. As a consequence, all our definitions are consistent in using only the notion of c-command.
This concludes our discussion of the theory of the ECP. In the next sections, we will consider in detail how the theory we propsed works for classical cases of extractions. We will also turn to the last theoretical point we have so far left open, i.e., the formulation of a definition of L-marking.

---

52. We will explore the issues at stake in forthcoming work.

4.4 Case study

Before we proceed to show how the theory of ECP we proposed in the previous section accounts for the problems that the assumption that the Spec of CP is a [+HR] position raises (see section 4.1), we will go through the classic paradigm of extraction to show how this theory works. In particular, we will study in detail cases of subject extraction which we have so far not discussed. Here again, as in the previous section, we limit our discussion to violations of the ECP. We postpone discussion of Subjacency violations until section 4.7. For the moment, we will follow Lasnik and Saito (1984) and Chomsky (1986) and assume that, as we will justify in section 4.7.1, intermediate traces of argument-operator chains ([-HR] chains which terminate in a [+θ] position) can, in fact must, delete. On the other hand, we assume that traces of adjunct chains ([-HR] chains which terminate in a [-θ] position) and traces of [+HR] chains cannot delete. We will use examples from French, because, as we will show, the distribution of the *que/qui* phenomenon will provide crucial support for the theory of the ECP that we have developed. Clearly, however, we do not intend our theory to be language-specific; the reader can verify that it does extend to other languages, among which is English. We will conclude this section with a brief discussion of complementizer agreement in a variety of languages and show how our theory makes correct predictions in a number of these languages.

4.4.1 Extractions out of complement sentences

4.4.1.1 Subject extraction

Consider a typical case of subject extraction out of an interrogative sentence:

(98)
a.Qui crois-tu qui est parti?
  Who do you think who has left

b.[$_{CP}$ qui [$_{IP}$ t_1 [$_{IP}$ crois-tu [$_{CP}$ t_2[$_{C'}$ qui [$_{IP}$ t_3 est parti]]]]]]

There are several conditions that must be met for this sentence to be good. First, the subject trace must be antecedent-governed and head-governed. Recall that in our view h-government is a condition on ANTECEDENT government (as defined in the last section) and that both occur only under strict c-command. In (98), t_2 is not properly head-governed by I since it is not within the first projection which includes I. Consequently, the only available candidate for head-governing the subject trace is C°.

Since Pesetsky (1982), the *que/qui* phenomenon has been viewed as a phenomenon of complementizer agreement which allows a subject trace to meet the ECP.[53] Depending on theories of the ECP, agreement of C° has

---

53. In Pesetsky's analysis, the *que/qui* phenomenon allowed an escape from the NIC.

been interpreted as a way to render C° invisible (to Minimality, for instance) or to lift C° to the rank of appropriate antecedent-governor. In Rizzi's approach, C° agreement "reinforces" the C° head so as to make it an appropriate head-governor.

In the theory we have developed, the derivation given in (98) apparently violates both the h-government and the a-government requirement. Recall that we have assumed that IP is a barrier because it is not L-marked by C°. Thus, in (98) IP is a barrier both for a-government and for h-government of t_3 in the Spec of IP, violating condition (iii) of the ANTECEDENT government definition. The failure of a-government (condition (i)) can be avoided, however, by a further adjunction to IP; the corresponding derivation would be as in (99):

(99) [CP qui [IP1 t_1 [IP1 crois-tu [CP t_2[C' qui [IP2 t_3 [IP2 t_4 est parti]]]]]]]

In (99), t_3, adjoined to IP2, a-governs t_4, the subject trace. But since IP2 remains a barrier, t_4 cannot be properly h-governed by C°, from which it is separated by a barrier.[54] The ECP is thus violated, and the derivation in (99) is ruled out.

To solve this apparent problem, we propose that C°, when reinforced by agreement, can serve as an L-marker for the IP projection and can thus void the barrierhood of IP. As we see it, this possibility is limited: we propose that it is only by virtue of establishing a "chain of

---

54. Note that the t_3 adjoined to IP will be properly h-governed by C°, since the upper segment of IP is not a barrier for t_3: IP does not dominate t_3 and thus cannot count as a barrier for this trace.

agreement" that the barrierhood of IP can be voided. In (98) we have the following situation:

-t_1, the intermediate trace in the Spec of CP, induces Spec-Head agreement so that C^o agrees with t_2.

- t_1 agrees with t_2, the original trace of the subject, by virtue of being part of the same chain.

- t_2 agrees with INFL, of which it is the subject.

- consequently, by transitivity, INFL agrees with C. Since under standard assumptions maximal projections inherit the features of their heads, we conclude that CP agrees with IP. We propose that it is only under these conditions that the barrierhood of an IP can be voided.

(100) The agreement principle

Given the structure:

$[_{XP} \ldots X \ldots [_{YP} \ldots Y \ldots]]$

If X and Y agree, then YP is not a BC for any α dominated by YP

We will assume that, although not visible, the same mechanism is available abstractly in English and other languages.[55]

---

55. In our terms, the *que/qui* effect is the reflection of a very important process. Strong empirical support for this view comes from an extremely interesting fact in Haitian Creole. Like many creole languages, the French-based creole from Haiti has not maintained any of the morphological inflections of French (except for plural markings?). For instance, it has not preserved clitics or verb or adjective agreement. But HC, strikingly, has preserved the *que/qui* effect, and uses it as in French, to prevent an ECP violation. See Lefebvre et al.(1983??) for a study of this phenomenon. See also Deprez (1988) for a more recent approach. This provides strong support for the idea that the *que/qui* alternation is not just a superficial morphological process, but that it has deep and interesting motivations.

Recall that in our terms the Spec of CP is a [+HR] position. In the derivation in (98) this causes no particular problem: the barrierhood of IP is eliminated under the Agreement Principle so that movement can operate directly from one [+HR] position, the Spec of IP, to another [+HR] position, the Spec of CP. There is thus no violation of the improper movement condition. Adjunction to IP is not required for a-government to obtain: since IP is no longer a barrier, t_2 in the Spec of CP will directly a-govern t_3 in the Spec of IP. Note, in fact, that a derivation which would involve first an adjunction to IP and then movement to the Spec of CP, such as the one given in (99), violates the improper movement condition. It is thus excluded.

Consider now the case in which the *que/qui* alternation has not occurred:

(101) *Qui crois-tu que est parti?
Who do you think that has left

Taking the presence of *que* to be a reflection of the fact that the WH-extracted subject has not moved through the Spec of CP, we obtain the following structure :

(102) [$_{CP}$ qui crois-tu [$_{CP}$ t_1 [$_{CP}$ que [$_{IP}$ t_2[$_{IP}$ t_3 est parti]]]]]

Recall that in our approach adjunction to CP and IP is free. In this case, since IP is a barrier, we can assume that adjunction to IP is obligatory. Why is this sentence excluded? t_3 is a-governed by t_2, which itself is a-governed by t_3 and h-governed by C^0. Note, however, that since IP is not L-marked, it remains a barrier for the h-government of the original trace of the subject t_3. This trace thus violates the ECP, and being the original trace of the subject, it

cannot be deleted at LF.[56] (101) is thus appropriately ruled out by the ECP.

The agreement principle we have proposed, although quite natural, seems at first to be a stipulation needed only to account for the *that*-t effect. We will argue, however, that this is not the case. Recall that we have assumed so far that the intermediate functional projections TP and AGRP-O are never inherent barriers. We can in fact now unify these assumptions with the Agreement Principle and propose the following definition of L-marking:

(103) α L-marks β iff

(i) α c-commands β and

(ii) α is a lexical X^o or

(iii) if α is a functional X^o, then α agrees with δ, δ the head of β

Following Chomsky (1986), we assume that all X^o which are part of the verbal inflectional system share an index as a result of either Head-to-Head movement or affix lowering. Moreover, we assume that intermediate auxiliary verbs share some index with other auxiliaries and the past participle, as proposed by Guéron and Hoekstra (1988). Given (103), it follows that TP and AGR-O will never be barriers, since either the verb moves to the head of their projection or else T or AGR-O lower to the verb. This definition of L-marking also provides a trivial answer to the question why VP is not an inherent barrier.

---

56. We will argue in the next section that the mechanism of trace deletion should be maintained, contrary to Rizzi's assumption.

Moreover, the assumption that C° cannot be an L-marker unless Spec-Head agreement has occurred, allowing C_o in turn to share features with the head it dominates, namely AGR-S, in fact simply follows from our definition of L-marking. Thus, the agreement principle is subsumed under the definition of L-marking, which appears to be independently required to ensure that the projections of T and AGR-O are not barriers even in a language such as English, where the verb does not raise and where the movement of the subject from the VP-internal position to the Spec of IP must be allowed. [57]

Before we turn to other cases of extraction, we would like to point out an interesting consequence of our assumptions for other cases of subject extraction. Returning to the complex inversion construction we discussed in section 4.1.1, we can now raise and answer a question which up to now has remained unanswered in most analyses currently offered in the literature. The question is the following: why can an object NP not occur in the Spec of CP? In other words, why is the sentence (104) excluded under the structure (105):

---

57. There is some support from historical linguistics that the *que/qui* phenomenon indeed involves some agreement. It is known that the form *qui* derives from the forms *que + il*, i.e., a complementizer associated with a subject clitic. This clearly resembles the processes of clitic inversion and complex inversion, which are still alive in the language in main clauses. There is, however, an interesting difference. The subject clitic adjoined to the V in C in main clauses has maintained a full agreement paradigm, showing person, gender, and number specifications. In embedded clauses, however, the form is frozen in the third person singular. In our view this is related to the fact that while the movement of the V in main clauses allows Case-marking in the Spec of CP, leading to the complex inversion cases discussed in section 4.1.1, the poverty of the complementizer inflection in embedded clauses does not permit the occurrence of an overt NP; that is, *qui* cannot Case-mark the Spec of its projection.

(104)
**Comment Paul-a-t-il Marie rencontré?
How Paul did he Mary meet
How did Mary meet Paul

(105) [$_{CP}$ Comment [$_{CP}$ Paul$_i$ [$_{C'}$ a-t-il [$_{IP}$ Marie rencontré t_i]]]]

In (105), the object NP *Paul* has moved into the Spec of CP. This severely ungrammatical sentence is ruled out in our view on several counts: 1) since the Verb in C does not agree with I (I agrees with the subject NP), IP will be a barrier. Thus, 2) Paul does not a-govern its trace, and 3) the V in C does not a-govern its trace, either. 2) could be fixed by assumimg that the NP *Paul* first adjoins to IP before moving to the Spec of CP. This derivation, however, leads to a violation of the improper movement condition, since the object NP moves from an adjoined position to the Spec of CP. It is thus also trivially excluded. The reader can easily verify that under any derivation, some violation of the ECP, the improper movement condition, or the Case filter (if the verb agrees with the subject and not with the object) will occur. Thus the sentence will always be ruled out. The discussion of sentences such as (104) points to an important consequence of our hypothesis, which we will see confirmed in the following sections of this chapter: in languages such as French and English (and more generally (see section 4.5.3)), only the subject sitting in the Spec of the functional-projection complement of CP can ever move to the Spec of CP. This conclusion is confirmed in French by the obligatory locality of the *que/qui* effect.[58] Consider the

---

58. We are not claiming that this must be the case in every language. Suppose in fact that in some languages, movement of the V to C is

following sentences:

(106)
a. Qui as-tu dit que Pierre pensait qui devrait venir à sa soirée
Who did you say that Pierre thought who would come to his party

b. *Qui as-tu dit qui Pierre pensait qui devrait venir à sa soirée
Who did you say who Pierre thought who would come to his party

In (106)b., the *que/qui* alternation has occurred twice, first in the lower CP, and second in the intermediate CP. This sentence is sharply excluded, while (106)a. is perfect. It is thus clear that the ungrammaticality of (106)b. must come from the second application of the *que/qui* alternation in the intermediate CP. In our view, (106)b. is excluded by the condition on improper movement: although the first *que/qui* alternation is required to satisfy the h-government requirement on the original subject trace, no such requirement forces the second application. In our view, once the subject has moved out of the lower CP it must proceed through successive adjunction, first to VP (since VP is a D-Minimality barrier), and then to IP. Since the I of the intermediate embedded clause does not agree with the moved constituent, the barrierhood of the intermediate IP cannot be voided by L-marking by C. Consequently, the extracted subject must adjoin to the intermediate IP, and further movement to the Spec of CP, a [+HR] position, constitutes a case of movement from a [-HR] position (IP-adjunction) to a [+HR] position (the Spec of CP), which is ruled out

---

a substitution, rather than an adjunction to the head of C as we have assumed to be the case for French and English (see section 4.1.1). In this case, the V in C could L-mark the projection of C directly under clause (i) of the L-marking definition. A number of consequences would follow, among which movement of an object to the Spec CP. VSO languages are possibly instances of this type of language, and maybe also certain Germanic languages manifesting the V/2 constraint.

as a case of improper movement. In (106)a., the subject has simply adjoined to the intermediate CP, so no improper movement occurs. Thus, we correctly predict the necessarily local character of the *que/qui* alternation.

4.4.1.2 Object extraction

(107) Quel livre crois-tu que Jean a acheté?
Which book do you think that Jean has bought

(108)[$_{CP}$ Quel livre [$_{IP1}$ t_1[$_{IP1}$ crois-tu [$_{CP2}$ t_2 [$_{C'}$ que [$_{IP2}$ t_3 [$_{IP2}$ Jean a [t_4[$_{VP}$ acheté t_5]]]]]]]]]

Let us follow carefully the derivation represented in (108). As we have assumed, VP is not a inherent barrier, since it is L-marked by I. But VP is a barrier by D-Minimality and thus needs to be adjoined to. IP2 is a barrier which can be escaped by adjunction. Next, CP2 inherits barrierhood from IP2, which is a BC even after adjunction, given our definition of inheritance in section 4.2.2. Thus, movement to the Spec of CP or adjunction to CP is required. But recall that under our definition of [+HR]/[-HR] positions, the Spec of CP is a [+HR] position. If the WH-element moves into the Spec of CP2, as represented in (108), this will be a case of improper movement: the trace t_4 in the [-HR] position adjoined to VP will be dominated by t_2 in the Spec of CP2, a [+HR] position. We conclude that the derivation represented in (108) is excluded.

There is, however, an alternative possible derivation, which involves the adjunction of the extracted object to CP. This derivation is represented in (109):

(109) [CP Quel livre [IP t_1[IP crois-tu [CP t_2 [CP que [IP t_3 [IP Jean a [VP acheté t_4]]]]]]]]

In (109), both a-government and h-government of every trace is met: t_4 is a-governed by t_3 and h-governed by V =*acheté*; t_3 is a-governed by t_2 and h-governed by C° = *que* since the upper segment of IP is not a barrier; t_2 is a-governed by t_1 and h-governed by V = *croire*; finally, t_1 is a-governed by *quel livre* and h-governed by C° = [+WH]. (109) thus represents the correct derivation of (107) in our view. In the theory we propose, extraction of a WH-complement never passes through the Spec of CP. In our view, this follows from the condition on improper movement. This is an improvement over previous theories for the following reasons: as we noted earlier, in Chomsky's approach to WH-movement there is a curious imbalance between adjunction and substitution. WH-movement is viewed as involving adjunction most of the time (to VP and any other barrier except CP), except when it passes through the Spec of CP, where it involves substitution. This is a consequence of the fact that in Chomsky (1986) the Spec of CP is viewed as an A'-position in the LGB sense. In our theory, this imbalance does not occur : WH-movement involves strict adjunction in all of its intermediate steps.[59]

---

59. Except for subject extraction, which, as seen above, involves first substitution and then adjunction. This is, however, a case of [+HR] movement followed by movement to a [-HR] position. As we will see in section 4.6.2 this succession is allowed quite freely.

In addition to allowing this conceptually pleasing uniformity in movement ([-HR] movement is adjunction, [+HR] movement is substitution), our view receives strong empirical support from the *que/qui* phenomenon. Under standard assumptions, the *que/qui* effect is interpreted as an agreement effect with a trace in the Spec of CP. The natural assumption is that this agreement is blind to the type of trace which is sitting in the Spec of C°. With this in mind, consider the following sentence:

(110) *Quel livre crois-tu qui Jean a acheté
Which book do you think who Jean has bought

(110) is sharply ungrammatical. Under the standard view, which assumes that WH-extraction passes through the Spec of CP, this fact remains unexplained. It has been commonly assumed that agreement with a trace in the Spec of CP is optional. This accounts for the fact that the *que/qui* effect is not needed with complement extraction. But it does not account for the fact that in cases of extraction other than subject extraction the *que/qui* alternation is in fact impossible.[60] In

---

60. One possible alternative would be to view the impossibility of *que/qui* agreement as a consequence of the Principle of Economy. In cases of non-subject extraction, C° agreement is not necessary, since the ECP can be met without it. The *que/qui* alternation can thus be considered as an unecessary additional step in a derivation, which is excluded by Economy. This reasoning, however, is only superficially correct. Recall that in Chomsky's *Barriers* theory, an overt C° is a Rigid Minimality inducer. Since *que* in French does not delete at S-structure, one must assume that some mechanism prevents a Minimality violation in cases of adjunct extraction at LF. One possibility is to assume that *que* deletes at LF, but there is no a priori reason why deletion of C° at LF should be considered less costly with respect to Economy than agreement with C° at S-structure.

Pesetsky (1989) argues that Economy is better viewed as some kind of Earliness Principle, the essence of which is to require satisfaction of UG principles as soon as possible. If this view

our view, there is an explanation for the impossibility of the *que/qui* alternation: (110) is ruled out as a case of improper movement. Our approach to the ECP and the hypothesis that the Spec of CP is a [+HR] position straighforwardly predicts the ungrammaticality of (110), which had so far remained mysterious.[61]

---

is correct, it is clear that the S-structure *que/qui* alternation would satisfy the ECP earlier than the LF deletion of C°. Under such a view, Economy should then require the *que/qui* alternation for adjunct extraction.

Under a view such as Chomsky's (1989), in which Earliness does not come into play, optionality and not impossibility of the *que/qui* alternation is expected. The fact that *that*-deletion is not required at S-structure seems to indicate that an Earliness account of Economy coupled with Rigid Minimality is too strong. *That* must delete at LF for the ECP to be met; *that* can delete at S-structure. If Earliness were correct, we would expect *that*-deletion at S-structure to be obligatory for adjunct and object extraction.

Note that given our theory of the ECP, this conclusion no longer follows. The fact that *that*-deletion is optional reflects a "true" optionality: structures with an overt C° and those with an empty C° are equivalent for Economy under either Chomsky's view or Pesetsky's Earliness view, since *that*-deletion is irrelevant for the satisfaction of the ECP for adjunct extraction.

In Rizzi's approach, in which an overt complementizer does not count as a Minimality inducer for any maximal-projection movement, the Economy view may hold. The theory remains too vague with respect to the barrier calculus to lead to any precise conclusion for the impossibility of the *que/qui* alternation with non-subject extractions. In other words, there is no account of the impossibility of the *que/qui* alternation which follows directly from any theory of the ECP so far proposed.

61. But cf. Taraldsen (1986) for an account in terms of Binding theory.

4.4.1.3 Adjunct extraction

(111)
Comment crois-tu que Jean se comportera??
How do you think that Jean will behave

The structure of (111) is given in (112):

(112) [$_{CP}$ Comment [$_{IP}$ t_1 [$_{IP}$ crois-tu [$_{CP}$ t_2 [$_{CP}$ que [$_{IP}$ t_3 [$_{IP}$ Jean se [$_{VP}$ comportera t_4]]]]]]]]

Here t_4 is the original trace of the manner adverb. Adjunction to IP and CP is allowed, thus there are no barriers and a-government is satisfied. Note also that here WH-movement involves pure adjunction in its intermediate steps. We assume with Rizzi that proper h-government of the original trace of the adverb t_4 is met by the functional projection immediately dominating the adjunction site of the adverb.[62] Moreover, t_3 is h-governed by the C° *que*, t_2 is h-governed by the matrix verb, and t_1 is h-governed by the empty C° of the matrix complementizer. H-government is thus satisfied.

As with object extraction, we predict that if the extracted adverb were to land in the intermediate Spec of CP, a violation of the improper movement condition would occur. This prediction is confirmed by the ungrammaticality of (113):

---

62. Given the more articulated structure that we have assumed for what is here abbreviated as IP, the h-governor of the trace of the adverb will more appropriately be T if the adverb is adjoined to AGR-O, or Agr-O is the adverb is attached to VP.

(113)
*Comment crois-tu qui Jean se comportera?
How do you think who Jean will behave

4.4.2 WH-islands

Let us now turn to WH-islands.

4.4.2.1 Subject extraction

(114)
*Qui te demandes-tu où va loger?
Who do you wonder where will live

(115) [CP Qui [IP t_1 [IP te demandes-tu [CP t_2 [CP où[CPt_3 [C'[IP t_4 [IP t_5 va loger]]]]]]]]]

In subsection 4.2.2.1, we argued that in cases such as (114) the WH-element does not occur in the Spec of CP; the subject is thus free to move into it. Given this posssibility, we apparently wrongly predict that the extraction of a subject will lead to a mere Subjacency violation, since the ECP is met for the original trace of the subject. To correct this, we could assume that unlike the [-WH] complementizer *que*, a [+WH] complementizer is not an agreeing complementizer. If so, IP would remain a barrier and t_4 would violate the ECP, since it is neither h-governed nor a-governed and can thus not be ANTECEDENT-governed. This solution, however, appears somewhat stipulatory. Moreover, there is empirical evidence against it: as is well known,

certain dialects of French, in particular the Québecois dialect, allows for apparent doubly-filled complementizers. Thus, (116) is a good sentence in the Québecois dialect:

(116) Je me demande où que tu vas loger
I wonder where that you will live

Note that the second complementizer has the form of a [-WH] complementizer in both standard French and in Québecois, so we expect complementizer agreement to be possible. That this must be the case is confirmed by the following sentence:

(117) Qui qui a telephoné
Who that telephoned

The first *qui* in (117) is a WH-phrase, while the second *qui* is the agreeing complementizer.[63] Thus (117) shows clearly that the second complementizer can agree. Let us now return to WH-islands. Consider the example in (118):

(118) *Quelle fille te demandes-tu où qui va loger?
Which girl do you wonder where who will live

In (118), the *que/qui* alternation has occurred. The sentence, however, remains impossible.[64] This shows that the *que/qui* agreement does not

---

63. See Lefebvre (198..) for an interesting study of this phenomenon in Québecois.

64. Thanks to Robin Coté for the judgments. One must be careful to distinguish the *que/qui* effect from a resumptive reading, which makes the sentence perfect.

(i) Quelle fille est-ce-que tu te demandes où qu'a va loger?
Which girl do you wonder where that she will live

In (i), *a* is a resumptive pronoun for the extracted subject.

permit subject extraction from a WH-island. We must then assume that even in standard French, where the *que/qui* effect is not apparent, the empty C may undergo agreement. If so IP can be L-marked. What then accounts for the ungrammaticality of (114)? Recall that in our view, the Spec of CP is a [+HR] position. As we will argue in section 4.7.1, traces in [+HR] positions cannot delete. Now, t_3 in the Spec of CP is separated from its antecedent t_2 by the WH-phrase adjoined to CP: Operator Opacity is violated and the trace in the Spec of CP cannot be a-governed. Since the Spec of CP is a [+HR] position, the trace there cannot delete, and the ECP is violated.

4.4.2.2 Object extraction

(119)
a. ?Quel livre te demandes-tu où acheter?
Which book do you wonder where to buy

b. [CP quel livre [IP t_1 [IP te demandes-tu [VP t_V[CP t_2 [CP où [C'[IP t_3 [IP PRO [VP acheter t_4]]]]]]]]]]

(119) is a typical case of an Operator Opacity violation, where t_3 fails to be a-governed by t_2 because of the intervening WH-element. But since (119) involves an argument chain, the ECP will not be violated because the trace can delete. This is a Subjacency violation which we discuss in more detail in section 4.7.2.

4.4.2.3 Adjunct extraction

(120)
a. *Comment te demandes-tu où Jean est parti?
How do you wonder where Jean has gone

b. [$_{CP}$ Comment [$_{CP}$ t_1 [$_{IP}$ te demandes tu [$_{CP}$ t_2[$_{CP}$ où [$_{IP}$ t_3 [$_{IP}$ Jean est [$_{VP}$parti t_4]]]]]]]]

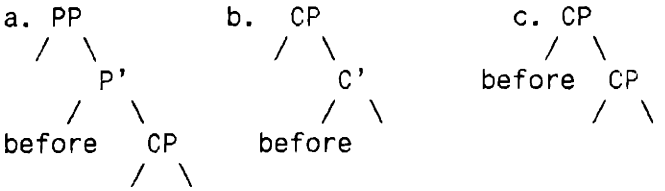
This is also a typical case of an Operator Opacity violation, where t_3 is not a-governed by t_2 because of the intervening WH-element. Since (120) involves an adjunct chain, no trace may be erased: consequently, (120) violates the ECP.

4.4.3 Adjunct Islands

4.4.3.1 The structure of adjunct clauses

The question of the structure of adjunct clauses has often been discussed. Three possible structures are given in (121):

(121)



Structure (121)b. can be easily rejected on the basis of French data. In French, adverbial complements take an overt complementizer both when they are tensed and also sometimes when they are untensed:

(122) Tensed

a. Jean est parti avant que Marie n'arrive
   Jean left before that Marie arrived

b. Jean est arrivé après que Marie ne soit partie
   Jean arrived after that Marie left

c. Jean est parti parce que Marie le lui a demandé
   Jean left because that Marie it of him asked

d. Jean est parti sans que personne ne le retienne
   Jean left without that anyone held him back

Untensed

a. Jean est parti avant de finir son travail
   Jean left before prt finishing his work
b. Jean est parti après avoir fini son travail
   Jean left after finishing his work
c. Jean est parti sans avoir fini son travail
   Jean left without having finished his work

There are also some empirical arguments against structure (121)a. In (121)a. the adverb is taken to be a head which takes a sentential complement. It has often been noted, however, that many types of adverbs, unlike their adjectival counterparts, are unable to take complements:

(123)

a. That man is proud of his daughter

b. *That man behaved proudly of his daughter

c. Cet homme est fier de sa fille (=a)

d. *Cet homme s'est comporté fierement de sa fille (=b)

This generalization appears to be contradicted precisely by the adverbs such as *before* and *after*, i.e., the type of adverbs which can precede adverbial adjunct clauses. Indeed, these adverbs superficially seem to be able to take both sentential and NP complements.

(124)

a. He left before you arrival

b. He arrived after your departure

c. Il est parti avant ton arrivée (=a)

d. Il est arrivé après ton départ (=b)

Because of these properties, these adverbs have often been analyzed as prepositions (Emonds (1976) Rouveret (1977) for French). There is, however, one problem with this proposal: in certain respects, these "adverbial prepositions" do not behave like other prepositions. This is most obvious in English. In English, most prepositions can be stranded after WH-extraction. Clearly, however, this is impossible in these cases.

(125)

*Which party did you leave before

*Which show did you call the manager after

*Which paper did you go on vacation without

Kayne (1983) analyzes the well-known impossibility of preposition stranding in French as involving a distinction between the properties of French and English prepositions. In his view, prepositions are h-governors in English but not in French. Thus, preposition stranding in French leads to an ECP violation. There are several ways in which this proposition could be reanalyzed.[65] Since preposition stranding is not

---

65. One possibility is to assume that preposition stranding requires incorporation (interpreted here along the lines suggested by M. Baker (1986)) of the prepositional head into the V. Under this view the adverbial prepositions like *before* and *after* could be barred from incorporating because they occur in some sense too far

our concern here, we will simply assume Kayne's proposal to be correct without trying to deduce it from deeper principles (See preceding footnote for a proposal). Note that under this view, we are led to the conclusion that "prepositions" like *before* and *after* are not h-governors. This explains why preposition stranding with these adverbial prepositions is never possible. Assuming a structure such as (121)a., there is no reason to assume that a-government cannot be satisfed in cases of WH-extraction, since the extracted constituent can either move to the Spec of PP or adjoin to the PP, as represented in (126):

(126) which party$_i$ did you leave [$_{PP}$ t'$_i$ [$_{PP/P'}$ before t$_i$]]

Further adjunction is permitted since in our view adjunction is free. Thus a-government is met. The failure of the ECP can only come from a failure of h-government. We thus assume, extending Kayne's hypothesis, that the adverbial prepositions are not h-governors, and thus that (121)a. is not the right structure.

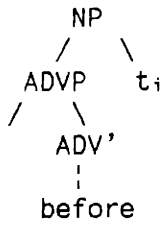
Alternatively, we could assume, along the lines of (121)c., that the apparent complement of these adverbial prepositions is not really a complement. Concretely, this would mean that these adverbial prepositions are in fact adverbs and occur adjoined to the clause or the NP they modify. The advantage of such a hypothesis is that when these adverbials occur alone, we do not need to posit an empty category in their complement position. If these adverbials were

---

away from the V; incorporation of these prepositions would thus lead to an ECP violation. This hypothesis allows the assumption that no preposition is ever a h-governor.

interpreted as prepositions, as in (121)a., they would have to be interpreted as prepositions which are sometimes "transitive" and sometimes "intransitive". Under the hypothesis that they are always adjoined, such an assumption is unnecessary . Moreover, the adjunction hypothesis provides an immediate explanation for why these adverbial prepositions never h-govern the NP they modify. Assuming, as is natural, that they head their own projections, the structure would be as in (127):

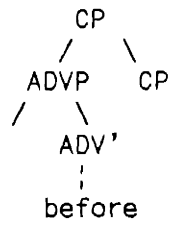
(127)



Under an intepretation of c-command which takes the first projection to be relevant, and not the first branching node, the adverb *before* could never c-command the trace of the NP, and thus it could never h-govern it. This accounts straightforwardly for the ungrammaticality of (125).

Returning now to the case of sentential complements, we can assume as well that these prepositional adverbs are adjoined to the clause, and thus that adjunct clauses have the structure in (121)c., or more exactly the structure in (128) below:

(128)



We will assume this structure to be correct and turn to the properties of extractions out of these adverbial adjunct clauses.[66]

4.4.3.2 Subject extraction

(129)
a. *Qui es-tu parti avant que n'arrive
Who did you leave before that arrived

[CP qui [IP t_1 [IP es-tu parti [CP t_2[CP avant [CP que [IP t_3 [IP t_4 n'arrive]]]]]]]]

In the derivation given in (129), landing in the Spec of CP has not occurred. Thus, the original trace of the subject t_4 is not h-governed and the ECP is violated: but exclusion of derivations such as (129) is not sufficient to account fully for the ungrammaticality of subject extraction out of adjunct clauses. Given the structure we have assumed

---

66. Although the following section will be developed with the assumption that adjunct clauses have the structure in (128), nothing crucial relies on this hypothesis. The structure in (121)a. could be assumed as well and everything would remain unchanged, as the reader can check for himself. The only important assumption for our account is that adverbs such as *before*, *after*, and *without* are not proper h-governors, an assumption which is independently motivated by the preposition-stranding facts in (125).

for adjunct clauses, we need to consider a second derivation, namely a derivation in which landing in the Spec of CP has occurred. First, note that the *que/qui* alternation does not improve the sentence.

(130)
*Quelle fille es-tu parti avant qui n'arrive
Which girl did you leave before who arrived

The derivation of (130) is given in (131):

(131)

[$_{CP}$ quelle fille [$_{IP}$ t_1 [$_{IP}$ es-tu parti [$_{CP2}$ t_2[$_{CP2}$ avant [$_{CP2}$ t_3 qui [$_{IP2}$ t_4 n'arrive]]]]]]]

In (131), landing in the Spec of CP has occurred, triggering complementizer agreement and voiding the barrierhocd of IP2 under L-marking. Thus, t_4 is properly governed and the ECP is satisfied. Again, we seem to predict wrongly that (130) should be only a Subjacency violation. Recall, however, that we have assumed that the Spec of CP is a [+HR] position and that traces in [+HR] positions cannot delete. We thus also need to check if t_3 in the Spec of CP meets the ECP. Clearly, however, this is not the case. As we argued in the above section, the adverbial *avant* cannot serve as a h-governor for t_3. Moreover, since CP2 is not L-marked by the matrix INFL (the head of CP2 does not agree with INFL), it is a barrier. Consequently, INFL will not be able to h-govern the trace t_3 in the Spec of CP. Thus the ECP is violated and the ungrammaticality of (131) is accounted for. Moreover, given our assumption that adverbs introducing adjunct clauses are OPERATORS (in the sense given in section (xx)), t_2 will not a-govern t_3 in the Spec of IP. Here again the ECP is violated. Note that the hypothesis that the Spec of CP is a [+HR] position is

crucial to this account. If we are correct in assuming that the structure of adjunct clauses is indeed as in (128),[67] then we know of no other theory of the ECP which is able to account for the fact that the *que/qui* alternation does not improve the extraction of a subject out of an adjunct clause.[68] In our view, however, this follows straightforwrdly from our hypothesis with no additional stipulation. This in turn provides strong support for our view.

4.4.3.3 Object extraction

(132)
a. ??Quel livre es-tu parti avant de lire
Which book did you leave before reading

b. [CP Quel livre [IP t_1 [IP es-tu parti] [CP t_2[CP avant [C' de [IP t_3 [IP PRO [VP lire t_4]]]]]]]]

(132), in our view, is another typical case of an Operator Opacity violation. We have proposed that the adverbial adjoined to CP may count as an OPERATOR for the extracted WH-element. This leads to a

---

67. Incidentally, note that the same conclusion follows in our view if the structure of adjunct clauses is assumed to be as in (121)a. However, as we have shown, structure (121)b., which would exclude the *que/qui* effect altogether, is not possible for French.

68. Note that in this case, an analysis under which the *que/qui* alternation is precluded with objects extraction in declarative sentences because of the Economy principle (cf. footnote above) cannot account for why the *que/qui* alternation, although not impossible, has no effect on the grammaticality of sentences such as (130). As conceived by Chomsky (1989), Economy considerations are always secondary to the satisfaction of UG principles such as the ECP. Thus the que/qui alternation is in fact expected to occur and to reduce the violation in (131) to a subjacency violation, an incorrect prediction.

failure of a-government: the trace t_3 adjoined to IP is not a-governed by t_2. As we will argue in the next section, intermediate traces of WH-extracted arguments can delete: thus in this case, no ECP violation will result. We take up the violation of Subjacency in section 4.7.[69]

4.4.3.4 Adjunct extraction

(133)
a. Comment est-il parti avant que Marie n'arrive
   How did he leave before Marie arrived

b. [$_{CP}$ comment[$_{IP}$ t_1 [$_{IP}$ est-il parti [$_{CP}$ t_2 [$_{CP}$ avant [$_{C}$ que [$_{IP}$ t_3 [$_{IP}$ Marie [$_{VP}$ n'arrive t_4]]]]]]]]]

As with the case of object extraction, t_4 is a-governed by t_2. It is also properly h-governed by I. But for t_3, a-government fails due to the intervention of the adverbial *avant*, an Operator Opacity inducer.

---

69. There appears to be some support in French for the idea that the adverb may be some kind of operator on the adjunct clause, since it affects the clause it modifies. First, these adjunct clauses must take the subjunctive when tensed. Second, the *ne* of negation appears in these sentences without a negative sense:

(i) Jean est parti avant que tu n'arrives.
Jean left before you arrived

This negative particle has been called by traditional grammarians an "expletive" negation. The distribution of this expletive negation remains somewhat obscure; note, however, that it appears in the complement of verbs with negative meaning.

(i) Je doute qu'il ne vienne
I doubt that he will come

In other words, its distribution is somewhat related to the distribution of negative polarity items. This provides evidence that some type of quantification may be occurring, which in turn provides support for the analysis of these adverbs as operators which induce Operator Opacity effects.

In this case we have an adjunct chain, so no trace may be erased. Consequently the ECP is violated.

4.4.4 Subject islands

4.4.4.1 Subject extractions

(134)
*Quel homme est-ce (le fait) qu'a rencontré Marie ennuie Pierre?
Which man (the fact) that met Marie bores Pierre

As in the above case of subject extraction out of adjuncts, there are in fact two derivations we need to consider and rule out: a derivation in which the extracted subject has not moved into the Spec of CP and a derivation in which the extracted subject has moved into the Spec of CP. Let us consider the first possibility. The structure is given in (135):

(135) [$_{CP1}$ quel homme est-ce que [$_{IP}$ t_1 [$_{IP1}$ t_2 [$_{CP2}$ t_3 [$_{CP2}$que [$_{IP2}$ t_4 [$_{IP2}$ t_5 a rencontré Marie]]] ennuie Pierre]]]]

In (135), the extracted subject has first adjoined to IP_2, then to CP_2, without going through the Spec of CP. This derivation is ruled out as a straightforward ECP violation. Since passage through C^o has not occurred, IP2 remains a barrier, and the trace t_4 in the Spec of IP2 will not be h-governed. Since t_4 is the original trace, it will not delete at LF.

Let us turn now to the second derivation, where the extracted subject has moved through the Spec of CP. As shown in (136), subject extraction does not get any better, despite the fact that the *que/qui* alternation has applied:

(136)
*Quel homme est-ce que [qui a rencontré Marie] ennuie Pierre
Which man that who met Marie bores Pierre

(136) has the structure in (137):

(137) [$_{CP1}$ quel homme est-ce que [$_{IP1}$ t_1 [$_{IP1}$ [$_{CP2}$ t_2 [$_{CP2}$ t_3 [$_{C'}$ qui [$_{IP2}$ t_4 a rencontré Marie]]] ennuie Pierre]]]]

In (137), since the *que/qui* alternation has applied, IP2 is not a barrier. The Agreement Principle is satisfied: the trace in the Spec of CP2 agrees with C°, which in turns agrees with the head of IP2. Thus, t_4 can be both properly a-governed by t_3 in CP2 and properly h-governed by *qui*. The ECP is thus satisfied, and the sentence should be only a Subjacency violation. Clearly, given the severe ungrammaticality of (136), this is the wrong result.

We see two possible ways of solving this problem. First, we could assume that the Spec of CP is in fact occupied by some invisible element which is the silent equivalent of NPs such as *the fact*, commonly appearing with subject sentences. In this case, passage through the Spec of CP would be forbidden, so the only possible derivation would be (135). As we have seen, this derivation is ruled out by the ECP. Despite the plausibility of this solution we will not adopt it.

The second way, which we will adopt, is to use our assumption that the trace in the Spec of CP, t_3 in (137), cannot delete. Let us observe the result that this assumption will have. t_3 is dominated by CP2, which is a barrier, since it is not L-marked. Thus t_3 will fail to be both a-governed and properly h-governed. Since, as we have assumed, CP can be adjoined to, the a-government failure can be avoided by a further adjunction to CP2. This further adjunction is represented in (137) by the trace t_2. This does not help, however: t_3 still fails to be h-governed. Recall that the definition (xx) of h-government disqualifies the I of the matrix as a potential h-governor for t_3; t_3 is not contained in the first projection dominating I, namely I'. The C^o of the matrix sentence cannot h-govern t_3, either, because it is separated from this trace by the IP1 barrier. IP1 is a barrier by inheritance from CP2 and because the agreement principle is not satisfied: the WH-element in CP1 does not agree with the head of IP1, since it is not its subject.t_3 is thus not properly h-governed. If, as we have proposed, t_3 cannot delete, this will lead to an ECP violation. The non-deletion of a trace in the Spec of CP at LF is a consequence of our hypothesis that the Spec of CP is a [+HR] position and that traces in [+HR] positions do not delete. Nothing more needs to be added, and subject extraction out of subject sentences will be ruled out as ECP violations under either of the two derivations we have considered.

Note that this explanation carries over straightforwardly to English and the *that*-deletion phenomenon. Subject extraction from subject

islands is impossible in English as well. Parallel to the French cases, however, the deletion of *that* does not improve the sentence.[70]

(138)

a.*Which man$_i$ did Mary think that [that t_i will leave] bothers Mary.

b.* Which man$_i$ did Mary think that [t_i will leave] bothers Mary

(138)a. is a typical instance of the *that*-t effect. But in (138)b., *that* has deleted so that one would expect a-government to be satisfied for the original trace of the subject. Consider the structure of (138)b.

(139) ...thinks[$_{CP1}$ [c' that [$_{IP1}$ [$_{CP2}$ t_1 [c' Φ [$_{IP2}$ t_2...

In the *Barriers* framework, t_1 can a-govern t_2, so sentence (138) should lead to a mere Subjacency violation. Clearly, this is the wrong result. The same wrong result obtains in Rizzi's theory. Indeed, deletion of *that* in his view means that an agreeing complementizer occupies the head of C°. If so, h-government of the original subject trace will obtain, so again the sentence should violate only Subjacency. In our approach, the fact that *that*-deletion does not help follows straightforwardly from the assumption that the Spec of CP is a [+HR] position and that traces in [+HR] positions cannot delete and thus must satisfy the ECP.

---

70. The ungrammaticality of the English case may be ruled out independently because, as is known, the *that* in a subject clause must always be present. If, as suggested by Stowell, *that*-deletion is limited by the ECP, then no matter what happens with subject extraction the ECP will be violated. Note that the idea that the the deletion of the complementizer is governed by the ECP is in conflict with Chomsky's proposal that the deletion of *that* follows from the principle of Full Interpretation (see KNOL).

4.4.4.2 Object extraction

(140)
?? un livre que je pense que lire le soir donne des cauchemars
a book that I think that reading at night gives nightmares

(141)... [$_{CP}$ que [$_{IP1}$ t_1 [$_{IP1}$ [$_{CP2}$ t_2 [$_{CP2\Phi}$ [$_{IP2}$ t_3 [$_{IP2}$ PRO lire t_4 le soir]]]] donne des cauchemars]]]

In (141) t_4, the original trace of the object extraction, is both a-governed and properly h-governed. Consequently, the ECP is satisfied. The marginality of this sentence is due to a Subjacency violation. We discuss this in section 4.7.

4.4.4.3 Adjunct extraction

This case has been studied in detail in the previous section. We refer the reader to this previous discussion.

4.4.5 CNPC island

We will limit our dicussion to cases of nominal complements since, as we mentioned earlier, the case of relative clauses is in fact mostly parallel to cases of WH-islands.

4.4.5.1 Subject extraction

As above for cases of extraction out of adjunct islands and subject islands, we need to consider two derivations: one which involves no passage through the Spec of CP and leads to the ungrammatical sentence (142), and one which involves a passage through the Spec of CP, triggering C_o agreement, and leads to the ungrammatical sentence (143):

(142)
*Qui as-tu entendu la rumeur que allait déménager
Who did you hear the rumor that was going to move

(143)
*Qui as-tu entendu la rumeur qui allait déménager
Who did you hear the rumor who was going to move

Recall that we have assumed that the sentential complement is adjoined to the NP, following a proposal by Stowell (1981) and Grimshaw (forthcoming). We have also assumed, following a proposal by Belletti and Rizzi, that the upper segment of an adjunction could inherit barrierhood. We repeat here the definition given in 4.2.2.5:

(144) τ,a segment of an XP, is a barrier for β iff
τ dominates δ,
δ a non-L-marked BC for β.

The derivation of (142) is given below:

(145) [$_{CP1}$ qui [$_{IP1}$ t_1 [$_{IP1}$ as-tu entendu [$_{NP}$ t_2 [$_{NP}$ [$_{NP}$ la rumeur] [$_{CP2}$ t_3 [$_{CP2}$ que [$_{IP2}$ t_4 allait déménager]]]]]]]]

In (145), the Agreement Principle is not satisfied, since no trace in the Spec of CP triggers C° agreement. Consequently, C° cannot L-mark IP2, IP2 remains a barrier, and t_4 will fail to be properly h-governed. The ECP is thus violated.

The derivation of (143) is given below:

(146) [$_{CP1}$ qui [$_{IP1}$ t_1 [$_{IP1}$ as-tu entendu [$_{NP}$ t_2 [$_{NP}$ [$_{NP}$ la rumeur] [$_{CP2}$ t_3 [$_{CP2}$ t_4 qui [$_{IP2}$ t_5 allait déménager]]]]]]]]

In (146), passage through the Spec of CP has occurred; the Agreement Principle is satisfied so that the trace t_4 is a-governed and properly h-governed, since IP2 is not a barrier. But recall that, as we have assumed above (following Chomksy (1989)), traces in [+HR] positions cannot delete. We thus need to check whether the ECP is satified for t_3, the trace in the Spec of CP. Both the upper segment of the NP and CP2 are barriers. Consequently, t_3 cannot be properly h-governed by the matrix verb, and the ECP is violated. Here again an ECP violation occurs under either derivation of the subject extraction.

The proposed analysis has interesting consequences for English. As in the case of subject extraction from subject sentences, we predict that the deletion of the *that* complementizer should have no effect on the ungrammaticality of the sentence. This prediction is corroborated by the facts and has, to our knowledge, so far remained unexplained.

(147)

* who did you hear the rumor that left

* who did you hear the rumor left

This again follows from the hypothesis that the Spec of CP is a [+HR] position and that consequently a trace in the Spec of CP cannot be deleted.

It could be argued that the deletion of *that* is independently forbidden and that the severity of the ungrammaticality of the subject extraction is simply due to the added and independent marginality of the *that*-deletion. But consider the following paradigm:

(148)

a. ?Who did you hear the rumor that John saw

b. ?? Who did you hear the rumor John saw

c. *Who did you hear the rumor saw John

As we show below, in the case of object extraction out of CNP, the ECP is satified, so that in (148)a. and b. only Subjacency is violated. (148)b. is distinctly worse than (148)a.; this more severe marginality is no doubt due to the deletion of *that*. However, (148)b. remains distinctly better than (148)c. This contrast remains unexplained if it is assumed that (148)c. violates only Subjacency and whatever condition is assumed to prevent *that*-deletion in this context. Under our analysis, the contrast is expected, since (148)c. violates the ECP.

4.4.5.2 Object extraction

(149)
Qui crois-tu la rumeur que Jean a rencontré?
Who do you believe the rumor that Jean met

(150) [$_{CP1}$ qui [$_{IP1}$ t_1 [$_{IP1}$ crois-tu [$_{NP}$ t_2 [$_{NP}$ [$_{NP}$ la rumeur] [$_{CP2}$ t_3 [$_{CP2}$ que [$_{IP2}$ t_4 [$_{IP2}$ Jean a rencontré t_5]]]]]]]]]

In (150) the ECP is satisfied, since t_5 is a-governed by t_4 and properly h-governed by the V = *rencontré*. This case of object extraction is parallel to the case of object extraction out of subject islands.

4.4.5.3 Adjunct extraction

This case has been discussed in detail in section 4.2.2.4. We thus refer the reader to this section and do not repeat the analysis.

4.4.5.4 Conclusion

This concludes our discussion of the theory of the ECP.

Let us summarize briefly. In this section we have made two new proposals:

1) We have proposed a new definition of L-marking, which we repeat here for convenience:

(151)

α L-marks β iff

(i) α c-commands β and

(ii) α is a lexical X° or

(iii) if α is a functional X°, then α agrees with δ, δ the head of β

2) We have proposed that a trace in the Spec of CP does not delete because it is a trace in a [+HR] position. This hypothesis, which is a natural consequence of our assumption that the Spec of CP is a [+HR] position, allows us to explain why the *que/qui* alternation and *that*-deletion do not prevent an ECP violation in the case of subject extraction from all the cases of islands we have reviewed, a fact left unexplained in other approaches to the ECP. In turn, the elegance of this explanation provides strong support for our hypothesis that the Spec of CP is a [+HR] position. In the following section, we return to the problems we raised at the beginning of this chapter, and we show that the theory of ECP we have developed correctly accounts for these problems.

4.5 Consequences

4.5.1 Solutions

In the previous sections we outlined the theory of the ECP we propose. We saw that the hypothesis that the Spec of CP is a [+HR] position creates no particular problem for regular cases of extraction. On the

contrary, we showed that this assumption allows a natural explanation for the impossibility of the *que/qui* effect in object and adjunct extraction from interrogative clauses, a fact which up to now has remained unexplained. Moreover, we showed that typical cases of island violations are elegantly accounted for by our proposed definition of the ECP and that the fact that the *que/qui* effect does not prevent an ECP violation in cases of subject extraction from islands follows naturally from our assumptions.

It is time now to return to the three problems we raised at the beginning of this chapter (section 4.1.2) and to assess the consequences of our assumptions for a solution to these problems. We repeat them here for convenience.

(152)

A. The intermediate trace problem

B. The weak crossover problem

C. The long-distance NP movement problem

Problem A. is elegantly solved, as we have seen. As we proposed in chapter 2, the movement from an adjoined position, i.e., a [-HR] position, to a [+HR] position is a case of improper movement. In our theory, IP is a barrier and must be adjoined to in all cases of complement extraction and adjunct extraction. Thus in all these cases, landing in the Spec of CP is made impossible since the movement always proceeds from an adjoined position. Movement will involve an adjunction to CP instead. Movement to the Spec of CP is only allowed from a [+HR] position, namely the immediate subject position dominated by the CP to which movement occurs. Here movement does not violate the

improper movement condition since this is a case of movement from a [+HR] position to another [+HR] position.

As we noted in the previous section, this view makes the correct prediction that the *que/qui* phenomenon (and obligatory *that*-deletion) will only occur in the complementizer immediately dominating the subject extraction site. Further movement of a subject across a complementizer will involve adjunction to, not a passage through, the Spec of CP. Passage through the Spec of CP would in this latter case also violate the improper movement condition. Thus we correctly predict the grammaticality of (153)a. and the ungrammaticality of (153)b.

(153)
a.Qui as-tu dit <u>que</u> Pierre pensait qui devrait venir à sa soirée
Who did you say that Pierre thought who would come to his party

b.*Qui as-tu dit <u>qui</u> Pierre pensait qui devrait venir à sa soirée
Who did you say who Pierre thought would come to his party

Problem B. also has an easy solution. We have come to the conclusion that long-distance movement (except for subjects) always involves an adjunction to CP. Consequently, there will be no intermediate trace in a [+HR] position which can serve to bind a pronominal and repair potential WCO violations. The question thus no longer arises for adjunct and object extractions.

With regard to subject extraction, the question is slightly different. Subject extraction, as we have shown, does involve passage through the Spec of CP. The only case where this could create a problem would be

that in which the extracted subject binds a pronoun in an adverbial expression adjoined to IP. Let us consider such a case:

(154)
Qui$_i$ crois-tu qui, avant que sa$_i$ mère ne vienne, va ranger sa chambre.
Who do you think who, before his mother comes, will clean his room

The prediction of our approach is that (154) should not be a WCO violation. It seems to us that this is correct and that the possessive pronoun *sa* can be bound by *qui*. However, the adverbial is clearly a parenthetical expression. It remains unclear how binding functions in this type of expression; the fact that the coreference is quite may be due to independent factors. A study of the binding properties in parentheticals is beyond the scope of this dissertation; we will thus not discuss it any further. So although (154) does not provide reliable evidence for our point of view, it does not contradict it, either.

We have not yet discussed problem C. We turn to it immediately.

4.5.2 Super-raising

Let us first consider regular cases of Super-raising such as (155).

(155)
* Jean semble qu'il a été considéré être intelligent.
Jean seems that he was considered to be intelligent

(156) [$_{IP1}$ Jean semble [$_{CP}$ t_1 [$_{CP}$ que [$_{IP2}$ il a été [$_{FP}$ t_2 [$_{VP}$ t_3 considéré [$_{IP3}$ t_4 être intelligent]]]]]]]

Recall that in section 3.3.5.1, we opted for a version of Minimality which we have called Dynamic Minimality. Since we are concerned here with NP movement, D-Minimality is relevant. Consider first the movement of the NP from its original subject position to the intermediate clause. There are two cases to consider: 1) a case where the past participle has raised to some functional projection, presumably AGR-O,[71] and 2) a case where the past participle remains inside the VP. In the first case, since we assumed that the trace of a moved head does not count as a D-Minimality inducer, VP will not be a barrier and the NP can move to the specifier of the functional projection to which the past participle has moved. In the second case, VP will remain a barrier. Note, however, that in (156), as in any case in which NP movement is allowed, the Spec of VP is not a θ-position. Thus, movement to the Spec of VP is possible and D-Minimality can be circumvented. The second case is the one represented in (156), with t_3 in the Spec of VP. Consider next the movement from the Spec of VP or the Spec of AGR-O to the Spec of CP. In (156), since the head of AGR-S is occupied by a lexical auxiliary, AGRP-S is a D-Minimality barrier for t_3. Moreover, since the specifier of AGR-S is occupied by the overt expletive, landing in this Spec is impossible and the D-Minimality barrier cannot be escaped. Consequently, the trace t_3 cannot be antecedent-governed. Since t_3 is in a [+HR] position, it cannot delete, and the ECP is violated. This accounts straightforwardly for the ungrammaticality of (155).

---

71. See Chapter II 2.4.2, for arguments that this is possible in French.

Although this account of super-raising follows simply from our assumptions, it raises a number of delicate and interesting questions about the relevant definition of D-Minimality. Consider a language in which the verb does not raise to AGR-S at S-structure, such as English. If every functional projection counts as a D-Minimality inducer, then an account of super-raising along the lines of the one given for the French cases is available in English as well: AGR-S will induce a D-Minimality barrier which cannot be escaped, since its specifier is occupied by the expletive. On the other hand, if we assume a definition of D-Minimality in which only the the lexically filled heads count as D-Minimality inducers, D-Minimality will not account for the English cases of super-raising. It thus appears that an account of super raising forces a choice in the definition of D-Minimality toward an approach in which all functional heads, whether lexically filled or not, induce a D-Minimality barrier. Recall, however, that in our discussion of the definition of D-Minimality in section 3.3.5.1, a crucial hypothesis is that traces of heads are never D-Minimality inducers. Given this hypothesis, which is crucial to our account of object movement, there is a possible derivation under which D-Minimality, even under a definition which takes all functional heads to be relevant, would not suffice to account for cases of super-raising in a language like English where the V does not raise. Consider the following case of super-raising in English:

(157) *John$_i$ is likely [$_{CP}$ it seems [$_{IP}$ t_i to be intelligent]]

The verb *seems*, as has been established by Pollock (1989) on the basis of the placement of negation, does not raise to I at S-structure in

English. Consider now a derivation where affix-lowering has occurred in the intermediate clause prior to the raising of the NP *John*.[72] In this case, only traces of functional projections remain, and consequently none of the functional projections will induce a D-Minimality barrier. This derivation is schematically represented in (158):

(158) [$_{CP1}$ [$_{IP}$ John$_i$ is likely [$_{CP2}$ t"$_i$ [$_{C'}$ Φ [$_{AGRP\text{-}S}$ it [$_{AGR\text{-}S'}$ t$_k$ [$_{TP}$ t$_j$ [$_{AGRP\text{-}O}$ t$_h$ [$_{VP}$ [seems AGR-O$_h$/T$_j$/AGR-S $_k$] [$_{IP}$ t'$_i$ to be intelligent]]]]]]]]]]

Now, the question arises of what blocks super-raising in these cases. There is, in fact, an answer to this problem. Recall that we have assumed that an empty C° (represented in (158) by Φ) can L-mark the complement of CP (here AGRP-S) if and only if it agrees with head of this complement. This condition is not fulfilled in (158): AGR-S agrees with the expletive *it*, while C° agrees with the trace of *John* in the Spec of CP. AGR-S and C° can thus be assumed to bear different indexes. As a consequence, C° will not be able to L-mark AGRP-S, which will remain a barrier, thus preventing t" in the Spec of CP2 to antecedent-govern t'$_i$, the original trace (or the trace in the Spec of

---

72. At first glance it appears that such a derivation could be blocked by the strict cycle: lowering occurs in the intermediate clause prior to the movement of the NP form the lower clause. This is not necessary, however. The NP could move from the lower clause to the Spec of the VP headed by *seems* prior to affix-lowering. Under such a derivation the strict cycle is not violated and the derivation outlined in (158) remains possible. In (158) we have ommitted the trace of the NP within the VP so as not to complicate the representation of the derivation. This possibility, however, must be kept in mind.

the VP *seems*) and t'(i) to be head-governed. Thus again ECP is violated.

The conclusions we can draw from this discussion are the following: While it seemed at first that an account of the super-raising cases in a language like English would favor a "rigid definition" of D-Minimality (where all functional projections count as D-Minimality inducers), it appears, given the derivation in (158), that even such a definition is insufficient to make D-Minimality the sole factor of the ungrammaticality of cases of super-raising such as (157). We conclude that cases of super-raisng remain in fact neutral with regard to the problem we raised in section 3.3.5.1 à propos of the definition of D-Minimality. Assuming that functional projections which are not lexically filled do not count as D-Minimality inducers, we can still rule out the cases of super-raising because of the lack of agreement between AGR-S and C°. In other words, it is clear that whatever the choice in the definition of D-Minimality (all functional heads count as D-Minimality inducers or only lexically filled heads), super-raising cases of the type in (156) and (158) can be excluded in our theory. They are excluded either by D-Minimality or because of the lack of agreement between C° and AGRP-S.

It is apparent, then, that the hypothesis that the Spec of CP is an [+HR] position raises no particular problems for an account of the classic cases of super-raising.

As we noted above, however, this assumption seems to raise problems for simpler cases of raising which are usually accounted for in terms of improper movement. Let us consider such a case:

(159) a.* Jean est impossible de dormir

b.* John is impossible to sleep

As is clear from the French example and in English sentences taking a tensed complement, *impossible* subcategorizes for a CP. The structure of (159) is as in (160) with Φ being an empty infinitival complementizer:

(160) [IP John is impossible [CP t_1 [C'Φ [IP t_2 to sleep]]]]

Here again, movement to the Spec of CP does not violate the improper movement condition, since it is movement from a [+HR] position to a [+HR] position. It is, however, natural to assume either 1) that agreement between the trace in Spec CP and the complementizer is impossible because of the nature of the infinitival complementizer (there is no equivalent of the *que/qui* alternation in the infinitival complementizer system in French (with *de*)), or 2) that agreement with I will not occur since an infinitival I lacks agreement features. Under either assumption, the consequence will be that IP in the infinitival sentence remains a barrier, since the Agreement Principle is not satisfied and C° cannot L-mark IP . The ECP is violated, since t_2 will be neither antecedent-governed nor properly head-governed. This accounts straightforwardly for the ungrammaticality of (159).

This concludes our consideration of the three problems raised at the very beginning of this section. We can conclude that the hypothesis that the Spec of CP is a [+HR] position does not cause any particular problem, given the proposal we have made for a theory of ECP. The analysis we have proposed for (160) has an interesting consequence: it allows us to account for the ungrammaticality of sentences such as (161) as violations of the ECP.

(161) *Who is it possible to sleep

Under standard assumptions, sentences such as (161) are accounted for as violations of the Case filter, reinterpreted in terms of the Visibility Condition on theta-assignement: the trace of the extracted subject does not have Case.[73] Given our proposal, the assumption that a variable needs Case becomes redundant for these structures. Consider the derivation of (161):

(162) Who is it possible [$_{CP}$t$_1$ [$_{CP}$ [$_{C'}$ Φ [$_{IP}$ t$_2$ [t$_3$ to sleep]]]]]

The natural assumption we made above, that the Agreement principle cannot be satisfied in infinitival sentences, leads to the conclusion that the trace t$_3$ will violate the ECP. IP will remain a barrier and t$_3$ will not be properly head-governed by the infinitival C°. Given this analysis, we no longer need to assume that the absence of Case on t$_3$ is responsible for the ungrammaticality of (161). In a section (xx), we discuss further the question of the Case requirement on variables. This will lead us to propose a new analysis of Kayne's

---

73. See Epstein (1987) for a different account.

examples of Case assignment into the Spec of CP. Before turning to these cases, however, we will discuss some interesting facts which provide cross-linguistic support for the Agreement principle, our definition of L-marking, and its interaction with the hypothesis that the Spec of CP is a [+HR] position.

4.5.3 Agreement in CP across languages

A number of facts very similar in nature to the *que/qui* effect are explored by Tarald Taraldsen (1986). In this section, we will briefly review some of these facts. As we will see, they provide cross linguistic evidence for our proposal about the *que/qui* effect and its interaction with the the assumption that the Spec of CP is a [+HR] position.

Scandinavian languages manifest a subject/object asymmetry in standard cases of extraction. Consider the following paradigm:

(163)
a. Vi vet hvem som snakker med Marit
We know who that talks with Mary

b.* Vi vet hvem snakker med Marit
We know who talks with Mary

(164)
a. *Vi vet hvem som Marit snakker med
We know who that Mary talks with

b. Vi vet hvem Marit snakker med
Vi know who Mary talks with

(162) shows that subject extraction is ungrammatical if the element *som* is not present in what we would analyze as the head of the

complementizer. (164), on the other hand, reveals that object extraction is incompatible with the presence of *som* in the complementizer. As noted by Taraldsen, these facts exactly parallel the *que/qui* effect. As we have seen above, the switch from *que* to *qui* is required in subject extraction and excluded with any other type of extraction. The analysis we have proposed for the *que/qui* effect can be straightforwardly extended to the Norwegian cases. Taking *som* to be the overt Norvegian version of an agreeing [+wh] C°,[74] we propose that the presence of *som* is required in cases of subject extraction to permit L-marking of the IP barrier and head government of the subject trace by C°. The structure for (162) is thus:

(165) ....[CP hvem [CP t' som [IP t snakker med Marit.]]]

Recall that IP is a barrier. For the subject trace to be properly head-governed, it is necessary that the barrierhood of IP be eliminated: adjunction to IP would satisfy antecedent government even if IP were to remain a barrier. But since we have assumed that head government is a condition on ANTECEDENT-government, this will not be sufficient. Head government of the original subject trace is also nedeed, and cannot obtain across a barrier. In our view, the presence of *som* allows L-marking of the IP barrier and thus permits the subject trace to be head-governed and to satisfy the ECP. The non-occurrence of *som* with object extraction can be accounted for very simply. Let us consider the structure of this extraction case:

---

74. Rizzi notes that *som* does not appear in cases of long-distance subject extraction out of a declarative sentence. We will assume, following Rizzi, that the agreement is simply not overt in cases of [-WH] complementizers.

(166) ....[CP hvem [T [IPt1 [IP Marit snakker med t2]]]]

In this structure, t_2 is properly antecedent-governed by t_1 and properly head-governed by the prepositional head *med*. Antecedent government and head government are satisfied, so there is no need to eliminate the barrierhood of IP through L-marking. Recall that a trace adjoined to IP in our system will be accessible for head governement by C°. The ECP is satisfied without the presence of *som*. This accounts for the grammaticality of (164)a. Moreover, the impossiblity of *som* with object or adjunct extraction is accounted for in parallel with the impossiblity of *qui* in French: it is a case of improper movement. Since IP is a barrier, adjunction is necessary. Movement from an adjoined position to the Spec of CP is ruled out by the condition on improper movement. Note that even if the object were to land in the Spec of CP, the agreement principle would not be satisfied, and IP would remain a barrier.

In his article, Taraldsen reports facts similar to those of Norwegian for Swedish and West Flemish. The West Flemish cases differ in an interesting way from those of Norvegian and Swedish in that the complementizer agreement phenomenon appears to be optional for subject extraction. The West Flemish paradigm is as follows:

(167)
a. Den vent die gekommen is.
The man that come is.

b. Den vent da gekommen is.

(168)
a.*Den vent die Pol getrokken heet
the man that Paul drawn has

b. Den vent da Pol getrokken heet

As we can see, *da*, the non-indexed complementizer, can appear with subject extraction. Taraldsen suggests (following Bennis and Haegeman) that this apparent optionality in fact reflects the possibility of extracting the subject directly from a position governed by the verb. West Flemish manifests a process of free subject inversion in *da*-clauses. In our view, this hypothesis also makes the right prediction. When the subject is extracted from a governed position, the head government requirement is satisfied on the original subject trace. Adjunction to IP is then freely permitted and there is no need for IP to be L-marked. The apparent optionality of the *die* complementizer for subject extraction can thus be straightforwardly accounted for in our approach.[75]

Other languages manifest overt agreement of their complementizers. Among these, well-known examples are Irish (Mc Closkey, Mc Closkey and Chung), the Bantu languages (Shneider-Zioga, Carsten and Kinyalolo),

---

75. Pushing things a little further, we can suggest that since Italian does not manifest agreement in its complementizer, extraction form the post-verbal position will always violate the ECP. Consequently, in this language extraction from the post-verbal position (possibly the VP-internal subject position) is always obligatory.

and Palauan (Georgeopoulos, Haik). The agreement pattern of these languages is apparently more complex and a careful study of the possibilities and impossibilities is beyond the scope of this dissertation. One fact, however, is worth mentioning. In Irish, complementizer lenition, although possible with both subjects and objects, is impossible for adjuncts. Under our view, this would be explained if arguments involved movement to the Spec of CP, while adjuncts such as *how* and *why* involved adjunction to CP.[76] This suggestion makes interesting predictions for the possibilities of extraction. We leave this topic for further research.[77]

In this section we have discussed the three problems that we raised at the begining of chapter 3. We concluded that the assumption that the Spec of CP is a [+HR] position raises no particular problems either for the improper movement condition or for weak crossover. We also showed that complex cases of super-raising violate the ECP, either because of D-Minimality or because of the lack of agreement of an infinitival complementizer. In the following section, we turn to other potential problems for the assumption that the Spec of CP is a [+HR] position, namely constructions where it has been argued that Case can be assigned to the Spec of an embedded CP by a matrix verb.

---

76. See footnote (xx) for a suggestion as to how this may be possible in some languages. We suggest that the difference between French and English with respect to the passage of the object in the Spec CP is related to the VSO order of Irish.

77. One surprising fact is that lenition is obligatory when a temporal adjunct is extracted. Extraction of locatives, on the other hand requires non-leniting complementizers. We have no explanation for this fact. As suggested by Ken Hale, it is possible that lenition is due to the relation between temporal adverbials and tense.

4.6 Case Marking into CP

Kayne (1984) argues that in some infinitival relative constructions, Case can be assigned to the trace of an extracted subject into what would currently be analyzed as the Spec of CP. In our view, the Spec of CP is a [+HR] position. If, as argued by Kayne, Case can be assigned to this position by a matrix verb, we would expect, contrary to fact, that an NP should be able to occur in the Spec of an embedded CP. Clearly, this possibility must be ruled out in both English and French. In this section, we discuss the constructions argued by Kayne to involve Case-marking into the Spec of CP. We argue that these constructions do not involve Case-marking into the Spec of CP, and that the Case requirement is satisfied differently. Our analysis makes crucial use of the Agreement Principle, which in our view is subsumed under the definition of L-marking proposed in section 4.4.1, and thus in turn provides support for this hypothesis.

4.6.1 Infinitival relative

In French, unlike English, epistemic verbs do not license infinitival complements with lexical subjects. Thus, sentences such as (168) are excluded:

(168)
a. * Je crois Jean avoir été arrêté par la police.
I believe John to have been arrested by the police

b. * J'ai estimé Jean avoir eu tort.
I think John to have been wrong

Control, on the other hand is possible:

(169)
a. Je crois avoir réussi mon coup
   I believe to have succeeded in my enterprise

b. J'estime avoir raison
   I think to be right

Infinitival complements of epistemic verbs differ from other standard control structures, however, in allowing WH-extraction of their subjects. Compare (170) with (171):

(170)

a.*l'homme que j'ai essayé d'être arrêté par la police

b.*l'homme que j'ai décidé d'être arrêté par la police

c.*l'homme que j'ai ordonné d'être arrêté par la police

the man that I tried/dedided/ordered to be arrested by the police

(171)
a. l'homme que je crois avoir été arrêté par la police
   the man that I believe to have been arrested by the police

b. l'homme que j'estime avoir été injustement condamné
   the man that I consider to have been unjustly condemned

c. une femme que l'on a dit avoir été très belle
   a woman that they have said to be very beautiful

The contrast between (168) and (169) shows that the embedded subject position of epistemic infinitivals is not accessible for government and Case-marking by the matrix verb: an overt NP cannot occupy this position, since it is not assigned Case. But PRO, which by definition must remain ungoverned, can. The inaccessibility of the subject position is expected if epistemic infinitivals, like control infinitivals, are dominated by an empty complementizer, that is, if they have the following structure (where Φ = empty C^0):

(172) Je crois [CP Φ [IP PRO avoir raison]]

But given the structure in (172), the grammaticality of examples like (171) is surprising. To account for the possibility of (171), Kayne (1983) proposes to exploit the notion of successive cyclicity. The structure of (171) differs crucially from that of (172) in that WH-movement, operating successive cyclically, has left a trace in the Spec of the intermediate CP:

(173) l'homme que je crois [CP t_1 [IP t_2 avoir été arrêté par la police]]

While it remains impossible in this structure to assign Case to the original trace of the subject, Kayne proposes that in this configuration, Case can be assigned to t_1, the trace in the Spec of the intermediate CP.

The ungrammaticality of (170) is assumed to be due to the presence of the complementizer *de. De*, like an overt *that* complementizer in English, blocks antecedent government of the trace in subject position by the trace in the Spec of CP. The sentences in (170) are excluded as a straightforward violation of the ECP.

We proposed in the previous section that extraction of a subject out of a control infinitival can be excluded under the ECP. Recall that we argued that in examples like (174)a., the ECP is violated because an infinitival complementizer fails to satisfy the agreement principle and thus cannot L-mark the IP projection of an infinitival complement. The same reasoning applies to examples of the type (174)b.:

(174)

a.*Who$_i$ is it possible [$_{CP}$t'$_i$ [$_{C'}$ Φ [$_{IP}$ t$_i$ to go

b.*Who$_i$ did Mary [$_{CP}$ t'$_i$[$_{C'}$ Φ[$_{IP}$ t$_i$ try to go

As we have argued, an ECP violation results from the failure to satisfy the Agreement Principle. We proposed that an infinitival complementizer does not agree and consequently cannot L-mark IP. In both (174)a. and b. the initial trace of the extracted subject fails to be antecedent-governed or head-governed, since IP remains a barrier.[78] If this analysis is on the right track, we do not need Case theory to account for (170). We thus do not expect Case theory to be crucially relevant to an account of (171), either.

Moreover, the assumption that Case can generally be assigned into the Spec of CP is problematic for an account of the ungrammaticality of sentences like (168). Indeed, (168), although ruled out under the structure in (175), should be possible with a structure such as the one given in (176).

---

78. Incidentally, this analysis also accounts for the *for-to* effect, which cannot be accounted for simply as a Case violation.

(i) *Who$_i$ is it possible for t$_i$ to go

In (i), under standard assumptions, the trace t$_i$ is Case-marked by the *for* complementizer. (i) is thus not ruled out by the Case filter, but by the ECP. We simply propose to generalize this approach to other cases of infinitivals under the assumption that the complementizer of an infinitival complement, overt or empty, differs from that of a tensed complement in that it does not permit agreement. This hypothesis is in a sense simply an extension of the idea that there is a relation between the complementizer and the tense properties of the sentence (Cf. Chomsky (1981)).

(175) Je crois [CP [IP Jean.....]]]]]]

(176) Je crois [CP Jean [C' [IP t .....]]]]

In (176), unlike (175), the Case filter is satisfied, since *Jean* can receive Case in the Spec of CP from the matrix verb. In the standard framework, this structure might be excluded under the assumption that the NP *Jean* is in an A'-position and consequently does not head a well-formed A-chain.[79] This possibility, however, is not available to us, since we have assumed that the Spec of CP is a [+HR] position. Generally speaking, it is crucial to our approach to ban Case assignment to the Spec of CP by a matrix verb. If Case could be generally assigned to the Spec of CP, we would wrongly predict that a sentence such as (177), where an NP has moved to the Spec of CP, should be grammatical:

(177)
** Je crois [CP Marie$_i$ qui [IP t$_i$ a rencontré Pierre.]]
I believe Marie who has met Pierre

However, (177) and the structure in (176) can be straightforwardly ruled out if we assume that Case cannot be assigned to the Spec of an embedded CP by a matrix verb.[80] We will thus adopt this hypothesis[81]

---

79. Kayne (1984) assumes, however, that it is not the nature of the position, but rather the nature of the element occupying the position which matters for the distinction between A- and A'-chains. Consequently, it is not very clear why the structrure in (176) should be ungrammatical.

80. Plausibly this impossibility is due to the fact that verbs can only assign one structural Case, and this Case is taken up by the whole sentence. This would in turn suggest that in ECM constructions, Case is not given to the NP directly by the matrix verb, but is rather mediated by the INFL node of the infinitival.

Cf. Reuland (1983) for such a proposal, which we adopted in section 3.4.4.

81. There is however one construction in French were it seems that Case is assigned to the Spec of CP. This construction has been called the pseudo-relative construction, because it superficially resembles a relative clause. In fact this type of construction only occurs in very restricted environements in French , i.e. most commonly as a complement to the verbs of perception (it can also be a complement of the verb *avoir* and of some prepositions):

(i) J'ai vu Marie qui sortait du cinéma
I saw Mary who was coming out of the movie theater.

Superficially, this construction resembles a relative clause. It has however very distinct properties. First, contrary to relative clauses, it can be headed by clitic.

(ii) a.Je l'ai vue qui sortait du cinéma
I saw her who was coming out of the movies
b.*Je lui ai parlé qui sortait du cinéma
I talk to her who was coming out of the movies

Second, it is limited to subject cases:

(iii) a. Je l'ai vue qui embrassait Henri
I saw her kissing Henri
b.*Je l'ai vue que Henri embrassait
I saw her that Henri kissed

Note that if we assume that this pseudo relative involves an NP in the Spec of CP and has the structure (iv), these and other properties (which we do not have the time to discuss see M.T Guasti (1989)) follow straigtforwardly from our assumptions.

(iv) J'ai vue [$_{CP}$ Marie$_i$ [$_{C'}$ qui [$_{IP}$ t_i sortait due cinema]]]

Cliticization is expected to be possible since the Spec of CP in our view is a [+HR] position. Moreover, the fact that only the subject and not the object can move to the Spec of CP is straightforwardly accounted for as a violation of improper movement. Recall that in our view, IP is a barrier unless it is L-marked under agreement with C. Consider the structure of (iii)b.:

(v) J'ai vue [$_{CP}$ Marie$_i$ [$_{C'}$ que [$_{IP}$ t_i [$_{IP}$ Henri embrassait t_i]]]]

In (v), C does not agree with I, because C agrees with the object *Marie* and I agrees with the subject *Henri*. Thus IP is a barrier. To avoid the barriehood of IP an extracted object must adjoin to it. But this adjunction prevents it from further moving to the Spec of CP, since this would be an instance of movement from a [-HR] position to a [+HR] position and would violate the condition

Note that the hypothesis that the Spec of CP is a [+HR] position makes a prediction for complement types such as (171): although the impossibility of Case-marking into the Spec of CP should prevent movement of an NP <u>to</u> this position, it should not prevent movement of an NP <u>through</u> this position. This prediction is verified by the following sentences noted by Kayne (1984) and Pollock (1985 pp. 308-311):

---

on improper movement. Thus noting need to be added to account for the restriction of this construction to subjects. What remains mysterious is why in this particular construction Case can be assigned to the Spec of CP. Note that complements of perception verbs are the only constructions which allow an ECM complement in French.

(vi) J 'ai vu Marie sortir
I saw Mary leave.

We will thus assume that whatever the property which permits an ECM construction with perception verbs, it will extend to the pseudorelaltive construction. See Deprez (1985) for a more detailed study of the properties of perceptual reports in French and in English.

(178)
a.? Jean a été déclaré être apte à participer
Jean was declared fit to participate

b.? Pierre a longtemps été supposé avoir téléphoné ce soir-là
Pierre was long assumed to have telephoned that evening

c.? Marie a été longtemps considérée avoir tué son amant[82]
Marie was long considered to have killed her lover

Under the hypothesis that the Spec of CP is an A'-position, which in Kayne's approach is crucial to exclude cases like (168), these sentences should be excluded as instances of improper movement. The fact that they are possible provides additional support for the hypothesis that the Spec of CP is a [+HR] position.[83]

Since we have made the hypothesis, contrary to Kayne's, that an NP in the Spec of CP cannot be Case-marked, we must offer an alternative analysis for the grammaticality of (171). The essence of our proposal

---

82. As noted by Pollock (1985), other sentences similar in structure to (178) do not permit NP movement:

(i)
??(*)Jean a été nié être intelligent
Jean was denied to be intelligent

Generally speaking, however, it seems that the verbs which do not permit NP movement are verbs which take factive complements. As has been observed, factive complements are generally islands. We suggest that the ungrammaticality of (i) is to be related to the factivity of the complement, whatever the analysis which is ultimately adopted to explain the islandhood of factive complements (See Zubizarretta (1981) and Rouveret (1980) for some proposals).

83. The marginality of these sentences is unexpected under our approach. Possibly, it may be due to Case requirementss on the sentence: passive past participles can only assign inherent partitive Case. This may not be fully sufficient to satify the Case requirement on the complement sentence, or there might be a slight definitness-effect violation. See Deprez (1984) for arguments that epistemic infinitivals are in some sense "nominal" in character and require Case-marking.

is that the infinitival complements of epistemic verbs differ from standard control constructions in allowing the satisfaction of the Agreement Principle and consequently permitting (under appropriate circumstances) the IP of the infinitival complement to be L-marked.

Let us first consider the ungrammatical sentences in (170). The structure of (170)a. is given in (179):

(179) l'homme [$_{CP}$que [$_{IP1}$ t_1 [$_{IP1}$ j'ai essayé [$_{CP2}$ t_2 [$_{C'2}$ d' [$_{IP2}$ t_3 être .....]]]]]]

This sentence violates the ECP: t_3 cannot be head-governed by the infinitival complementizer because the Agreement Principle is not satisfied and IP remains a barrier. We propose that it is a lexical property of the infinitival complementizer which is responsible for the failure of the Agreement Principle. In our view, an infinitival complementizer (*de* or 0) does not agree. It is then comparable to the English *that* complementizer in dialects which manifest the *that*-trace effect. Clearly, there is no equivalent of the *que/qui* effect with the French infinitival complementizer *de*. Under this hypothesis, the Agreement Principle cannot be satisfied: the head of an infinitival C will never agree with the head of I. As a result, the embedded IP remains a barrier and the subject trace is neither properly head-governed nor antecedent-governed.[84]

---

84. Alternatively, we could assume that it is a property of the infinitival INFL which is responsible for the failure of the Agreement Principle. Infinitival complements, although they have Tense (see Stowell (1981) for arguments), do not manifest agreement. A rather natural assumption would be that in this case, coindexation with the element in subject position does not occur. If so, even if it were possible for the trace t_2 in the Spec of

Let us turn now to (171). Superficially, the structure of this sentence seems identical to that of (170). There are, however, a number of properties which distinguish the infinitival complements of epistemic verbs from those of verbs like *try*. One of the most apparent differences is that the former never take an overt complementizers while the latter always do.

(180) Je crois (*de) avoir réussi
I believe (**de*) to have succeeded

(181) J'essayé *(de) réussir
I tried *(*de*) to succeed

Another difference concerns the tense properties of these complements. Infinitival complements of epistemic verbs are always better with a compound tense or a modal construction than with simple tenses. This tendency is reversed in complements of *try* verbs.

(182)
a. Elle croit être arrivée en retard
She believes to have arrived late

b.??Elle croit arriver en retard
She believes to arrive late

c. Elle croit devoir tout expliquer
She believes to have to explain everything

d.??Elle croit tout expliquer
She believes to explain everything

CP2 to agree with the complementizer, since t_2 does not agree with I, IP2 remains a barrier and t_3 cannot be properly head-governed.

(183)
a. Elle a essayé d'arriver à l'heure
She tried to arrive on time

b.?? Elle a essayé d'être arrivé à l'heure
She tried to have arrived on time

c. Elle a essayé de tout expliquer
She tried to explain everything

d.?/*Elle a essayé de devoir tout expliquer
She tried to have to explain everything

Although it is not our goal to attempt to explain each of these differences, we take them to be indicative that the structure of these two types of infinitival complements differ. In the particular case under discussion, we propose, following Deprez (1984) and Guéron and Hoekstra (1987), that sentences of the type in (170)b. involve the raising of the infinitival I to C°. This proposal is justified in Deprez (1984) on the basis of similarities between the French epistemic construction and the AUX-to-COMP infinitival complement of epistemic verbs in Italian and the inflected infinitival complement of epistemic verbs in Portuguese.[85]

In both languages, overt movement of the verb to C is required to license a lexical subject (See Rouveret (1980) and Raposo (1986)). This suggests that a similar process of AUX-to-COMP may be available in the French construction. Support for this hypothesis comes from

---

85. Guéron and Hoekstra (1987) use similar comparisons to justify their proposal. They propose the following (descriptive) principle:

(i) If the non-finite complement of an epistemic verb such as *pensar/ritinere/croire* etc. denotes an event, then the epistemeic verb must govern the tense morpheme of its complement.

This will force raising of INFL to C in cases of WH-extraction.

contrasts between subject extraction with compound tenses and subject extraction with simple tenses. As mentioned above, infinitival complements of epistemic verbs are usually better with compound tenses; however, complements with simple tenses are not excluded. The following example is perfect:

(184) Je crois rêver
I believe to dream
I believe I'm dreaming

Subject extraction in this case, however, is distinctly worse than in cases such as (171):

(185) ?*l'homme que je crois t rêver
the man that I believe to dream

Pollock (1989) has argued that in infinitivals, only auxiliaries and, more marginally, modal verbs can move to I. Evidence for this is given by the position of negation with auxiliaries. Negation can either precede or follow an auxiliary: when it follows it, AUX must be in I.[86] As we will argue shortly, the movement of I to C is essential to license subject extraction. This hypothesis in turn suggests the following explanation for (185). Suppose that a lexical I is required to head-govern the trace of the subject extraction; then in sentences such as (171) the AUX may have raised to C. In (185), however, only an empty I has risen to C, so head government is in some sense weaker. The existence of the contrast between (185) and (171) provides some

---

86. Cf. Pollock (1987) for a detailed discussion.

support for an AUX-to-COMP phenomenon in French epistemic infinitival complements.[87]

Our analysis of (185) makes a prediction: if AUX-to-C° is involved in subject extraction, subject extraction should be better when the negation follows the auxiliary than when it precedes the auxiliary, since AUX-to-C° will involve prior movement to I over the negation. This seems to be generally correct, although the contrast is sometimes quite subtle:

(186)
a.une jeune fille que je croyais n'avoir pas été invitée
a young girl that I believed to have not been invited

b.?? une jeune fille que je croyais ne pas avoir été invitée
a young girl that I believed not to have been invited

(187)
a. un homme que j'estime n'avoir pas eu de chance
a man that I consider to not have had any luck

b. ?? un homme que j'estime ne pas avoir eu de chance
a man that I consider not to have had any luck

---

87. The contrast between compound and non-compound tenses remains with unaccusative verbs, indicating that the severe marginality of (185) is attributable to more than the lack of Case, as in the contrasts mentioned by Pollock (1985).See below

(188)
a. un livre qu'il a jugé n'être pas suffisamment complet
   a book that he judged to be not sufficiently complete

b. ??un livre qu'il a jugé ne pas être suffisamment complet
   a book that he judged not to be sufficiently complete

The contrasts in (186)-(188) provide additional support for the movement of I to C_0 in the epistemic infinitival construction.[88] The structure of (171)a. is then as follows:

(189) l'homme [$_{CP}$que [$_{IP1}$ t_1 [$_{IP1}$ je crois [$_{CP2}$ t_2 [$_{C'2}$ être [$_{IP2}$ t_3 t_I. ...]]]]]]

In this configuration, IP2 will be L-marked by *être*, which has raised into C. Consequently, it will not be a barrier and t_3 will be properly antecedent- and head-governed and will satisfy the ECP. This accounts for the grammaticality of (171).

The account we have offered of the contrast between (170) and (171) does not rely on Case theory. Some support for this position comes from cases comparable to (171) in English involving the verb *assure*. As noted by Kayne (1984), *assure* does not license a lexical subject in its infinitival complement:

(190) *I assure you John to be the best student in my class.

However, parallel to the French epistemic complements, extraction of the subject of the infinitival complement is fine:

---

88. see Deprez (1984) for arguments and a more detailed analysis of various types of infinitivals. See also Pollock (1985), Guéron and Hoekstra (1987)

(191) John, who I assure you [t [t to be the best student in my class]

Kayne proposes to account for the grammaticality of (191) on a par with cases of (171). Under this view, the verb *assure* assigns Case to the variable in the Spec of the CP complement. This proposal, however, is somewhat problematic. Indeed, contrary to other verbs, such as the classic *give*, the verb *assure* is not a double-object predicate. Case-marking of the second "object" is always prepositional.

(192)

a. I assured David of my gratitude.

b.*I assured David my gratitude.

c.*I assured my gratitude David.

From the ungrammaticality of (192)b-c, we can conclude that the verb *assure* does not assign a second Case. It thus seems doubtful that the grammaticality of (191) could be due to the assignment of Case by *assure* to a trace in the Spec of CP.[89] It could be assumed that *assure* can assign inherent Case to a second object but that for some unknown reason, inherent Case must be realized as prepositional. Notice that this does not solve the problem posed by Case assignment into C, however. As viewed by Chomsky (1986)a., inherent Case assignment is tightly linked to θ-assignment. But Case assignment to the Spec of CP involves no θ-relation. We thus expect inherent Case assignment to the Spec of CP to be impossible. Under our approach, Case theory does not come into play. It is sufficient to assume that

---

89. Thanks to David Pesetsky for pointing this out.

the complement of *assure* has the same structure as the complement of the French epistemic verbs: namely it allows I to move to C. Some support for this hypothesis comes from the fact that the contrast with the placement of negation noted in (187) seems to obtain in English also:[90]

(193))

a. ?This boy, who you assured me to not have ever been arrested

b. ?/*This boy, who you assured me not to have ever been arrested

Our proposed analysis permits accounting for one fact which is unexpected under Kayne's (1984) view. The complement of *assure* differs from the French epistemic infinitivals in one respect: it does not permit control.

(194) *I assure you to have been the best student in my class.

This is unexpected under Kayne's analysis. Under his view, the infinitival complement of the verb *assure* is a CP. Consequently, the subject of this infinitival complement is ungoverned, so that PRO is expected to be possible. Under our proposal, the ungrammaticality of (194) can be seen to follow from the assumption that I-to-C raising is obligatory in this infinitival complement.[91] Assuming this to be

---

90. Thanks to D.Pesestky (pc) for noting this fact

91. Alternatively, we could propose that the complement of the verb *assure* is simply IP. In our approach this does not cause any problem because we do not assume that *assure* can assign a second structural Case. Consequently, an overt NP will not be allowed in this construction. Moreover, PRO will not be allowed, either, because the Spec of the IP of the clausal complement of *assure*, although not Case-marked, will be governed.

correct, PRO is excluded because it is governed by the I in C. We predict correctly that the only possible way an infinitival complement will be acceptable with *assure* is if the subject has been extracted. This is supported by the facts.[92]

The account we have proposed of the contrast in (170) vs. (171) and of (190) does not involve Case theory. Moreover, in the structure we have proposed in (189), the subject variable is not assigned Case, yet the structure is well-formed. We are thus led to the conclusion that there exist well-formed Operator-variable contructions in which the variable has no Case. This conclusion goes against currently accepted assumptions and raises the question whether it should be assumed that variables generally need Case. The assumption that variables need Case has been questioned in the literature (Borer (1982), Epstein (1987)). This question is important, because it is relevant to the proper formulation and the interpretation of the Case filter and of the level at which it applies. We will thus discuss this question in the next section.

---

92. A complete parallelism between the French epistemic infinitival and the English construction under *assure* could obtain if we were to assume obligatory I-to-C in the French infinitivals as well. The difference between French and English with respect to the Control structure would then have to receive a different explanation. A possible view would be to assume, as proposed in Deprez (1984), that the empty category in the French epistemic infinitival is pro and not PRO. The comparable English structure would then be excluded on the basis that pro is never licensed in English. As argued by Deprez (1986 , 87, 89) and in Pollock (1986), French allows a restricted form of "pro-drop", namely what we have called expletive-drop. English, on the other hand, does not.

4.6.2 Variables and Case-marking

As first conceived by Vergnaud (1980) and Chomsky (OB), the Case filter matters essentially for overt NPs:

(195) *NP if NP is overt and has no Case.

Under such a view of the Case filter, variables are not expected to need Case, since they are not overt. Chomsky (1981), following a suggestion by Aoun (1981), proposed that the Case filter be reinterpred as a Visibility Condition on θ-assignment at LF. Under this view, every argument, including empty categories and sentences, is expected to need Case.[93]

The following two examples have been at the center of the debate.

(196) *Qui semble-t-il [t être intelligent]

*Who does it seem [t to be intelligent]

(197) Who do you believe sincerely [t to be intelligent]

---

93. There are two major problems with this hypothesis: 1) expletives, being non-arguments, are expected not to need Case. As is well known, however, this is not the case: overt expletives (at least in English) must appear in Case-marked positions. 2) PRO, which is an argument, does not require Case. Various solutions have been proposed to solve 1). To our knowledge, no solution is offered to 2). Chomsky simply stipulates that PRO has inherent Case features. Clearly this cannot be the right solution: overt pronouns in English also have inherent Case features, which are, in fact, overtly visible. This does not, however, allow them to be licensed in the subject position of infinitival sentences.

The structure of (196) is as follows:

(198) [CP who [CP does [IP1 t1 [IP1 it seem [IP2 t2 to be intelligent]]]]]

Under standard assumptions, the infinitival complement of *seems* is an IP. This IP is not a barrier for NP movement in standard cases of Raising, so there is a priori no reason to assume that IP should be a barrier in this case. If IP is not a barrier, t_2 will be antecedent-governed by t_1 and head-governed by *seems*, so the ECP will be satisfied. Nothing excludes this sentence except the assumption that variables need Case. Sentences of this type are thus the major argument for the hypothesis that traces need to be Case-marked. Sentences of the type in (197) constitute the major arguments for the opposite assumption. Consider the sentence in (199):

(199) *I believe sincerely John to be intelligent

The ungrammaticality of (199) is standardly assumed to be due to a violation of the adjacency requirement on Case. *Believe* cannot assign Case to the NP *John* because of the intervening presence of the adverb *sincerely*. Let us tentatively assume this to be correct. Under the assumption that variables must be assigned Case, the grammaticality of (197) is mysterious. Consider the structure of (197):

(200) [CP who [do [IP you believe sincerely [IP t1 to have left]]]]

In (200) t_1 is separated from the Case-assigning verb by the adverb *sincerely*. We must conclude that t_1 cannot receive Case. Under the assumption that variables need Case, the sentence (197) should be

excluded, contrary to fact. A solution to this apparent conflict has been proposed by Epstein (1987). Epstein proposes that variables are subject to a weaker contraint than overt NPs. He formulates the following principle.

(201) A variable α must be governed by a Case-assigner.

Under standard assumption Case is assigned under government. But the government relation necessary for Case-marking is only a subset of the more general relation of government. Unlike Case assignment, government is not constrained by an adjacency requirement. Thus, it is commonly assumed that a dative complement, for instance, is governed by the verb, even though it is not adjacent to it. Under Epstein's proposal the ungrammaticality of (196) is straightforwarly accounted for. *Seems* is not a Case-assigner; consequently even though the variable is governed by it, the structure (198) will be adequately excluded. In (197), on the other hand, the variable is governed by the Case-assigning verb *believe* so that the sentence is correctly predicted to be grammatical. Given the structure we have proposed for (171), the principle in (201) could be fulfilled in two ways. First, we may consider I to qualifiy as a Case-assigner even when it does not actually assign Case, as in infinitival. If so, the subject trace will be governed by I and adequatly licensed under (201). Alternatively, we may take the trace in the Spec of CP to be the variable which is governed by the verb *croire*, as in Kayne's proposal. But clearly either of these ways of satisfying (201) would work equally well for cases like (170). To avoid this problem, we could reinforce Epstein's principle (201) and require that variables need to be canonically

governed by a (potential) Case-assigner. In this case, the movement of I to C which we have proposed for epistemic complements may be taken to fulfill the strenghthened requirement without creating overgeneration of unwanted extractions.[94] This seems to be somewhat redundant, however, given the account we proposed above. We assumed that I moves to C to satisfy the Agreement Principle and, indirectly, the ECP. We could assume that I-to-C also satisfies the Case filter. There are, however, considerations pertaining to subtle differences among cases of subject extraction from epistemic complements, to which we turn shortly, which suggest that this is not correct and that the Case filter is satisfied differently. For this reason, we will explore an alternative to Epstein's proposal.

Let us assume tentatively that variables do not need Case. If so, we need an alternative account for the ungrammaticality of (196). A sentence such as (196) may have two different structures. The verb *seems* subategorizes either for a CP complement or for an IP complement. Let us first look at a structure in which *seems* has subcategorized for a CP. (196) in this case will have the following structure:

(202) [Who does it seem [$_{CP}$ Φ [t to have left]]]

In our view, this derivation will be straightforwardly ruled out by the ECP. Recall that we have proposed that an infinitival C^{o} is not an agreeing complementizer and thus cannot L-mark IP. Consequently, the

---

94. The strengthening of the requirement causes problems for regular subject extraction, however.

IP of the embedded sentence remains a barrier and C_0 is unable to properly head-govern the subject trace. This leads to an ECP violation. In this case again, the hypothesis that variables need Case is not needed for the explanation of the ungrammaticality of this derivation. Accounting for the case in which *seems* subcategorizes for an IP, however, is more difficult. Let us consider this structure:

(203) [Who does[t_2 [it seem [$_{IP}$ t_1 to have left]]]]

Under standard assumptions, the IP complement will be L-marked by the verb *seems* so that it will not be a barrier. If so, t_1 will be antecedent-governed by t_2 and properly head-governed by *seems*. The ECP is satified and the sentence is predicted to be grammatical, contrary to fact.

The solution we propose for this problem relies on the expletive replacement hypothesis made by Chomsky (1986). Chomsky argued that expletives, not being meaningful at LF, must be eliminated. This is, in his view, a direct consequence of the principle of Full Interpretation, which requires that every element represented at LF contribute to semantic interpretation. Assuming that this is correct, in a sentence such as (196) the expletive *it* will have to be replaced at LF. As is well known, expletives are lexically specified for the elements they can be replaced by. Thus, *there* in English is lexically specified to associate with NPs and *it* is lexically specified to associate with sentences. Pushing this a little further, I propose that *it* can only be replaced by a CP and crucially not by an IP. We can view this as an extension of the fact that *it* cannot associate

with elements bearing phi-features. Both NP and IP, which presumably bears the features of its subject, are elements with phi-features. Given this assumption, the derivation in (203) will be excluded as a violation of the principle of Full Interpretation: *it* will not be replaced at LF and the sentence will be uninterpretable.

Assuming that the analysis we have proposed for (196) is plausible, we do not need the assumption that variables need to be Case-marked. On the other hand, we have not shown that this hypothesis is wrong. It may still come into play elsewhere. Our claim is essentially that it is not needed to account for cases such as (196). There is, however, independent evidence which suggests that Case might still be generally needed for variables.

Pollock (1985) noted that not all cases of subject extraction out of epistemic infinitival complements are equally good: sentences with passive constructions or unnaccusative verbs in the complements of the epistemic verbs allow subject extractions more easily than sentences with purely intransitive or transitive verbs. We reproduce some of Pollock's contrasts in (204) and (205):

(204)
a. L'homme que je croyais être arrivé/entré/avoir disparu
   the man who I thought to have arrived/come/disappeared

b.?*L'homme que je croyais avoir téléphoné/toussé/plongé
   the man who I thought to have telephoned/coughed/dived

(205)
a. L'homme que je croyais avoir été injustement
condamné/accusé/executé
the man who I thought to have been unfairly
condemned/accused/executed

b?*L'homme que je croyais avoir injustement
condamné/accusé/éxécuter ces prisonniers
the man who I thought to have unfairly
condemned/accused/executed these prisonners

On the basis of these and similar contrasts, Pollock proposes the following rule:

(206) NP --> [+Case] when governed by V*

V* = unaccusative verbs, passive participles and *être*

More recently, a similar proposal has been made by Belletti (1986). Belletti proposes that verbs of the type V* are verbs which can assign a partitive inherent Case to their D-structure complements. To account for the contrast in (204) and (205), Pollock argues that subject extraction out of epistemic infinitivals is possible if Case can be given to the lowest trace by a V*. As he argues, this suggests that Case assignment into the Spec of CP, if possible at all, is a very marked option. Assuming this to be correct, and pursuing the analysis we proposed above, we obtain the following structure for sentences of the type in (204)a. and (205)b.

(207) l'homme [$_{CP}$ que [$_{IP1}$ t_1 [$_{IP1}$ je croyais [$_{CP2}$ t_2 [$_{C'2}$ I [$_{IP2}$ t_3 avoir été accusé t_4]]]]]]

In this structure, we have a [+HR] chain minimally including[95] the

---

95. There are more traces in the Spec of the functional projections dominating the past participle and the auxiliary, which we have

traces (t_2, t_3, and t_4), followed by a [-HR] chain minimally including the operator *que* and t_1 adjoined to the upper IP. Under the Case condition on chains, however, the [+HR] chain is excluded because the head is not Case-marked. As for the [+HR] chains, the following question arises: which of the three traces t_2, t_3, and t_4, each part of an [+HR] chain, is to be taken as the relevant variable for the operator *que*? Under standard assumptions, it is the head of a [+HR] chain which is taken to be the relevant variable. Taking the trace in the Spec of CP to be the head of the [+HR] chain, we come to the conclusion that variables do not need Case, since these sentences are well-formed. But this position leaves the contrast in (204) and (205) unaccounted for.

We propose that instead of taking the highest position in a [+HR] chain to be the relevant variable, we take the entire sequence of [+HR] positions to function as the relevant "extended variable". This proposal has two advantages: first, it bypasses the violation on the condition on the well-formedness of [+HR] chains. Second, it allows us to maintain the assumption that variables need to be Case-marked. Our variable, the chain (t_2, t_3, t_4), contains both a θ-marked position and a Case-marked position. It thus satisfies the Visibility version of the Case filter. Given the proposed notion of extended variable, the contrasts in (204) (205) can now receive a very natural explanation. Grammatical instances of subject extraction are cases in which the extended variable meets the Case requirement. Ungrammatical

---

not included to keep the structure manageable. Their presence is irrelevant to the present discussion.

instances of subject extraction are cases in which the Case requirement is not met.[96]

In a sense, what we are proposing here can be viewed as an extension of the condition on the well-formedness of [+HR] chains.

(208)

$(a_1....a_n)$ is a well-formed [+HR] chain iff (a) or (b):

(a) a_1 is a Case-marked position
and a_n is a θ-position

(b) a_1 is the foot of a well-formed [-HR] chain and
a_n is a Case-marked position.

Given (208), it is possible for the head of a [+HR] chain to fail to be Case-marked if and only if the head of this otherwise well-formed [+HR] chain (every link is properly antecedent- and head-governed) is properly bound by an operator; that is, the head of the [+HR] chain is the foot of a well-formed [-HR] chain. The head of the [+HR] chain

---

96. Our conclusion seems too strong, since there are speakers for whom tne contrasts in (204) and (205) do not exist. Pollock suggests that for these speakers, Case-marking in the Spec of CP must be a possibility. We suggest instead that the speakers in question have a more robust system of inherent Case-marking. For these speakers, inherent Case-marking is not limited to the verbs of the V* class. If this is correct, we expect these speakers to accept impersonal constructions with intransitive or even transitive verbs. Sentences such as those in (i) have been reported in the literature to be acceptable for a number of speakers:

(i)

a. Il a telephoné beaucoup de gens aujourd'hui.
There have telephones many people today
Many people have tlephoned today.

b. A une certaine époque, il a mangé beaucoup de linguistes dans ce restaurant
At a certain time, there ate many linguists in this resturant
There was a time when many linguists ate in this restaurant.

will not have to be either a Case-marked position or a θ-marked position. If so, it is odd to regard the head of the [+HR] chain as the variable bound by the operator. In our view, the variable will be the entire sequence of [+HR] positions, including a_n, which will be a Case-marked position and, in the examples (204)a. and (205)a., a θ-marked position. In Chomsky (1981), the notion of argument is defined as a chain. Our proposal similarly defines a variable as a chain. That is, it is either a chain of length 1, as is usually assumed, but it can also be a chain of length n. Consider the following abstract sequence:

(209) OP [-HR] [-HR] [-HR]... [+HR] [+HR] [+HR]

Under current assumptions, this sequence is only well-formed if the head of the [+HR] chain is Case-marked. Under our proposal, this is no longer necessary. It is sufficient that the foot of the [+HR]-sequence be Case-marked and that the [-HR] chain be well-formed for an extended variable to be well-formed.

Let us consider some more complex examples of what we have called extended variables. Consider the pair of sentences in (210):

(210)
a.*le jeune garçon que tu as essayé d'être susceptible d'être le meilleur étudiant de la classe.
the young boy that you tried to be likely to be the best student in the class.

b. le jeune garçon que tu crois être susceptible d'être le meilleur étudiant de la classe.
the young boy that you believe to be likely to be the best student in the class.

(210)a. is ungrammatical, although it seemingly meets the requirement that we have posited for a [+HR] chain followed by an [-HR] chain. (210)b., on the other hand, is grammatical with apparently the same structure. Let us consider in detail the derivation of each of these sentences:

(211) ...[$_{CP1}$ que [$_{IP1}$ t_1 [$_{IP1}$ tu as essayé [$_{CP2}$ t_2 [$_{C'}$ d' [$_{IP2}$ t_3 être susceptible d'[$_{IP4}$ t_4 être [t_5 le meilleur étudiant de la classe]]]]]]]][97,98]

In (211), t_3 is not Case-marked; under the definition given in the previous section, this is not a problem if a well-formed [-HR] chain can license the [+HR] chain. The trace t_5, which, in our terms, is the foot of the extended variable, is appropriately Case-marked by the V* *être*. As we assumed above, however, in such structures the Agreement Principle cannot be met and thus L-marking of IP does not occur. *De* in

---

97. We have not detailed the derivation between t_3 and t_4. This must be a proper [+HR] chain under any assumption, as shown by the grammaticality of (i):

(i)
Jean est susceptible d'être le meilleur élève de sa classe.
Jean is likely to be the best student in his class.

98. Cf. Ruwet(1982) for arguments that *susceptible* is a raising adjective in French.

CP2 is not an agreeing complementizer, and consequently it cannot L-mark IP. As a result, t_3 will not be head-governed, which leads to an ECP violation.

The proper structure for (210)b. is (212), in which I has raised to C_2.

(212) ...[$_{CP1}$ que [$_{IP1}$ t_1 [$_{IP1}$ tu as cru [$_{CP2}$ t_2 [$_{C'2}$ être [$_{IP2}$ t_3 t susceptible d'[$_{IP4}$ t_4 être [t_5 le meilleur étudiant de la classe]]]]]]]]

In this structure, the ECP is satisfied, since the Agreement principle is met and IP is not a barrier. The extended variable ($t_2..t_5$) is Case-marked, since t_5 can receive inherent Case from *être*.[99] The structure is grammatical, as we predict.

The analysis we have proposed to account for the contrasts found by Pollock in the French infinitival relative construction predicts that similar contrasts should be found in the English constuctions in which a variable seems to fail to be Case-marked. Recall that we have mentioned two constructions in which variables seem to fail to be Case-marked: the infinitival complement of the verb *assure* and cases

---

99. See D.Couquaux for arguments that the structure of copular sentences involves an SC structure like the one given in (212). There is only one case of an impersonal construction where *être* accepts an NP: *Il était une fois* ("Once upon a time", lit. "There was a time"). We do not have an explanation for why sentences of the type *Il est un homme intelligent* ("There is an man intelligent") are not very good in comparison with *Il est arrivé un homme* ("There arrived a man"). Sentences with a different expletive, though, are very common: *C'est un homme intelligent* ("It's an intelligent man"). Perhaps the expletive *il* is incompatible with an Small Clause.

of subject extraction out of a *believe* clause over an adverb. We repeat both cases here for convenience:

(213)

a. John, who I assure you to be the best student in the class.

b. Who do you believe sincerely to be the best student in the class?

Interestingly, the prediction is fulfilled for the construction in (213)a. We find contrasts between subject extraction out of complements containing a V* verb as opposed to complements containing an intransitive or a transitive verb:

(214)

a. ?* John, who I assure you to have bought my oldest car

b. ?* Mary, who I assure you to have danced very well

c. ? This secretary, who I assure you to arrive always on time

d. a man who I assure you to be well appreciated by everyone

Under our view, sentences like (214)b. and c. are also examples involving extended variables. The foot of the [+HR]] chain is Case-marked by a V* verb.[100]

---

100. Although example c. is still somewhat marginal, it is better than examples a. and b. Examples with *be*, such as (213)a., are perfect. Although we do not have an explanation for the fact that sentences with *be* are better than sentences with other verbs of the V* class, it might plausibly be related to the fact that *be* seems to be able to assign accusative Case, standardly assumed to be a structural Case in some instances, as for example in the very commonly-used expression "It 's me" repeatedly and uselessly condemned by prescriptive grammarians. The marginality of (214)c. may be due to the assignment of inherent Case and the definitness effect usually associated with it (Cf. ? Who did there arrive yesterday).

The prediction is not fulfilled, however, in sentences of the type in (213)b. For these sentences, there is no distinction between complements containing V* verbs and those containing other verbs.

(215)

a.Who do you believe sincerely to have bought the ugliest antique

b.Who do you believe sincerely to have arrived late

c.Who do you believe sincerely to be appreciated by many people

This suggests that these examples do not involve extended variables and that their grammaticality should be accounted for differently. Interesting examples noted by Postal show that an adverb having scope over the matrix predicate may marginally occur after an exceptionally Case-marked NP:

(216)

a.I believe John sincerely to be the best student in my class.

Given this possibility, sentences of the type in (213)b. might have a different structure from the one we gave above. We had assumed that the position of the WH-trace occurred following the adverb, but given the existence of sentences like (216), this may be incorrect. The structure may be rather as in (217):

(217) who$_i$ do you believe t_i sincerely to be intelligent.

If so the trace would be adequately Case-marked by the verb, and the absence of contrast in (215) would follow. Moreover, this removes the main argument against the Case requirement for traces.[101]

Another example of extended variables, which we have seen several times, is the [+HR] chain created by regular subject extractions showing the *que/qui* alternation. Recall that in our view the Spec of CP is a [+HR] position. Moreover, it is not a Case-marked position. Thus we have a [+HR] chain between the Spec of IP and the Spec of CP which does not conform to clause a. of the well-formedness condition on [+HR] chains. A literal interpretation of the Case condition on [+HR] chains would rule out any subject extraction under the analysis we have proposed. But given the notion of extended variable we have

---

101. The plausibility of this view is in our sense reinforced by the fact that examples similar to (216) occur commonly in Scandinavian languages. Cf. Holmberg, Thrainsson, etc. Holmberg suggests an analysis of these examples in relation to the rule of Object Shift that he proposes. For him the stucture of similar sentences in Icelandic would be:

    [I believe John sincerely [t to have left]]

    The object has first moved out of the ECM clause by Object Shift and has adjoined to the VP of the higher clause. The fact that the V appears higher is due to movement to I. It may be that (216) is a remnant of an earlier period of English where V-to-I occurred, and perhaps Object Shift as well. For modern English, we could assume, following Pesetsky, that *believe* has raised to μ (or AGR-O) and that both the NP and the adverb *sincerely* are adjoined to VP (alternatively, the NP might be in the Spec of a lower μ). This would entail, however, that a verb moved into μ can assign Case (contrary to what is assumed by Pesetsky) as long as the NP is adjacent to it. This suggests the possibility of successive Case-marking for double-object constructions, similar to the proposal in Larson (1988), but from a Functional Projection. We will not pursue such a suggestion.

developed above, the [+HR] chain formed in cases of subject extraction by the movement of the subject to the Spec of CP is fully licensed.

Given the definition of extended variables we have proposed in the previous section, we expect sentences with extended variables to be ill formed if the [-HR] chains of which they are the foot are themselves ill formed. This is in fact confirmed by an interesting phenomenon first observed by Pesetsky (1983) which has become known in the literature as the "Surprising subject object asymmetries". Pesetsky (1983) observed that there are some unexpected subject/object asymmetries in cases of extraction from sentential complements embedded within islands. Consider the following cases:

(218)
a.*Qui est-ce que tu te demandes si Pierre crois qui viendra
Who do you wonder whether you believe will come

b.?Qui est-ce que tu te demandes si Pierre crois que Marie rencontrera.
Who do you wonder if Peter believes that Mary will meet

Under standard assumptions, both of these sentences should lead to a subjacency violation. Indeed, antecedent government (and head government) is met for the object extraction and for the subject extraction in the same way as it is met in regular cases of long distance extraction (que /qui). It is only at the next level of embedding that an island is crossed. Since ECP can be satisfied for both (218)a. and (218)b., the fact that the former is much worse than the latter remains unexplained. In our framework, there is a crucial difference between (218)a. and (218)b. In (218)a., there is an intermediate trace in the Spec of the most embedded CP. Since this trace is in an [+HR] position, it cannot be deleted. Note that this

trace meets the ECP at the level of S-structure since it is antecedent governed by a trace adjoined to the VP of "croire" and head governed by the V "croire". But as we have concluded above the chain which involves the trace in the Spec of CP and the trace in the Spec of AGR-S is a Case of Extended variable; the head of this chain (the trace in the Spec CP) is not in Case marked position. This chain can thus only be licenced under the condition b of the definition of well formedness of [+HR] chains. This condition requires that an extended variable be the foot of a well formed [-HR] chain. Here the [-HR] chain is not well formed since there is a least one link which does not meet antecedent government. As a consequence, the extended variable formed by subject extraction is not well formed, leading to an increased level of ungrammaticality. In the case of object extraction, on the other hand, the variable is well formed since it is a simple variable which occur in a Case marked position. This accounts for the more marginal status of (218)a. compared to (218)b and for other cases of "Surprising subject/object asymmetries [102] [103] This concludes our

---

102. See Pesetsky (1983), Browning(1986), Koopman & Sportiche (1988) and Rizzi (1989) for different accounts of these asymmetries and more extensive discussions.

103. Support for our analysis comes from other types of extractions which involves extended variables. Consider the following contrast:

(i)
a.?des gens que je me demande pourquoi Pierre a invités
b.*des gens que je me demande pourquoi Pierre a tous invités

To our ears the sentence (i)b. is significantly worse than the sentence (i)a. Recall that we have suggested in chapter III section 3.4.6 that in French, Floating quantifiers which occur with objects extractions are in the Spec of AGR-O . If this is correct, (i)b. involves an other case of extended variables. In this case as well, the extended variable is not licenced, since

discussion of the Case requirement on variables.

To summarize: in this section, we have considered constructions which have been claimed to involve Case marking into the Spec of CP. In our framework, since the spec of CP is a [+HR] position, this possibility must be excluded or else we would wrongly predict that NP can occur in the Spec CP of embedded sentences. We have proposed an account of these contructions in terms of the ECP. Our account has lead us to raise the question of the requirement of Case marking on variables. We have proposed that this requirement be maintained but that the notion of variable be extended so as to include [+HR] chains of length n, in which the Case requirement is not satisfied necessarily on the head of the chain. As we have seen, the notion of extended variable allows us to account for interesting contrasts among subject extractions from infinitival sentences which remain unaccounted for under the hypothesis that Case is assigned in the Spec of CP. Moreover, it also allows an account of the "Surprising subject/object asymmetries discovered by Pesetsky (1983). To conclude our study of [-HR] chains, we turn in the next section to a discussion of Subjacency.

4.7 Subjacency

In this section we turn to Subjacency, a question we had left open up to now for reasons of expositional clarity. A question often asked about Subjacency is whether or not it is governed by the same principles as the ECP. In *Barriers*, Chomsky (1986)b. establishes a

---

the [-HR] chain is not well formed. Thus the lower acceptability of (i)b. is in fact expected in our view.

parallel between Subjacency and the ECP, stating that the same barriers count for both types of violations. Lasnik and Saito (forthcoming) similarly establish a relation between the ECP and Subjacency by using the notion Subjacent in the definition of antecedent governement. Rizzi (1989), on the other hand, considers Subjacency to be entirely different from the ECP. We will make the strong (and tentative) claim that the ECP and Subjaceny are in fact identical. We propose to view Subjacency as the ECP applied at a level at which all traces are present. ECP then reapplies at LF, taking into account only traces left after trace deletion has applied. Subjacency, as we see it, involves one additional mechanism, which is what we will call the calculus of the "force" of a barrier. In our view, the calculus of the force of a barrier replaces the calculus of the total number of barriers being crossed in assessing the strength of a particular Subjacency violation. Since the concept of trace deletion is central to our proposal, we will discuss it before we turn to the issue of Subjacency. We will discuss Chomsky's (1989) proposal that trace deletion follows form the principle of Full Interpretation and argue in favor of this hypothesis and against the alternative view proposed by Rizzi (1989) which make use of the notion of referential indices. Moreover we will argue that the trace deletion hypothesis has empirical consequences and allows an explanation of the fact that Floating Quantifiers are incompatible with [-HR] chains, a fact which up to now has remained unaccounted for.

4.7.1 Trace deletion

In the *Barriers* framework, following the original proposal by Lasnik and Saito (1984), the mechanism of trace deletion essentially serves to distinguish WH-argument chains from adjunct chains;[104] it is instrumental in distinguishing Subjacency violations from ECP violations. In an argument-operator chain, if an intermediate trace cannot be τ-marked because of a failure of antecedent government, this trace may be erased. If the trace is erased, there will be no ECP violation at LF. A faulty trace in an adjunct chain, however, cannot be erased, so it will always lead to an ECP violation.

Chomsky (1989) proposes viewing the deletion of traces in operator-variable chains as a consequence of the principle of Full Interpretation. In Chomsky's view, the Principle of Full Interpretation determines what counts as a legitimate object at LF. At LF, the elements relevant for interpretation can all be viewed as chains of the type $(\alpha_1 \ldots \alpha_n)$, including one-membered chains. The list of the legitimate LF objects is given below:

(219) (From Chomsky (1989) p 63)

1) Arguments: each element is an A position, α_1 Case marked and α_n theta-marked, in accordance with the chain condition.

---

104. It has been extended in Chomsky (1989) to traces of certain functional projections. We will not be concerned with these cases in the present discussion.

2) Adjuncts : each element is in an A-bar position

3) Lexical elements: each element is in an X^o position.

4) Predicates, possibly predicat chains if there is predicate raising, VP movemnt in overt syntax, and other cases

5) Operator-variable constructions, each a chain (α_1, α_2), where the operator is in an A-bar position and the variable α_2 is in an A-position.

Chomsky proposes a condition of homogeneity on legititmate LF chains: NP-chains and adjunct chains are considered to be homogeneous chains because they involve strictly A-positions and strictly A'-positions respectively. Op-variable chains, on the other hand, are considered to be non-homogeneous chains because they include an intervening adjunct chain which is not needed for interpretation. The intuitive idea behind the principle of Full Interpretation is that any element which does not contribute to semantic interpretation or the licensing of some semantic interpretation at LF must be eliminated. Elements which are legitimate interpretable LF objects, on the other hand, must be present. Under this view, it follows that intermediate traces of operator-variable chains must obligatorily be eliminated, since they are not part of a homogenous chain; traces of NP and Adjunct chains must obligatorily be present and thus can never delete. Consequently, if an intermediate trace in an A-chain or an Adjunct-chain fails to be antecedent-governed, an ECP violation will result, since the faulty trace cannot delete.

We have eliminated the notion A vs A'. Clearly however, Chomsky's proposal can be reinterpreted easily in our framework by simply

subtituting the notions [+HR] and [-HR] for the A/A' terminology. This substitution has one interesting different conceptual consequence. In Chomsky's view, adjunct chains are homogeneous because they involve strictly A' positions. But the nature of these positions is not completely homogeneous. Consider the case of a wh-extracted adjunct. Since in Chomsky's view the final landing site of a wh-movement is the Spec of CP, a wh-extracted adjunct will be a chain which involves adjoined positions in its initial and intermediate steps but not in its final step, which is a substitution. In our view, on the other hand, since we have argued that the final step of a wh-extraction is an adjoined position, i.e a [-HR] position, the chain of an extracted adjunct will involve only adjoined positions. Thus it will be completely homogeneous. We obtain the following picture: AN-chains involve only [+HR] positions, adjunct chains only [-HR] positions and operator-variable chains a mixture of [+HR] (the initial step) and [-HR] (intermediate and final step) positions. Deletion can then be interpreted as a tendency to uniformize chains wrt the type of positions they include. In the case of an operator variable chain, deletion of the operator in a [-HR] position will be prevented by other principles of UG such as for instance Binding theory[105].

---

105. Consider a classical case of Object extraction. Suppose that the operator is deleted. Then the empty category required by the projection Principle will failed to be identified. It cannot be PRO, if as standardly assumed PRo must be not be governed. It cannot be an anaphor since bindng by another NP would violate the θ-criterion. In some languages, it may be pro, if in these languages empty objects are licenced (may be in the way suggested by Rizzi (1986). But in languages like English, empty pro in object positions are not licenced. Thus the sentence will be ruled out by independent principles.

4.7.1.1 Referential Indices

Rizzi (1989) proposes an alternative way of distinguishing between argument chains and adjunct chains. He argues that traces of θ-marked arguments bear referential indices but that traces of adjuncts do not. When antecedent government cannot obtain between two intermediate links of a WH-chain, chain formation does not occur. In this case, elements bearing a referential index can resort to another type of antecedence relation, namely Binding. Binding is defined as follows:

(220) X binds Y iff

(i) X c-commands Y, and

(ii) X and Y have the same referential index

As implied by the definition in (220), Binding can obtain at any distance. This alternative antecedence relation is obviously not available for elements not bearing referential indices, namely adjuncts and, more generally, non-theta-marked elements. Since only elements which are theta-marked may receive a referential index, the notion of Binding supplants and replaces the notion of lexical or theta-government. But Binding differs in an important respect from lexical government: it applies equally to subjects and objects, while in the standard view, lexical government is restricted to objects. Rizzi's distinction between argument extractions and adjunct extractions is crucially based on the hypothesis that adjuncts do not bear referential indices. Consider, however, the following paradigm, due to D. Pesetsky (pc):

(221)

a.Bill spoke quickly$_i$ because Mary spoke that way$_i$

b.*Bill spoke that way$_i$ because Mary spoke quickly$_i$

c.Because Mary spoke that way$_i$, Bill spoke quickly$_i$

The contrast between (221)a. and (221)b. suggests that adverbial expressions are subject to principle C of the Binding theory. In (221)b., the coreferring expression *that way* c-commands the adverb *quickly*, so the sentence is excluded. In (221)c., on the other hand, the coreferring expression *that way* does not c-command the adverb *quickly*, and the sentence is acceptable. This paradigm is predicted if adverbial expressions are subject to principle C and must generally be free. If so, this paradigm casts serious doubts on Rizzi's hypothesis that adjuncts cannot bear referential indices.

While Rizzi's notion of Binding distinguishes operator-variable chains of arguments from adjunct chains, it cannot straightforwardly account for the fact that NP-movement chains, which are chains involving theta-marked arguments, behave like adjunct chains with respect to antecedent government. In NP chains, as in adjunct chains, if antecedent government is not met for every link of the chain, an ECP violation results. The necessity for a requirement of strict antecedent government for NP chains is shown by the following example, originally due to Mark Baker:

(222) *John seems [that it was told t] that Bob will leave.

In this example, *John* has raised from the complement position of *told* to the subject position of *seems* in one move. In Rizzi's theory,

antecedent government will be blocked by the presence of the expletive *it*, a potential binder for the trace of *John*. Relativized Minimality is violated, so the chain (*John*, t) will not be constructed.[106] But this is not sufficicent to exclude (222). Although antecedent government is not met, Binding should obtain. The trace is lexically governed by a θ-assigning verb; it will receive a referential index and is appropriately c-commanded by the NP *John*, its binder. Since the example is sharply excluded, it is clear that Rizzi must assume that in this case, as opposed to other cases where antecedent government is blocked for argument extraction, Binding does not suffice to make the sentence grammatical.

Rizzi argues that the ungrammaticality of (222) is due to a violation of the theta-criterion. Recall that in his view, a chain cannot be formed if antecedent government is not met. Thus, in (222) no chain can be formed. As defined in Chomsky (1981, 1986), the theta-criterion applies to chains. Similarly, it is to chains that theta-roles are assigned. Since in (222) no chain has been constructed, the theta-role cannot be assigned and (222) will violate the theta-criterion.

---

106. This slightly more complex case of super-raising is accounted for as an ECP violation in our theory:

(223) *John seems that it was told t that Mary had left.

We have assumed that English has a rule similar to the *que/qui* alternation which holds only for the empty complementizer of tensed sentences.(See xx quoted in Rizzi for this assumption, which is supported by the fact that some English dialects have it with an overt *that*). In fact, this case turns out to be equivalent to the simple cases of super-raising: the barrierhood of the intermediate IP is not suppressed because the transitivity of the agreement relation does not obtain. See section 4.5. for a discussion of simple cases of super-Raising.

Rizzi's account of (222) crucially relies on the hypothesis that no chain can be formed whenever antecedent-government is not met in one step of the movement. Under this view, the ECP appears to be a condition on the process of chain formation rather than a well-formedness condition on formed chains. This entails that examples of WH-extraction across islands, as well as any other examples of movement in which one link does not meet the ECP, do not allow the construction of a chain. As we will show, however, several facts suggest quite strongly that, contrary to Rizzi's hypothesis, a chain must be formed at some level of representation in cases of extractions out of islands. These facts involve phenomena standardly assumed to be licensed by chains, such as reconstruction (or more appropriately Chain-binding (Cf. Barss (1984) (1986))) and parasitic gaps.

Let us first turn to facts involving reconstruction. Consider the following example:

(224) Which picture of himself does John think that Bill likes?

This sentence is a typical example of what are traditionally called connectivity or reconstruction effects. The hallmark characteristic of sentences such as (224) is that the anaphor is outside the c-command domain of its understood antecedent. In this particular example, the anaphor can additionally be ambiguously bound both by *John* and by *Bill*. On the basis of the ambiguous interpretation of such sentences, Barss (1986) has argued for and developed a theory of Binding which in cases such as (224) is parasitic on the existence of well-formed chains. We will not give a summary of Barss' formulation of the conditions on chain binding nor of his arguments. The reader is

invited to consult Barss (1986) for a detailed account. Suffice it to say that the type of binding effects found in sentences like (224) crucially presupposes the existence of a chain at some level of representation. Consider now a typical case of extraction out of a WH-island:

(225) ??Which picture of himself does John wonder where Bill hung.

Notice first that this sentence has only the marginality of a typical island violation. There is no additional binding theory violation which would considerably degrade the sentence and put it on a par with sentences like (226):

(226) John thinks that Mary likes himself

Consequently, in (225) the anaphor must be adequately bound. The second interesting fact about (225) is that the sentence allows for an ambiguous interpretation. *Himself* can refer either to John or to Bill. Although the reference to Bill may be a little harder to get in (225) than in (224), it remains clearly possible.[107] The fact that these typical connectivity effects obtain accross a WH-island strongly suggests that a chain must be formed at some level of representation in this example.[108] Examples with extraction out of other types of

---

107. The increased marginality of the reference to Bill may plausibly be attributed to the fact that the chain is somewhat deficient: one link of it does not respect appropriate locality conditions.

108. As noted by Howard Lasnik (pc), the fact that the anaphor can refer to John may provide an argument that the WH-element containing the anaphor must have adjoined somewhere between the complementizer and the subject. There are, in fact, several possible landing sites: CP, VP, and any functional projection which may occur in between. But if, as we suggested above,

islands point in the same direction.[109]

(227)

a. Which picture of himself did John leave the country before Peter could publish

b. Which picture of himself did John hear the rumor that Peter was going to make public.

A second argument pointing to the same conclusion can be made on the basis of parasitic gaps. A standard assumption is that parasitic gaps are licensed by operator-variable chains at S-structure.[110] Bearing this in mind, consider the examples in (228):

(228)

a.??This is the girl that everybody who meets wonders why John likes.

b.??Which book does John wonder why Mary filed before reading

(228)a. and b. are examples of parasitic gaps which are licensed by WH-extractions which have crossed WH-islands. Although the sentences in (228)a. and b. are marginal, due to the WH-island violation and the usual marginality of parasitic gaps, they are far more acceptable than

---

neither VP nor any of the intervening functional projections is a barrier, there is no reason why adjunction to any of these categories should occur. CP, on the other hand, as we have argued, is a barrier. Plausibly, then, CP is the intermediate landing site where the WH-element containing the anaphor has adjoined. If so, sentences of the type in (225) provide some support for CP adjunction.

109. The fact that anaphor reconstruction is allowed across a WH-island is noted by Rizzi. Adopting an idea put forth by Cinque(1989), he argues that anaphor reconstruction is possible for elements having an "intrinsic referential quality." See Rizzi (1989) for details

110. Cf. Cinque(1983) for a different analysis.

constructions in which no chain licenses a parasitic gap. (228)b. suggests that a chain must be constructed at least up to the WH-complementizer of the WH-island so as to license the parasitic gap in the adjunct clause. (228)a. suggests that a chain must also be constructed above the WH-complementizer. The relative acceptability of the chain-Binding effects in (225) and of the parasitic gaps in (228) provides some strong arguments in favor of the existence of chains at some level of representation in constructions in which antecedent-government fails. Parasitic gaps are licensed at S-structure. Chain-binding, as conceived by Barss (1986), is also licensed at S-structure.[111] We thus conclude, contrary to Rizzi's proposal, that the failure of antecedent government does not prevent the building of a WH-chain.

If the conclusion of the preceding paragraph is correct, it argues in favor of a theory which builds WH-chains at S-structure even in cases where antecedent government does not obtain. This conclusion thus provides an indirect argument in favor of Chomsky's hypothesis that WH-traces are first constructed at S-structure and then subsequently

---

111. A possible argument that Chain Binding must be an S-structure effect is given by WH-in-situ constructions. Consider (i):

(i) To whom does Bill think that John will give which picture of himself?

Under standard assumptions, the WH-in-situ element will raise at LF to the matrix complementizer. It is not possible, however, to construe *himself* as referring to Bill in (i). Thus, binding must have applied prior to the formation of the LF chain.

deleted at LF. It also casts some doubts on Rizzi's view of Binding as an alternative to chain formation.[112]

Given the problems faced by Rizzi's proposed alternative to distinguish argument chains from adjunct chains, we will adopt Chomsky's view on trace deletion. One drawback of Chomsky's proposal is that the notion of homogenous chain remains conceptually rather vague with respect to the difference between operator-variable chains and adjunct chains. A possible interpretation may come from the following considerations: the fundamental difference between adjunct chains and NP movement chains on the one hand and operator-variable chains on the other is that for the former kind of chains, a chain of length one can be a fully interpretable LF object: it is an argument or it is an adverb. Chomsky also proposes that predicate chains are homogenous. Again, predicate chains of length one are fully interpretable LF objects. Pushing this a little further, we note that NPs, Adverbs, and verbal (or adjectival) predicates are all elements which participate in the fundamental relation of predication. NPs are predicates of individuals, while adverbs and verbs are predicates of events. Possibly we can then view the notion of homogeneous chains as being restricted to predicative chains. In the case of operator-variable chains, it never makes sense to consider a chain of length one: two elements are minimally needed, an operator and a variable.

---

112. The argument we have raised is, of course, not absolute. One could well imagine that the conditions on Chain Binding and on parasitic gaps are stated on the Binding relation rather than on the Chain relation. But this would make the Binding relation so similar to chains that we would face an inelegant redundancy.

Thus, there are no operator-chains of length one which are interpretable LF objects. Operator-variable chains also differ maximally from the other type of chains in that they participate in the fundamental relation of Quantification.

Although these considerations do not provide justification for obligatory trace deletion in operator-variable chains, they establish a link between NP-chains and adjunct chains which Rizzi's theory (or Lasnik and Saito's) does not naturally capture. In Rizzi's terms, NP-chains and adjunct chains are maximally distinguished, since the former are θ-chains while the latter are non-θ-chains par excellence. The fact that it is only this two types of chains which are subject to a strict antecedent government requirement is thus rather surprising. Under Chomsky's view, the similarity between argument chains and adjunct chains is quite natural. The fact that these two types of chains pattern together with respect to antecedent government is thus not surprising.

In addition to this conceptually pleasing consequence, we will argue in the next paragraph that there is in fact an interesting empirical consequence of the hypothesis of trace deletion. This empirical consequence involves the licencing of Floating Quantifiers.

4.7.1.2 Quantifier Float and operator variable chains

In chapter II section 2.3.2, we argued for the validity of the following descriptive generalization:

(229) Floating quantifiers can occur in intermediate positions of an AN-chain but are incompatible with the intermediate position of argument Vbl-chains.

Throughout Chapter III, we have seen that this generalization appears to have wide cross-linguistic validity. Up to now however, this generalization has remained unexplained. Under Sportiche (1988) theory of quantifiers, the presence of a floating quantifiers is assumed to signify the presence of a NP trace. We have argued in the previous section that the hypothesis of trace deletion in operator variable chains seems more appropriate to the treatment of the distinction between argument and adjucnt chains than the alternative hypothesis proposed by Rizzi (1989). In particular, we have adopted Chomsky's proposal that the intermediate traces of operator variable chains must delete to create an admissible LF interpretable object which satisfy the Principle of Full interpretation. This hypothesis offers in fact a simple explanation for the generalization (229). Let us assume that the presence of a trace at all levels of representation is necessary to licence the presence of a floating quantifier. Given this assumption, the absence of floating quantifiers with operator-variable chains ([-HR] chains of arguments) will simply follows from the requirement of trace deletion: intermediate traces of operator-variable chains must delete to satisfy FI; they can thus not support floating quantifiers. On the other hand, the possibility of floating quantifiers with [+HR] chains is prediced since traces in [+HR] positions do not delete and can thus support floating quantifiers. Recall that in Chapter I section 1.1.2.3, we have suggested, on the basis on the impossibility for an FQ to quantify over the understood

subject of a passive construction, that FQ can only modify traces which are syntactically reppresented. We can now make this suggestion more precise, and say that FQ cannot modify an ec. not present at LF.

Recall that in Chapter II we had noted that the distribution of Floating quantifiers in French seemed to contradict the generalization (228). Floating quantifiers are apparently not incompatible with some instances of operator-variable chains in French. We repeat the examples showing this apparent contradiction below:

(230) a. ces livres, que j'ai tous lus sont facinants
these books, that I have all read are fascinating

b. les enfants que tu as tous grondés sont partis en pleurant
the children that you have all scolded left crying

c. les chaises que tu as toutes repeintes cet après-midi
the chairs that you have all repeinted this afternoon

We argued however that this apparent contradiction was due to some other factor, which we have so far left unexplained. If the proposal we made to explain the absence of quantifier float with operator-variable chain is correct, we are lead to the conclusion that the floating quantifiers in (230) must be licenced by traces in [+HR] positions, i.e traces which do not delete at LF. Recall that we have argued that in French movement of an object NP to the Spec of the intermediate functional projection AGR-O is not possible, because AGR-O is not a Case-assigning functional projection. But since French is a language in which Verb raising takes place at S-structure, nothing prevents movement through this [+HR] position. In fact the occurence of past participle agreement with clitic and with Wh-movement shown in

(231) provides independent evidence that movement of objects through the spec of AGR-O is possible.

(231) a. Je les ai repeintes
I repeinted them

b. Les chaises que j'ai repeintes cet après midi....
The chairs that I repeinted this afternoon.

We thus suggest that the examples of floating quantifiers in (230) simply reflect the fact that movement has first proceeded through the Spec of AGR-O before moving further to [-HR] adjoined positions. Under this view, the structure of examples like (230) is as in (229):

(232) [les chaises$_i$ [$_{C''}$ que [$_{I''}$ t_i j'ai [$_{AGR\text{-}O''}$ [toutes t_i] [$_{AGR\text{-}O'}$ repeintes$_k$ [$_{VP}$ t_k t_i]....

As required, the floating quantifier is in an [+HR] position, and is licenced by a trace in [+HR] position which does not delete at LF. This type of [+HR] chains are other examples of what we have called "extended variables" in the previous section. They are thus licenced under condition b. of the definition (xx). The impossibility of such chains in English simply follows from the theory of object movement we have elaborated in Chapter III. Our definition of Dynamic Minimality (Chapter III section 3.3.5.1) precicts that movement of an object through the Spec of an intermediate functional projection is not possible in English[113] since Verb movement does not occur. VP is a D-minimality barrier in English and consequently any movement of object must proceed by adjunction or first through the Spec of VP, the latter

---

113. With the exception of passive constructions and other derived subject constructions for which we assume that movement to the Spec of VP occurs first, thus avoiding the D-minimality barrier.

possibility being restricted to cases where the Spec of VP is not a θ-position.

The simplicity of the account of the distribution of floating quantifiers that Chomsky's (1989) notion of chain homogeneity permits in turn provides strong empirical support for this hypothesis. After this digression we now return to the main topic of this section, namely Subjacency.

4.7.2 Subjacency

4.7.2.1 ECP vs Subjacency

As is well known, the types of violations which result from extraction out of islands vary both with the element extracted (adjuncts vs. arguments) and with the type of islands from which the extraction has occurred (object extractions out of adjunct islands or subject islands are usually perceived as more "severe" violations than object extractions out of WH-islands, especially infinitival WH-islands). The first type of variation, adjuncts vs. arguments, has been analyzed in the *Barriers* framework as a difference between ECP and Subjacency violations.[114] The second type of variation, namely the severity of a

---

114. Cf. also Huang (1983) and Lasnik and Saito (1984) for an earlier account along these lines

Subjacency violation relative to a particular type of island, is expressed in Chomsky (1986)b. by a calculus that takes into account the number of barriers being crossed by a particular movement:

(233)
"our intuitive idea is that movement should become "worse" as more barriers are crossed, the best case being the crossing of zero barriers." (Chomsky 1986b. p. 28)

We developed an approach to the ECP in the previous sections. Moreover, in section 4.7.1, we argued for the notion of trace deletion as relevant to the distinction between argument-operator chains (which are non-homogeneous chains in Chomsky's view) and adjunct and [+HR] chains (which are homogeneous chains). We are thus now ready to turn to an examination of the "weaker" types of violations, namely violations which in the *Barriers* framework fall under the principles of Bounding theory.

In *Barriers*, two notions are important for Subjacency: the notion of the crossing of a barrier and the number of barriers being crossed. In Chomsky's view, the crossing of a barrier occurs when a trace α is included in a maximal projection which is a barrier and which excludes the antecedent β of α. Recall that exclusion is defined as in (233):

(234) Exclusion

α excludes β if no segment of α dominates β

Government in turn is defined in terms of exclusion:

(235) α governs β iff α m-commands β and there is no τ, τ a barrier for β, such that τ excludes α.

Given these definitions, consider the following structure:

(236) $...[_{\delta}...\alpha ...[_{\tau}...\beta...[_{\tau}...\sigma...]]]$

In (236), σ is dominated by τ, β is not-excluded by τ, and α is excluded by τ. Thus, β governs σ, since β m-commands σ and β is not excluded by τ, which dominates σ. Assume that τ is a barrier for σ. If β is an antecedent for σ, no barrier is crossed in the movement from the σ to β. On the other hand, if α is the antecedent of σ, since α is excluded by the barrier τ, the movement from the postion of σ to that of α crosses a barrier, namely τ. Recall that in the framework we have developed, we have assumed free adjunction to every maximal projection. Consequently, in our view no barrier will ever be crossed (in the technical sense given to this term in *Barriers*). Consider, for instance, an example of argument extraction from a WH-island:

(237) who do you wonder where Mary met

(238) $[_{CP1}$ who $[$do $[_{IP1}$ t_1 $[_{IP1}$ wonder $[_{CP2}$ t_2 $[_{CP2}$ where $[_{IP2}$ t_3 $[_{IP2}$ Mary met $t_4]]]]]]]]$

CP2, although L-marked, is a barrier by inheritance from IP2. But since adjunction to CP is generally allowed, there is no barrier which is crossed in the movement from the position of t_3 to the position of t_2 in the structure in (238). CP2, although a barrier for t_3, does not exclude t_2. We have assumed, following Rizzi, that in this case antecedent government fails anyway, due to a violation of Relativized Minimality: t_3, although not separated from t_2 by a barrier, is separated from it by an intervening potential binder. This is enough to induce a failure of antecedent government of t_3 by t_2.

We could reanalyze the failure of antecedent government in (238) in terms of crossing by defining a new type of barrier as follows.

(239) τ is a barrier for β if

τ is the first segment of a maximal projection including a Operator Opacity inducer for β.

As a second step, we would also need to redefine the notion of crossing. Indeed, under (239) the first segment of CP2 in (238) will be a barrier. But this does not help if crossing is defined in terms of exclusion rather than inclusion: CP2 does not exclude t_2, so unless Subjacency barriers are defined in terms of inclusion rather than exclusion, no barrier will be crossed. Thus, it seems that for the purpose of Subjacency, we need to redefine crossing in terms of inclusion rather than exclusion. If we do this, (238) will involve the crossing of a barrier, namely the first segment of CP2, which includes t_3 but does not include t_2. Note, however, that if we redefine crossing in terms of inclusion, this will create serious problems elsewhere. In particular, a barrier would be crossed in each sentence where IP is a barrier; that is, even a simple matrix question would involve one barrier-crossing at IP. Consider the following example:

(240) What did you see?

[CP What [C·did [IP t_1 [IP you see t_2]]]]

t_2 is included in IP. t_1 is not included in IP. Thus, crossing is defined in terms of inclusion, then IP will be crossed in this simple object extraction. Redefining crossing in terms of inclusion amounts to supressing the benefit of the mechanism of adjunction to escape the

barrierhood of a given category. Consequently, every non-L-marked BC and every XP which is a barrier by inheritance will be a barrier counting for Subjacency. Clearly, this is not a desirable result. One of the unwanted consequences of such a redefinition would be that the same number of barriers would be crossed in a WH-island as in a regular extraction out of a complement sentence. In other words, object extraction from a regular sentential complement would be expected to be just as difficult as object extraction out of a WH-island. This is clearly the wrong result.

Moreover, in our theory there are cases of extraction out of islands which do not involve any failure of antecedent government, such as extraction out of a subject island. Consider an example of argument extraction from a subject island.

(241) Which book do you think that reading at night will give you nightmares?

(242) [$_{CP1}$ which book ...you think [$_{CP2}$ that [$_{IP2}$ t_1 [$_{IP2}$ [$_{CP3}$ t_2 [$_{CP3}$ Φ [$_{IP3}$ t_3 [$_{IP3}$ PRO reading t_3 at night]]]] will give you nightmares]]]]

The only violation in (242) is the failure of head government for t_3, the trace adjoined to the subject IP3. Recall that t_3 is not properly head-governed by the head of IP3, since it is not within the first projection of this head. It is not properly head-governed by C_0 in CP3, either, because IP3 is a barrier, since it is not L-marked. Since antecedent government of intermediate traces is not at stake in this type of example (all traces are antecedent-governed, since adjunction

is free), it is clear that the notion of barrier crossing is irrelevant for this case.

We conclude that an approach to Subjacency such as the one advocated in *Barriers* is unavailable in the theory we have developed. We thus need an alternative approach. In this section, we will outline a tentative solution to the problem of Bounding theory. Our suggestions for now will remain sketchy: a full development of these ideas is beyond the scope of this dissertation.

Let us begin by noting that in the two cases we have just reviewed, as well as in all other cases of extraction of an argument out of an island, the ECP would be violated if the intermediate traces of arguments were not deleted. This observation immediately suggests a possible approach to Subjacency in terms of the ECP. Let us assume that ECP applies both at S-structure and at LF. If so, argument extraction from an island will violate the S-structure version of the ECP. At LF, however, traces of argument-operator chains will delete and no ECP violation will occur. We propose to reduce the cases of argument extraction handled by Subjacency in *Barriers* to violations of the ECP at S-structure. Under this view, Subjacency violations and ECP violations are not due to two distinct sets of principles, but are in fact due to the same principle applying at two different levels. "Subjacency" violations are in our terms ECP violations at S-structure and "ECP violations" are in our terms ECP violations at both S-structure and LF. For adjunct chains and [+HR]-chains, since we have assumed following Chomsky (1989) that no trace deletion occurs, ECP will be violated both at S-structure and at LF.

Let us investigate in more detail a possible formalization of this intuitive idea. We will adopt Lasnik and Saito's mechanism of value assignment to a τ-feature and assume a view of the ECP as a filter defined as follows:

(243) *[$t_{[-\tau]}$]

All traces satisfying the ECP at S-structure will be assigned a positive value for their τ-feature. All traces failing the ECP at S-structure will be assigned a negative value for their τ-feature. The τ-filter applies first at S-structure, and traces marked for a negative value will violate the S-structure τ-filter. As assumed by Lasnik and Saito, a feature value, once assigned to a trace, cannot be modified. Recall that following Chomsky (1989), we have assumed that at LF all intermediate traces of argument-operator chains (those not in [+HR] position) will be deleted. Consequently, for complement extraction, the only remaining traces will be the traces in the original site of extraction, which will have been marked for a positive τ-value at S-structure and will thus satisfy the LF τ-filter. Thus, argument-operator chains will only violate the S-structure τ-filter. For adjunct chains and [+HR] chains, on the contrary, no trace deletion occurs. Traces which have been marked negatively for their τ-feature at S-structure will remain negatively marked at LF. Consequently, the adjunct chains and [+HR] chains which contain negatively marked traces will satisfy neither the S-structure τ-filter nor the LF τ-filter.

In the Lasnik and Saito framework, as well as in the *Barriers* framework, this view of the repeated application of the ECP is not possible, essentially because of a problem with the absence of the *that*-t effect with adjunct extractions. In both the L&S framework and the *Barriers* framework, the fact that the presence of an overt complementizer does not induce an ECP violation for adjunct extractions is rather surprising. In *Barriers*, Chomsky proposes that the presence of the complementizer induces a Minimality barrier for subject extraction. Crucially, however, the Minimality barrier must be suppressed in cases of adjunct extraction out of regular sentential complements. Chomsky assumes, following Lasnik and Saito, that the ECP need not be satified for adjunct chains prior to LF. At LF, the complementizer deletes and thus no longer induces a minimality barrier: consequently, the intermediate trace of a long-distance extracted adjunct in the Spec of CP can properly antecedent-govern its trace.

In our framework, we have argued that both adjunct and object WH-phrases adjoin to an intermediate CP on their way to a higher COMP. Thus, the presence of an overt [-WH] complementizer is never expected to induce any ECP violation. Consequently, in cases of adjunct extractions out of sentential complements, the ECP can be satisfied directly at S-structure in our framework. Adjunct chains can thus be assumed to be created at S-structure and checked for the ECP at both S-structure and LF.

Note that our view of the difference between "Subjacency violations" and "ECP violations" as resulting from a successive application of the

ECP at S-structure and at LF allows us to partially derive the notion of "stronger" violation usually attached to ECP violations. In the standard theory, there is no a priori reason why a violation of the ECP should be more "severe" in terms of native speaker judgments than a violation of Subjacency. Both in some sense involve a notion of locality of antecedence. In our view, however, there is a simple sense in which the stronger violation might be derived: traditional ECP violations induce ECP violations at two levels of representation while traditional Subjacency violations induce ECP violations at only one level of representation. The "stronger" violations can be viewed as resulting from cumulative ECP violations. We thus derive the strength of ECP violations partially from a cumulative effect, a rather natural assumption.

The qualification "partially", however, is important. It is clear that some types of violations which have been standardly regarded as ECP violations involve no cumulative violations: this is true in particular of cases of WH-in-situ adjuncts.[115] Consider the following case:

(244)* Who bought this why ?

In the Lasnik and Saito (1984) framework, the ungrammaticality of this sentence is interpreted as an ECP violation. The LF representation of (244) will be schematically as in (245):

---

115. It is also true of any ECP violation in a language with WH-in-situ, for instance Chinese, Japanese, Hindi, etc.

(245) [Why$_j$ [who$_i$ [t$_i$ bought this t$_j$]]]

The intuitive idea behind Lasnik and Saito's approach is that the trace of the adjunct cannot be antecedent-governed by the WH-element *why* because the subject WH-element occupies the c-commanding complementizer position. In our view, (245) can be seen as an Operator Opacity violation at LF. *Who*, a potential binder for the trace t$_j$, is a closer binder for t$_j$ than its legitimate antecedent *why*. Thus, the ECP is violated; but since the movement of *why* has occurred only at the LF level, the ECP will be violated only at one level of representation, namely LF. A strictly cumulative approach to the ECP would predict that the type of violation found in sentences such as (245) is on a par with typical Subjacency violations. As reported in the literature however, this type of violation is usually perceived as more severe than the violation of an argument extraction out of a WH-island, for instance, such as the example given above in (237). If this is correct, then (245) shows that a single violation of the ECP at LF is inherently more severe than a single violation of the ECP at S-structure. Consequently, the cumulative effect of S-structure plus LF ECP violations only partially derives the severity of ECP violations for overt syntactic movement.

Some independent reasons might possibly be found to explain this fact; a thorough investigation of WH-in-situ constructions would be required to investigate such a possibility, however, and such an investigation is beyond the scope of this dissertation. We thus leave this topic for further reasearch. For the moment, we will simply admit that for reasons that remain unclear, single violations of the ECP at LF induce

more severe grammaticality judgments than single violations of the ECP at S-structure. Note, however, that the strict division between LF and S-structure ECP violations is in many cases an oversimplification. Clearly, argument movement also occurs at LF, but, as has been claimed in the literature, such movement violates neither ECP nor Subjacency. Consider the following case:

(246) Who wonders where we bought what? (= L&S () p.)

According to L&S, this example is two-ways ambiguous: the ambiguity is due to the freedom of movement of *what* at LF, where it can take scope over either the embedded sentence or the matrix sentence. This is schematically represented in (247) below:

(247)

a. [who wonders [what$_i$ [where we bought t_i]]]]

b. [what$_i$ [who wonders [where we bought t_i]]]]

Assuming that this is correct, sentences of the type in (246) are a problem for our approach. In our view, both derivation (247)a. and derivation (13)b. violate the ECP at LF. In both cases, we have a classical example of Operator Opacity violation. If, as we have suggested above, ECP violations at LF result in stronger violations than ECP violations at S-structure, we predict that both derivations should lead to strong ECP violations. This is clearly the wrong result.[116]

---

116. The possibility that LF argument movement leaves no intermediate traces is not available to us. The original trace of *what*, t_i in (247) needs to receive a positive value for its τ-feature so as to satisfy the LF τ-filter. This will be possible only if t_i is

One way of getting out of this problem would be to assume, as first proposed by Chomsky (1982),[117] that there are in fact two LF levels, LF and LF'. If so, we can further assume that LF argument extractions take place at LF in a regular fashion, i.e., leaving intermediate traces wherever needed. The original trace of t_i in (247) is now able to receive its positive value for the τ-feature. Next, trace deletion occurs, erasing all intermediate traces of the WH-argument chains between the levels of LF and LF'. At LF', finally, all chains are checked again for ECP violations. Under this view, the process of ECP checking and trace deletion repeats itself between the level LF and LF', just as it occurred between the levels of S-structure and LF. In a sense, our need for a level of LF' is just an expositional device. We can just as well imagine that the process of trace deletion occurs whenever needed to eliminate all traces which are part of a non-homogeneous chain at any given point in the derivation between S-structure and LF, conceived now as the final level of representation. ECP is simply checked after all trace deletion has taken place and rules out any homegenous chain which contains any trace with a negative value of its τ-feature. Recall that in Chomsky's view, which we have adopted, homogenous chains are adjunct chains and [+HR] chains for which trace deletion never takes place. It is only at this final

---

antecedent-governed at some level of representation. Given our assumption that IP is a barrier, the WH-element raised at LF will never be able to directly antecedent-govern t_i. Thus, if LF movement of arguments proceeds without leaving intermediate traces, t_i will receive a negative value for its τ-feature and the sentence will be wrongly excluded.

117. Cf. also Safir LI 1987) for additional arguments.

level of representation, where all trace deletion has taken place and only homegenous chains are left, that ECP violations will lead to "strong" violations.

Assuming that this view of the distinction between strong ECP violations and weak ECP violations is plausible, we still need to deal with two important problems. First, as has been argued, LF movement of arguments does not lead to Subjacency violations. Second, as is now well established, Subjacency violations are variable in strength. At this point in our proposal, we can only distinguish two degrees of violations: ECP violations, which occur at the final level of LF representation and which we have called "strong" ECP violations, and ECP violations which occur at intermediate points in a derivation between S-structure and the final LF level of representation, which we have called "weak" ECP violations. In our terms, strong ECP violations correspond to ECP violations in the *Barriers* framework, while weak ECP violations correspond by and large to all Subjacency violations in the *Barriers* framework, with the exception of LF argument movement, which is not usually taken into account in most theories of Subjacency.

With regard to the first problem, it has been argued by D.Pesetsky,[118] contrary to L&S, that an argument WH in situ within islands leads to some degree of variation in judgments. If this is correct, it may in fact provide some support for our approach. Again, a solution to this

---

118. Following Choe and Nishigauchi Ref?? Cf. also Logombardi (unpublished) ref??

problem would require an extensive investigation of WH-in-situ constructions, and is thus beyond the scope of this dissertation.

4.7.2.2 Weak vs Strong Subjacency

Let us turn to the second problem, that of the variation in the strength of various types of typical Subjacency violations. Recall that in our view no barrier is crossed, under the definition of barrier crossing which is used in the *Barriers* framework. We can thus not rely on the numbers of barriers being crossed to evaluate the strength of a Subjacency violation. In every case of island extraction we have considered, however, a barrier dominates a trace which violates the ECP either because of a failure of antecedent government or because of a failure of head government or both.

A possible alternative to counting barriers for the evaluation of the strength of a weak ECP violation would be to take into account what I will call the "force" of a barrier. Let us examine more precisely the four cases of islands that we have considered, namely adjunct islands, subject islands, WH-islands, and complex NP islands. Consider first the case of adjunct islands. We have the configuration sketched in (248):

(248) $[_{CP1}$ WH $[_{IP1}..[_{I'}....[_{CP2}$ ADV $[_{IP2}$ t_1 $[_{IP2}.....]]]]]]$

Recall that as we proposed in the previous section, an adverb in adjoined to CP will induce an Operator-Opacity violation. Under the

definition in (239) above, CP2 is a barrier for t_1 on two counts. First, it is a barrier by inheritance, since it dominates IP2, a blocking category,[119] and second, it is a barrier because it is not L-marked and it is a BC.

Let us now turn to a case of extraction out of a subject.[120] The schematic structure is given in (249):

(249) [$_{CP1}$ WH [$_{IP1}$ [$_{CP2}$ t_1 [$_{CP2}$ [$_{IP2}$ ...]]] [$_{I'}$ I]]]

IP1 is a barrier for the head government of t_1. As with adjuncts, IP1 is a barrier for t_1 on two counts: it is a barrier by inheritance from CP2, and it is a barrier because it is itself a non-L-marked blocking category.

Next, let us turn to the case of WH-islands:

(250) [$_{CP1}$ WH_i [$_{IP1}$ [$_{VP}$ V [$_{CP2}$ WH_i [$_{IP2}$ t_i [$_{IP2}$ ....]]]]]]

---

119. Recall that under our view, CP inherits barrierhood from IP even after adjunction.

120. Although we have not considered cases of extractions out of NPs, our theory makes the right prediction for the extraction of a complement out of a subject NP. Assuming adjunction to NP is possible, we would get the following structure:

(i) Wh [$_{IP}$ t_1[$_{IP}$ [$_{NP}$ t_2 [$_{NP}$ N t_3]]]]

t_2 adjoined to the subject NP will fail to be properly antecedent-governed. IP is the barrier dominating t_2. It is a barrier both by inheritance and by lack of L-marking. It is thus a STRONG barrier, as we will define them below. Thus, extraction out of a subject NP leads to a strong Subjacency violation as in *Barriers*.

By (239), CP2 is a barrier for t_i. Note, however, that CP2 is only a barrier on one count, namely by inheritance from IP.

Finally, consider the case of CNP islands, which are noun-complement structures:

(251) [$_{CP1}$ WH_i [$_{IP1}$ [$_{VP}$ V [$_{NP}$ [$_{NP}$ N] [$_{CP2}$ t_1 [$_{CP2}$ [$_{IP2}$ ....]]]]]]]

As we proposed earlier in this chapter, the upper segment of IP is a barrier for the head government of t_1. Note that here again, as in the case of WH-islands, NP is only a barrier on one count, namely by inheritance from CP2.

From the observation of these four cases, we come to the following conclusion: the barriers for islands which are usually considered to induce strong Subjacency violations are all barriers on two counts, both by lack of L-marking and by inheritance.They are conjunctive barriers with respect to the definition of barriers given earlier, in the sense that they meet both conditions on the definition of barriers. In islands which are usually considered as inducing weak Subjacency violations, on the other hand, the barriers dominating the faulty traces are only barriers by inheritance. In other words, the barriers of weak Subjacency islands are disjunctive barriers. We can then define the force of a barrier as follows:

(252) τ is a STRONG barrier for β iff (a) **AND** (b)

a. τ immediately dominates δ, δ a BC for β

b. τ is a BC for β and τ dominates β

(253) τ is a WEAK barrier for β iff (a) **OR** (b)

a. τ immediately dominates δ, δ a BC for β

b. τ is a BC for β and τ dominates β

Given the distinction we have made between two types of barriers, we can now say that a strong Subjacency violation occurs whenever a faulty (negatively marked for a τ-feature) trace is dominated by a STRONG barrier, while a weak Subjacency violation will obtain whenever a faulty trace is dominated by a WEAK barrier.

Let us turn to one final case. Recall that (as discovered by Obenauer and reinterpreted by Rizzi) certain VP adverbs induce a Operator Opacity violations. Consequenlty these adverbs induce an ECP violation for adjunct extraction. As noted by Rizzi, however, they do not induce Subjacency violations for object extraction. As we have assumed, VP is L-marked by I. Thus, VP is a barrier on no counts with respect to the definitions of barriers given above: it is L-marked and it not a barrier by inheritance. We thus conclude that it is completely transparent for Subjacency.

We have given a technical definition of "strong" and "weak" barriers. It is possible, however, that the notion of force can also result from other factors, such as the semantic force of the operator contained in a Spec of CP. Thus, for instance, a tensed CP is conceivably a

stronger barrier than an infinitival CP.[121] Similarly, *whether* may be a weaker operator than a fully informative question word such as *who* or *what*. In other words, we are tentatively suggesting that in addition to a technical definition of barrierhood force, there might be other factors, such as tense and lexical choices, entering the notion of the force of a barrier. We leave the exploration of the plausibility of such an approach to Subjacency open for further research.[122]

4.7.3 Summary of chapter III

We have suggested that Subjacency is in fact a consequence of the ECP applying at S-structure.[123] The way we see it is rougly as follows: τ-marking applies at S-structure for all types of extractions. At S-structure, the effect of an ECP violation depends on what we have called the force of a barrier. The assessment of the force of a barrier takes into account the barrier dominating a negatively τ-marked trace. This determines whether the violation is weak or strong in reference to a first scale taking into account only S-structure ECP violations. At LF, intermediate traces of non-homogeneous chains

---

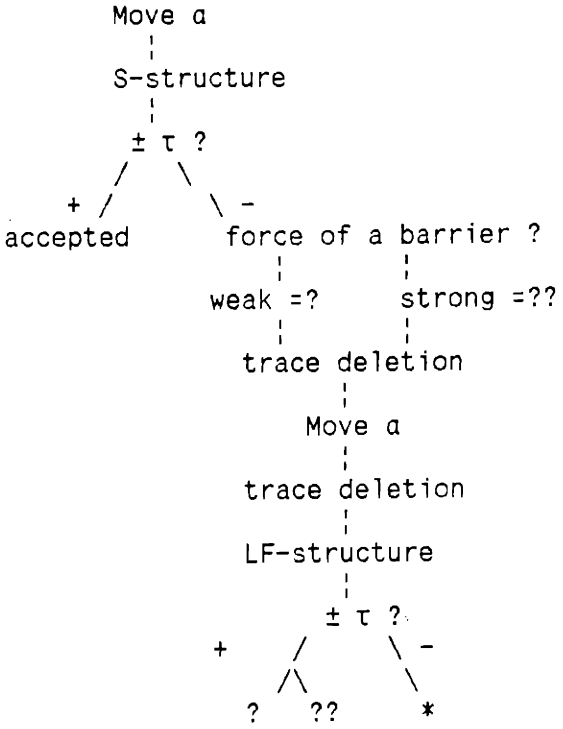
121. Possibly only a [+Tense] IP makes CP a barrier by inheritance.

122. An important aspect of a theory of Subjacency which we have not considered is its parametric variation. As claimed by Sportiche and Rizzi, there are differences in the types of Subjacency violation found in English and those found in French and Italian. We leave this problem open for futher reasearch.

123. This suggestion bears some similarities to Kayne's proposal in "ECP Extensions".

delete. Remaining traces are checked again for their value with respect to their τ-features. Extractions creating chains which violate the ECP both at S-structure and at LF are the strongest type of violations. Schematically, this theory can be represented as follows:

(254)



REFERENCES

Abney,S. (1986) "Functional Elements and Licensing," presented at GLOW conference, Barcelona

--------(1987) The English Noun Phrase in Its Sentential Aspect Phd dissertation, MIT

Aoun, Y (1985) A grammar of Anaphora, MIT Press, Cambridge

Authier, J-M. P. (1988) The Syntax of Unselective Binding Phd dissertation UCSD

Baker, M. (1988) Incorporation: A Theory of Grammatical Function Changing, The University of Chicago Press, Chicago

Barss, A.(1986) Chains and Anaphoric Dependencies, Phd dissertation, MIT

Belletti, A. (1988) "The Case of Unaccusatives", Linguistic Inquiry 19, pp 1-34.

Belletti, A. and Rizzi, L.(1986) "Psych-verbs and Theta-theory" Lexicon Project Working Papers 13, MIT, Cambridge

Bennis, H (1986) Gaps and Dummies Foris, Dordrecht

Bennis, H. and Hoekstra, T. (1985) "Gaps and Parasitic Gaps," The Linguistic Review: 4, 29-87

Borer, H.(1982) Parametic Syntax, Foris, Dordrecht

--------(1986a) "I-subject" Linguistic Inquiry 17, 375-416

--------(1986b) "Anaphoric AGR," ms Universty of California, Irvine

--------(1984) Parametric Syntax, Foris, Dordrecht

Brandi, L. and Cordin, P. (1989) "Two Italian Dialects and the Null Subject Parameter," In Jaeggli and Saphir eds

Brody, M. (1984) "On the Contextual Definitions and the Role of Chains", Linguistic Inquiry 15.3, 355-380.

Browning, M (1987)Null Operator Constructions Phd dissertation, MIT

Burzio, L. (1986) Italian Syntax: A Government and Binding Approach, Reidel, Dordrecht

Carsten, V. and Kinyalolo (1989) " AGR, Tense,Aspect and IP structure: Evidence from Bantu" Paper read at the GLOW Colloquium in Utrecht

Chomsky, N (1981) Lectures on Governement and Binding Foris, Dordrecht

--------(1982) Some Concept and Consequences of The Theory of Governement and Binding, MIT Press, Cambridge

--------(1985) The Logical Structure of Linguistic Theory, Universty of Chicago Press

--------(1986a) Knowledge of Language: Its Nature, Origin and Use, Praeger, New York

--------(1986b) Barriers, MIT Press, Cambridge

--------(1989) "Some Notes on Economy of Derivation and Representation" MITWPL 10

Chomsky,N. and Lasnik, H. (1977), "Filters and Control", Linguistic Inquiry 8, 425-504

Christensen, K.K (1986) "Norwegian Ingen: a Case of Postsyntactic Lexicalization. In Dahl, O. and Holmberg, A. (eds) Scandinavian Syntax, Department of General Linguistics, University of Stockholm

Christensen, K.K. and Taraldsen, K.T. (1989) "Expletive Chain Formation and Past Participle Agreement in Scandinavian Dialects", in Paola Beninca eds Dialect Variation and The Theory of Grammar, Foris

Davidson, D. (1966): "The logical form of action sentences," in D.Davidson 1980:105 122.

Davidson, D. (1980): *Essays on Actions and Events*, Oxford: Clarendon Press.

Deprez, V. (1984) "Sur la Syntaxe des infinitives du Francais et le Complémenteur "de" ms, University of Paris

--------(1986) "Stylistic Inversion and Verb Movement" ms, MIT

--------(1988) "Raising in Haitian Creole: A case of NP movement" ms, MIT (to appear in Chicago University Press)

--------(1988a) "Verb movement and Stylistic Inversion" in Proceedings of the East Coast Conference on Linguistics

--------(1988b) "Two ways of Moving the Verb in French Interrogatives",

paper presented at the Univeristy of Pennsylvania

--------(1988c) "L'inversion du Sujet en Moyen Français" to appear in La Revue Québéquoise de Linguistique, numéro special sur la Syntaxe diachronique

Deprez, V. and Hale, K. (1985) "Resumptive Pronouns Objects in Irish" in Paul Jefferis and Williams, J. Mahon, The Proceedings of the Harvard Celtic Colloquium, Vol 5, p 38-47, Harvard University.

Deprez, V. and Pierce, A. (forthcoming) "Verb movement in Child Language: A Comparative Study of French and English"

Diesing, M. (1989) "Verb Movement and the Subject position in Yiddish" NLLT 7

Emonds, J. (1976) A Transformational Approach to English Syntax, Academic Press New York

--------(1978) " The Verbal Complex V'-V in French" Linguistic Inquiry 9, 151-175

Epstein, S. D. (1987) Empty Categories and Their Antecedent, Phd dissertation, University of Connecticut, Storrs

Frampton, J. (1987) "Parasitic Gaps and the Theory of WH-chains", ms Northeastern University, Boston.

Felix,S (1984) Parasitic gaps in German, In Abraham, W ed, Erklarende Syntax des Deutchen. Tubingen. Narr

Fukui, N. (1986) A Theory of Category Projection and Its Applications, Phd dissertation, MIT

Fukui, N. and Speas, M.(1986) "Specifiers and Projections", MITWPL 6

Goldsmith, J (1978) "Complementizer and Root sentences" LI,Cambridge

Guéron, J. and Hoekstra (1988) "T-chains and the constituent structure of Auxiliaries ms. Université de Paris VIII, University of Groningen.

Guasti, M.T (1988) " La Pseudo Relative et Les Phénomènes d'Accords" ms, to appear in Rivista di Grammatica Generativa,13

Haider, H. and Prinzhorn, M. (1986) eds Verb Second Phenomena in Germanic Languages, Foris, Dordrecht

Haig, J. (1980), "Some observations on Quantifier Floating in Japanese", Linguistics 18.

Higginbotham, J. "Pronouns and Bound Variables" Linguistic Inquiry 14, 395-420

------- (1985) "On Semantics", Linguistic Inquiry 16, 547-593.

--------(1986) "Elucidation of Meaning", ms. MIT, Cambridge, Massachusetts

Hoji, H. (1985) Logical Form Constraints and Configurational Structures in Japanese, Phd dissertation, University of Washington, Seatle.

Holmberg, A. (1986) Word Order and Syntactic Features, Phd dissertation, University of Stockholm

--------(1987) Why is "John a book bought not well formed?" Working paper in Scandinavian Syntax

--------(1985) "Icelandic Word Order and Binary Branching" Working Paper in Scandinavian Syntax 11

Huot, H. (1981) Constructions Infinitives du Français, Droz, Genève-Paris

Jackendoff, R. (1972) Semantic interpretation in Generative Grammar, MIT Press, Cambridge

--------(1977) X'Syntax: A Study of Phrase Structure MIT Press, Cambridge

Kayne, R.S.(1972) "Subject Inversion in French Interrogatives" in J.Cassagrande and B. Sciuk (eds.) Generative Studies in Romance Languages

--------(1977) Syntaxe du Français: Le Cycle Transformationnel, Le Seuil, Paris

--------(1984) Connecteness and Binary Branching, Foris, Dordrecht

--------(1986) "Connexite et Inversion du Sujet" in La Grammaire Modulaire eds Ronat,M et Couquaux,D.

--------(1989) "Facets of Romance Past participle Agreement" in Paola Beninca eds "Dialect Variation on the Theory of Grammar" Foris

--------(1989) "Notes on English Agreement" ms, Graduate Center CUNY

Kayne, R. and Pollock, J-Y (1978) " Stylistic Inversion, Successive Cyclicity, and Move NP in French", Linguistic Inquiry, 9. 595-621.

Kitagawa, Y (1986) Subject in Japanese and English, Phd

dissertation, University of Massachusetts, Amherst

Koopman, H. (1984) The Syntax of Verbs, Foris, Dordrecht

Koopman, H. and Sportiche, D. (1988) "Subject", ms UCLA

Koopman, H. and Sportiche, D (198) "Multiple Variable Binding" Linguistic Inquiry,

Koster, J. (1987) Domains and Dynasties, Foris, Dordrecht

Koster, J (1986) "The Relation between Pro-drop, Scrambling and Verb Movement" ms University of Groningen

Kuroda, S-Y (1985a) "Whether You Agree or Not", unpublished manuscript University of California, San Diego.

Larson, R (1988) "On the Double Object Construction," Linguistic Inquiry 19.3. 335-392

Lasnik, H (1989) "On Case theory...." ms, MIT / University of Connecticut

Lasnik, H. and Saito, M. (1984) "On the Nature of Proper Governement", Linguistic Inquiry 15, 235-289.

Lasnik, H and Saito, M. (to appear) Move α

Lebeau, D.S. Language Acquisition and the Form of the Grammar, Phd dissertation, University of Massachusetts, Amherst

May,R. (1985) Logical Form, MIT Press, Cambridge

Mahajan,A. K. (1989) "Agreement and Agreement Phrases" MITWL 10.

--------(1988a) "Word order and Negation in Hindi", ms. MIT

--------(1988b) "Scrambling Weak Crossover and Clause Structure", ms. MIT

Myagawa (1989) Structure and Case Marking in Japanese, (forthcoming) Academic Press

Obenauer, H-G (1983) "Une Quantification Non Cannorique: La "Quantification à Distance" ", Langue Française: 58

--------(1984) "....." GLOW Newsletter 12

--------(1985) "On the identification of Empty Categories", The Linguistic Review:4, 153-202

--------(forthcoming) Aspect de la Syntaxe A', Thèse de doctorat d'Etat, Université de Paris VIII

Mc Connell-Ginet, S. (1982): "Adverbs and Logical Form" Language, 58, pp 144-184

Milner, J-C (1982) Ordre et Raisons de Langues, Le Seuil, Paris

Platzack, C.(1986) "COMP, INFL and Germanic Word Order", in Hellan, L. and Christensen K.K Topics in Scandinavian Syntax Reidel, Dordrecht.

Platzack, C. and Holmberg, A. (1989) "The role of AGR and Finiteness" Working Papers in Scandinavian Syntax 43

Pollock, J-Y (1989) "Verb movement, Universal Grammar and the Structure of IP", Linguistic Inquiry 20, 365-424

--------(1986) "On Case and the Syntax of Infinitives in French," in Guéron, J., Obenauer, H-G. and Pollock, J-Y, eds., Grammatical Representation, Foris, Dordrecht.

--------(1985) " Sur la Syntaxe de *en* et le Paramètre du Sujet Nul" In Ronat, M. and Couquaux, D.,eds., La grammarie Modulaire, Minuit, Paris

Pesetsky, D. (1982a) "Complementizer-trace Phenomena and the Nominative Island Condition", The Linguistic Review 1:297-343

--------(1982b) Paths and Categories, Phd dissertation MIT

--------(1984) "Extraction Domains and a Surprising Subjec/Object Asymmetry", GLOW Newsletter, 12: 58-60

--------(1987) "Wh in situ: Movement and Unselective Binding", in E.Reuland and A.Ter Meulen, eds. The Representation of (In)definitness, The MIT Press, Cambridge pp 98-129

--------(1989) "Language particular Processes and the Earliness Principle", paper presented at the GLOW Colloquium, Utrecht

Pierce, A. (1989) On The Emergence of Syntax Phd dissertation, MIT

Raposo,E. (1987) "Case Theory and Infl-to-Comp: The Inflected Infinitive in European Portuguese," Linguistic Inquiry 18: 85-109.

Reuland, E. (1983) "Governing -ing", Linguistic Inquiry 14:1

Rizzi, L (1982), Issues in Italian Syntax, Foris Publications Dordrecht, Holland

--------(1986) "On chain Formation," in H. Borer, ed, The Syntax of Pronominal Clitics, Syntax and Semantics, volume 19, Academic

Press, New York

--------(to appear) Relativized Minimality , MIT Press

Rizzi, L. and Roberts, I.G (1989) "Complex Inversion in French" ms University of Geneva, Geneva.

Roberts, I.(1985) "Agreement parameters and the development of English modals auxiliaries" Natural Language and Linguistic Theory Reidel

Rouveret, L (1980) "Sur la notion de proposition finie. Gouvernment et Inversion, Recherches Linguistiques, Université Paris VIII, Vincennes

Ruwet, N. (1982) Grammaire des insultes et autres études, Le Seuil, Paris

Safir, K.(1984) " Multiple Variable Binding", Linguistic Inquiry 15.4, 603-638

Saito, M. (1985) Some Assymetires in Japanese and Their Theoretical implications, Phd dissertation, MIT

-------- (1987), "Scrambling as Semantically Vacuous A'-movement", talk presented at NYU Worshop on Phrase Structure (1986), ms., University of Tsukuba.

Saito, M. and Hoji, H. (1983) "Weak Crossover and Move-α in Japanese," Natural Language and Linguistic Theory 1:2.

Schlyter, S. (1972): "Une hiérachie d'adverbes en francais" Recherches Linguistiques 1, pp 139-157

Schwartz, B.D and Sten, V.(1989) "All Verb Second Clauses are CPs" Working Papers in Scandinavian Syntax 43

Speas, M. (1986) Adjunctions and Projections in Syntax, Phd dissertation, MIT.

-------- (1989) "Projection from the lexicaon and Generalized Transformations" Talk given a the MIT Lexicon seminar

Sportiche, D. (1988) " A theory of Floating Quantifiers and its Corrolaries for Constituent Structure" Linguistic Inquiry 19,425-449

Stowell, T. (1981) : Origins of phrase structure, Doctoral Dissertation MIT.

Stowell, T. and Lasnik, H. "Weakest Crossover" ms, UCLA and University of Connecticut

Torrego, E (1984) " On Inversion in Spanish and Some of Its

Effects" Linguistic Inquiry 15, 103-129

Taraldsen, K.T (1983) Parametric Variation in Phrase Structure: A Case Study, Phd dissertation University of Tromso

--------(1986) "Som and The Binding Theory" in Hellan,L. and Christensen K.K Topics in Scandinavian Syntax Reidel, Dordrecht.

Thrainsson, H. (1984) "Some Points on Icelandic Word Order", Working Paper in Scandinavian Syntax 8.

--------(1986) "On Auxilliaries, AUX and VPs in Icelandic" in Hellan, L. and Christensen K.K Topics in Scandinavian Syntax Reidel, Dordrecht.

Travis, L. (1984) Parameters and Effects of Word Order Variation, Phd dissertation, MIT

Webelhuth, G. (1989) Phd dissertation, University of Massachusetts, Amherst

Williams,E and Di Sciullo (1987)On The Definition of Word, MIT Press Cambridge

Van Wyngaert (1989) Talk given at the SCIL (1989 conference