THE SYNTAX OF OPERATORS

by

Isabelle Haïk

vol. I

Maîtrise, Université de Créteil, Paris XIII
(1976)

Diplôme d'Etudes Approfondies, Université de Vincennes, Paris VIII
(1979)

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

September 1985

© Isabelle Haïk 1985

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

Signature of Author .................................................................
Department of Linguistics and Philosophy
September 18, 1985

Certified by ................................................................. Noam Chomsky
Thesis Supervisor

Accepted by ................................................................. Morris Halle
Chairman, Departmental Committee

ARCHIVES
MASSACHUSETTS INSTITUTE
OF TECHNOLOGY

OCT 31 1985
LIBRARIES
THE SYNTAX OF OPERATORS

by

ISABELLE HAIK

Submitted to the Department of Linguistics and Philosophy
on September, 24th, 1985 in partial fulfillment of the
requirements for the Degree of Doctor of Philosophy

ABSTRACT

The aim of this thesis is to explore the implications that the
existence of Logical Form has, both for the derivation of sentences and
for the interaction of subtheories of Universal Grammar. Given that the
behavior of lexical anaphors can be reduced to that of NP traces, as in
Chomsky (1985a), it will be shown that principles A and B of the Binding
theory can be derived from Theta theory. Arguments will be represented
abstractly as chains, whose formation is governed by Principle A and the
Empty Category Principle, as formulated by Kayne (1981a). In addition
to argument movement, certain predicates are shown to move at LF, to
permit th-marking of their arguments. This movement will be similarly
constrained.

A major claim of this thesis is that a bound pronominal confers
operator status on the category which contains it, and hence must be
assigned scope. This claim receives independent support insofar as it
explains an apparent counterexample to the hypothesis above, that
Principle A must hold between links of a chain. Furthermore, this
property of bound pronouns will play a central role in the availability
of certain readings in sentences involving sloppy identity, and certain
structures involving VP-deletion and parasitic gaps. In these
structures, it is just the assignment of scope to the category
containing the bound pronoun which gives rise to the appropriate logical
forms.

In addition to standard types of LF movement, i.e. movement to COMP
(as in wh-movement in Chinese), and adjunction (as in Quantifier
Raising), it will be argued that a third type exists. This involves the
identification of the moved category with its target, yielding a
structure in which subtrees are represented on distinct planes, which
meet at the merged (i.e. identified) node. The creation of such
coordinate structures will account for the properties of parasitic
gaps, which become across-the-board gaps at LF. Moreover, sloppy
identity obtains only in coordinate structures, thus making it
unnecessary to appeal to λ-abstraction to account for it.

In addition to permitting movement, LF licenses the insertion of
material missing at S-structure. This enables various 'deletion' constructions to be properly interpreted.

Thesis Supervisor: Noam Chomsky
Title: Institute Professor
Acknowledgements

Being a student at MIT has been an extremely rewarding experience, and I would like to express my gratitude to the faculty, who gave me this unique opportunity. I had longed for the sense of belonging to an intellectual community and being integrated into the department surpassed my expectations.

This thesis has been a pleasure to write, and I thank my committee, Noam Chomsky, Ken Hale and Haj Ross, for having participated in its making. I thank my adviser, Noam Chomsky, not only for his contribution to this thesis, but more importantly, for making the study of linguistic theory such a challenging and aesthetically pleasing endeavor.

Ken Hale has always been very supportive, and I thank him for his genuine interest in discussing ideas in general. His opinion often led to clarifications on my part, and certain times, his reaction made me realize that certain analyses could not hold. Other times, his respect for them encouraged me to defend them, which often helped me keep my interest alive.

Haj Ross was a source of deep linguistic knowledge. His sharp vision of the overall accounts I was proposing is clear by the number of problematic sentences that pop up along the text, some of which are still unsolved. I also thank him for offering a course in Relational Grammar, which raised lots of interesting questions.

My thanks also go to Morris Halle, for his supportive attitude. With Morris, I felt that it was acceptable to stumble on difficulties, so long as one had a true desire to learn.

I thank Jim Higginbotham for contributing to my interest in the problems of Logical Form. His classes were always challenging.

Linguistics started for me at Creteil, with a delightful class taught by Ulrika Dubos, on modern poetry and stylistics. Her contagious enthusiasm eventually led me to my master's thesis, a stylistic study of a poem by Dylan Thomas. The other highlight of these early years was Jean Yves Pollock's class in linguistics, where I first came to realize that we follow rules when we speak. I thank them for generating such enthusiasm and interest.

In Vincennes. Richie Kayne's class was extremely inspiring, his inventiveness and at the same time his clarity turned the hardest facts into disciplined objects. By being able to find a coherent explanation for any problem, he taught me not to give up on the unsolved ones, and to pursue one's idea to their furthest consequences. I also benefited greatly from discussions with Alain Rouveret, Richard Carter, Jaqueline Guérin and Hans Obenauer.

I would like to give special thanks to Diane Massam, Kyle Johnson and Juliette Levin, my friends at MIT, with whom I shared so many things and had such great fun. Diane lent a sympathetic ear to matters both linguistic and personal, and we spent much time pondering such questions. I appreciated Juliette's communicative energy and her willingness to discuss linguistics and share ideas. Kyle always knew when to give me encouragement, and I would often turn to him for sound advice.

Rita Manzini, Barry Schein, David Law, Richard Sproat, Mark Baker, Michael Brody, and Irene Heim helped me in many ways.

- 4 -
For help at crucial moments I am grateful to Dominique Sportiche.
Thanks to Marc Ryser for his support throughout the writing of this dissertation. Our discussions often generated ideas and clarified certain issues. (Thank you for the title, too.)
I would also like to thank the following people for good conversations and good arguments, which constitute a large part of the fun of being a linguist: Maria Luisa Zubizarreta, Ian Roberts, Anne Rochette, Ken Safir, Danilo Salamanca, Andy Barss, Jim Blevins, Jean Roger Vergnaud, David Feldman, Alicja Gorecka, Betsy Ritter, Richard Larson, Peter Ludlow, Grant Goodall, Anna Szabolcsi, Ruth Kempson, Tim Stowell, Nigel Fabb, Malka Rappaport, Osvaldo Jaeggli, Hyeon Sock Choe, Margaret Magnus, Andrea Calabrese, Jim Huang, Mike Hammonds, Hilda Koopman, Tanya Reinhart, Joe Emonds, Carme Picallo, Daniel Couquaux, Beatrice Lamiroy, Pierre Pica, Joseph Aoun, Wyn Chao, Suzan Rothstein, Alessandra Giorgi, Esther Torrego, Patrick Bellier, Luigi Rizzi, Dan Finer, Donca Steriade, Maggie Browning, Lisa Travis, Mamoru Saito, Howard Lasnik, Edwin Williams, Peter Coopmans, Diana Archangeli, Hagit Borer, Eric Reuland, Mario Montalbetti, Doug Saddy, John Lumsden, Carol Tenny, Peggy Speas, Tova Rapaport, Naoki Fukui, Betsy Sager, John Truscott, Robert May, Luigi Burzio, Alec Marantz, Adriana Belletti, Dany Jaspers, Katalin Kiss, Pino Longobardi, Bonnie Schwartz, and Jaklin Kornfilt.
From over the ocean, my friends Muriel and Arnaud did not cease to be close to me.
Lastly, my thanks go to my family: my parents, Gaston and Yvette; my brothers Aldo and Yvon; and my sister Pauline. Their love and their trust have been a constant source of support for me.
A mes parents
# Table of Contents

Abstract 2

Acknowledgments 4

Introduction 11

Chapter 1 LOGICAL FORM AND THE BINDING THEORY 14

1.1 Introduction 14
1.2 Chomsky's theory of lexical anaphors 14
1.3 Illicit movement 17

1.3.1 The problem 17
1.3.2 Chomsky's account 20
1.3.3 Condition A as constitutive of the definition of chain-links 24

1.4 Deriving principle B from th-Theory 26
1.5 Wide scope reciprocals 29

1.5.1 The problem 29
1.5.2 Bound pronouns acting as operators 41

1.6 Thematic and binding principles 49

1.6.1 The theta-criterion 49
1.6.2 Certain consequences 53
1.6.3 Governing categories 54
1.6.4 Government 59
1.6.5 Anaphors subjects of NPs 61
1.6.6 A'-chains 63

1.7 Consequences for S-structure 64

1.7.1 S-structure chains 64
1.7.2 Raised subjects 71
1.7.3 Leftward-tous and restructuring at Logical Form 74
1.7.4 One problem with leftward-tous 77
1.7.5 Passive in restructuring constructions 79

1.8 Two possible GCs 80

1.8.1 S-bars as predicates of small clauses 80
1.8.2 Other sentences 94
1.8.3 Benefactives 96
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4.3 Coreferentiality explained</td>
<td></td>
</tr>
<tr>
<td>3.4.4 Empty anaphors</td>
<td>203</td>
</tr>
<tr>
<td>3.4.5 Locality condition between a moved S' and the antecedent trigger</td>
<td></td>
</tr>
<tr>
<td>3.4.6 Coreference between subjects only</td>
<td>205</td>
</tr>
<tr>
<td>3.4.7 Binding of pronoun is similar to anaphoric binding</td>
<td></td>
</tr>
<tr>
<td>3.4.8 VP-deletion in embedded questions</td>
<td>209</td>
</tr>
<tr>
<td>3.5 Sloppy identity allowed by pronouns treated as operators</td>
<td></td>
</tr>
<tr>
<td>3.6 No governed empty pronominals</td>
<td></td>
</tr>
<tr>
<td>Chapter 4 ACROSS-THE BOARD VARIABLES</td>
<td></td>
</tr>
<tr>
<td>4.1 Parasitic gaps</td>
<td></td>
</tr>
<tr>
<td>4.1.1 Introduction</td>
<td>231</td>
</tr>
<tr>
<td>4.1.2 Parasitic gaps are empty pronominals</td>
<td></td>
</tr>
<tr>
<td>4.1.3 Parasitic-gap constructions at LF</td>
<td></td>
</tr>
<tr>
<td>4.1.4 Parasitic gaps obey subjacency</td>
<td></td>
</tr>
<tr>
<td>4.1.5 Movement of the adjunct</td>
<td></td>
</tr>
<tr>
<td>4.1.6 Connectedness effects</td>
<td></td>
</tr>
<tr>
<td>4.1.7 Movable clauses</td>
<td></td>
</tr>
<tr>
<td>4.1.8 LF parasitic gaps</td>
<td></td>
</tr>
<tr>
<td>4.1.9 Parasitic gaps in LF conjuncts</td>
<td></td>
</tr>
<tr>
<td>4.1.10 Parasitic gaps inside subjects</td>
<td></td>
</tr>
<tr>
<td>4.1.11 The relativized NP must correspond to a subject</td>
<td></td>
</tr>
<tr>
<td>4.1.11.1 A difference between overt parasitic gaps and LF parasitic gaps</td>
<td></td>
</tr>
<tr>
<td>4.1.12 The real gaps and the adjunct have the same GC</td>
<td></td>
</tr>
<tr>
<td>4.1.13 LF-conjuncts are S-structure coordinates or predicates</td>
<td></td>
</tr>
<tr>
<td>4.1.14 Two differences between parasitic gaps and ATB-gaps</td>
<td></td>
</tr>
<tr>
<td>4.1.14.1 Reconstruction in parasitic gaps and ATB-gaps</td>
<td></td>
</tr>
<tr>
<td>4.1.15 Parasitic gaps cannot be subjects</td>
<td></td>
</tr>
<tr>
<td>4.1.16 Licensing parasitic gaps</td>
<td></td>
</tr>
<tr>
<td>4.1.17 No weak crossover effects with parasitic gaps</td>
<td></td>
</tr>
<tr>
<td>4.2 Pronouns of laziness</td>
<td></td>
</tr>
<tr>
<td>4.2.1 The similarity between parasitic gaps and pronoun of laziness</td>
<td></td>
</tr>
<tr>
<td>4.2.2 Analysis of pronouns of laziness</td>
<td></td>
</tr>
<tr>
<td>4.2.3 Why do pronouns of laziness form an A'-chain</td>
<td></td>
</tr>
<tr>
<td>4.2.4 Pronouns of laziness and locality effects</td>
<td></td>
</tr>
<tr>
<td>4.2.5 Pronouns of laziness in adjuncts and subjects</td>
<td></td>
</tr>
<tr>
<td>4.2.6 No connectedness effects</td>
<td></td>
</tr>
<tr>
<td>4.2.7 Two alternatives</td>
<td></td>
</tr>
<tr>
<td>4.2.8 Paycheck-sentences</td>
<td></td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>4.2.9</td>
<td>Pronouns of laziness and the condition on semantic conjuncts</td>
</tr>
<tr>
<td>4.2.10</td>
<td>An indirect binding approach to pronouns of laziness</td>
</tr>
<tr>
<td>4.2.11</td>
<td>Pronouns of laziness as resumptive pronouns</td>
</tr>
<tr>
<td>4.2.12</td>
<td>Coordination across discourse</td>
</tr>
<tr>
<td>4.3</td>
<td>Sloppy identity</td>
</tr>
<tr>
<td>4.3.1</td>
<td>The condition on alphabetic variance</td>
</tr>
<tr>
<td>4.3.2</td>
<td>The representation of bound pronouns</td>
</tr>
<tr>
<td>4.3.3</td>
<td>Sloppy identity and VP-deletion</td>
</tr>
<tr>
<td>4.3.4</td>
<td>The syntactic domain of sloppy identity</td>
</tr>
<tr>
<td>4.3.5</td>
<td>LF VP-movement is wh movement</td>
</tr>
<tr>
<td>4.3.6</td>
<td>The sloppy binder and the empty VP must have the same GC</td>
</tr>
<tr>
<td>4.3.7</td>
<td>Sloppy identity and crossing</td>
</tr>
<tr>
<td>4.3.8</td>
<td>Why no sloppy reading inside a parasitic gap</td>
</tr>
<tr>
<td></td>
<td>4.3.8.1 Comparative deletion</td>
</tr>
<tr>
<td>4.4</td>
<td>Free relatives</td>
</tr>
<tr>
<td>4.4.1</td>
<td>LF union of phrase-markers</td>
</tr>
<tr>
<td>4.4.2</td>
<td>When -ever is missing</td>
</tr>
<tr>
<td>4.4.3</td>
<td>Condition on scope assignment is a condition on representation</td>
</tr>
<tr>
<td>4.4.4</td>
<td>Extraction from extrapoosed sentences</td>
</tr>
<tr>
<td>4.4.5</td>
<td>Free relatives from PPs</td>
</tr>
<tr>
<td>4.5</td>
<td>Bare XP comparatives</td>
</tr>
<tr>
<td>4.5.1</td>
<td>Bare XP analysis</td>
</tr>
<tr>
<td>4.5.2</td>
<td>The analysis</td>
</tr>
<tr>
<td>4.5.3</td>
<td>The bare XP is extracted from the comparative PP</td>
</tr>
<tr>
<td>4.5.4</td>
<td>Crossing effects</td>
</tr>
<tr>
<td>4.5.5</td>
<td>Exceptions to the crossing violations</td>
</tr>
<tr>
<td>4.6</td>
<td>Reanalysis as a union of phrase-markers on a single plane</td>
</tr>
<tr>
<td>4.6.1</td>
<td>Introduction</td>
</tr>
<tr>
<td>4.6.2</td>
<td>Union of phrase-markers on a single plane</td>
</tr>
<tr>
<td>4.6.3</td>
<td>Tough-constructions</td>
</tr>
<tr>
<td>4.6.4</td>
<td>Consequences</td>
</tr>
<tr>
<td>4.6.5</td>
<td>Other empty-operator constructions</td>
</tr>
<tr>
<td>4.6.6</td>
<td>Clitic-movement in tough-constructions in French and Italian</td>
</tr>
</tbody>
</table>
Introduction

This thesis is a study of phenomena pertaining to Logical Form (LF) and an investigation of the theoretical consequences of the existence of this level within the Government and Binding theory.

LF has been recognized as the level at which operators are assigned scope, an operation which has as one of its functions that of disambiguating sentences, with respect to scope of quantifiers. We will have little to say about this aspect of quantification here, and, often the cases under study will involve one quantifier only. Rather, the treatment of the category under consideration as a quantified phrase will, itself, be the point of interest.

It may happen that S-structure is incomplete in some respect, and consequently cannot yield the interpretation that it is assumed to have. One of the themes of this thesis will be that certain S-structures are anomalous or defective for theory-internal reasons and that LF will reshape the structure, thereby eliminating the potential violation of certain conditions. Thus, the cases under investigation will not only provide extensive support for the existence of LF, but will yield insight into its specific nature.

Chapter 1 will be concerned with the interaction of two subtheories of the grammar, binding theory and theta theory. It will be suggested that, if principle A defines chain-links, then principle B can be
derived by appeal to th-theory. This is natural in a theory like Chomsky’s (1985a), in which the behavior of lexical anaphors is reduced to that of NP-traces.

Certain constructions suggest that the th-criterion be relaxed at S-structure. This will permit certain categories to be distant from their binders at S-structure, and thus not form a chain (as in tough-constructions (chapter 4, section 6) and leftward-tous (chapter 1)). Only at LF will the relation between the binder and the bindee become local, permitting the appropriate chain to form, satisfying the th-criterion, or the principle disallowing vacuous quantification.

In general, various mechanisms are employed for achieving locality at LF. Leftward-tous appeals to reconstruction. VP-deletion (chapter 3) and tough-constructions appeal essentially to Quantifier Raising. Finally, parasitic gaps and sloppy identity (chapter 4) appeal to the formation of coordinate structures, via a process of merging nodes.

It was suggested above that LF saved certain defective S-structures. These defects are of various types. In the adjective-operators of chapter 2, a predicate is associated with the wrong argument. In certain cases of antecedent-contained VP-deletion (chapter 3), an operator binds a type of category it shouldn’t. Other such cases of VP-deletion have S-structures which would give rise to a governed empty pronominal, if LF-movement did not apply. This is true of parasitic gaps as well (chapter 4).

As mentioned earlier, this thesis will propose one technical
innovation, which is the formation of conjuncts at LF, developed in chapter 4. It will also propose a theoretical innovation, which is that bound pronouns make the referential constituent which immediately dominates them (S', S or NP) a quantified phrase. This implies that this constituent may be assigned scope. Metaphorically, a bound pronouns acts, in those cases, like an operator. This property of bound pronouns will be used in various constructions to explain certain semantic facts or certain basic contrasts, like the possibility of the wide scope reading of a reciprocal (chapter 1:1.5.2), and the necessity of coreference between two subjects in certain VP-deletion sentences (chapter 3, section 4).

Lastly, sloppy identity, which arises in special syntactic environments, will be considered. Sloppy identity obtains in coordinate structures, given a certain assumption about the interpretation of pronouns. With this assumption, reference to \( \lambda \)-abstraction is no longer necessary for an account of this phenomenon.
1.1 Introduction

In this chapter, we will see that Chomsky's (1985a) recent theory of anaphors as LF-clitics can be considered a first move toward eliminating principle B of the Binding theory. We will see that such an approach differs minimally from that of Chomsky (1981) or (1985a) and that it is aimed at capturing the fact that movement itself is limited by the binding theory to avoid having to stipulate that chain-links are local. Our view will be close to that of Brody's (1984a) and (1984b), who derives the distribution of empty categories from the definition of chains. In the discussion, the definitions of governing category (henceforth GC), the th-criterion and the binding conditions are those of Chomsky (1981), until we discuss them explicitly.

1.2 Chomsky's Theory of Lexical Anaphors

Principles A and B of the Binding theory apply respectively to lexical anaphors and NP-traces on the one hand and to pronominals on the
(1) \textbf{Governing category} (Chomsky, 1981: p.220)
\[ \beta \text{ is a GC for } \alpha \text{ if and only if } \beta \text{ is the minimal category containing } \alpha, \text{ a governor of } \alpha \text{ and a SUBJECT accessible to } \alpha. \]

(2) \textbf{Binding Conditions}
A: an anaphor is bound in its GC.
B: a pronominal is free in its GC.

The question is whether this is a coincidence. In Chomsky's (1985a) analysis of anaphors, the correlation is not accidental. Instead of the anaphors being seen as prototypical and NP-traces analogical to them, the behavior of NP-traces is seen as prototypical, and that of lexical anaphors as deriving from that of NP-traces. Briefly, Chomsky's analysis is that lexical anaphors are subject to one condition: they must be governed by some antecedent at LF. In order to be so governed, they have the option of adjoining to INFL, in which case the subject can be a proper antecedent for them. But, since there is movement, and given that this movement is analogical to clitic-movement, (namely, the lexical anaphor forms some kind of th-chain with its trace), one binding condition now applies to the trace, condition A of the Binding theory.

In other words, the local domain in which the lexical anaphor can find

\[ ------- \]

1. For the discussion to follow, we can consider that an accessible SUBJECT is the subject of a clause or of an NP, since we do not yet enter into much detail about the formulation of the binding theory.

2. They can also adjoin to VP, in which case internal arguments of the V may be proper antecedents for the anaphors. See Chomsky (1985a) for details.

3. See Lebeaux (1984), who shows that the movement of each only and not the whole phrase each other can explain certain differences in the distribution of reciprocals versus reflexives in terms of the ECP.
its antecedent is defined as the domain in which a trace can find its antecedent. This analysis makes the binding condition on lexical anaphors of Chomsky (1981) derivative from that of NP-traces: condition A no longer applies to both lexical anaphors and NP-traces, it only applies to NP-traces.

So, now, the Binding theory on [f]-categories has to take care of two elements only, and no longer three: traces and pronouns. This is the first move towards simplifying the Binding theory, counterbalanced with the new condition on lexical anaphors, namely, that they have to be governed by their antecedent. Note, though, that the reason why pronouns and NP-traces are now in (near) complementary distribution is still not explained, it is simply reflected in the complementary statement of the Binding principles. The proposal here will be that the complementary distribution between NP-traces and pronouns can be reduced to th-theory, if we incorporate condition A in the definition of chain-links. Before discussing chains, let us examine a particular case relevant for our purposes, discussed by Lasnik (1985).

4. [f]-categories are categories with a minimal number of features, like [pronominal], [anaphor], [N] and features for gender, number and person. For lack of time, no study of long-distance anaphors will be provided here, but it seems that a promising approach is to characterize their LF movement as wh movement. This would explain why they can occur embedded, why they are bound by subjects only, and why they may not occur in embedded subject position. Their behavior would be similar to incroyable, versus wrong, studied in the next chapter. My thanks to Marc Ryser for a discussion that lead to this conclusion.

5. That condition A "is part of the mechanism for assigning th-roles" is expressed by Chomsky (1985a: section 5.2.3).
1.3 Illicit Movement

1.3.1 The problem

Lasnik (1985) points out that the Binding theory on NP-traces is too weak in certain cases, which he identifies as strong crossover, such as in (3):

(3) John, seems that he likes t

The problem with this sentence is that it does not violate any constraint, the chain (John, t) is well-formed with respect to th-theory, and the trace properly obeys principle A, since it is bound by the pronoun in its governing category. In order to rule such sentences out, Lasnik appeals to the locality condition on chains, which is a potential violation of strong crossover (see Koopman and Sportiche, 1983), which holds of A'-chains such as in (4): 6

(4) *Who does he like t?

The locality condition on chains has the effect that (a,b) is a

6. Local binding, defined in Chomsky (1981), means binding by the nearest binder. So, if the first binder of a category is in an A-position, the category is locally A-bound, and if it is in an A'-position, the category is locally A'-bound. The locality condition as a well-formedness condition on chains has been shown by Obenauer (1983) to explain the fact that intervening quantifiers seem to prevent a moved quantifier like combien 'how many' in French from relating to its trace. Rizzi (1983) presents new arguments for such a view, and Barss (1984) too. I will not deal with these questions in this work.
chain-link if and only if a binds b and there is no coindexed c such that c c-commands b and a c-commands c. Note that the locality condition on chain-links can derive from other well-formedness conditions, in a theory which defines empty categories contextually. For example, Koopman and Sportiche rule (4) out on the grounds that, since t is locally bound by he, it cannot be defined as a variable and hence is PRO, since it has a th-role independent from that of its antecedent. Since PRO may not have a governing category, (4) is ruled out. However, given certain arguments against the contextual definitions of empty categories (cf. Brody (1984a), Sportiche (1983) and Epstein (1984)) Lasnik opts for a theory in which categories are freely assigned pronominal and anaphoric features and are then filtered in or out by the binding principles and the th-criterion. In that case, the locality condition is no longer derivative.∗

Going back to the examples, in neither (3) nor (4) is the trace locally bound by the element which belongs to its chain, and this gives some strength to Lasnik's claim that the locality condition should apply to A-chains as well as A'-chains. One problem with this analysis, however, is that it suffers from counterexamples.

Consider (5)-(6):

7. See Lasnik (1985) for a discussion of the arguments against the contextual definitions of empty categories, like the fact that PRO may be locally A'-bound by an operator, but should not be defined as a variable.
(5) They strike each other as t intellgent

(6) ?John was shown t himself in the mirror

In each of these sentences, the trace is locally A-bound by an element, respectively each other, himself, which does not belong to the chain headed by the subject NP (lest it violates the th-criterion). Rizzi (1983) argues that each other does not c-command the trace in (6)b, on the face of examples like (7), where the quantifier cannot be interpreted as binding the pronoun:

(7) Luc strikes everyone as more intelligent than him

This is probably because there is something special with comparatives. If the complement is a true complement of the A, then every may bind a pronoun:

(8) Lucy strikes every boy as angry at him

Note also that the following sentence is bad with coreference between John and him, which could be explained if the pronoun does in fact bind:

8. Kyle Johnson pointed out to me that the contrast is even clearer with a gerundival complement:

(i) *John strikes every boy as being more intelligent than him
(ii) ?John strikes every boy as being angry at him

Alicja Gorecka (personal communication) notes that the binding theory is violated in comparative clauses, in a similar adjectival environment:

(iii) John considers Peter more intelligent than himself

If the comparative clause is a complement of the AP, it is in the domain of the subject of the AP, Peter, and hence himself should be bound within that domain, which it does not seem to be.
c-command the name:

(9) Mary strikes him as more intelligent than John

If this is correct, the sentences in (5)-(6) should be excluded, under Lasnik's locality condition on chains.

Chomsky presents another argument for rejecting the locality condition, based on LF movement of anaphors, which we discuss in section 1.5. We will present another analysis of the facts above, which directly accounts for them, but not for Chomsky's additional argument. We will thus address the treatment of Chomsky's argument in that section.

The analysis that we will eventually favor aims at capturing the fact that the original motivation for making NP-traces subject to condition A is that this condition is supposed to constrain movement. And sentences like (3), where a binder happens to license illicit movement, are in fact accidental. We consider now two other possible accounts of improper movement.

1.3.2 Chomsky's Account

One solution, proposed by Chomsky (1985a), invokes the prohibition of NP-traces from being Case-marked, as a particular interpretation of the Case condition on chains (this is not Chomsky's ultimate formulation, but it is sufficient for the present discussion):

---------

9. Thanks to Dominique Sportiche, who reached the same conclusion, for a fruitful discussion on this subject.
(10) **Case condition**
A chain receives one Case in one and only one way.

The following requirement is also added (again, this is informal):

(11) At S-structure, all arguments must receive Case (by being in a Case-marked position). (PRO has inherent Case, and is "Case-marked" in its S-structure position.)

This requirement implies that the moved NP will always be in a Case-marked position. So, the Case condition forces its trace to occur in a non-Case marked position. Now, examples parallel to (3), which display the locality effect on A-chains as described by Lasnik, seem to be impossible to construe with a non-Case-marked trace without violating some other condition, like the th-criterion.¹⁰

(12) **Theta-Criterion (definition)** (Chomsky, 1985a: (83))
Each argument α appears in a chain containing a unique visible th-position P, and each th-position is visible in a chain containing a unique argument α.

It is thus hard to tell whether the Case condition is the right approach for the problem under discussion. Chomsky (Class lectures, Fall 1984) also discussed cases in which this was not true, that is, cases where the trace is not necessarily Case-marked without being straightforwardly excluded as violations of something else, as when it is the object of an adjective, as in (13):

--------

10. Dominique Sportiche pointed this out to me. Below, I consider in detail some of his examples.

11. A visible position is a position in a chain which contains a Case-marked position.
(13)a. *John_i seems that he_i is proud t_i  
   b. *John_i seems that he_i likes a picture t_i

Chomsky (1965a) rules (13)a-b out by the Uniformity condition, which requires that inherent Case, that is to say, Case assigned by a noun, a preposition or an adjective, also be assigned to the element that it th-marks. In (13)a-b, the element th-marked by proud and picture is not Case-marked by it, a violation of the Uniformity condition.

If (13)a-b are ruled out by the Uniformity condition, we have to check whether the Case condition accounts for all other possible cases. The only other circumstances where Case is not assigned to an object are when the predicate is passive, or a raising predicate or an ergative one, and also, but more controversially, when the object is an S' and not an NP.12 I will neglect the discussion of whether S' receives Case or not, acknowledging the difficulty of answering this question.13 As for the NP cases, they will give rise to a configuration where the local binder of the NP-trace is in a non-th-position. An example of passive with the abstract structure under study is (14):

12. Ergative predicates (cf.Burzio (1981)), or unaccusative, in relational grammar terms (cf. Perlmutter and Postal (1983)) are predicates which take objects which appear on the surface as subjects, like arrive.

(14) *John₁ is believed that he₁ is liked t₁

The relevant chain is (John, t), which is well-formed with respect to th-theory, and with respect to (10), the Case condition on chains, since passive morphology suppresses Case-assignment to the object position. Now the question concerns he. If the property of passive morphology of retaining the th-role of the external argument is optional, then (14) cannot be excluded as a th-criterion violation, since the pronoun he would occupy a th-marked position. If Burzio's (1981) generalization that predicates which do not Case-mark their object do not th-mark their subject is a necessary state of affairs, then passive morphology will suspend th-marking to the subject position, which permits to rule (14) out on thematic grounds.¹⁴

As for the ergative and the raising cases, they are like (14), except that the position of he is unquestionably non-th, contrary to the passive cases, as in (15), so the th-criterion violation is unquestionable:

(15)a. *John₁ is believed that he₁ arrived t₁

b. *John₁ is believed that he₁ seems t₁ to be reading

In conclusion, in sentences where the trace is not Case-marked, the subject position will be non-th (Burzio's generalization). Given that there must be an argument in that non-th-position, this argument ends up

---

¹⁴ For properties of passive, see Perlmutter and Postal (1983a) and Williams (1981b). Certain theories aim at deriving this dethematization property of passives, as in Jaeggli (1984); Baker, Johnson and Roberts (1985) and Roberts (1985).
without a th-role, violating the th-criterion. 15

Under this analysis, Lasnik's (1985) cases of improper movement are excluded in a complementary manner by th-theory and Case theory, and one may prefer to unify them.

1.3.3 Condition A as constitutive of the definition of chain-links

Our account will be different. As we have seen, condition A alone is not sufficient to rule (3) out, because the condition that the binder of the trace should be the moved NP is not inherent to the binding condition on NP-traces. However, this is the result which should be achieved: in some way, it should be stated that the only appropriate binder for an NP-trace with respect to principle A is the moved NP (or intermediary traces), and not some accidental binder. This requirement would be inelegant if it had to be stipulated for NP-traces only, but under Chomsky's (1985a) analysis, lexical anaphors give rise to a chain at LF. So, the behavior of lexical anaphors and moved NPs can be unified at LF. Let us thus assume informally for the moment that condition A is

15. There is one special case, when the subject position is occupied by an expletive. If the expletive is taken to be a proper binder, then improper movement is predicted to be good, but sentences are ungrammatical:

(i) Many people seem that there arrived

So, we will follow Williams (1984) and assume that there is the scope-marker of the indefinite NP, in which case, this NP cannot move higher than it. But see Chomsky (1985a) for a differing conception. Chomsky considers the possibility that chain-links have to obey condition A, to further reject it, on the basis of sentences like (1). See Chomsky (1985a: chapter 5).
a condition inherent to chain-links at LF, to the effect that in a chain-link \((a, b)\), \(a\) binds \(b\) in its governing category. And let us consider again Lasnik's examples and their counterexamples, represented by (3) and (6):

(3) \(\ast\)John\(_1\) seems that he\(_1\) likes \(t_1\)
(6) (?)John\(_1\) was shown \(t_1\) himself\(_1\) in the mirror

In (3), if the trace forms a chain with John, the link \((\text{John}, t)\) does not obey condition A, since the trace is not bound by John in its GC. Note that if it forms a chain with the pronoun, the sentence will be excluded by the th-criterion for various obvious reasons (no two th-positions in a chain and no argument without a th-role). As for (6), the trace can form a chain with John without violating any condition. The contrast between (3) and (6) thus indicates that Lasnik's cases of illicit movement should in fact be accounted for by the binding theory.

As noted by Chomsky, one could reject this analysis because of the following conceptual problem: reducing illicit NP-movement to the binding theory must be effected technically by rendering principle A a well-formedness condition on chains, rather than a condition on traces. However, the Binding theory comprises principle B, which applies to pronominals and which expresses the almost complementary distribution between pronominals and anaphors. It is thus strange and probably undesirable that the two principles would apply to two different types of objects: a chain-link, which represents a relation, and a pronominal, which is a single formative. Under the theory which makes condition A apply to th-chains, the relevant observation is that pronouns should
never form a th-chain with their binders, since the two should have
independent th-roles from each other. It is thus possible to derive
principle B from th-theory, if pronouns are forced to form a th-chain
with their binder in case the binder belongs to the GC of the pronoun.

1.4 Deriving principle B from th-Theory

We will later deal with the exact definition of governing category
(GC), which should correspond to that of Chomsky (1981). For the
purpose of the discussion, let us assume for the moment that the GC of
an element A is the minimal S which contains A.

Let us define a chain-link:

(16) **Chain-links (Definition) (to be revised in (19))**

\((\alpha, \beta)\) is a chain-link if and only if \(\beta\) is in an A-position and
is bound in its GC by \(\alpha\).

This definition of chain-links should derive the main effects of
principle B, since pronouns bound in their GC will be forced to form a
chain with their antecedents, and this will violate the th-criterion in
most of the cases.

This definition also makes crucial use of the fact that anaphors
create chains at LF. If (16) was operative at S-structure, where lexical
anaphors are still in their A-position, sentences like (17) would
incorrectly be excluded in the way that pronominals bound in their GC
are excluded at LF:
(17) John\textsubscript{i} met himself\textsubscript{i}

We thus state the condition that chains are necessarily formed at LF:\textsuperscript{16}

(18) Chains are formed at LF, where chain-links are defined as in (16).

This analysis also hinges on the fact that pronouns do not move at LF, as opposed to anaphors. Incidentally, the definition of chain-links captures the effects of principle B, but also part of principle C. So, names bound in their GC are also ruled out by the th-criterion, and we can view the definition of chain-links as applying to indices, and not necessarily empty categories or pronouns.\textsuperscript{17}

One last remark. (16) suffers from the fact that it forces chain-links to be formed whenever principle A is met between a binder and a bindee. However, it may happen that a trace has more than one...

\textsuperscript{16} As for chain-links formed by expletives, they should also fall under condition A, so condition A applies to all A-chains, that is to say, chains where the head is in an A(rgument)-position. We rule out an expletive trace which is not in a chain, as in (i), versus (ii), by a general condition excluding empty expletives, which is independently needed to exclude (iii), and not by the ECP:

(i) *It is probable [e] to seem that S
(ii) It seems [e] to be likely that S
(iii) *It seems that John is in a rush without PRO appearing that he will leave

In (i), [e] is not bound in its GC by anything and it does not have a th-role, so it is an empty expletive, which is excluded as a principle. In (ii), owing to S'-deletion, [e] is governed by seem, and is bound in its GC by \textit{it}, so it forms an A-chain with it. [e] is thus not an empty expletive, so the structure is well-formed.

\textsuperscript{17} My thanks to Kyle Johnson for noticing this, and for his comments on this point.
binder in its GC, as in configurations with moved lexical anaphors, where both the antecedent and the anaphor bind the trace of the anaphor in its GC, and as is the case in (5) and (6). Such structures are well-formed only when the trace forms a chain with the moved anaphor, and ill-formed with respect to the th-criterion if it forms a chain with the antecedent of the anaphor or with both, so we have to make it possible for a trace to freely belong to one chain only, still forcing at least one chain to be formed. The new conditions replace definition (16):

(19) **Conditions on Chain-links.**

(i) If $\beta$ is in an A-position and $\beta$ is bound in its GC by $\alpha_1, \alpha_2, \ldots, \alpha_n$, then $\beta$ forms a chain-link with $\alpha_i$.

(ii) If $(\alpha, \beta)$ is a chain-link, then $\alpha$ binds $\beta$ in its GC.

These conditions apply to elements in A-positions in order to avoid making cliticized lexical anaphors form a th-chain with their antecedent, which would be excluded by the th-criterion.

Our hypotheses have a number of consequences and problems, which we consider in later sections. For the moment, let us deal with a case that May, cited in Chomsky (1985a), presents as evidence against the locality condition on chains proposed by Obenauer (1983), Lasnik (1985) and Rizzi (1983). This example is also taken as a counterexample to the claim that chain-links have to obey condition A, so we have to solve it before we proceed.
1.5 Wide Scope Reciprocals

1.5.1 The problem

Let us consider (20), from Higginbotham (1982), pointed out by May for the particular question of improper movement. Chomsky (1985a) analyzes the ambiguity of (20) with respect to the reciprocal's interpretation as being due to either matrix or embedded scope assignment to the reciprocal:

(20) Peter and Mary thought they should visit each other.

Under the so-called embedded scope interpretation of the reciprocal, Peter and Mary have the same thought, that they should visit each other, namely, the thought that Peter should visit Mary and that Mary should visit Peter. Under the matrix scope interpretation, each of them has the thought that he/she should visit the other, where the pronoun is interpreted as a bound variable. The problematic interpretation is the latter, in which the reciprocal associates with Peter and Mary, since it is supposed to be derived from an LF in which the reciprocal each other moves to the matrix S, as in (21):

(21) Peter and Mary [each other] thought they should visit.

According to Chomsky, and assuming Chomsky's (1981) Binding theory, this LF is well-formed because the [+anaphor] trace of each other is bound in its GC by some element, here they. In other words, this is a
well-formed case of "illicit movement", to use Lasnik's terminology. This example would be problematic for us if we agreed with the LF treatment of the matrix scope interpretation of the reciprocal, since the LF movement of (21) plainly goes against the hypothesis that chain-links must obey condition A.

The question is then to explain how the reading "Peter and Mary each think that they (he\she) should visit the other" can be expressed without having to give Peter and Mary as an antecedent to each other. 18 The structure of the argument will be the following. First, we will see that we could analyze the S' embedded under thought as the union of the two thoughts attributed to the referents of the subject NP. But, then, this idea will run into problems which will speak in favor of the wide-scope analysis. However, an additional piece of data will suggest that the wide scope analysis cannot cover similar cases, so we will give a condition which will save the hypothesis of the union of thoughts. And in the next section, we will comment on this condition. Before we give the account of the relevant reading of (21), let us consider the notion of the sum-interpretation of plurals.

A predicate applying to at least two expressions of pluralities can

---

18. For discussion on the interpretation of reciprocals, see Carlson, (1980) and references cited there, and Schein (1984) and (1985). It is not possible to extend the interpretation of each other to all de re interpretations, as was suggested to me by Jacqueline Guérin, since in general, expressions which have a de re reading can unboundedly escape the domain of the epistemic predicate, as in (i), where John does not think (i) about Bob, but about someone who happens to be Bob:

(i) John thinks that Mary said that someone should meet Bob
be interpreted as a sum of predicates if it applies to the union of the sets denoted by its arguments, and not necessarily to each element of these sets: this is the weak distributive reading. (See Carlson (1980); Schein (1985); Frey and Kamp (1985).) For example, a sentence like (22) is true if John lit the candle and Mary lit the lantern, without it having to be true that John lit both of them and Mary too:

(22) John and Mary lit the candle and the lantern

In (22), the predicate lit the candle and the lantern is interpreted as the sum of the predicates lit the candle and lit the lantern, applied to the union of \{John\} and \{Mary\}, namely \{John, Mary\}.

Now, consider the two thoughts: "Peter should visit Mary" and "Mary should visit Peter". Suppose it is possible to form the union of these two thoughts in the way one forms the union of two predicates: by making the arguments of visit denote the union of two sets, here \{Peter\} and \{Mary\} for both the object and subject position. We get: "Peter and Mary should visit Mary and Peter". Now, the way they should visit is such that it can be expressed by a reciprocal: the truth conditions associated with each other are that all members of a set must have a certain relation R with another, different, member of the same set.

Given that the truth conditions associated with each other are met here, with relation R being visit, the union of the two thoughts may be expressed by the sentence: "John and Mary should visit each other". So, up to now, we have shown that, if it is possible at all to express in one sentence the union of two thoughts, then the sentence "John and Mary should visit each other" can be the expression of the union of the two
thoughts "Mary should visit Peter" and "Peter should visit Mary".

We have obtained this sentence, meant to express the union of two thoughts, themselves corresponding syntactically to two clauses, by construing the arguments of the predicates as unions of sets. The question now is to determine whether such a sentence as "Peter and Mary should visit each other" may in fact be interpreted as a union of two thoughts. If it cannot, this will mean that something syntactic blocks this interpretation.

A predicate is interpreted as a sum only if it applies to a plurality, weakly distributively. So, let us consider the plurality denoted by *Beth and Malcolm* in (23):

(23) Beth and Malcolm think that Peter and Mary should visit each other.

Given what was said above, it should be possible to interpret the matrix predicate of (23) as a sum of predicates applying to the subject plurality. (23) then should be able to mean: "Beth thinks that Peter should visit Mary and Malcolm thinks that Mary should visit Peter".

Hence, even though each one of Beth and Malcolm has a different thought, their two thoughts should be able to be expressed in "Peter and Mary should visit each other". Now, the problem is that this sentence cannot mean this at all. It must unambiguously be interpreted with Beth and Malcolm having the same thought: "Peter and Mary should visit each other".

Looking at the data more closely, one finds that the sum interpretation of plurals is limited to one single clause, a
generalization reminiscent of the syntactic property of reciprocals: the
expressions of plurality must be interpreted within one single clause
too. Here are some examples:

(24)a. Beth and Malcolm will visit Lyn and Robin
    b. Beth and Malcolm think (that) Lyn and Robin are angry
    c. Beth and Malcolm think (that) someone hates Lyn and Robin
    d. Beth and Malcolm believe Lyn and Robin to have left
    e. Beth and Malcolm believe someone to hate Lyn and Robin

The sum interpretation is impossible in (24)b,c and e, that is to say in
the environment in which the second expression of plurality, Lyn and
Robin does not have Beth and Malcolm in its GC. This strongly indicates
that the syntactic condition on the sum interpretation of a predicate or
a proposition is that their plural arguments be in a binding relation
with each other. However, let us modify this view.

The properties of such constructions are very similar to what obtains
also with expressions like the same N', different N', and similar
expressions.¹⁹ All NPs of this form must be associated with an
expression of plurality which occurs in their governing category, as
shown in (25):

¹⁹. Locality facts have been noted independently by Stump, cited by
Dowty (1985); Schein (1984); Carlson (1985); Heim (1985); Szabolcsi
(1984); and others working on the interpretation of comparatives and
superlatives, as well as plurals. Not all of them agree with limiting
the domain to a GC.
(25)a. Beth and Malcolm have the same address
   b. Beth and Malcolm believe that a different person screamed
   b. Beth and Malcolm think (that) someone will steal the same car
   c. Beth and Malcolm believe a different person to have screamed
   d. Beth and Malcolm believe someone to have kissed the same person

The reading which interests us is the one in which *same or different is evaluated with respect to the subject NP *Beth and Malcolm*, which Stump calls the internal reading of the adjective. Again, it seems that this reading is possible only if the NP containing the adjective is bound in its GC by the plural subject NP. One first hypothesis is thus that these NPs, and the expressions of plurality in (24) above are LF reciprocals, which makes them *behave* like reciprocals.

There is, however, one case in which true reciprocals differ from these NPs, namely, when they find themselves in the subject position of the sentence containing the other expression of plurality. There, reciprocals are excluded, but not these NPs, as shown here:

(26)a. *Each other hit Bill and Bob*
   b. The *same person* hit Bill and Bob

This state of affairs can be expressed by saying that NPs containing these distributive adjectives and the NPs involved in the sum interpretation must have the same scope as the expression of plurality.20 Scope itself being limited to one minimal S, it follows

20. This has been suggested by Irene Heim.
that the two NPs must occur in the same S. 21 The locality effects on the two NPs is thus not due to some direct binding relation between the two NPs but rather to the fact that they are scope takers, that they take the same scope, and that scope is clause-bound.

These examples have been brought into the discussion to suggest that pure reciprocals are subject to the same conditions as the plurals in a sum-of-plurals interpretation, or as NPs containing comparative adjectives. So, now, the question is even worse than before: given some more evidence that other NPs involved in some reciprocal interpretation cannot escape from their clause (cf. the examples above), how come the reciprocal is allowed to escape its clause in (20) ("P. and M. thought that they should visit e.o")? Let us consider an additional fact, which seems to speak in favor of Chomsky's approach, that is, in favor of wide scope movement of each other saved by local binding by a coreferential pronominal: in the sentences with the sum-of-plurals interpretation, it is also the case that an intervening pronominal renders the interpretation possible, as shown here:

(27) Beth and Malcolm think that they hate Lyn and Robin

This devastating example seems to indicate that the NP Lyn and Robin should be analyzed as a reciprocal element, whose "antecedent" is the matrix NP, and whose movement to the matrix clause is saved by local

21. The claim about the limitation of scope to one GC is supported by examples like (i), in which everyone cannot take scope over the matrix indefinite:

(i) Someone thinks that Mary like everyone
binding by the pronoun *they* (presumably, indirect binding), and hence saved from a violation of condition A when it moves to the matrix clause, as in the following tentative LF:

(28) [Beth and Malcolm] [Lyn and Robin] think that they hate t

Crucially, such an approach is contradicted by one type of sentences, where a pronoun coreferential with the matrix NP licenses the other plural NP not only when it locally binds it, but also when it does not, so long as it belongs to the same minimal S as that expression of plurality, as in (29):

(29)a. Beth and Malcolm think that Lyn and Robin will hit them

b. *Beth and Malcolm think that Lyn and Robin said that they are fools

The star in (29)b indicates that the sum interpretation of the two pluralities is not possible.22 For the moment, the important example is (29)a, since it shows that it is the very presence of the pronominal which saves the structure and not its local binding relation with the plural NP.23

22. We also expect *same* to be saved by a pronoun, which it is not, as shown in (1) (at least, for some speakers):

(1) Beth and Malcolm think that the same person hit them

See the discussion of the next chapter, section 2.1.3.

23. Unless local binding has access to A'- as well as A-positions. In that case, the pronoun and the plural NP could both move by QR, and the pronoun could be assigned wider scope than the NP, hence locally binding it, as in (i):
Now, we can make an analogy between the interpretation of true reciprocals and that of the sum of plurals in noting that both obey the condition that they are possible in the long distance situations only if the licensing pronoun is interpreted as a bound variable.

(30)a. Peter and Mary think that they should visit each other
   b. Beth and Malcolm think that Lyn and Robin will hit them
   c. Beth and Malcolm think that they hate Lyn and Robin

In all of these, the different propositions which are summed up in the interpretations of these sentences are those in which the pronoun is assigned the value of the variable attributed to the matrix subject: Peter thinks that he should VP, and Mary thinks that she should VP, etc.

Given the similarities between the behavior of reciprocals and that of plural NPs in the sum-of-plurals interpretation, it is tempting to say that they fall under the same generalizations. Given that a pronoun acts as licensing in the latter cases even when no local binding obtains, we can assume that it is not the local-binding property which licenses the wide scope interpretation of the reciprocals in (30)a (i.e. 20). Obviously, one then wonders why this pronoun has to c-command the reciprocal, as shown by the ungrammaticality of (31):

\[(1) \text{B and M think that } [\text{them}_j [[[L \text{ and } R]_1 [t_1 \text{ will hit } t_j]]]\]

I leave this as an unworked out possibility.
(31) *Peter and Mary think that each other should visit them

Nothing surprising here given that each other must be coindexed with its antecedent, (31) is excluded because the pronoun them is bound in its GC by each other (cf. the contrast between (26)a and b). Relevant examples, in which the pronoun is not bound in its GC by each other are impossible to construe, since, as we saw for (29)b above, the licensing pronoun must itself belong to the GC of the reciprocal element. This forces the pronoun to be in a c-commanding and non-c-commanded position with respect to the reciprocal inside the GC of the reciprocal. So, we can extend Heim's suggestion to all reciprocals: reciprocal elements must have the same scope as their antecedents. This is why they move at LF. Since they are coindexed with these antecedents, they cannot c-command them, given the effect of principle B. So, in the majority of cases, the antecedent will end up c-commanding the anaphor on the surface.

This does not solve our problem, namely how to state the licensing effect of the pronoun in the cases of long-distance reciprocals of (30)a-c. The effect of a bound pronoun inside a constituent makes the constituent interpretable as a function of the NP which binds the pronoun.24 Now, note that one other relevant fact is that the constituent which is to be translated as a function of the antecedent of the reciprocal must itself be bound in its GC by this antecedent, as

24. See, e.g. Engdahl (1980) and (1984), for the claim that question words range over functions, in examples like:

(1) Which book did every author recommend?
   - His first book.
shown by the impossibility of the wide scope interpretation of the reciprocal in a sentence like (32):\(^{25}\) (The constituent interpreted as a function is the most embedded S, as a stipulation, for now):

(32) Peter and Mary think that Beth said that they should visit each other

In conclusion, for the moment, let us give the following descriptive generalization. We will consider why (33) has to hold in the following section:

(33) **Condition on the interpretation of reciprocals and other distributive elements:** The Same-Game condition

\(\alpha\) must have the same scope as \(\beta\), where

(i) \(\alpha\) is a reciprocal element.

or

(ii) \(\alpha\) is inside a constituent interpreted as a function of the antecedent which itself has the same scope as \(\beta\).

(iii) \(\beta\) is the antecedent.

(34) A reciprocal element is a reciprocal, or an NP containing same\!
different with the internal interpretation, or a plural NP interpreted in a sum of plurals.

"Antecedent" here loosely means the NP associated with the reciprocal or with the NP containing same or different or with the plural NP in the sum interpretation. If (ii) holds, then (i) does not have to hold, that

\[\text{25. It is not clear how this could be accounted for by the LF-movement theory of each other onto the matrix clause which does not require of movement itself that it be clause-bound, since its trace would be properly locally bound by they. This was pointed out by Edwin Williams. Note that (i) is well-formed, if the two embedded pronouns are bound:}\]

(i) P and M think that they said that they should visit each other

The most embedded one is saved by the intermediary one.
is to say, a reciprocal is licensed if it obeys the Same-Scope condition directly or if the most immediate referential constituent (i.e. S' or NP) which contains it obeys it, under the condition that this constituent be interpreted as a function of the antecedent, namely, if it contains a pronoun interpreted as bound by the antecedent. As far as the interpretation goes, two constituents which are expressions of pluralities and which have the same scope may enter into the sum interpretation. The claim is that, in Peter and Mary think that they should visit each other, one plurality is Peter and Mary and the other must be the S' that they should visit each other. We return to this immediately.

Summarizing this section, we have seen that condition (33) and the added proviso (ii) account for the facts. Obviously, one should try to explain why (ii) would hold. Still, we have seen that, if it is correct to generalize the behavior of reciprocals to that of the other NPs characterized in (34), then the account of the reciprocals in terms of local binding by the pronoun does not extend to those, which we took as meaning that local binding by the pronoun is not what allows the wide scope reading of reciprocals in the first place. The conclusion of this section is that the reciprocal anaphor does not have to move to the matrix sentence to express the wide scope reading. This interpretation derives from the fact that the sentence which contains the reciprocal is interpreted as a sum of thoughts, which is made possible by the fact that this S' has the same scope as the matrix subject, and that it contains a pronoun bound by this subject. So (20) is not a counterexample to the claim about chain-links that they must obey.
condition A.

1.5.2 Bound pronouns acting as operators

One could argue that the wide scope interpretation of reciprocals is more trivial than in our account or Chomsky's. Suppose that one claims that, in a sentence like (35), the pronoun is in fact the proper antecedent of the reciprocal, and that the fact that it has the bound reading is irrelevant for the binding of the reciprocal:

(35) Peter and Mary think that they should visit each other

The claim that the proper antecedent of the reciprocal is the pronoun is even confirmed by the ill-formedness of (36), in which the pronoun turns out not to be a proper antecedent for the reciprocal, because it is a singular:

(36) *Everybody promised that he would visit each other

(36) cannot be excluded on the grounds that everyone itself is not a proper antecedent for the reciprocal, since (37)a is not too bad, or (36) with a plural bound pronoun, instead of a singular, as shown in (37)b: 26

26. Obviously, (36) is a problem for the wide scope analysis of the reciprocal, as well as a problem for our analysis, which allows wide scope interpretation whenever the reciprocal's GC is interpreted as a function of the antecedent, but we will adopt the idea that, after all, it is the embedded pronoun which is the real antecedent of the reciprocal, so this pronoun must denote a plurality.
(37)a. Everybody visited each other
b. Everybody promised that they would visit each other

Nevertheless, the claim that the embedded pronoun is the antecedent of the reciprocal is too weak if nothing is added, because it predicts that the binding of reciprocals by a pronoun interpreted as a bound pronoun can be effected, however embedded these two elements are. But this is not the case: the pronoun and the reciprocal cannot be deeply embedded with respect to the antecedent, as shown below, if the reciprocal is to have the "wide scope" reading:

(38) John and Peter both hoped that Mary would think that they should visit each other

It is impossible to interpret (38) with John and Peter having the sum of the two hopes: "Mary thinks that I should visit the other". And this cannot be due to some locality restriction on bound pronouns themselves, since bound pronouns may actually occur deeply embedded with respect to their antecedent, as shown below:

(39) John and Peter both hoped that Mary would think that they are funny

These facts show that the mere binding of the reciprocal by the plural bound pronoun is not enough to account for the phenomenon.

Our analysis will be the following. We will assume that the syntactic antecedent of the reciprocal is the plural pronoun they. So, they are subject to the Same Scope condition, which they respect in the original example. However, when the pronoun is interpreted as a bound
pronoun, the wide scope reading of the reciprocal is possible only if
the constituent which dominates it participates in a sum-of-plurals
interpretation with the binder of the pronoun. So, in Peter and Mary
thought they should visit each other, it must be the case that the
complement S' is interpreted as in a sum of plurals with the plural NP
Peter and Mary. Now, let us assume that a constituent may participate
in the sum-of-plurals interpretation with another one if it denotes a
plurality, and that a constituent denotes a plurality if it is itself an
expression of plurality or if at least one of its immediate arguments is
an expression of plurality. This means that a clause can be interpreted
as such if at least one of the arguments of the predicate of the clause
denotes a plurality, and this is what the clause in the wide-scope
reciprocal cases is. In Peter and Mary thought that they should visit
each other, the embedded subject they is the expression of plurality
which allows treating the embedded S' as an expression of plurality
too.

Now, we have shown independently that arguments which participate in
a sum-of-plurals interpretation are subject to the Same Scope condition,
i.e. must have the same GC. So, the matrix subject and the embedded
clause must have the same GC, and hence the clause cannot be embedded
twice, accounting for the impossibility of the wide-scope interpretation
in (38) above. The LF of (20) is the following:

\[(40) \quad [\text{Peter and Mary } [[\text{that they } [\text{each other}]_i \text{ should visit } t_i]_j
[\text{x thought y}]]]]\]

In (40), each other is adjoined inside its S, and binds its trace in its
GC, and is assigned a proper antecedent, the plural pronoun they. The embedded S' is extracted, and also binds its trace in its GC. So the representation is well-formed with respect to the locality conditions on movement.

This stipulated movement of the S' explains why the wide scope interpretation of the reciprocal is unavailable when this S' is more deeply embedded: in that case, the S' will be allowed to move only one sentence up, which is too far away from the matrix NP. Now, consider the second property of these sentences, namely, that the S' which is assigned scope is the one which minimally contains the pronoun bound by the antecedent, and not some higher clause, as shown in (38), repeated here:

(41) John and Peter both hoped that Mary would think that they should visit each other

The impossibility of the wide scope reading can be accounted for according to the following hypothesis. There is a parallel between all constituents which are assigned scope: they all must contain an operator, and they are all assigned scope by May's (1977) condition on Analyzability. This condition makes constituents which dominate operators be assigned scope themselves, instead of being stranded from their operator. This is what happens in the LF of (41), if we assume that the operator is the pronoun.

A note of caution: that bound pronouns behave like operators is not a logical claim. This claim is a descriptive one, and it purports to the syntax of the constructions. It seems that, as far as the syntax is
concerned, bound pronouns behave like operators, in the sense that they make the constituent which contains them a quantified phrase, subject to rules of scope assignment. It is only because of this property that we make a parallel between bound pronouns and operators. The deeper reason for this behavior will still remain to be explained, a task which is not tackled in the present work. However, we will adopt the following descriptive generalization, in order to account for the assignment of scope to certain constituents:

(42) Bound pronouns may behave like operators.

In assigning scope to an operator, one assigns scope to some maximal projection which dominates this operator, as is overtly the case with wh movement.27 In the spirit of May (1977) and Guérin (1984), let us assume the following definition:

(43) Quantified phrases
A quantified phrase is a maximal projection which constitutes a GC for an operator, or a category dominating this GC which does not immediately dominate a subject.

(44) Scope assignment
The scope of an operator \( \alpha \) is the scope of some quantified phrase \( \beta \) defined with respect to \( \alpha \).

27. For example, to explain certain binding facts between quantifiers and pronouns, Guérin (1984) proposes that the scope of a quantifier is the c-command domain of the operator constituent which contains it, where an operator constituent corresponds to the maximal projection which directly dominates this operator. See Bresnan (1976) for the definition of Analyzability, which determines what constituents are subject to wh movement and May (1977) for the definition of quantified phrases along these lines. Analyzability is a concept introduced by Chomsky. See, for example, Chomsky (1965, p.98).
So, S and NP are GCs, so they can be quantified phrases. But S' and PP can also be quantified phrases, since they dominate S and NP, and do not immediately dominate a subject.

Returning to the wide scope reciprocals of (35) and to pursue the comparison between bound pronouns and operators, the scope of the pronoun is the scope of a quantified phrase that dominates it. Since it is dominated by S', the S' may be quantificational. Now, if the pronoun is inside an NP, its scope will not be able to extend to S', so the pronoun will not make the S' quantificational. In that case, only the NP will be able to participate in a sum-of-plurals interpretation, but the NP is too far from the matrix NP, and this makes the wide scope interpretation of the reciprocal impossible, as in (45):

(45) John and Peter think that their parents should visit each other.

To summarize so far, we assume, first, that reciprocals are assigned at LF the same scope as their antecedent. Then, the phenomenon of wide scope reciprocals is due to the fact that the embedded pronoun is a bound pronoun, which forces the sentence that contains the reciprocal to participate in a sum-of-plurals interpretation with the matrix subject. (Why this requirement should hold is not clear.) The reciprocal still needs to be assigned a direct antecedent denoting a plurality, evidence being given by a sentence like (36) (everybody promised that he would visit each other, in which the pronoun can only denote a singular. The sentence expressing the thought is able to take scope because the bound pronoun they makes the sentence quantificational. All reciprocal elements are subject to the Same Scope condition, and given condition A
on chain-links, they must belong to the same GC.

Note that, in this account, the syntactic antecedent of the reciprocal is the plural pronoun, not the higher subject. Under the analysis where scope is assigned to the embedded clause, an alternative account would be to allow the reciprocal to be extracted from the moved clause and take the matrix subject as its direct antecedent, as in (46):

(46) \[
S \\
/ \ \\
each other \ /
\]

S' \\
/ \ \\
NP \ VP \ NP \ VP

\[
they \ V \ NP \ J.M. \ V \ S' \\
/ \ \ / \ \\
visit t_1 \ think t_j
\]

However, if S' is moved, then extraction of each other out of it violates condition A. S could be moved instead, but let us assume that S' and S are inseparable. The reason why the direct relation between each other and the matrix subject should not be allowed is that this relation is not permitted with NPs containing comparative adjectives (see chapter 2, section 2.1.3 for discussion). Also, plural NPs in the embedded subject position may have a wide scope reading and participate with the matrix NP in a sum-of-plurals interpretation. The extraction analysis would have to make them violate the ECP. So, the conclusion is that, in the wide scope interpretation of wide scope reciprocal elements, there is no direct relation between the reciprocal and the matrix NP.
Again, the reason why a pronoun interpreted as a variable seems to behave like an operator is presumably that bound pronouns put the constituents which dominate them in the scope of the binder of the pronoun: they make them referentially dependent on the binder. Even referential expressions which resist being construed as in the scope of a plurality, like definite NPs, can be so construed when they contain a bound pronoun, as shown in the contrast below: 28

(47)a. Many people thought about the book
   b. Many people thought about the book they wrote
   c. Many people thought about their book

To summarize, the last three sections have shown that Lasnik’s cases of improper movement can be captured by requiring that Principle A be a condition on chain-links. 29 The wide-scope reciprocals have been shown

28. Also, certain sentences behave like definite referential constituents, in that they do not accept binding into them, as in (i), where it is impossible to construe the indefinite in the scope of the plurality. But, when a bound pronoun is added in it, this interpretation is possible, as in (ii):

(i) Everyone ignores that Mary will meet someone
(ii) Everyone ignores that he will meet someone

29. In fact, the condition on chain-links cannot capture the cases of downward followed by upward movement discussed in Sportiche (1983) and Brody (1984):

(i) *John seems [e] 2 to believe [e] 3 to turn out that S

Neither the th-criterion nor the binding theory can prevent the derivation of (i) with movement of John from position 2 to position 3 and then position 1, since the chain (John, e, e) is well-formed in both respects. However, if movement itself is conditioned by principle A, then the movement from 3 to 1 is impossible. So, in order for chains to represent movement, one should form chains according to the order in which each category has been occupied by an argument, rather than by
not to necessarily be cases of illicit movement. In such sentences, the S' which contains the reciprocal participates in a sum-of-plurals interpretation with the antecedent of the reciprocal. This S' can be assigned scope because it contains a bound pronoun, and these have the syntactic behavior of operators.

1.6 Thematic and binding principles

1.6.1 The theta-criterion

Let us now consider the thematic principles we assume. The purpose of the th-criterion is to ensure that arguments receive at least one th-role and that all th-roles are assigned to arguments, at LF. And the Projection principle ensures that the structure remains the same in the course of the derivation, except for adjunction. And, by the Projection principle, the th-criterion has to hold at D- and S-structure.

(48) Theta-Criterion (Chomsky (1985a, p.137))

Each argument \( \alpha \) appears in a chain containing a unique visible th-position \( P \), and each th-position \( P \) is visible in a chain containing a unique argument \( \alpha \).

local binding. Such a chain could be defined in terms of links, as in \([(e3, e2), (John, e3)]\), where \((a,b)\) is a link. In that case, the link (John, e3) violates condition A. Brody (1984) solves these cases by requiring that D-structure be a representation of GF-n, that is to say, of GFs (Grammatical Functions) which are the last members of chains. Since it must be the case, by the Projection principle, that D-structure is a pure representation of GF-th, a GF\( \alpha \) must contain an argument at D-structure.

30. See also Chomsky (1981, p.335).
(49) **Projection principle** (Chomsky (1981: (38)))

(i) if \( \beta \) is an immediate constituent of \( \gamma \) in \([\gamma \ldots \alpha \ldots \beta \ldots]\) or \([\gamma \ldots \beta \ldots \alpha \ldots]\) at \(L_i\), and \( \gamma = \overline{\alpha} \),
then \( \alpha \)-th-marks \( \beta \) in \( \gamma \).

(ii) if \( \alpha \) selects \( \beta \) in \( \gamma \) as a lexical property, then \( \alpha \) selects \( \beta \) in \( \gamma \) at \(L_1\).

(iii) if \( \alpha \) selects \( \beta \) in \( \gamma \) at \(L_1\), then \( \alpha \) selects \( \beta \) in \( \gamma \) at \(L_j\).
where \(L_i\) is a syntactic level of representation.

In this work, we adopt the view of chain-formation and the relation between the various levels as follows:

1. D-structure is a pure representation of GF-th (cf. Chomsky (1981)), so all arguments are in th-positions and all th-positions are occupied by arguments.

2. Chains are formed at LF.

3. If a chain contains more than one position, then the argument has moved from one position to another, which implies that the chain must contain at least one empty position. Well-formedness conditions on chains will ensure that this empty position is the one that the argument comes from. Also, if an empty category is defined as an argument at D-structure, then it cannot disappear in the derivation.

Let us look at the th-criterion as stated in (48). Given the definition of D-structure as a pure representation of GF-th, with all arguments in th-positions, and all th-positions filled by arguments, and given the condition of the th-criterion that a chain cannot contain more
than one argument, it is not necessary to state that chains contain only one th-position. Suppose that a chain contains two th-positions. Given that the th-criterion rules out chains which contain more than one argument, only one of these positions can be filled with an argument. The other argument should find another chain, but, supposing that the other chains are well-formed, these already contain one argument and cannot contain more than one argument. So, this argument cannot end up in a th-chain. The th-criterion can thus be formulated without this requirement (to be modified in (53)):

(50) **Theta-Criterion (tentative)**
Each argument $\alpha$ appears in a chain containing some visible th-position $P$, and each th-position $P$ is visible in a chain containing a unique argument $\alpha$.

Let us now consider whether the conditions on chain-links as in (19) are not too powerful, given this statement of the th-criterion. The consequence of the formulation of condition (19) is that, by being bound in its GC, an element enters a th-chain with the binder. Given that two arguments cannot belong to the same chain, by the th-criterion (50), all cases of pronominals or names bound in their GC by an argument or by a position linked to an argument will be ruled out by the th-criterion, such as (51):

(51) *John$_i$ saw him$_i$

Or in the following case, pointed out to me by Chomsky:

(52) John$_i$ seems to him$_i$ t$_i$ to VP

Whatever happens with the empty category in the embedded subject
position, the chain \((\text{John, him})\) violates the \(\text{th-criterion}\) since it contains two arguments.

As we have seen above, there seems to be a redundancy between the requirement that a chain cannot have more than one \(\text{th-position}\) and that it cannot contain more than one argument. The other possibility, that is, that a chain may contain more than one argument but that it cannot contain more than one \(\text{th-position}\), depends on another assumption. For example, in the case of \((52)\), it would depend on the stipulation that chains cannot intersect.\(^{31}\) If chains could intersect, and in this particular case, if they could have the same head, then, given the hypothesis that the \(\text{th-criterion}\) allows two arguments to be in the same chain, it would follow that \((52)\) would be ruled in, with the two chains \((\text{John, him})\) and \((\text{John, t})\). However, there is something unsatisfactory in the condition that a chain cannot contain more than one argument, and this can be seen in the fact that this condition excludes \((51)\) and \((52)\) on the same grounds, whereas they are of a different shape with respect to \(\text{th-structure}\). The sentence \text{John saw him} is analyzed with the chain \((\text{John, him})\), where both positions are \(\text{th-positions}\). And the sentence \text{John seems to him t to VP} with the two chain possibilities: 1. \((\text{John, him})\) and \((\text{t})\) or 2. \((\text{John, him, t})\). If case 2 is chosen, the chain looks like that of \((51)\) in that it has the same \(\text{th-structure}\): it contains two \(\text{th-positions}\). If case 1 is chosen, then the chains are different from that of \((51)\), and, as we have seen, the sentence has to be ruled out for

\(^{31}\) This holds only of \(A\)-chains. In chapter 4, we are going to allow multiple \(A\)-chains, as in Chomsky's (1982) analysis of parasitic gaps.
containing two arguments, and the exclusion has nothing to do with the th-structure of the chain. It seems, however, that a more elegant account is one which excludes chains because of their th-structure rather than because of what is in them. We will thus return to a th-criterion which prevents chains from having two th-positions, and will deal with the problem of the overgeneration of (52) by stating that chains cannot intersect.

(53) **Theta-Criterion**
Each argument \( \alpha \) appears in a chain containing a unique visible th-position \( P \), and each th-position \( P \) is visible in a chain containing at least one argument \( \alpha \).

(54) **Condition**
A-Chains may not intersect.

Given (54), a sentence like (52) *(John seems to him t to VP)* is ruled out in a similar fashion to (51) *(John saw him)*. Owing to condition (54), the only possible chain is *(John, him, t)*, which is excluded by the th-criterion, since it contains two th-positions.

1.6.2 Certain consequences

This view of chains is close to that of Brody's (1984) because Brody derives the fact that traces must be governed from the fact that they must be non-heads of chains, which derives from a version of Case theory.\(^{32}\) Moreover, he does not distinguish between government and

\(^{32}\) However, we still use the ECP as a condition on A'-chains, stated as in Kayne (1981), to capture the effects of subadjacency in the definition
proper government, but between government and absence of government.

In our definition of chains, since chain-links are defined according to condition A, and since condition A refers to the concept of governing category, then chains will be formed only with elements which have a GC, and never with those which do not have one. So, this view of chains borrows the idea that government is a necessary factor for chain-formation, achieving the result that only governed categories can be NP-trace-like, and that ungoverned elements can only be pronominal-like. The two views differ in the sense that Brody draws particular results from the condition that chain-formation applies to any empty categories: for example, all governed empty categories must be non-heads of chains. This requirement, which is derived from his definitions, has the consequence that empty governed pronouns are precluded, owing to the th-criterion, as in (55):

(55) * John \(_1\) thinks that Leo likes e \(_1\)

In (55), \(e\) is a governed empty category, and hence must form a chain with a binder, here John. This results in a violation of the th-criterion, since the chain (John, e) contains two th-positions. We will see in section 3.5.1 (chapter 3) that control theory, as stated in Manzini (1983), and discussed in Brody (to appear), can arrive at the result that only pronouns without a GC will be licit in a structure.\(^{33}\) Also, one consequence of our definition of chains is that of A'-chains.

\(^{33}\) In the spirit of Manzini (1983), the condition will be that empty pronouns must be bound in their domain-governing category, as defined
we make no use of the feature distinctions of categories like [anaphor] and [pronominal], since chains are potentially formed with any category.

1.6.3 Governing categories

Let us now turn to the definition of governing categories. The specific problem that we have to deal with is the following: Chomsky's (1985a) Binding theory eliminates the notion of Accessible subjects, in favor of the ECP for anaphors in subject position of tensed clauses, and of the notion of "potential binder" for all other cases, which captures the exceptions to the complementary distribution between anaphors and pronominals (i.e. when these are inside NPs). Chomsky's (1985a) definition of a GC allows it to vary depending on the element that it is defined for, anaphor or pronominal to the effect that an NP is a GC for a pronoun and not necessarily for an anaphor. We could transpose this definition into theta-theory, instead of leaving it in Binding theory, since chain-formation comprises binding theory, but since there is no distinction between a pronominal and an anaphor in our system, this is not possible.

Moreover, we are concerned with keeping to the idea that the ECP (as applying to NP-traces) should be derived from th-theory, to the effect that elements which transmit a th-role can only occur in governed

in Manzini (1983), where a domain-GC is a GC, or the GC of the maximal projection immediately dominating the element. Given that pronominals should be free in their GC, they will only be able to occur in non-governed positions. See chapter 3.3.5.1.
positions, hence the definition of GC should still be defined in terms of a governor. Thirdly, we will see in section 1.7.2 that the notion of subject may be relevant for clitic extraction from complements of causatives, where it seems that the GC of a clitic is extended when the domain of the subject itself is extended, as in Rouveret and Vergnaud (1980) and Aoun (1981). So, we still need to refer to subjects, contrary to Chomsky (1985a).

Let us take the following definition:

(56) **Governing category** (definition)

A is a Governing Category (GC) for x if and only if A is the c-command domain of the nearest c-commanding subject to x such that A contains a governor of x.

Let us consider subjects. The GC of a subject of a tensed clause must be the tensed clause itself, as shown by the ungrammaticality of (57):

(57) *They think that themselves should leave

As Chomsky construes it, this case is no longer treated by the Binding theory: it is excluded by the ECP, at LF, when the anaphor moves to the matrix INFL, leaving a non-properly governed trace. However, following Brody's view or Aoun's, we may derive the effects of the ECP from the binding theory, that is, from theta-theory, since the binding principles are part of the definition of chains. Suppose first that subjects are accessible to themselves, in other words, that c-command is a reflexive relation, so that a subject is always included in its c-command
(58) \textbf{C-command} (definition, from Aoun and Sportiche, 1983:(44))
\[ x \text{ c-commands } y \text{ iff } \forall z, z \text{ a maximal projection, } z \text{ dominates } x \]
only if it dominates \( y \).

Let us add: \( x \) does not contain \( y \), and \( x \) may equal \( y \).

Let us borrow Chomsky's idea that AGR is a governor when INFL is
tensed. Then, a tensed sentence is a GC for its subject, since the
subject is in its own domain, and since that domain contains the
governor of the subject, namely, INFL. So, the fact that INFL is a
governor in tensed sentences makes the distinction between tensed and
non-tensed sentences, without having to say that AGR is an accessible
subject. And in a sentence like (59), where the subject position is not
governed, because INFL does not contain AGR, and because the matrix verb
is not an S'-deleter, the subject has no GC:

(59) John wants e to go

And in a sentence like (60), where the embedded subject is governed by
the matrix predicate, which is an S'-deleter, the GC of the embedded
subject is the matrix S, since it is the domain of the nearest subject
which contains its governor:

(60) John seems e to be sick

The definition of GC also accounts for the possible elements in
object positions.

\[-----------------------------\]

34. Binding should not be reflexive, though, otherwise all anaphors
would be self-bound. Thanks to Andy Barss for pointing this out to me.
Let us consider anaphors and pronominals inside NPs. For Chomsky (1985a), the presence of implicit arguments in subject position of NPs is what makes the distribution of anaphors and pronominals non-complementary, an analysis that we adopt here. So, one can assume that, for certain nouns, the implicit subject of the NP acts as a specified subject, as in (61):

(61)a. They like stories about them  
b. They like stories about themselves

Considering that there is an implicit argument in the Spec position of the NP, then the NP is a GC for the post-prepositional object. The presence of this implicit argument is not necessary, and in fact it should not be always syntactically represented, since it is allowed to be understood as distinct from _themselves_ in (61)b. In the case of the anaphor, it is not represented, so the NP is not defined as a GC, since there is no subject in it. The closest subject being the sentential subject, the GC of _each other_ is the whole clause. Now, in order for the antecedent of the anaphor to govern it, the anaphor can move to INFL, and it will form a chain which properly satisfies condition A.

More problematic cases are those in which the anaphor or the pronoun is in the subject position of the NP, but we have to define the GC of these elements, so, before we study the examples, let us consider the theory of government that we will adopt.

35. Cf. Giorgi (1984b) for recent discussion of binding facts inside NPs.
1.6.4 Government

In Chomsky (1985a), government is allowed to cross a maximal projection down to the specifier position of that maximal projection, and does not percolate down to the complement of the head of that maximal projection. We propose the following definition to achieve these results (but see Chomsky (1985a) for the original formulation and additional concepts):

(62) \[ \text{Government (definition)} \]
\[ \phi \text{ governs } \alpha \text{ (i) if the first maximal projection which dominates } \alpha \text{ also dominates } \phi, \alpha \]
\[ \text{ (ii) if the first maximal projection } \gamma \text{ which dominates } \alpha \]
\[ \text{ is canonically governed by } \beta, \text{ and } \alpha \text{ is not canonically governed in } \gamma. \]

The phrase canonical government originates in Kayne (1983):

(63) \[ W \text{ and } Z (Z \text{ a maximal projection, and } W \text{ and } Z \text{ immediately dominated by some } Y) \text{ are in a canonical government configuration iff} \]
\[ \text{ a. } V \text{ governs NP to its right in the grammar of the language in question and } W \text{ precedes } Z. \]
\[ \text{ b. } V \text{ governs NP to its left in the grammar of the language in question and } Z \text{ precedes } W. \]

36. The notion of directionality of government is also suggested in Zagone (1982), a notion crucially used in Stowell (1981) for the elimination of phrase-structure rules. Koster (1984) claims that proper government, in Kayne's (1983) sense, and by extension, the definition of g-projections, entails that the series of governors is uniform with respect to directionality of government, thus accounting for the distribution of preposition-stranding in Dutch (but see Huybregts and van Riemersdijk (1984) for some objections). The parameter of directionality of th-role assignment has been added to the theory, along with that of government by Travis (1984) and Koopman (1984), accounting for various word order phenomena.
In English, canonical government of heads is to the right, so an element which is to the left of the head is not canonically governed by it, even though it is governed by it. Let us consider the following configuration:

\[(64)\]

\[
\begin{array}{c}
\text{Spec}_1 \\
\text{Spec}_2 \\
Y \\
Y' \\
\end{array}
\begin{array}{c}
X' \\
X'' \\
b \\
X \\
YP \\
\end{array}
\]

In (64), X governs Spec\(_1\) and YP. Spec\(_2\) is governed by X, by (ii) of the definition, because Spec\(_2\) is dominated by a maximal projection which is canonically governed by X without itself being canonically governed inside that maximal projection.\(^{37}\)

Now, by definition, the GC of a category \(a\) is the domain of the closest subject which contains the governor of \(a\). A situation of particular interest is when the category has more than one governor. In this case, our claim is that more than one GC may be defined, and that each can be optionally chosen to satisfy principles of the theory. This turns out to be crucial for the account of anaphors subjects of NPs, to which we now turn.

---

\(^{37}\) See Chomsky (1985b), and the idea that the notion of closest governor plays a role in government. The use of canonical government reflects this.
1.6.5 Anaphors subjects of NPs

Let us consider (65):

(65)a. They like each other's stories
b. They like their stories

At LF, the anaphor of (65)a moves onto INFL, and should properly bind the empty category left in the Spec position of the NP. In (65)a-b, the subject position of the NP is governed by the noun stories and by the verb like, and this subject of NP belongs to its own domain, following from the definition of c-command given above in (58). This means that the GC of the subject of the NP can either be the NP, if the relevant governor is taken to be the noun, or it can be the S, if the relevant governor is taken to be the verb. Since the choice is open, it will yield good results for both anaphors and pronouns, in a complementary fashion. We will see in section 1.8 that other cases exist in which an element has two possible GCs by being doubly governed, and that only one of the choices yields a well-formed result.

Let us now consider the situation where each other's pictures is in an embedded subject position, the cases which necessitated the notion of accessibility in terms of the i-within-i condition in Chomsky (1981):

38. In fact, the reciprocal moves to S, since it is treated like a quantifier, subject to the Same-Scope condition with its binder. This does not change the main lines of the argument.
(66) They think that each other's pictures are on sale

The analysis will require some stipulations. Each other adjoins to S, yielding the LF:

(67) [each other [they think that [t_1's pictures_j]_j AGR_j are on sale

The account is similar to the i-within-i analysis, in that it uses the coindexing between AGR and the subject NP. Suppose that the head noun pictures is coindexed with its maximal projection, as can be readily assumed (cf. Williams (1980)). Then the subject of the NP is governed by two coindexed elements, the head noun, and AGR. Suppose now that a category is identified as having a GC only if the governor is unambiguously represented in the sentence. Let us assume the following:

(68) X unambiguously governs α if X governs α and there is no Y, such that Y governs α and Y is coindexed with X.

(69) GC (definition)

φ is a GC for α iff φ is the c-commanding domain of the nearest c-commanding subject to α such that φ contains an unambiguous governor of α.

This implies that if a category is ambiguously governed, then it does not have a GC. In (69), the trace of each other does not have a GC, since it has an ambiguous governor. So this trace is an empty category which cannot be bound in its GC, since it does not have one. It bears a th-role of its own, so it is PRO-like. Having a PRO-like element in such a position is not a problem, its interpretation derives from control theory, which itself can be expressed in terms of identification.
within a certain domain as in Manzini (1983). 39

The second problem is the attachment of the element each other to something. We will assume that each other does not have to be in a chain. We thus have to assume that it behaves like an adverbial, contributing to the semantics of the sentence in the same way that it does if it actually is in a chain with the argument position that it is coindexed with.

1.6.6 A'-chains

In order to be licensed, an operator must form a chain with at least one variable, which must be an empty category in English. 40 We assume Kayne's (1981a) definition of the ECP to be a condition on such chains: 41

39. N. Chomsky pointed out to me that an anaphor in such a position behaves like a long-distance anaphor (e.g. it has to be bound by a subject), as Giorgi (1984a) shows, as in (1). If such properties do not always correlate with the binding properties of PRO, this is a problem for our analysis:

(1) *I persuaded the boys that each other's pictures were on sale

40. If the operator binds more than one variable, the other variable(s) may be overt elements, as in our analysis of pronouns of laziness, chapter 4. See Sells (1984) for the study of resumptive pronouns, which are phonetically overt variables.

41. Maybe some local binding relation is at stake again, as shown by Obenauer (1983) and then Barss (1984) and Rizzi (1985).
(70) **Condition on A'-chains**

(i) If \( X \) is in an A'-position which has an entire \( S \) as its immediate scope, then \( X \) forms an A'-chain with \( Y \), where \( X \) is in a percolation projection of \( Y \).

(ii) If \( X \) is in an A'-position which does not have an entire \( S \) as its immediate scope, then \( X \) forms an A'-chain with \( Y \), where \((X,Y)\) obeys condition A.

If \( X \) is adjoined to \( S \), it does not have scope over an entire \( S \), since its scope does not comprise the S-node which dominates it. This is to state that QR is clause-bound. Kayne's (1981a) definitions are the following:

(71) **Empty Category Principle (ECP)** (Kayne, p.57-58)

An empty category \( \phi \) must have an antecedent \( \alpha \) such that

1. \( \alpha \) governs \( \phi \) or
2. \( \alpha \) c-commands \( \phi \) and there exists a lexical category \( X \) such that \( X \) governs \( \phi \) and is contained in some percolation projection of \( X \).

(72) A is a percolation projection of \( B \) if \( A \) is a projection of \( B \), or \( A \) is a projection of \( C \), where \( C \) bears the same superscript as \( B \) and governs a projection of \( B \), or a percolation projection of \( B \).

Coindexation between governors usually reflects complementation (see Kayne (1981a) and (1983) for details). For us, the ECP is a condition on chains, and chains are formed with respect to the operator, and not the empty category. We will study A'-relations in more detail in chapter 4.42

42. Going back to the impossibility of anaphors in embedded subject positions, our binding account does not make the same predictions as that of Chomsky (1965a) and Lebeaux (1984), which relies on the ECP. When the subject is properly governed but is not bound inside the embedded \( S \), as with Stylistic Inversion (Cf. Kayne and Pollock (1978)), extrapolosed anaphors should be OK in the ECP account, and bad in the
1.7 Consequences for S-structure

1.7.1 S-structure chains

Conceptually, this theory makes principle A part of the theory, since it is used in the definition of A-chains. As for principle B, it follows from the th-criterion. However, such a theory makes crucial use of Chomsky's analysis of lexical anaphors at LF, which has as a consequence that chains cannot be formed at S-structure, since the th-criterion would be violated in most cases, as in (51) (*John saw him, *John saw himself, with a th-chain at S-structure). This means that the th-criterion cannot hold at S-structure. But this is a strong claim, binding account, which they are:

(i) *Où ont-t-ils dit qu'étaient arrivés les uns les autres?
   'Where did they say that each other arrived?'
(ii) *Voici le livre qu'ils ont dit qu'a lu la même personne
   'Here is the book that they said that the same person had read'
(iii) *Voici le livre qu'ils m'ont dit que liront Muriel et Arnaud

In (ii), same cannot have the internal reading and in (iii), no sum interpretation is possible. Condition A explains that a post-verbal reciprocal is not allowed a freer behavior than a preverbal subject, and the ECP accounts for it if the movement of the anaphor is not successive cyclic, as N. Chomsky has noted. Another environment is the subject position of infinitivals introduced by for, in which the subject can be a lexical anaphor, but cannot be an empty category:

(i) They want for each other to be happy
(ii) Who do they want for to be happy?

In (i), each other has the matrix S as its GC, since this category contains its governor, for, and a subject c-commanding it. But when movement is precluded from the embedded subject position by the ECP, if, along the lines of Kayne (1981a), the matrix verb does not form a p-projection of a COMP filled with for.
which immediately overgenerates. Consider a sentence like (73):

(73) John seems to Mary t to VP

According to the present assumptions, and given that the th-criterion does not have to hold at S-structure, we predict that arguments may switch places from D-structure to S-structure, so long as D-structure and LF obey the th-criterion. For example, (73) can be analyzed as having the following derivation:

(74) D-structure:  - seems to John [Mary to VP]
    S-structure: John seems to Mary [t to VP]  (mapping onto PF)
    LF: Mary seems to John [t to VP]

This should certainly not be an acceptable derivation, because the sentence "John seems to Mary to VP" is not interpreted as Mary being the subject of the embedded clause, and John as the perceiver, contrary to the interpretation of its D-structure and LF. Given that the D-structure argument-predicate relation is restored at LF, and given that S-structure is uninterpretable in terms of such relations, this derivation does not violate the th-criterion. Nor does it violate the Projection principle, since the structure remains the same in the derivation. It does, however, violate the Isomorphy principle of Sportiche (1983). Sportiche defines the Isomorphy Principle as another condition on th-structure, which requires that arguments do not change their th-roles during the derivation. So, to avoid derivations like (74), the Projection principle and the Isomorphy principle together must hold at all levels. The Projection principle is a condition which
applies between predicates and the structure they occur in: the predicates must select their complements in a unique fashion throughout the derivation. The Isomorphy principle is the principle that we need to ensure that the arguments, namely the categories plus their lexical content, keep their relation with respect to their predicates throughout the derivation. We take the following formulation:

(75) **Isomorphy principle** (definition, Sportiche, p.23)
Thematic structure is syntax invariant.

In other words:

(76) **Isomorphy principle** (definition)
If an argument is selected in a position P1 at a level L, then it is not selected in a position P2, at a level L', L ≠ L', P1 ≠ P2.

This principle rules out ill-formed derivations like (74), with Mary the faulty argument, which occurs in two th-positions in the derivation. Note that John is not faulty in this derivation, because it S-structure position is invisible in the th-structure, since it is not a th-position. It thus cannot be said to change th-roles. With this principle, the th-criterion does not have to hold at S-structure. For example, in a derivation like (77), it is not necessary for the predicate to know that it takes John as its subject, at S-structure: so long as there is a position there, the Projection principle will be satisfied. As for the Isomorphy principle, it is not violated because no argument switches th-position:
In (77), the fact that John and the empty embedded subject position are not related until LF is not a violation of the Projection principle, nor the Isomorphy principle. So, these two principles must be operative in a whole derivation. As for the th-criterion, it is met at D-structure and LF, but not S-structure.

Do these considerations imply that chains are not formed until LF? Sentences like "John saw himself" show that chains cannot be formed at S-structure, because the binder is in a th-position, or linked to a th-position. As for NP-traces, nothing would go wrong if their chains were construed at S-structure. And this is the same with clitic traces. So, one could say that chains are allowed to be formed at S-structure, but are not obligatory, and that they are necessarily formed at LF.

There is some empirical justification for claiming that NP-chains are not formed at S-structure, and that clitic-chains are, the relevant distinction between the two types of chains being the position of the head: this position is either an A-position (NP-chains) or a non-A-position (clitic-chains). The Principle of Full Interpretation, of Chomsky (1985a), is a general principle requiring that the presence of all elements in a sentence be justified. This principle applies at
LF, in Chomsky (1985a), but let us assume that it applies at all levels, and in particular, at S-structure. The main characteristics of the licensing principles is that they are relational in nature, they do not rely on intrinsic properties of elements. For example, it is not sufficient for an NP to be admitted in a sentence if, say, it has phonological content, because this is only an inherent property.

The formation of chains at S-structure can be viewed as a necessity for certain elements to be licensed there. Following Chomsky (1985a), what licenses an NP is either being predicated of (by being a subject), or bearing a th-role. As for clitics, it seems reasonable to assume that they are not in a position which licenses them, so they need to be related to a th-role, in order to be licensed. For example, the absence of expletive clitics can be explained on these grounds, if we make one single assumption. Suppose that expletives do not form a chain with the argument understood as related to them, as il and the extraposed S in (78):

(78) Il est clair que Marie est intelligente
    'It is clear that Marie is smart'

Then, if a clitic must be licensed, and if licensing involves either being in a licensed position or being in a th-chain, an expletive clitic cannot exist, since the clitic is neither in a licensed position (an A-position) nor in a th-chain (its chain does not have a th-role), as shown in (79):

(79) *Ils le trouvent clair que Marie est intelligente
    'They find it clear that Marie is smart'
So, consider again the fact that NP-movement is to an A-position.43 Since A-positions are licensed, a moved NP will be licensed at S-structure, making it unnecessary to relate it to its trace, at least as far as the licensing principle is concerned.

As for clitics, they are found in adjoined positions, which are not licensed positions. In order to be licensed, clitics must be related to the structure somehow. The second option comes into play: to be licensed at S-structure, an argument in a non-A-position has no other choice than being licensed by th-marking. This entails that clitics, which are arguments, must form a chain with some argument position, already at S-structure. So, we can assume that, at S-structure, chains are formed from the point of view of the head of the chain: if the head is not already licensed by being in a licensing position, then it must belong to a licensing chain.

Note that this makes chain-formation at S-structure optional. But the result that we want for the cases discussed in the next section is that chain-formation is possible only for clitics and is not allowed for NPs in A-positions. So we will rely on some principle of least effort, which will prevent forming a chain when this is not necessary.

Concerning chain-formation at LF, we could view it as a manifestation of the th-criterion, that is to say, the necessity for arguments to receive at least one th-role. However, in order to get principle B as a

43. See Johnson
(1985) for a study of properties of adjunctions and movement in general.
consequence of chain-formation, we must make chain-formation obligatory at that level, which implies that this algorithm has no functional justification.

To summarize so far, we have assumed that, at S-structure, chains are formed only if necessary. What makes chain formation necessary at that level are licensing principles. We have to assume that the th-criterion does not necessarily hold at S-structure, so the head NP of an NP-chain does not have to form a chain with its trace at S-structure, and should not, given the condition that chains are formed only if they are necessary. Then, to prevent the ill-formed derivation of (74), we have to maintain the Projection principle and the Isomorphy principle at all levels. On the other hand, clitic-chains must be formed at S-structure in order for the clitic, which is in a non-licensed position, to be integrated in the structure by entering a th-chain. Now, let us consider a fact which shows that there is a distinction between clitic-chains and NP-chains.

1.7.2 Raised subjects

These facts concern causative constructions in French. We assume that clitic-chains have to obey principle A. For another characterization of the local domain in which clitics must bind their traces, see Borer (1981) and Zubizarreta (1982), who claim that government has to hold between the clitic and its trace, and that the relation between the two is not constrained by binding principles. See also Jaeggli (1980), Aoun (1981), Manzini (1983) and Goodall (1984). We
will however leave such a possibility aside, and, rather, claim that the locality condition between clitics and their traces is due to condition A.

Rouveret and Vergnaud (1980) discover that it is possible to cliticize subcategorized PP complements onto the causative verb only if the embedded subject is also cliticized on the matrix verb (the analysis presented here has been proposed in Aoun (1981)). Consider (80) (from Rouveret and Vergnaud: (304)a-b):

(80)a. *Paul y a fait comparer cette sonatine à ses amis

b. Paul leur y a fait comparer la sonatine

Paul made -to it- compare the sonata to his friends\them

'Paul made his friends\them compare the sonata to it'

Rouveret and Vergnaud attribute the ungrammaticality of (80)a to Opacity, namely, to the fact that the trace of the clitic is not bound in the domain of the subject *ses amis*, which is the embedded S. And since the cliticized counterpart, (80)b, becomes acceptable, they propose that the domain of a subject is the union of the domain of its manifestation, that is to say, the union of the c-command domain of *leur* with that of the trace of *leur*, so that the trace of the clitic y is bound in that domain.44

Now, any theory which accounts for the contrast in (80) in terms of an extension of the domain of the subject has to account for the fact

44. See Aoun (1981), for a formulation of these results inside a theory with chains.
that this does not happen with A-movement. When the subject moves to an
A-position, its domain of binding does not extend to the higher clause,
as shown in French, again, where raised subjects do not allow
clitic-climbing:45

(81) *Paul les semble [ t₁ lire t₂ ]
    Paul -them- seems to be reading
    'Paul seems to be reading them'

In our terms, following Rouveret and Vergnaud (1980), the domain of a
subject can be the domain of the highest member of the chain formed by
the subject without adding any condition to this. If clitic-chains are
formed at S-structure, then the contrast in (82) falls out, since the
domain of the subject clitic will be the higher clause. If A-chains are
not formed until LF, then the domain of a raised subject will not extend
the possible domain of the clitic until LF. Since clitic-chains are
formed at S-structure, they still have to appear embedded, which
explains (81).

So, our hypothesis that NP-chains are not formed at S-structure
receives some empirical confirmation, if the account of the contrast in
(82) is on the right track. There exists another set of minimal pairs

45. See Aoun (1981), who raises the same problem, and solves it by
characterizing A-chains and clitic-chains differently. Rouveret and
Vergnaud implicitly answer this question by providing the condition
that the moved subject must be an argument of the predicate which it
moves to, in order for its domain to extend to the matrix VP. For them,
the causative verb and the embedded verb form a thematic complex, in
which case, the embedded subject is an argument of the whole complex,
and hence obeys their condition. This does not happen in raising
construction, where no thematic complex is formed, so the domain of the
subject cannot expand there.
that our view of LF-chains can explain.

1.7.3 Leftward-tous and restructuring at Logical Form

Kayne (1975) observed that tous 'all' can move leftwards and place itself before the verbal element which governs the object position that tous is associated with, as in:

(82) Paul les a tous rencontrés
Paul them has all met
'Paul has met them all'

As an aside, note that all the examples will contain clitics, associated with tous, not full NPs. This is obligatory with leftward movement of tous, but we will tacitely assume that this particularity is unrelated to the characteristics of the movement.

For simple sentences like (82), it seems that the movement of tous could easily be syntactically captured by identifying it to clitic-movement. However, the data are contrary to this expectation, as shown below:

(83)a. Paul a tous voulu [PRO les oublier]
    Paul has all wanted them to forget
    'John wanted to forget them all'

    b. *Paul les a (tous) voulu [PRO oublier]

In complex sentences, tous can raise, but, as shown in (83)b, the clitic cannot.46 So, it may be the case that tous behaves like a wh word. But

46. Tous is not forced to raise in front of vouloir in (83)a, as shown in (1):
this is false too, as shown below:

(84) *Paul a tous pensé que je les aurai oubliés
'John thought that I would have had forgotten them all'

Some work on the characterization of the syntactic environment of Leftward-\(\text{tous}\) has been done by such authors as Quicoli (1976), Pollock (1978), Rochette (1980), Taraldsen (1983) and others. The crucial features of the environments are, first, the infinitival character of the embedded clause, and the identity of the matrix predicate: that is to say, Leftward-\(\text{tous}\) can occur only from infinitivals, only under certain verbs. Following general concensus on this, let us identify the environment in which Leftward-\(\text{tous}\) is permissible as the typical environment for restructuring.\(^{47}\) Restructuring is a syntactic process which occurs overtly in Italian, among other languages, whereby a matrix verb combines with the embedded verb, as shown by Rizzi (1978) and Burzio (1981). Roughly speaking, restructuring has the effect that the matrix predicate and the embedded one form a verbal complex, such that the embedded sentence boundary and the embedded subject seem to disappear from the structure, with respect, mostly, to binding theory.\(^{48}\) One of the effects of restructuring is that all clitics which correspond to embedded arguments may climb up on the matrix predicate

\[\text{(i) Paul a voulu tous les oublier}\]

47. See, for example, Taraldsen (1983). Certain speakers accept \(\text{tous}-\)movement from subjunctives under verbs like \text{vouloir}, which means that restructuring is parameterized as to the mood of the embedded clause.

48. This description is in fact Rizzi's (1978) syntactic analysis of this phenomenon.
(clitic-climbing), as in (85):

(85) Paolo lo vuole leggere
    'Paolo wants to read it'

The paradox for French is that restructuring, which does not apply, given the absence of clitic-climbing (see (83)b above), still applies, given the possibility of Leftward-tous in the same environment ((83)a). This contradiction may be resolved in a theory which adopts the existence of Logical Form. Given such a level, it is possible to postulate that restructuring applies at LF, in the unmarked case. So, movement of tous is an overt indication of restructuring. Clitics must form a chain at S-structure, so, because restructuring has not applied yet, they will be unable to climb, accounting for (83)b. As for the analysis of tous, it requires some stipulations.

Let us assume that tous moves into an adverbial position, in syntax, or that it is base-generated there. And let us assume that the licensing condition on adverbials is that they govern a VP. That is to say, an adverbial is licensed in its position, in the manner of subjects. This seems natural, since one can interpret adverbials as

49. The scope of quantifiers is a large piece of evidence for the claim that restructuring applies at LF in all languages. Consider the difference between (i), which is not a restructuring environment, and (ii), which is:

(i) Someone will think that Mary has met everyone
(ii) Someone wants to meet everyone

The wide scope reading of the universal quantifier over the indefinite is hard to obtain in (i) and easy in (ii). This can be accounted for if QR is limited by condition A, and if restructuring applies at LF.
predicates which take the VP or INFL as their argument. If this is the case, that is, if it is enough for tous to be in an adverbia
tion position in order to be licensed at S-structure, then it can occur next to any verb. However, since tous must be interpreted as a quantifier, it must form a chain with a variable at LF. And given that tous does not have an entire S in its immediate scope, its chain is subject to condition A. Since this chain is formed only at LF, and since restructuring applies there, then the chain is well-formed. Note that the variable that the quantifier binds is a chain: the clitic-chain. I assume that tous becomes the head of the new chain (tous, cl, t) at LF, like Guérin (1984).

1.7.4 One problem with leftward tous

In examples (82) and (83)a above (Jean les a tous rencontrés and Jean a tous voulu les oublier), the object that the verb which tous is connected with is a clitic, but it is also possible to find no overt element corresponding to the object position, when the moved quantifier is an NP like tout, or rien ('everything' and 'nothing'), as in (86):

(86)a. Elle a tout laissé, et elle est partie
'She let everything down, and she left'

b. Emma a tout voulu comprendre
'Emma wanted to understand everything'


51. Kayne (1981b) is the first to have made a distinction between leftward movement of tous and clitic-movement in terms of the distinction between quantifiers and arguments. Our analyses differ technically, but they are somehow similar in spirit.
In the earlier analysis, it was important not to have to form a chain between *tous* and the clitic-chain at S-structure, to explain that the relation between the two did not have to obey condition A before LF. As we said, something in an adverb position is licensed by being in a licensed structural position with respect to the VP, so *tous* does not have to be licensed by th-marking, hence it does not have to form a chain with the empty category in object position. Now, given that chain-formation at S-structure in fact depends on whether what would be the head is licensed or not, then, from the point of view of the empty category, a chain is not needed either. Given that a chain between *tous* and the empty category is not needed, then it is impossible, by the principle of least effort (cf. section 1.7.1 above). This means that, at S-structure, the empty category is pro-like, in that it is an empty category with a th-role of its own. However, remember that categories are not identified at S-structure, since the th-criterion does not hold then. For the particular case of (86), this means that the empty category in the object position is not defined yet. It is in a licensed position, which is enough.  

At LF, however, if *tous* was not present, the empty category would correspond to a pronominal, and it would be excluded (see (90) chapter 3:3.5.1). This problem does not arise, since *tous* forms a well-formed chain with it, given restructuring.

To conclude, if it is right that the licensing principles are the

---

52. In chapter 4, we claim that governed empty categories must be licensed by being bound. So, the presence of *tous* at S-structure is forced, in order to license the empty category, but note that it does not have to form a chain with it.
motivations for forming chains at S-structure, then clitic-chains and chains headed by tous can be minimally distinguished with respect to these principles, in a theory where restructuring applies at LF. Now, given the account of clitic-climbing in terms of the binding theory, and given the idea that NP-chains are not formed until LF and the hypothesis that restructuring applies at LF, we make a prediction, which turns out to be false.

1.7.5 Passive in restructuring constructions

We predict that it should be possible to passivize a restructuring verb with respect to an embedded object, as in (87):

(87) *Cette maison a été voulue démolir
    'This house was wanted to demolish'

At S-structure, both cette maison and the empty category in object position are in licensed positions, so they do not have to form a chain, so it does not matter that their relation violates condition A:

(88) [cette maison]₁ a été voulue [PRO démolir t₁]

Now, if, in restructuring sentences, the GC of an embedded argument becomes the matrix clause, then, if restructuring applies at LF, it should be possible to form a chain between the moved subject and the empty category, without violating the condition on chain-links and the th-criterion at LF, in (88). So, such sentences as (89) are predicted to be well-formed.
Note, however, that the overt counterparts are also predicted to be well-formed in Italian. But, as noted by Burzio (1981), they are not:

(89) *Quel film sara' voluto vedere t
   'This movie will be wanted to see'

As Burzio claims, the impossibility of passive in restructuring sentences does not have anything to do with the relation between the raised subject and its trace. Rather, it is due to the fact that the PRO in the embedded subject position does not have a proper controller, when the matrix verb is passivized. The explanation thus relies on control theory, rather than on strict Binding theory. 53

1.8 Two possible GCs

1.8.1 S-bars as predicates of small clauses

In this section, we will see that the choice between one of two or more governors in the definition of governing category plays a role in certain small-clause constructions in French. Small clauses are constructions which are the combination of a predicate with a subject, without the presence of an INFL node relating the two. There is disagreement as to whether a small clause (henceforth SC) forms a constituent or whether it consists in the two nodes sisters to a subcategorizing predicate. (See Stowell (1981) for extensive work on

53. One other type of explanation would be one relying on the morphological impossibility to passivize a derived complex verb.
SCs, Chomsky (1981), Kayne (1984b), Mouchaweh (1984). For the second proposal, see Schein (1983) and Williams (1983). I will here adopt a view where SCs are predicates with a base-generated Chomsky-adjoined subject position, as in (90):

(90) People consider [AP [NP the President] [AP stupid]]

Given a definition of government which allows government into a specifier position, through a maximal projection, the subject of the SC is governed by the matrix predicate and can thus receive Case from it, if the predicate is a Case-assigner. Government into a maximal projection gets rid of Stowell's (1981) problem that the category label of the SC had to be a non-maximal projection of a maximal projection, quite a deviance from the principles of X-bar theory. Moreover, the subject position of the SC is defined as an A-position, since it is a sister to a predicate, and, by coindexing, its subject, as in Williams (1980). As Stowell suggests, the predicate of a SC is subcategorized for by the matrix predicate, in terms of category-selection. However, in terms of semantic selection, the matrix predicate selects a proposition,

54. This accounts for the fact that not all matrix verbs can take the same predicates (cf. the contrast between John wants Mary elected president and *John believes Mary elected president.
that is to say, the combination of a predicate and a subject. We can express these two requirements in one formulation. In order for a SC structure to satisfy the lexical requirement of a SC-taker, only two things need to be represented in the subcategorization frame of these predicates: the category of the predicate subcategorized for, and the annotation that this predicate is a predicate and not a simple complement. Williams uses superscripts to unambiguously indicate the predication relation, which we borrow. So, verbs like consider, which take SCs, have a subcategorization frame like: [- XP^1].

Now, the licensing principle which requires that predicates take subjects will force a subject position to be base-generated together with this predicate. And the presence of this subject does not have to be redundantly indicated in the subcategorization frame of the SC-taker, it is forced by the predication index borne by the predicate.

Williams (1980) argues that all categories can be predicates, and that certain categories, like S', may be predicates if they contain an open variable position in them. I am not sure whether, for him, it is crucial that the S' contains this open position, in order for it to be defined as a predicate, but this should not be a necessary requirement, given the facts below.

There are two types of constructions in French where the predicate is an S', and in which this S' is not a relative clause: under perception verbs, like voir 'see', entendre 'hear', regarder 'look', écouter 'listen to', apercevoir 'glimpse', traver 'find' (in the sense of "discover"), découvrir 'discover', etc, and the verb avoir 'have', and
some other existential constructions, which we discuss in the next section. We will see that the two constructions differ minimally in that one is a SC in which the subject is PRO, and the other is a SC in which the subject is the overt NP. These are illustrated below, but before we consider these examples, let me acknowledge Taraldsen's work, whose ideas have had a great impact on those expressed here.

(91)a. Paul a entendu Pierre qui jouait du violoncelle
'Paul heard Pierre (who was) playing viola'
b. Paul a vu Pierre qui sortait du cinéma
'Paul saw Pierre leaving the theatre'
c. J'ai aperçu Michel qui courait après son chien
'I glimpsed at Michel running after his dog'
d. J'ai trouvé Léna qui jouait avec le feu
'I found Lena playing with fire'
e. J'ai rencontré Sylvie qui parlait avec sa sœur
'I met Sylvie talking with her sister'

(92)a. Paul a sa fille qui est malade
litt: Paul has his daughter who is sick
'Paul's daughter is sick'
b. Paul a la tête qui tourne
litt: Paul has his head that feels spinning
'Paul feels dizzy'

(91)b is from Kayne (1981b). Let us ponder on what the structure of

55. See also Burzio (1981:6.44) for the study of reduced relatives as SCs.

56. We reach the same basic conclusions concerning the constructions, the most important ones being that the chain involved in the constructions in question is an A-chain and that the structure is that of a SC. The A-chain idea was proposed in Taraldsen (1984) for constructions in Norwegian with the complementizer som.

57. Kayne (1975) introduces these examples in the literature and notes their peculiar subject-object asymmetries. Kayne (1981b) analyzes them
the sequence NP + S' is. First of all, this sequence could be analyzed as a relativized NP. But not all NPs can be relativized, such as proper names and NP restricted with a possessive determiner. However, this does not prevent such NPs above to be associated with the S', which indicates that the S' is not a relative clause. Taraldsen (1983) considers another revealing fact, which is that the head NP can cliticize onto the matrix verb, being detached from the clause, as evidence against the relative clause analysis, since this cannot be done with true relatives:

(93)a. Je l'ai vu qui sortait du cinéma

b. *Je l'ai frappé qui sortait du cinéma
   'I hit him leaving the movie theatre'

Now, let us consider the last set of examples above: (92)a-b. It may be that the object of the perception verb is a th-marked object. But it is harder to argue for this in (92)a-b, for the following syntactic reason: if the NP was an argument of avoir, then it should be an indefinite NP, as shown in the ungrammaticality of *Paul has the daughter.58

Thematically speaking, Paul has a benefactive th-role, and the state of affairs expressed by the S' is what affects Paul. So, this S' is the theme argument of avoir. This is compatible with the requirement that the object of avoir be indefinite, if S' can have features for definiteness, or at least does not have the feature [+def]. The

as NPs followed by an S' containing a controlled PRO, to account for this asymmetry and the fact that the S' behaves like a predicate of the NP.

58. See Szabolcsi (1984) for an interesting account of the definiteness effect.
question is then to determine what the NP heading the S' is doing in this position. Our hypothesis, which converges with Trædal's, is that the whole construction is a SC, in which the predicate S' is subcategorized by avoir. Semantically, the state of affairs which affects Paul is described by the predication of Paul by the S'. In the general case, predicates of SCs th-mark their subjects, as in (94):

(94) John considers [AP [Mary] [AP drunk]]

They do not, in case the predicate is a past participle, as in (95):

(95) John wants [[the law] [violated t by these people]]PrP

The cases of (92)a-b are similar to those of (95), in that the S' does not th-mark the subject. What allows the position in (92)a-b to be created is what allows it in any SC, namely, the predication relation. Let us consider the D-structure of (92)a:

(96)
The subject position of the SC is non-thematic. In the general case, raising to a non-th position is due to the fact that the raised NP would not get Case if it stayed in its D-structure position. In these examples, the movement of the raised NP should be explained. The reason why the embedded subject moves to the subject position of the SC is that empty expletives are excluded (see section 1.3.2, fn.16). Raising is thus forced to apply, and the S-structure of (96) is:

(97)

\[
\begin{array}{c}
\ldots \text{C}'' \\
/ \quad / \\
\text{NP} \quad \text{C}'' \\
/ \quad / \\
\text{sa fille} \quad \text{COMP} \quad \text{INF}'' \\
/ \quad / \\
\text{que} \quad \text{NP} \quad \text{INFL} \\
/ \quad / \\
\_\_ \quad \text{INFL} \quad \text{VP} \\
\_\_ \quad / \\
\_\_ \quad \text{est malade}
\end{array}
\]

At LF, the NP sa fille must have a th-role, so it should form a chain with a th-position. Given the definition of chains, it should enter a chain with an element which is bound by it in its GC. The question is thus to determine whether the trace is properly bound by sa fille, in

59. The following sentence is acceptable:

(i) Paul a que sa fille est malade

It is possible that (i) is the version of (96) with the NP in situ. Alternatively, the complement of the verb avoir could be a simple S'. Semantically, it sounds like an answer to (ii), which translates as "what's the matter with Paul?"

(ii) Qu'est-ce que Paul a?

In the text, we analyze the non-raised structure as ill-formed, so (i) is not the non-raised version of (96).
(97). The governor of t is INFL, hence its GC is the S which contains it. Since the NP sa fille is outside of S, the trace cannot form a chain with it. This means that (97) is excluded by the th-criterion. This is in fact true, the sentence Paul a sa fille que est malade is ungrammatical. It is grammatical only if the complementizer is qui.

It has been proposed, as in Pesetsky (1981), that the complementizer qui has the property of being a proper governor, and that que is not a proper governor. The relevant property here is not proper government, but simple government. So, qui is a licit governor. Then, given that government obtains to the specifier position through one maximal projection, the subject position is governed by qui, if qui occupies the COMP of the S'. Under these circumstances, the trace in subject position is governed, not only by INFL, but also by qui. Then, the nearest subject which contains qui is sa fille, the subject of the SC. So, the chain between sa fille and the trace can be constructed, without violating condition A, hence without violating the th-criterion. Note that our definition of ambiguous governor and the requirement that the definition of GCs dismisses ambiguous governors prevent us from coindexing qui with the subject, since the subject is already coindexed with the other governor, INFL. And, as mentioned in the above footnote, in these examples, the complementizer qui is not the wh-word. If it was the wh-complementizer, the derivation should be different, with the

60. Taraldsen (1983) differs technically, in making qui part of the chain comprising sa fille and the trace. It is qui which locally binds the trace in subject position, and which, by being the non-wh version of the word, makes the trace an anaphor and not a variable. He defines GCs such that S' is the GC of a subject and S is that of an object.
structure and the interpretation that of a relative clause.

This analysis makes a prediction, which turns out to be correct: if the constructions are derived by raising of an NP belonging to the S' predicate onto the subject position of the SC, then the relation between the moved NP and its trace must obey condition A. Supposing that the NP which raises is not a subject, then, given that the verb will be its governor, its GC will be the embedded S. Given that objects do not have another possible governor which would extend their GC, this GC will always be the embedded S. This S does not contain the raised NP, so such SC constructions should be impossible to obtain with non-subjects. This is shown true below: 61

(98)a. *Paul a sa voiture que quelqu'un a volée
litt: Paul has his car that someone stole
'Paul had his car stolen'
b. *Paul a les yeux que quelqu'un lui a fermés
litt: Paul has his eyes that someone closed him
'Paul had his eyes closed by someone'

This analysis also predicts that, given that the trace cannot be a variable, the constructions are impossible with true wh-words in COMP:

(99)a. *Marie a son amie à qui Paul téléphone sans arrêt
litt: Marie has her friend to whom Paul always gives phone calls
b. *Paul a les pieds sur lesquels son frère a marché
'Paul had his feet stepped on by his brother'

61. Taraldsen (1984) proposes that the subject-object asymmetry in questions with som in Norwegian are due to the [+en] nature of the empty category bound by som, due to the [-wh] character of this complementizer. For the French facts, the [+en] nature of the empty category is forced by the fact that the NP head of the construction needs a th-role.
Let us consider, now, whether the sentences with perception verbs should be analyzed the same way, or whether they differ in some respects. The first thing that we notice is that it is not clear whether the NP governed by the perception verb is th-marked by it or whether it is the S' which is. Usually, perception verbs can take both types of complements. For example:

(100)a. John saw his brother
   b. John saw that his brother got a haircut

However, some verbs are helpful to make the distinction, in that they do not take an S' and take an NP, like apercevoir, 'glimpse', in French as well as in English:

(101)a. Paul a aperçu son père
      'Paul glimpsed his father'
   b. *Paul a aperçu que son père semblait pressé
      '*Paul glimpsed that his father seemed to be in a hurry'

apercevoir can take the constructions in question, revealing that the th-marked argument of the predicate is the NP, in at least the apercevoir case.

62. See Massam (1985) for a discussion, and the conclusion that the complement is the S'.

63. It is very likely that perception verbs in fact may take the two types of complements. The interest of the apercevoir case is that it exemplifies the other option than with avoir. Taraldsen (1983) analyzes the SCs under the perception verbs as taking a SC headed by an S', as with avoir and not the NP governed by the matrix verb. See Massam (1985) for these SC-constructions, accounting for various types of raisings across languages.
Given that the S' can still not be analyzed as a relative clause (a proper name like Marie cannot be relativized, in the general case), this means that, in this structure, the S' must be a secondary SC predicate. Chomsky (1985a) proposes that secondary predicates are generated without a subject position of their own, and that they th-mark the NP which is sister to them, here Mary. This allows one NP to receive two th-roles, which is permitted as the th-criterion is stated there (a th chain may receive more than one th-role, so long as there is only one th-position in the chain). If the th-criterion is stated in such a way that a position cannot receive more than one th role, then another possible structure for (102) would be one in which the predicate of the SC forms a SC with an empty subject, PRO, as in Chomsky (1981). For the moment, let us consider the first analysis, that in which a th-position may be assigned more than one th role. In fact the PRO analysis will have to be the correct one, given the definition of the th criterion. so we will later adopt it. The D-structure of (102) is (103), tentatively:

(103) Paul a aperçu Marie₁ [₉, qui τ₁ montait les escaliers]  

64. The distinction between a secondary SC and a primary SC is that the latter is th marked by the matrix verb and the second is not, as in (i) and (ii), respectively:

(i) John finds Peter amusing  
(ii) John met Peter drunk
Since Marie is th-marked by the matrix verb, there is no need for this NP to be related to the empty category in the embedded S'. So, it is possible to analyze this S' as having the structure of a relative clause, namely with a wh-word in COMP and a variable in the subject position, but this yields a relative-clause interpretation, so there must be another available structure. If the empty category cannot be a variable, then it must be an NP-trace, since it cannot be pro-like (it would be excluded similarly to all governed empty pronouns). So, the only correct possibility left is the one in which the empty category forms a th-chain with the head Nt of the construction, Marie. Evidence for this is the impossibility of the trace in any other position than the matrix subject:

(104)a. *Paul a aperçu Marie, que je croyais qui t montait les escaliers.
   'Paul glimpsed Mary who I thought was climbing the stairs'

   b. *Paul a vu Pierre, qu'un fou attaquait t.
     'Paul saw Peter being assaulted by a madman'

The S-structure should thus be the following:

(105) Paul a aperçu Marie, [s, qui t montait les escaliers]

   th-chain

But this is a plain violation of the th-criterion: the chain (Marie, t) contains two th-positions. It is thus ill-formed, unless we change the

65. Obviously, relative clauses are not excluded in such positions, as in (i):

(i) Paul a vu quelqu'un qui montait les escaliers
    'Paul saw someone climbing the stairs'
formulation of the th-criterion, or unless PRO is analyzed as the subject of the S'. If the predicate of the adjunct SC had a subject of its own then this subject would be able to form a proper chain with the trace in subject position, without it being a violation of the th-criterion. So, we assume that the SC has a PRO subject, and that the structure of (104) is the following, with qui, instead of que, as the complementizer, in order for the trace to be governed by it, so that it is bound in its JC by PRO (cf. the analysis of (94)a above):

(106)

\[
\begin{array}{c}
\text{S} \\
/ \ \\
\text{NP} \quad \text{VP} \\
/ \ \\
\text{Paul} \quad \text{VP} \quad \text{S} \\
/ \ \\
\text{V} \quad \text{NP} \quad \text{NP} \quad \text{S} \\
/ \\
\text{a aperçu Marie, PRO, COMP} \quad \text{S} \\
/ \\
\text{qui NP VP} \\
/ \\
\text{t, montait les escaliers}
\end{array}
\]

The presence of this PRO like element should be allowed in such a position. It is, if it is not governed. We must assume that non-canonical government cannot cross a segment of a maximal projection, that is to say, a node in an adjunction configuration so that qui does not govern PRO, being separated by the S' segment. 66 And it is not governed by V because V does not canonically govern the Chomsky-adjointed

---

66. See May (forth.) and Chomsky (1985b) for the original definitions of segment and the definition of government and c-command making use of this notion. We may diverge from them here. A segment is one of the nodes of the same category, in an adjunction structure.
SC. It would govern it if the SC was a direct complement of the V, but it is not, since it is a secondary SC.

Note that such types of sentences may not be the only ones which show that the th criterion should stick to requiring that chains may not have more than one th position. As we have just seen, the situation in which a th chain violates the th criterion is one which threatened to end up with two th-positions, due to the presence of a secondary predicate with an NP-trace-like element in it. In our examples, this NP-trace-like element is the trace in subject position. Other NP-trace-like elements are the objects of past participles as in (107):

(107) He came home followed by the Feds

The problem also occurs when the predicate is inside an NP, th-marking it, but this is also combined with a configurational problem, raised by Hellan (1985), as in (108):

(108) [The joke made t to John’s expense] was so bad that nobody understood it

In (108), a PRO subject would solve the two problems.\footnote{See Stowell (1983) for arguments for claiming that all categories have a subject position.} The problem of relating the trace to some constituent that would bind it and the problem of the th-criterion, since, if the containing NP is defined in some way as the antecedent of the trace, then this NP would improperly have to form a chain with the trace, and hence this chain would have two th-positions.

\footnote{67. See Stowell (1983) for arguments for claiming that all categories have a subject position.}
Before we continue on these constructions, let me mention the reason why such S's as predicates are not found in English:

(109) *John had his daughter who\that was sick

We saw that the only way that such constructions could be derived was by having a complementizer govern the subject position in order for the trace to be able to form a \h-chain with the subject of the S'-predicate. In English, as shown by the that-trace effect, the complementizer that is not a proper governor. Here, we will assume that it is not a governor at all. This eliminates the possibility of making the GC of the subject position greater than it is, hence impossible for the subject of the S'-predicate to receive a \h-role, a violation of the \h-criterion.

Let us now pursue the study of these constructions, for their intrinsic interest.

1.8.2 Other sentences

We have seen that the small clauses embedded under perception verbs were secondary SCs (at least with a predicate like apercevoir, and the SCs under avoir were arguments of the verb. Avoir can also take such SCs without \h-marking the subject at all. In these cases, the constructions are existential (with il y a 'there is') or presentational (see Guéron (1980), for a definition of these concepts), or absolute:

(110)a. Existential
Il y a Pierrot qui cherche sa plume.
'There is Pierrot looking for his quill-pen.'
Presentational
b. Bertrand ne se sentait déjà pas rassuré, et cet homme qui
n'arrêtait pas de le suivre...
'Already Bertrand did not feel safe, and this man, constantly
following him...'
c. Voici Pinocchio qui dit encore un mensonge
'Here is Pinocchio cooking a lie again'
d. Absolute
Avec Félicie qui passe son temps à regarder par la fenêtre, il y a
de quoi devenir fou.
'With Felicie spending her time looking through the window, one
can be driven nuts.'

All of these are of the type described above, evidence for which is
given by the impossibility for the predicated NP to correspond to the
object, as in the following:

(111)a. *Regarde, il y a sa sœur qu'elle laisse entrer
'Look, there is her sister that she lets in'

b. *Elle s'était perdue dans cette nouvelle ville, et sa fille que
j'avais laissée toute seule...
'She got lost in this new city, and her daughter, who I had left
all alone...'

c. *Voilà cet idiot que Pierre a encore frappé
'Here is this fool being beaten by Pierre again'

d. *Avec Schubert que Pierrot sabote, il n'y a pas moyen de se
concentrer
'With Schubert being massacred by Pierrot, there is no way to be
able to concentrate'

For some reason, (111)b is not so bad, but it may be that the S' in
(111)b is interpreted as a non-restrictive relative. We will leave it
as a problem otherwise. All of (111)a-d are primary SCs, with the SC a
complement of the existential predicate y avoir in (111)a, the
conjunction et in (111)b, the presentational element voilà in (111)c,
and the preposition avec in (111)d.
1.8.3 Benefactives

Let us study in more detail the SCs under avoir, when the subject of this verb gets a benefactive th-role. The interesting thing about it is that the embedded sentence must not only contain a pronominal coreferential with this subject, but also this pronominal must be contained in the subject of the SC, and cannot be found elsewhere:

(112)a. Paul a sa fille qui est malade
   'Paul has his daughter sick'

b. *Paul a Suzie qui est malade
   'Paul has Suzie sick'

c. *Paul a Pierrot qui aime sa fille
   'Paul has Pierrot in love with his daughter'

d. *Suzie a Fred qui lui parle toujours de travail
   'Suzie has Fred always talking to her about work'

The sentences get improved, or good, if the SC contains a benefactive clitic coreferential with the matrix subject, as in (113)a-b, or if the VP denotes a benefactive action, whether or not the sentence contains a pronoun coreferential with the matrix subject, as in (113)c-d:

(113)a. Paul a Pierrot qui veut lui voler son vélo
   'Paul has Pierrot wanting to steal his bicycle from him'

b. Paul a Pierrot qui dit qu'il va lui voler son vélo
   'Paul has Pierrot saying that he's going to steal his bicycle from him'

c. Paul a le meilleur docteur en ville qui le soigne
   'Paul has the best doctor in town taking care of him'

d. Paul a Marie qui s'occupe de sa fille
   'Paul has Marie to look after his daughter'
Note that if the SC is headed by an adjective, the facts are the same: the subject must contain a pronoun coreferential with the matrix subject:

(114)a. Je ne pouvais pas venir hier, parce que j'avais ma fille malade
   'I could not come yesterday, because I had my daughter ill'

b. *Je ne pouvais pas venir hier parce que j'avais Suzie malade

Informally, it seems that a certain locality condition holds between the subject of the matrix verb and the subject of the SC (rather, its determiner position). Let us consider the S-structure of (114)a:

(115)  
\[ \text{S} \]
\[ / \]
\[ NP \quad VP \]
\[ / \]
\[ Paul \quad V \quad j \quad est \quad malade \]
\[ / \]
\[ a \quad NP \quad C'' \]
\[ / \]
\[ sa \quad fille \quad COMP \quad INFL'' \]
\[ / \]
\[ qui \quad NP \quad VP \]
\[ / \]
\[ t \quad est \quad malade \]

Given the definition of government, the matrix verb governs the specifier of the embedded subject. The definition ensures that a maximal projection Z which is canonically governed by X blocks government by X if the element which is under consideration is canonically governed by the head of Z, but not if it is not. The pronoun sa is not canonically governed inside NP, so NP does not intervene for an outside governor. As for the higher S' (C''), the NP that it dominates is not canonically governed inside it, so this NP may be governed from outside the S'. Inside the higher S', qui does not
govern it because the lower S' intervenes, constituting a segment which may not be crossed by non-canonical governors. So, neither of the two maximal projections which separate sa from the matrix verb block government, so the matrix verb governs the pronoun. This means that the pronoun is bound in its governing category, when this GC is defined by using the matrix V as the governor of the pronoun.

These remarks perfectly integrate within Guérin's (1984b) account of inalienable possession in French. Guérin claims that the NP referring to the possessed may form a chain with the NP referring to the possessor. So, if we assume that such a chain has to be formed, the only possible position for the pronoun in order for such a chain to respect condition A is the embedded subject position or its specifier position, accounting for the facts. However, the question is to justify the claim that the matrix subject must form a chain with some other element.

Let us assume that there is something special about the benefactive th-role: suppose that the benefactive th-role is assigned at D-structure, but that it needs to be licensed by something in order to be properly linked to an argument. And this licensing is checked at LF and/or at the interpretive level, as a requirement on proper interpretation. It is often the case that the benefactive interpretation depends on the meaning of a verb and its complement(s). For example, in (116), the referent of lui (dative pronoun) is understood as affected by the action denoted by the V':

- 98 -
(116) Il lui a pris tout son argent  
'He took him all his money'

or: Il lui a pris tout l'argent qu'il y avait sur la table  
'He took him all the money there was on the table'

But not all actions are understood as benefactive (or malefactive), in which case a benefactive lui sounds strange, as in (117):

(117) *Je lui ai mis un bouquet de fleurs sur cette table  
'I put a bunch of flowers for him on this table'

We will assume that the distinction between (116) and (117) is effected at the interpretive level: the benefactive th-role is ruled in as compatible with the meaning of the VP in (116) and ruled out as incompatible with it in (117).

Now, when the benefactive NP also corresponds to a possessor, it is licensed. For example, when a possessive occurs in a VP which does not by itself denote a benefactive action, a benefactive clitic is acceptable, as in (118)a-b, two possible counterparts of (117):

(118)a. Je lui_1 ai mis son_1 bouquet de fleurs sur la table  
b. Je lui_1 ai mis un bouquet de fleurs sur sa_1 table

So, the benefactive th-role assigned to lui seems to be licensed because this NP is assigned the possessor th-role. Lastly, a third way for the benefactive th-role to be properly assigned to an NP is to be the complement of the preposition for, which assigns it. So, let us
assume the following principle:

(119) Conditions on the benefactive th-role

A Benefactive th-role is licensed if

(i) the VP of which the benefactive argument is an argument is interpreted as denoting an action benefactive for the referent of that NP.

or (ii) if it is assigned by a benefactive th-role assigner, like for.

or (iii) It is borne by an argument with the possessor th-role.

Now, going back to the SCs under avoir, this verb has two potential th-roles to assign, either benefactive, or possessor (or both). When avoir subcategorizes for a clause, we assume that it assigns the benefactive th-role to the subject. This th-role, however, has to be licensed in the various ways which are expressed in (119).

That is to say, it is licensed at the interpretive level if the complement of avoir can be interpreted as benefactive in some way to the referent of the subject NP. And, as we saw, when the VP of the small clause under avoir denotes an action which can be interpreted as benefactive for the referent of the subject of avoir, the construction is well-formed, and there is no need for a possessive pronoun inside the S, as in (113)d, repeated here:

(113)d. Paul a Marie qui s'occupe de Patricia

This sentence is well-formed semantically if Patricia is someone who Paul cares for, and deviant otherwise.

68. See also Stowell's (1982) and Guéron's (1984) discussions of object-incorporation in dative-shift constructions. See also Green (1974) for a thorough study of the facts, brought to my attention by D. Massam.
If the action is not benefactive in some way, then the presence of a possessive pronoun is necessary, as we remember from (112)b above, or as shown here:

(120)a. Paul a sa fille qui parle à Pierrot
   b. *Paul a Marie qui parle à Pierrot

In (120)a, the benefactive th-role is licensed at LF by receiving a possessor th-role. It receives this th-role by transmission through a secondary th-chain. This analysis is very close to that of Hale (1981), who shows that part-whole relationships in Walpiri involve two arguments, one of which is redundant in the th-structure. Translating his analysis in terms of chain-formation, this may also be a case of a secondary chain formed between the two arguments. Zubizarreta (1983) claims that certain th-roles are assigned at LF only, and calls them adjunct th-roles, like those assigned to the subject position of predicates like promise, in the raising interpretation.69. These are secondary chains, which are formed owing to inherent properties of certain th-roles and not of predicates. Following Zubizarreta's ideas about secondary th-roles, or Guéron's (1984), a secondary th-chain superimposes on th-structure, and is not visible for the th-criterion. At LF, these NPs are seen each one in a th-position, and with no th-chain relating the two, which is well-formed with respect to the formulation of the th-criterion which requires that chains have one and only th-position. And, superimposed on this structure is the th-chain

69. See also Ruwet (1977) for a study of predicates like promise, and Johnson (1985).
representing the possessor th-role. The formation of this secondary chain is what allows its head, Jean, to satisfy the licensing condition on the benefactive th-role. Thirdly, the secondary chain, like all th-chains, has to obey condition A of the binding theory. And, as we saw above, the definition of GC allows us to say that the pronoun sa is bound in its GC by the subject Paul, since sa is governed by avoir.
Chapter 2

Adjectives and Logical Form

2.1 Same and Different

2.1.1 Introduction

These two adjectives have received much attention the past recent years, for example, by Stump, in a recent paper, Carlson (1985). Schein (1984) Dowty (1985) Heim (1985) and implicitly Szabolcsi (1984). Certain authors, like Heim, have studied them as a subcase of comparative constructions. and others have proposed an analysis for them independently of the problem of comparatives in general. I will here give an analysis specific to these adjectives. without bringing them into the larger picture of comparative constructions. Not to overcrowd the examples, these will involve same only, but the analysis carries over to different.

2.1.2 The interpretation of same and different

Consider (1):
(1) John and Peter were talking to the same person.

We will study what Stump calls the internal reading, that is to say, the reading which is licensed by an expression of plurality occurring in the same sentence as the adjective here John and Peter. (1) means that John was talking to the person that Peter was talking to. First, we note that there is a possible LF for this sentence which would yield the right truth conditions, but it would omit to express that same implies a comparison. This is when the same N is represented as an indefinite escaping the scope of the plurality: "There is a person y such that both of Peter and John were talking to that person". It is clear that it is insufficient if we want to express the meaning of (1) with different, instead of same. There is no first order formula which would represent the meaning induced by different. As the above-mentioned authors advance, such a sentence as (1) implies reference to functions: the NP the same person is interpreted as having the value of the function talk to applied to the subject. And the truth conditions associated with same are that this value is the same for all of the individuals denoted by the subject NP. whereas the truth conditions associated with different are that these values must be different for each of these individuals.

One question is what the mechanism is which yields this function. Heim (1985) makes the function appear straightforwardly at LF by \( \lambda \)-abstraction of the subject. And for the others authors, the function appears in the translation of the common noun associated with same or different, and its value is supplied by the context. We will see, in the treatment of sloppy identity, that we do not want to use...
\(\lambda\)-abstraction to create functions, but rather, following the latter analyses, these are created when expressions of plurality take scope over other NPs. We will assume that the interpretation of sentences like (1) is effected in two steps. First, the NP containing the adjective is construed like an indefinite in the scope of the expression of plurality. So, the translation of (1) is (1)'

(1)'
\[
\forall x \in \{J, P\} \exists y (y \text{ satisfies } "a \text{ person}" \land x \text{ was talking to } y)
\]

For short, the verb be will mean satisfy, in future LF translations.

What is important is that the individual which satisfies the indefinite is not to be a member of the set denoted by the common noun, but rather is interpreted as a member of the set of sets comprising one person: a person is satisfied by a set consisting of one person. So, when the indefinite is a plural, like two people, the expression is satisfied by sets consisting of two people. ¹

Returning to (1)', this formula logically entails that there is a function \(f\) such that \(x\) was talking to \(f(x)\), where \(f(x)\) is "a person". What the meaning of same does is provide the information that the value of this function is the same for all \(x, y\). So the truth condition associated with same should be added to the formula in (1)' as in (2): ²

¹ See Cormack (1985) for the interpretation of plural indefinites introducing set variables. It does not seem that plural and singular indefinites should be treated differently

² One question is whether the indefinite should directly be represented as a function of the subject, as Stump does, and Carlson (1985) or Dowty (1985), or whether this function should be added as an entailment, as we have done.
(2) "same f(x)" is true iff \( \forall x, y \in \{ \text{John, Peter} \} \ [f(x) = f(y)] \)

In making the function arise as a scope effect of a plurality, the NP containing same must be treated as a hidden indefinite (in the case of different, it is an indefinite overtly): definite NPs which do not contain bound pronouns do not fall under the scope of pluralities unless they are coreferential with NPs in the scope of a plurality.³

Thus, we expect NPs containing same to occur where indefinites may, e.g. in positions which only admit indefinites, as in (3), also noted by Anna Szabolcsi (personal communication): ⁴

3. This is shown here:

(i) a. John and Peter each saw the movie
   b. John and Peter each saw a movie, and they hated the movie

In (i)a, it is impossible to construe the definite in the scope of the plurality: John and Peter must have seen the same movie. Whereas, in (i)b, there may be two movies in the situation under description, since the definite NP is coreferential with an NP which itself may be construed as in the scope of the plurality.

4. See Szabolcsi (1984a) for the determination of possible environments for indefinites. However, not all environments which allow only indefinites allow the adjective same as shown in (i) and note that comparatives may not occur in certain of these environments either, as in (ii):

(i) *John and Mary have the same child
(ii) *John has the same child as Mary

It seems that the positions reserved for indefinites which accept same are those which can be relativized in an amount relative, studied by Carlson (1977), and Ross (1984), in a predicative sentence:

(iii) *The child he has is Peter
(iv) The noise there was was the ghost's chains
(3)a. John and Mary have the same sister
   
b. There was the same noise in the kitchen and in the attic

   Also, Lewis (1975) shows that certain adverbs are quantifiers which
   unselectively bind indefinites, as in (4):

(4) She often drinks a cup of coffee before going to bed

The meaning of (4) is obtained from an LF like:

(5) often [y is "a cup of coffee" ^ she drinks y before going to bed]

According to Lewis, adverbs of quantification range over cases, where a
   case is an n-tuple of its participants. Indefinites are represented as
   free variables, which get bound by the adverb. For example, (5) is
   interpretable as: "it is often the case that she drinks y, y a cup of
   coffee, before going to bed".

   We might expect NPs containing same to occur within the scope of such
   adverbs of quantification. So, if a sentence contains the adverb, say
   always, and an NP containing same, we expect it to be paraphrasable as:
   "it is always the case that if so and so obtains with some y, then y
   has the same value for all those cases". And this is in fact what
   happens:

(6) He always reaches the same conclusion

(6) is paraphrasable as: "it is always the case that he reaches some
Let us first consider how the LF can yield the proper interpretation. The cases in question are expressed by the sentence itself, and the "conclusion" is represented as a function of these cases:

\[ \forall x, x \text{ a case satisfying "he reaches } y, y \text{ "a conclusion"}, \]
\[ \exists y, y \text{ "a conclusion"}. \]

This logically entails that there is a function \( f \) which maps the set of cases onto the set of conclusions reached by him.

meaning of same: \( \forall x, y \text{ in the set of cases } (f(x) = f(y)) \).

To conclude, we analyze NPs containing same and different as indefinites in the scope of a plurality. This logically entails that there is a function which maps the plurality onto the set introduced in the discourse by the use of the indefinite, and the adjectives give information about the mapping between the two sets.

We will just ask one question about the syntax of such constructions, which we posed in the preceding chapter: we saw that the NP which contains the adjective and the expression of plurality must have the same governing category. We also saw that a pronoun bound by the

---

5. Adverbs of quantification seem to be constrained to positions which c-command the NP containing same/different at S-structure, and not to be able to take scope over this NP when they are left-dislocated, as shown, respectively, below.

(i) ??The same person always will sing past midnight
(ii) ??Always, John wears the same coat

6. See also Schein (1985), who could treat these examples by individuating events, and making them interact with the indefinite.
expression of plurality saves the structure and allows the embedded clause to be interpreted with respect to the matrix expression of plurality, when the embedded NP is a reciprocal, or a plural participating in the sum-of-plurals interpretation. However, a bound pronoun does not save the adjectives. The question of why this is so is addressed in the next section.

2.1.3 Bound pronouns and same

We saw, in section 1.5 of chapter 1, that the interpretation where a reciprocal seems to have wide scope was in fact due to the presence of a pronoun interpreted as a bound variable, which itself allowed the clause containing it to be interpreted in a relation with the matrix NP. But the pronoun did not save the interpretation of same and different, when they were embedded, as in (8) (cf. Section 1.51, fn.22):

(8)a. John and Igor think that Mary and Olga will betray them
b. *Peter and John think that they will betray the same person

The reason why a pronoun does not save the interpretation of same is that this adjective is not free to escape from applying to the noun that it accompanies. If the interpretation of (8)b was done along the lines of that of (8)a, then its LF would be the following, with the embedded S' in scopal relation with the matrix subject:

(9) [[P.and J.].x[[that they will betray the same person].k[x thinks k]]]]

Now, in the case of the sum of plurals, we stipulated that a constituent containing plurals could itself count as a plural. So, the moved S
could count as a proper element to interpret with respect to the plural subject by inheriting the plural property of its arguments. Now, what would it mean for the clause to inherit the property which relates the subject Peter and John and the NP the same person? It would mean that the meaning of same percolates up to $S'$. The thought "I will betray $y$, $y$ a person" would be interpreted as a function of $\{\text{Peter, John}\}$, and the adjective same would provide the additional requirement that the value of this function is the same for Peter and for John. And, if the thought "I will betray $y$" is the same for everybody, then it follows that $y$ is the same person for everybody too. So, the proper interpretation could actually be derived from such an LF: if the thought is the same, then the person which occurs in the thought is the same too. So, why can't the interpretation be obtained in this way? What prevents construing the interpretation in this way is that the adjective same does not apply to the thoughts. This adjective takes the NP, and not the clause, as its semantic argument. Given that the adjective is syntactically associated with the NP, the function whose values it takes as its arguments is one which maps $\{\text{Peter, John}\}$ onto the set of persons, not onto the set of thoughts. Remember that the syntactic condition for two elements to participate in these types of interpretations is that they have the same scope. Now, since the NP the same person is inside the clause, its scope is limited to that clause, even if the clause is moved up. So the NP containing same cannot be construed in a direct relation with the matrix subject.

This means that the difference between the interpretation of plurals and the interpretation of comparative adjectives is that the property of
denoting a plurality is compositional, but not that of being the argument of an adjective: a constituent like $S'$ may denote a plurality if its arguments denote pluralities. But the meaning of same does not percolate up from an NP to the clause that contains that NP. Such a contrast as the one between the sum-of-plurals interpretation and the interpretation of adjectives like same is important for the following reason: it indicates that the relevant relation with the plurals at LF is one between the subject and the clause. Otherwise, if it was between the subject and the plural NP Mary and Olga (or the reciprocal in the wide scope reciprocal interpretation), there would be no contrast between plural NPs and NPs containing same and different. Moreover, in sentences containing same and a bound pronoun, like (9)b, since the clause is able to move up to the matrix clause, because of the pronoun, and since the sentence is bad, then the assignment of scope to this clause is not what allows sentences to be well-formed, in the sum-of-plurals interpretation and in the cases of wide scope reciprocals. This means that it is indeed the case that the NP inside this clause cannot escape it, otherwise the sentences with same would be acceptable. This answers the question that we posed in section 1.5.2 (chapter 1), concerning the possible extraction of the NP out of the moved clause as the saving device for these sentences. 7

As for another environment in which the internal reading of the same

7. Also, if a bound pronoun allowed an embedded quantifier to have scope over the matrix $S$ in general, a sentence like (i) would be able to be interpreted with everyone having scope over someone, which it is not:

(i) Someone$_1$ thinks that she$_1$ will meet everyone
book is possible, consider (10):

(10) John and Peter told Mary where to get the same book

What explains the grammaticality of (10) must be that the direct association of the embedded NP with the matrix one is made possible with respect to condition A by applying restructuring. We said earlier that restructuring applies at LF. In this case, the sentence would not have to move up. In situ, it restructures with the matrix verb, allowing the matrix S to be a GC for the embedded NP. Hence, the NP may be assigned matrix scope without violating condition A.

We now turn to two other types of adjectives.

2.2 Incroyable

2.2.1 Introduction

We have just seen that the adjective same is an adjective of comparison, which relates two NPs with each other, such that the NP which contains this adjective is interpreted as a function of the other one. We saw that the relation between the two NPs is local, and that it cannot be extended if the adjective is contained inside a sentence which contains a bound pronoun, contrary to what is the case with wide scope reciprocals and plurals in the sum-of-plurals interpretation. We hypothesized that the reason of such a difference is that the adjective same can apply only to the NP that it is a constituent of, and that, if
the sentence were to be saved by the bound pronoun, it would also have to be the case that the adjective applies to the moved clause and not to the NP anymore. It is not a surprising result that this requirement cannot be met: adjectives apply to the NP that they are constituents of. Hence, the requirement that the adjective applies to the clause would be a violation of the syntax. In the next sections, we are going to study two adjectives, which have the particularity which is missing in same and different. Namely, they can apply to constituents that they are not associated with on the surface, and at the same time, they do not apply to the constituents that they are associated with. These two adjectives are incroyable 'incredible' and a class of semantic analogues, and wrong. They behave, we will argue, like operators at LF, and take as their argument a constituent of the sentence which is not the NP. These two adjectives form a contrastive pair in that they take different arguments: S' for incroyable and VP for wrong. This makes them superficially similar to wh words on the one hand and to clitics on the other. We will see that this is more than a superficial resemblance, and that these two syntactically behave like wh words and clitics. The theoretical result will be that, as for their properties, certain types of LF movement mimic syntactic movement. 8

2.2.2 Some facts

In French, it is possible to say (11), intending to mean (12):

8. But see Huang (1982) for the claim that LF movement is freer than overt movement.
(11) Paul est allé voir un film incroyable

(12) For a certain movie x, Paul went to see x, and it is unbelievable that Paul went to see x

(I will give a more precise rendering of the reading a little later.)

As we see, under this reading, the adjective does not modify the NP movie. Rather, it modifies the sentence which contains the NP: what is unbelievable is the fact that Paul saw that particular movie. Two questions then are raised. The first one is how it is possible for the AP not to take the NP as its argument, and the second one is how it can take something else than the NP, here the S', as its argument, since it does not occur in a proper configuration with respect to it.9

These two questions can be partly answered in the following manner. If, following Zubizarreta (1982), certain predicates, such as susceptible in French, 'likely', are allowed to assign their th-role at

9. These facts are somewhat exceptional, and, often, speakers do not accept these judgments. My judgments are that the readings under consideration are acceptable with certain types of sentences, but they do not generalize to all sentences of the form: NP V [...]incroyable. ..]. In general, the N of the indefinite must be semantically unmarked. For example, (i) is not acceptable:

(i) *Paul a volé une banque incroyable

Here, we are concerned with explaining how these adjectives can behave the way they do, leaving unsolved the reason why this is a limited phenomenon. English obtains comparable readings with incredible in the specifier, as Richard Sproat informs me, as in:

(i) Paul ate an unbelievable number of toasts

I ask the reader to correlate the sentences with the English counterparts, which I do not do systematically.
LF only, then it is also possible for the adjective incroyable not to th-mark the NP at all, and then find a way to assign its th-role at LF to the S'. This has two consequences. The first one is that the AP is never in a thematic relation with the NP. But, if, following Stowell (1981), phrase-structure rules are to be eliminated from the grammar, and replaced by well-formedness conditions which rule in all and only all the correct combinations of predicates and arguments, with all and only all the correct orders, then we have to account for the fact that the AP occurs inside the NP whereas no th-marking ensures that they can be associated syntactically. In other words, the adjective is in a syntactic relationship with the NP which it is not th-marking, so something must make them compatible with each other, and only with each other. A plausible answer to such a question is that the AP and the NP are put together in syntax owing to Agreement. Sportiche (1983) suggests that two formatives which agree with each other are subject to a government condition. So, if we assume that Agreement holds between adjectives and nouns, then the two can and must be found together. In particular, APs must agree with N or NP, to the effect that the AP must govern an N or an NP. In the sentences that we are going to deal with, the AP is inside an NP, so we are considering cases in which Agreement has to hold between AP and N. Moreover, adjectival Agreement does not

10. Since English does not show overt, morphological, Agreement between Ns and As, this means that Agreement is like Case: it is an abstract notion, which may or may not be reflected by the morphology.

11. Agreement between AP and NP holds in attributive constructions, as in small clauses, or in sentences with be (subject to debate for these, since if the AP is inside the VP, it does not govern the subject):
put S's together with adjectives, so the adjective cannot associate with 
the S' in syntax. This hypothesis about Agreement makes sense only if 
it is not checked at LF. So, we assume that Agreement is a property of 
S-structure, and maybe D-structure, but not LF.

The second consequence has to do with the structural conditions on 
th-role assignment. Following Williams (1980) and Rothstein (1983), 
predication takes place when the argument and the predicate c-command 
each other. 12

Since the AP occurs inside an NP, it cannot th-mark the sentence 
which contains the NP, since it does not c-command it. The only resort, 
if the structural conditions on th-role assignment remain unchanged, is 
that the AP moves at LF to a position such that it can c-command its 
argument, S', and such that S' c-commends it.

Before turning to the properties of this movement rule, let us 
characterize some semantic properties particular to the construction.

2.2.3 The argument of incroyable

The reading indicated for (12) is insufficient. Incroyable also says 
that the proposition is incredible of somebody. For example, in (12), 

(i) John considers [[this painting]NP[ugly]]
(ii) [This painting] is not [very beautiful], I'm afraid

12. When the AP occurs inside an NP, one should assume that the N is the 
argument of the AP, as in:

(i) [The [new]AP teacher]NP
Paul est allé voir un film incroyable, it is not only unbelievable that Paul went to see such a movie. but it is also unbelievable of Paul that he went to see that movie. So, the adjective also comments on the personality of some individual mentioned in the sentence. We could assume that this reading is obtained by th-role assignment to Paul of a th-role like, say, "topic". Then, the AP would take two LF arguments, a topic and a theme, S', which says something about this topic. However, the topic-comment distinction can be defined structurally, and not necessarily thematically, so it is not necessary to make the predicate assign any special th-role (overtly, as in topicalization structures). Let us consider that incroyable takes a relation as its argument, as instantiated by a sentence containing (at least) two argument positions. One of the arguments will be interpreted as the topic and the other as the focus: the latter being the indefinite NP containing the adjective, and the former an NP, usually subject of the sentence. In the example above, it is the relation between Paul and the particular movie that he went to see which is surprising. Let us assume that a relation corresponds syntactically to a constituent minimally involving a predicate and two arguments. Thematically, S' is the theme argument of the AP, with Paul and the movie interpreted as the two objects for which the relation is set.

One striking fact about such constructions is that the individual which is the topic of the sentence cannot be the individual denoted by

---

13 See Li (1976) for considerations on the notion of topic.
14. Thanks to Marc Ryser, for suggestions and comments on this question.
the indefinite NP that contains the AP, as shown in (13):

(13) *Un homme incroyable a lu toute la nuit
    'An unbelievable man read all night'

The star in (13) indicates that it cannot have the reading in which *incroyable comments on the man. This is not because the indefinite is in subject position. This is impossible as well in a sentence like (14), in which *incroyable cannot comment on the woman:

(14) Paul plait à une femme incroyable
    'Paul attracts an unbelievable woman'

Moreover, (13) above cannot simply mean that it is unbelievable that a certain man x read all night, which shows that the NP topic is obligatory in the sentence. Since the person which is commented upon cannot be the one denoted by the indefinite, and since no other individual is mentioned, the adjective cannot have the reading in question.

The reason why the indefinite cannot be the topic could be that when the AP is assigned scope, the indefinite is carried along, and is assigned scope over the sentence too. And scope assignment to the indefinite makes it be interpreted as a focus. Since focus and topic are in complementary distribution, the indefinite is ineligible for the assignment of the topic function, which forces the presence of another argument to receive this function.

There exists a third use of *incroyable, when it is synonymous with *extraordinary. With that sense, *incroyable can directly th-mark the
noun it agrees with, but we can distinguish this reading from the one under study in that the individual mentioned is unbelievable not because of the action denoted by the sentence, but because of a set of properties which are not necessarily expressed in the sentence. There, the content of the VP is irrelevant for determining the quality of being unbelievable.

2.2.4 Specificity of the indefinite

The second fact concerns the interpretation of the indefinite containing \textit{incroyable}. First, this NP is interpreted as specific. This correlates with the fact that the indefinite is assigned the focus function.

And second, what is unbelievable with respect to the referent of this NP NP is its identity, not the class of objects to which it belongs. For example, in (15), the interpretation is not that it is unbelievable of Paul that he went to see a movie, instead of a play, but rather that he went to see that particular movie:

(15) Paul est allé voir un film incroyable

2.2.5 Verbs of psychological attitude

The third fact is that the NP which contains \textit{incroyable} may occur in subject position, but only with verbs of psychological attitudes, as in (16):

(16)
(16)a. Une femme incroyable plait à Paul
    'Paul is attracted by an incredible woman'

    b. Une femme incroyable fait peur à Paul
    'An incredible woman frightens Paul'

These contrast with a sentence like (17).

(17) *Une personne incroyable a battu Fisher
    'An incredible person beat Fisher'

It is possible that the syntax of such verbs could explain this particularity. As first studied by generative semanticists, these verbs were considered to involve a flip rule, which switches the subject and the object arguments. But it is not certain that this particularity is syntactic in nature. Rather, it may very well follow from the fact that the adjective comments on the traits of character of some individual, and that such comments can be made on that individual if, either this individual is the agent of some surprising action, or the experiencer of some surprising emotion. So, the most acceptable sentences are those in which the NP topic is assigned the th-role of agent or experiencer in the sentence.

N. Chomsky pointed out to me that constructions with incroyable were very similar to those with floating each, as in (18), studied in detail by Burzio (1981):

(18) [One interpreter each₁] was assigned to [the diplomats]₁

Burzio argues that each occurs inside the indefinite NP, but refers back to the plural. So, both constructions involve two NPs, one of which
contains an element which has to be associated with the other NP. Burzio claims that each must occur in a position bound at D-structure by its antecedent, the plural NP. And, similarly to the adjective facts, each can occur in derived subject positions, and in subject positions of psychological verbs. This leads him to postulate a movement analysis with such verbs, in which the subject is a deep object, as with \textit{plaire a NP } 'please + dative object'. It could thus be the case that the indefinite must be bound at D-structure by the NP topic, in the \textit{incroyable}-sentences too.

Let us now turn to the syntactic properties of these constructions.

2.2.6 The scope of the AP

Given the condition that th-roles can be assigned only if the predicate phrase and its argument c-command each other, and given that this argument is the sentence interpreted as a relation, the AP must move onto a position in which mutual c-command obtains. Now, since the AP is contained inside an NP, the question is whether the AP, which is treated like an operator, carries with it the whole NP, or if it is extracted out of the NP. Theoretically, we would expect it to carry the NP with it, since the scope of an operator is the scope of the quantified phrase that immediately dominates it, and this quantified phrase is the maximal projection which constitutes a GC for it, namely NP. We will consider an empirical argument in favor of extracting the whole NP, and then the AP out of it, instead of extracting the AP only, in a few pages. So the LF of (11) is (19):
Let us now consider the possible scope of the AP. As we saw in earlier examples, the AP can take scope over the S' which dominates the indefinite NP, and the indefinite NP can be either a subject (of a psychological verb) or a direct object. It cannot, however, be the object of a preposition, as shown in the following minimal pair:

(19) \([\text{incroyable}_j][\text{un film } t_j][\text{[Paul], est allé voir } t_k]]\]

There is no direct explanation of such a fact, given that LF movement out of PPs is in general allowed. So, I will leave this as an open problem.

**Incroyable** can take scope over a matrix clause, as shown in (21) whether it is indicative or subjunctive or tensed:

(20)a. Paul a épousé une femme incroyable
   b. *Paul s'est marié avec une femme incroyable
   'Paul married an incredible woman'

(21)a. ?Paul veut que Marie regarde un film incroyable
    'Paul wants Mary to watch an incredible movie'
   b. ?Paul croit que Marie est allée voir un film incroyable
    'Paul believes that Marie went to see an incredible movie'
   c. Paul a pu regarder jusqu'à la fin un film incroyable
    'Paul could watch until the end an incredible movie'

Under the matrix scope reading of the adjective, (21)a means that it is ___

---

15. For example, negative elements like **personne** can be assigned scope from PPs:

(1) Paul ne s'est marié avec personne
    'Paul did not marry anyone'
unbelievable of Paul that, for a certain movie, Paul wants Marie to watch that movie; (21)b mean that it is unbelievable of Paul that, for a certain movie, he believes that Marie went to see that movie; and (21) means that it is unbelievable of Paul that, for a certain movie, he could watch that movie until the end. The fact that matrix scope is allowed for this AP shows that its LF movement is not limited by condition A on chain-links, and thus that this movement is more like wh movement than NP movement. So, we will assume that the NP is adjoined to S' (i.e. COMP"), so that its immediate scope is a full S.

Huang (1982) claims that LF wh movement is not limited by Subjacency. However, the scope of the adjective is limited by Subjacency, as shown in the following examples: 16

\begin{enumerate}
\item[(22)a] ?Paul m'a demandé à qui montrer un livre surprenant 'Paul asked me to whom to show a surprising book'
\item[b.] *Paul m'a demandé à qui Pierre a montré un livre surprenant 'Paul asked me to whom Pierre showed a surprising book'
\item[c.] *Paul aime un\l'homme qui connaît une femme incroyable 'Paul likes a\ the man who knows an incredible woman'
\item[d.] *Paul est resté après avoir regardé un film incroyable 'Paul stayed after watching an incredible movie'
\end{enumerate}

The judgments indicated are those for when the adjective takes scope over the whole sentence, and they show that the LF movement of the NP containing the AP is limited by the wh-island condition (cf. (22)a-b), the Complex NP Constraint (cf. (22)c), and the Condition on Extraction.

\begin{enumerate}
\item[(22)] All these facts extend to the expression of amount, like incredible number, as can be checked easily.
Domain (the CED) (cf. (22)d). These examples show that the LF movement of the adjective is closer to syntactic wh-movement than LF wh-movement.

As expected, there is a subject-object asymmetry, when the AP takes matrix scope. It is impossible for the AP to take matrix scope when the indefinite is in the embedded subject position, as shown below:

(23) *Paul veut qu'un homme incroyable comprenne

(23) cannot mean that it is unbelievable of Paul that, for a certain man, he wants that man to understand. The unavailability of this reading is explained by the ECP, which we take as a condition on A'-chains, if we assume that LF movement of the indefinite NP does not go through COMP. In that case, it does not belong to a percolation projection of the trace, since this projection stops at S'. Recall that we adopt Kayne's formulation of the ECP, with the requirement that the moved NP must bind a position in some percolation projection of that position. Note also that an inverted subject is more acceptable, which indicates that its position is properly governed:

(24) ?Paul veut que comprenne un homme incroyable

17. The CED, which states that movement cannot take place out of a non-properly governed constituent, has been proposed by Huang (1982), and can be seen as a particular case of the ECP, as defined by Kayne (1981a): for Kayne, domains of government percolate up along a line of properly governed constituents. Hence, if a constituent is not properly governed, percolation up cannot take place. Given his requirement that binders must belong to the domain of government of bindees, the ECP will rule out the examples that the CED does.
So far, we have seen that the movement of the indefinite NP has the same properties as syntactic wh-movement. Remember that I have assumed that the whole NP is extracted from the sentence before the AP also is, from the NP, excluding direct extraction of the AP.¹⁸ Let us now consider the fact mentioned earlier which shows that the NP is assigned scope, when the AP is.

The interaction of scope between a quantifier and the indefinite NP may indicate whether the NP is assigned scope when the AP is. We will have to rely on contrastive judgments, for this point to be made, because we will assume that the specific indefinite is able to be construed as in the scope of a plurality, in (25)a-b, which sometimes has to be forced. So, we will consider (25) as a starting point for the judgments, with the indefinite in the scope of the plurality:

(25)a. Pierre et Paul aiment une femme incroyable
  'Pierre and Paul like an incredible woman'

  b. Pierre et Paul veulent que Marie aille voir un film incroyable
     'Pierre and Paul want Marie to go see an incredible movie'

The readings of (25)a-b are obtained from the respective LFs:

---

¹⁸. Tim Stowell suggested to me that the reason why the AP is not extracted directly is Subjacency, if NP and S are bounding nodes, as shown below:

(1) \[[\text{incroyable}]_s \text{Paul est allé voir } [\text{un film } t_1]_{NP}]\]

According to our assumptions, the movement of the NP is due to the definition of scope assigned to an operator: scope is assigned to some quantified phrase that dominates it.
(26)a. \[[\text{incroyable}]_j[[\text{une femme } t_j]_i [P \text{ et } P \text{ aiment } t_1]]\]

b. \[[\text{incroyable}]_j[[\text{un film } t_j]_i [P \text{ et } P \text{ veulent que Marie voie } t_1]]\]

Now, there is an interesting difference between QR, which assigns respective scope to quantifiers, and the LF movement of the indefinite NP which contains the AP: QR is clause-bound, it is limited like NP-movement is limited, whereas the movement of the indefinite is not clause-bound.

Remark that the claim that QR is clause-bound is demonstrated by the fact that embedded quantifiers cannot take scope over matrix quantifiers, as in (27), where \textit{someone} cannot be construed in the scope of two banks:¹⁹

(27) Someone thinks that Mary is going to rob two banks

It is often argued that, when the order of the quantifiers is the other one, i.e. with the singular embedded and the plural in the matrix clause, this NP takes scope over the plural, when it is allowed to be satisfied with one individual only in the situation described, as in (28):

(28) Two people think that Mary is going to rob someone

This is not correct. That \textit{someone} is satisfied by a single individual in the situation described may also be due to the fact that two people

---

¹⁹ However, see May (forth.) for the claim that this is not due to some propery inherent to QR.
is construed as independent from someone, and not as in the scope of
someone. Since it is a singular, there is no obvious scope effect to
show that it does not take scope over two people.

Going back to the adjectives, given the difference of scope between
QR and the movement of the indefinite, if the AP carries with it the
indefinite NP, the special reading of the adjective should be blocked
when the indefinite is construed in the scope of an embedded plural,
since the plural cannot move higher up. Consider (29):

(29) Marie veut que Philippe leur offre un livre incroyable
'Marie wants Philippe to offer them an incredible book'

Two readings are incompatible in (29): the one in which the adjective
takes scope over the whole sentence, and the one in which the indefinite
is within the scope of the plurality. This can be explained, if the AP
carries with it the indefinite. Since the scope of the plural
quantifier cannot escape the embedded clause, the indefinite would have
to stay in the embedded clause if it was to be in the scope of the
plural. That it cannot be so interpreted when the AP moves up shows
that it also moves with it, escaping from the scope of the plural
quantifier.20

In conclusion, the account of the unbelievable-facts can be

--------

20. Note also that I treat any expression of plurality like a quantified
phrase. In French, almost all of them can take scope over other
elements, whether they are indefinite or not. This is why the pronoun
leur in (29) is analyzed as a quantified phrase. In English, only
indefinites and universal quantifiers like every easily take scope over
other NPs.
summarized as follows. **Incroyable** is allowed to assign an adjunct th-role. The adjective occurs inside an NP because Agreement puts the N and the AP together, but, because of the mutual c-command condition on predication, the AP must be assigned scope over its S' argument, so, the proper configuration is built at LF, by movement of the indefinite NP which contains the AP. This movement has the syntactic properties of overt wh-movement, in that it is not clause-bound, but obeys island-conditions and the ECP.

2.2.7 Another property of **incroyable**

**Incroyable** is not the only adjective which behaves like an operator. It only represents a semantic class, which comprises adjectives like **surprenant** 'surprising', **intolérable** 'intolerable', **inratable** and **ineffable** 'untellable', **inimaginable** 'unimaginable', **revelateur** 'revealing' and others. One may wonder whether this class can be characterized semantically.21 These adjectives must all involve an understood cognizer, and, in fact, this class does not comprise adjectives like **amusant** 'amusing', **extraordinaire** 'extraordinary' **incomparable** 'incomparable'. The difference between an adjective like **amusing** and an adjective like **surprising** cannot be given on pragmatic grounds only, since they both involve a cognizer. The difference between them must then be formal: the cognizer must be formally represented in the lexical entry of **surprising**, and not be represented in that of **amusing**. This is evidenced in the contrast below, where the

21. Thanks to Marc Ryser, for a useful discussion on this subject.
overt argument sounds more natural in (30)a than in (30)b:

(30)a. It is surprising to Mary that her daughter does not want to go to school

b. ??It is amusing to Mary that her daughter always wants to go to school

Let us assume that what distinguishes these two adjectives is that one has a th-role of experiencer specified in its thematic-grid, whereas the other does not. A th-grid, as defined by Stowell (1981) is the representation of the complement structure of a given lexical entry, which is carried with it in the derivation.

Second, all these adjectives take exclamatory complements. They do not, however, take interrogative or declarative complements. For example, none of the adjectives below has the reading in which it modifies the sentence:

(31)a. *Paul est allé voir un film insoupçonnable
   'Paul went to see an unsuspicous movie'

b. *Paul aime une femme indubitable
   'Paul likes an undoubtful woman'

(32)a. **Paul a lu un livre improbable
   'Paul read an improbable book'

b. **Paul déteste un film clair
   'Paul hates a clear movie'

((32)a is also ruled out because the adjective does not have a cognizer place in its th-grid.)

Now, these constructions seem to pose a problem for a theory of semantic selection.
2.2.8 A clash of selectional features

Consider a simple sentence, again:

(33) Paul est allé voir un film incroyable

This sentence has the following particularity: the whole clause is declarative, it merely provides information about Paul. But also, it is interpreted as an exclamative complement of the predicate *incroyable*. Grimshaw (1979) argues that semantic selection must be a different component from categorial selection, and that predicates have to specify both in their th-grid. Propositional arguments can be of three distinct semantic types: exclamatives, questions, and declaratives. So, in (33), the sentence is a declarative, and at the same time an exclamative argument, which goes against the fact that these features are mutually exclusive. In chapter 4, we will see, following Goodall's (1984) representation of coordinate structures, that structures which stand on different planes may be interpreted as the union of the phrase-markers which occur on each plane. A three-dimensional representation of the meaning of the sentence is required, in sentences with *incroyable*, so that, not only will this solve the problem of representing the sentence as of two distinct semantic types, but also will it enable us to represent the sentence as a coordination of two sentences, one, the declarative S' itself, which provides information: Paul went to see a movie x, and the other one, which is a comment on the particular fact expressed by the sentence. So, we have to read the LF of (33) with the moved AP on another plane than the rest of the sentence. On its plane,
the sentence will bear the feature [E] (for exclamative), and on the main plane, it will bear the feature P (for proposition):

(34)

\[
\text{incroyable} \quad <E> \quad S' \\
\text{un film} \quad t_j \quad COMP \quad S' \\
\text{Paul} \quad V \quad NP \\
\text{est alle} \quad t_j \quad voir
\]

To conclude, the analysis of the incroyable-facts, in French and in English with incredible amount, requires treating the AP as an operator, taking scope over its argument at LF. However, since it carries with it the NP that contains it, and since the AP is interpreted on a different plane than the main sentence, this sentence may bear two contradictory semantic features: E and P.

Furthermore, one other possible understanding of the facts could be that the structure formed by the moved adjective is a relativized NP, interpreted as a concealed exclamation, instead of being an S' with an NP and an AP adjoined to it. In other words, a sentence like (11) would be represented like the corresponding exclamative NP: It's unbelievable [the movie that Paul went to see]. However, given a certain number of technical problems, this suggestion will not be pursued here. Still, one should note that sentences like (11) and their exclamative-NP counterparts have lots in common.
2.2.9 Exclamative relativized NPs

Certain relativized NPs may be interpreted as exclamatives. These are concealed exclamations, in that they are NPs, with the content of the exclamation expressed by the relative clause.\textsuperscript{22} Let us call them relativized exclamatives, understanding that these NPs are interpreted as exclamatives propositions. We will compare the particular facts which hold with the exclamative adjectives and facts which hold of their paraphrases as relativized exclamatives. (11) is paraphrasable with a relativized exclamative, as in (35):

(35) C'est incroyable le film que Paul est allé voir
'It's incredible the movie that Paul went to see'

The sentences will be given with the form \textit{c'est} + \textit{Adj} + \textit{NP}, to make sure that the extraposed complement is an exclamative, given that NPs cannot be extraposed if they are not exclamatives.

Concealed exclamations in the form of a relative clause may be of two kinds: the content of the exclamation may bear on the amount of what the head NP denotes, or it may bear on the identity of the referent of that NP or the type of things that it denotes. For example, (36)a is an exclamation about the amount of work, whereas (36)b is an exclamation about the identity of the man, and (36)c about the type of things

\textsuperscript{22} See Baker (1968) for a study of concealed propositions, and Grimshaw (1979).
denoted by the head NP: 23

(36)a. It's amazing the work he did last week
   b. It's amazing the man he managed to interview
   c. It's amazing the people whose opinion he cares about

We classify (36)b and c together, and (36)a as an amount relative, following Carlson (1977). 24

Sentences containing the operator-like exclamative adjectives are on a par with relativized exclamatives of the second type. First, the type of the argument of the adjective is the same: the adjective takes a sentence as its argument. Semantically, this sentence is treated as a relation between two arguments, one of which is treated as a topic. This implies that there must be at least two arguments in the sentence, which are in the relation expressed by the verb, as shown in (37):

(37) C'est surprenant la musique sur laquelle il s'est endormi
   'It's surprising the music he fell asleep on'

Here, the arguments are il and la musique.

23. (36)b is better in French than in English, for some reason.

(i) C'est incroyable l'homme qu'il a réussi à interviewer

24. Carlson analyzes amount relative clauses as complements of the determiner, and shows that they are closer to comparative uses than to true relatives. One of the characteristics of the amount relatives is that the operator in COMP cannot be a wh word. This is also the case with amount relatives as exclamatives:

(i) ?It's amazing the people who he talks to

(i) can only be understood as being an exclamation about the types of people that he talks to, and not the number of such people.
As with the incroyable-facts, in the case of the relative exclamative, the constituent which is associated syntactically with the AP is the relativized NP, but the real argument of the adjective is the clause, not the NP.

When the relativized position is a subject position, then this subject position must be that of a verb of psychological attitude, similarly to sentences like (16)-(17), as in the following contrasts:

(38)a. ??C'est incroyable la personne qui a frappé Paul
   'It's incredible the person who hit Paul'
   b. C'est incroyable la personne qui plait à Paul
   'It's incredible the person who attracts Paul'
   c. *Une personne incroyable a frappé Paul
   d. Une femme incroyable plait à Paul

The NP which is syntactically associated with the AP must be unmarked, semantically. The more specific the class of objects denoted by the N', the less acceptable the sentence. This is true of both types of exclamatives, but the operator-like adjectives are much worse than the relativized exclamatives:

(39)a. *Paul a réussi à interviewer un général incroyable
   b. ??C'est incroyable le général que Paul a réussi à interviewer

Also, the NP that the AP is syntactically associated with cannot be the topic, as shown in (40)b:

(40)a. *Une femme surprenante a lu toute la nuit
   'A surprising woman read all night'
   *C'est surprenant la femme qui a lu toute la nuit
It's surprising the woman who read all night

In conclusion, the same array of facts holds for both types of exclamative constructions, but the way they are related is still mysterious.

2.3 Wrong

Wrong (and right) is another adjective which is similar to an operator. Like incroyable, wrong is allowed not to th-mark the NP that it is associated with on the surface. Wrong can convey the meaning that an action should not have been performed, or that a state of affairs should not hold, with respect to a certain individual, as in (41)a-b:

(41)a. The wrong book appeared
   b. John talked to the wrong person

In (41)a, the event of appearing should not have taken place with respect to the book, and in (41)b, John's talking to someone should not have been performed with respect to the person that he talked to.

Wrong can also convey an ethical meaning, as in (42):

(42) John likes the wrong people

In such a sentence, it seems that wrong behaves the same way as it does in (42) above, but that the ethical reading is rendered by the fact that the NP has the reading in which it denotes the kind of people, and not the identity of the individuals. In (42)a-b, it is the identity of the
book and the person, respectively, which is relevant, similarly to the sentences with *incroyable*.

We will focus our attention on the reading of the adjective in sentences like (42)a-b, that is to say, sentences in which it is the identity of the wrong thing which is relevant, and not the kind of thing that it is. So, the question is the same as with *incroyable*, namely, that of determining what the argument of the AP is, and how this predicate-argument relation is realized. Note that if the argument was the NP itself, then we would have to account for why the meaning of the linguistic environment enters into the interpretation of the AP. It could be that this linguistic context occurs as some kind of additional information about why the NP is "wrong". Then, the NP would be the real argument of the AP. However, in that case, one would have to explain why any NP can act as the argument of the adjective. Like *incroyable*, the adjective *wrong* does not impose any semantic restriction on the NP. For example:

(43)a. Mary received the wrong letter
   b. The wrong tree grows in this garden
   c. John took the wrong path

Even if some sentences, like (44)a-b, are not appropriate, it is not direct evidence for the claim that the adjective selects the NP:

(44)a. *John weighs the wrong kilos
   b. *John run the wrong two miles

If we paraphrase these sentences the way they are paraphrased above,
then we get something like: John weighs certain specific kilos, and he shouldn't weigh them, for (44)a, and John run certain two miles, and he should not have run them. So, one requirement on the NP is that it must be specific, whether it is interpreted as a type or as an individual, and this specificity requirement is not met with measure phrases as those above. However, specificity is not indicated in a lexical entry (at least, for some analyses of specificity):25 There is no semantic feature which distinguishes an NP like a book when it is specific from when it is not specific. What makes an NP specific may be the way it is used in a particular sentence, in a particular context of utterance. So, the specificity requirement is not a requirement on semantic selection. So, (44)a-b are not counterexamples to the claim that wrong imposes no selectional restriction on its argument, hence it can be maintained that wrong does not th-mark the NP.26

We will assume that wrong takes a VP at LF, making (42)a-b semantically equivalent to (45)a-b:

(45)a. A certain book wrongly appeared
   b. John wrongly talked to a certain person

25. See Ludlow (1984) for the claim that specificity is not a property of logical form, that is to say, it is not obtained by scope assignment or by semantic features.

26. The absence of selectional requirement of predicates is only an indication that they are not th-assigners, because it is not a necessary property of th-assigners that they semantically select their argument. That is to say, if the argument is not selected, it does not mean that the predicate is not a th-assigner. The converse, that non-th-assigners do not impose selectional requirements on their arguments is a necessary property.
Along the lines of _incroyable_, _wrong_ is to be treated as a predicate which assigns an adjunct th-role. It takes a VP, so we will assume that, like lexical anaphors, it adjoins to INFL' at LF, in order to govern the VP. We will assume that, in this movement, the AP carries with it the NP that contains it, because the NP is the quantified phrase obtained by treating the AP like an operator. Suggestive evidence for this is that such NPs are acceptable in the subject position of a SC with wide scope, as in (48) below. This would be impossible if the AP was extracted alone, since it would violate the Subject Condition (no element may be extracted from inside a subject), accounted for be the ECP. So, the LFs of (42)a-b are the following:

(46)a. \[ t_i \_ [\text{wrong}_j[[\text{the } t_j \text{ book}]] \_ \text{INFL} \_ [\text{appeared } t_i]] \]

b. John [\text{wrong}_j[[\text{the } t_j \text{ man}]] \_ \text{INFL} \_ [\text{talked to } t_i]]

(The trace in object position of _appeared_ is the NP trace of syntactic movement, since this is an ergative verb. _Wrong_ can occur inside any subject, so the fact that this is an ergative verb is not meaningful.)

Before we proceed, note that, in (46)a, in which the subject adjoins to VP, leaving an empty category, this empty category is properly bound by the moved NP, since no maximal projection intervenes in between the trace and the moved NP (INFL' is not a maximal projection). The configuration is the following:

(47) \[
\begin{array}{c}
\text{INFL}'' \\
\text{INFL}' \\
\text{INFL} \\
\text{INFL}'' \\
\text{INFL}' \\
\text{INFL}'' \\
\text{INFL}' \\
\text{INFL}'' \\
\end{array}
\]
Now, what are the arguments for claiming that VP is the argument of wrong, and not, say, the sentence? If the argument of wrong was the sentence, that is to say, INFL" or COMP" (i.e. S or S'), then sentences with wrong would be paraphrasable as 'it is wrong that so and so'. But in this case, wrong means 'false'. If wrong means 'false' when it takes a sentence as its argument on the surface, then it should also mean 'false' when it takes it at LF. Sentences with wrong inside an NP do not mean 'it is false that so and so', so wrong does not take an S (or S'). So, we will consider the VP the argument of the AP.

Let us now turn to the predictions made by the LF analysis of the adjective. Given that the AP, and with it the NP, moves to INFL', it should behave like a lexical anaphor, and its scope should be limited by principle A. We have to be cautious with the type of sentences that we are considering. As we saw earlier, wrong can be associated with an NP which is interpreted either as denoting a kind of things, or as denoting a specific set of individual(s). It is the specific reading that we consider. In that case, the verb believe combines in an interesting way with the adjective wrong. It is clear what the combination means: to believe something about the wrong person (or thing, etc) is to be mistaken with respect to the identity of the person (or thing, etc). So, consider (48):

(48) John believes the wrong man to have spoken to Mary

(48) is ambiguous between a reading in which John believes that the man
who spake to Mary should not have been that particular man, and a reading in which John is mistaken in his belief with respect to the identity of the man who spoke to Mary (it is not x who spoke to Mary, with x the person that John believes talked to Mary). Let us call the two readings, respectively the narrow scope reading of the adjective, and the wide scope reading, since, according to our assumptions, they are obtained from two distinct LFs', one in which the AP adjoins onto the embedded INFL', and the other in which it adjoins to the matrix one, where the argument of the AP is either the embedded VP or the matrix one.

Now, consider the following paradigm:

(49)a. John believes Mary to have met the wrong man  
   b. John believes that the wrong man met Mary  
   c. John believes that Mary met the wrong man

Consider the reading in which John is mistaken in his belief with respect to the identity of the man. None of these sentences can have that reading. In other words, the wide scope interpretation is not possible in either of these sentences. This means that the LFs from which such readings would be obtained are not well-formed. It is clear why: all of the positions in which the NP containing wrong appears are positions in which an anaphor cannot be bound by a matrix NP. Only the subject position of a small clause, or an infinitival, can be bound by a matrix NP, and in fact (49) is acceptable with the wide scope reading. So, if it is true that the NP moves to INFL' at LF, then, if this movement is constrained by condition A, we can explain why only (49) is
acceptable, of all the possible configurations for the embedded NP.

One question is to determine what differentiates the movement of incroyable, and the movement of wrong. There does not seem to be any intrinsic quality that would make them different: both adjectives are determined as operators and adjunct th-role assigners. They both adjoin to an A'-position. So, in a theory which feature-distinctions for empty categories, their traces should have the same status, and hence movement should have the same syntactic properties. In our terms, what distinguishes them is the scope that they take. Incroyable takes scope over a full clause, and it behaves like a wh word, whereas wrong takes scope over a VP, and it behaves like a clitic, or a lexical anaphor. Since the only difference between the two is the position to which the NP lands, conditions on chains must be sensitive to the position of the head, which is expressed by our statement that condition A applies only to chains where the operator does not have a full S as its immediate scope.27

At this point, it would be interesting to test whether Lasnik's cases of illicit movement are permitted here, that is to say, local binding by a coindexed element which would allow the adjective to have wide scope. But this cannot be tested, since the NP containing wrong should eventually be coindexed with the matrix subject, as is the case with lexical anaphors, leading to a violation of principle C, since the NP

27. Also, remember from chapter 1, fn.4, that lexical anaphors as well could be differentiated with respect to the nature of their LF movement: long-distance anaphors are like wh quantifiers, and short distance anaphors are like QRed quantifiers.
containing the adjective is a referential expression. Relevant examples are of the form of (50):

(50) *John\textsubscript{i} believes that he\textsubscript{i} saw the wrong man\textsubscript{i}

This closes this question. Let us now turn to another type of facts which also suggests that the NP the wrong man behaves like a quantifier at LF, indicating that this is due to the adjective, which behaves like an operator.

First of all, one could test whether an NP behaves like a quantifier by checking whether the NP creates weak crossover violations. In fact, these NPs are subject to weak crossover:

(51) a. *Its\textsubscript{i} publisher destroyed the wrong book\textsubscript{i}
    b. *Someone who had met him\textsubscript{i} yesterday talked to the wrong man\textsubscript{i}

Now, let us consider an argument of another type. Schein (1983) and Williams (1983) argue that the scope of a quantifier in a small clause is the matrix clause and cannot be the SC itself. If the NP containing the adjective behaves like a quantifier, then we expect the adjective not to be able to take the predicate of the SC as its argument, but to take only the matrix predicate. So, first, consider (52):

(52) The wrong man is angry at John

The relevant reading of this sentence is the one in which some person is angry at John, and for some reason, another person should be angry at John. We can imagine a context in which, for example, John is a child, and his uncle, but not his father, is angry at him for some misdeed of
his. The right man to be angry at him would be his father. Now, consider the scope of a quantifier as a subject of a SC:

(53) John considers nobody angry at Bill

(53) means that there is nobody, such that John considers that person angry at Bill. It cannot mean that John considers that there is nobody angry at Bill. This is because quantifiers cannot minimally take scope over SCs. Schein (1983) explains this by claiming that quantifiers adjoin to their governing category. Since the subject of a SC has only one governor, the matrix predicate, then it has only one GC, hence only one adjunction site, that of S, its GC. Now, consider (54):

(54) John considers the wrong man angry at Mary

(54) cannot be construed with the meaning that John considers that someone is angry at Mary who should not be angry at her. Rather, it seems that wrong can only have the wide scope reading, in which John considers wrongly of a person that that person is angry at Mary (he is mistaken with respect to the identity of the person). This shows that the NP containing wrong behaves like a quantifier at LF.

Note that this is also a test for the incroyable-facts, and they lead to a similar conclusion, as well as facts with same, as we can see:

(55)a. Paul croit une femme incroyable amoureuse de Jules
   'Paul believes an incredible woman in love with Jules'

b. John considers an incredible number of people drunk

c. Peter and John consider the same woman angry at Bill and Suzan
The wide scope reading of the adjectives in (55)a-c is much preferred. The paraphrases of the readings which are not available are the following. For (55)a: Paul believes that, for a certain woman x, it is unbelievable that that woman is in love with Jules. For (55)b: John considers that there is a huge number of people drunk. And for (55)c: Peter and John consider that the woman who is angry at Bill is the same as the woman who is angry at Suzan.

To conclude this part, adjectives like incroyable and wrong act as operators at LF. One adjective takes S' as its argument and hence behaves like a wh word, and the other takes VP as its argument, and hence behaves like a clitic. The rest of the chapter is devoted to constructions similar to these, where an NP contains a modifier which does not modify it, but the sentence that it occurs in. The difference between these adjectives and the other modifiers is that the argument of the modifier actually occurs in syntax, in the form of an empty category.

2.4 Exclamations with tough adjectives

2.4.1 Difficile à croire is like incroyable

Jean-Roger Vergnaud noted that the phenomenon displayed by incroyable was also displayed by the corresponding tough-adjective difficile à croire:
This sentence means (if the adjective does not modify the noun directly): Paul liked a certain movie x, and it is hard to believe that Paul liked that movie. Semantically, the class of tough-adjectives which accept this reading is, derivatively, the same as that of incroyable, that is to say, first, they are exclamatives: hard to believe is good, but not hard to tell. Second, they involve a cognizer, which is due to the fact that they all combine with a verb, and this verb is a verb of cognition, which takes a propositional complement. Here is another example, in which comprendre is not to be understood as taking a question, but as taking an exclamation, where it has the meaning which it has in "being understanding" (i.e. "sympathetic").

First of all, let us determine what the structure of sentences like (57) is. If it is possible for a predicate not to assign its th-role until LF, it must be the case that this th-role is an external th-role, because of the Projection principle. It would be contrary to the spirit of the Projection principle to allow a predicate to assign an adjunct

28. I say "derivatively", because I am talking of the meaning of the complex hard to believe, and not of the tough adjective alone, hard.

29. These two readings of understand have been pointed out by Steriade (1981), in sentences with if-extraposition, in which the extraposed sentence is construed as the complement of the predicate understand and at the same time as an adverbial, as in (i):

(i) I wouldn't understand if he came
th-role to an internal argument, since it would allow in principle the creation of an argument position during the derivation. So, we stipulate, following Zubizarreta (1982), that a predicate cannot assign an adjunct th-role to an internal argument. In that case, given that verbs like croire 'believe' take an internal argument, and given that the Projection principle requires the presence of this argument in the derivation, and given the ban on assigning an adjunct th-role to an internal argument, then it must be the case that these verbs do have an argument in the derivation. This argument is an empty category, as in the following S-structure:

(58) Paul a aimé [un film [difficile à [[Op] PRO croire [e]]]]

Semantically, the object of croire must be the sentence, since what is hard to believe is that Paul saw that movie. However, syntactically, this empty category occurs as a variable bound by an operator (OP) in the COMP of its S', in a construction which is predicated of the NP un film. The problem is that if the S' is predicated of an NP, then the variable should be of the NP-category, not S'. and it should have the semantic value of this head NP. There thus seems to be a contradiction between the fact that the empty category should have the semantic value of the S', given the meaning of the sentence, and at the same time should have the head NP as its antecedent, given that the NP is the head of the construction.30

30. In chapter 4, I provide an analysis of tough-constructions which does not involve the empty operator. However, it does involve a variable object of the embedded predicate. For the cases at hand, we assume Chomsky's (1981) analysis of tough-constructions.
There are two plausible solutions to this problem. The first one will be discussed in this section, and the second in the next section. The first solution is the one that we have already adopted for the incroyable facts. In syntax, non-proper semantic relations are constructed, but, at LF, the proper semantic relations are restored. In our examples, the following has to hold at LF: the complement of croire is the S' that contains it, and the NP un film escapes from any semantic relation with the complex adjective difficile à croire, or the verb croire. Given that the rule of predication applies at LF only (which identifies the value of the variable inside the embedded clause as being that of the head of the construction), there is no more assumptions to make than we have already made. The tough adjectives will be treated as operators, will be extracted from the NP which contains them, and will take the sentence as their external argument, via the rule of predication. This is to say that, at LF, "tough movement" applies to S's.

This is the analysis. So, at LF, the NP containing the AP moves to S' and the AP is extracted from it, yielding:

(59) \([\text{difficile à OP}_1 \text{PRO croire } \text{[e]}_1]_j [\text{S'}_i \text{[un film } \text{t}_j]_k [\text{Paul aime } \text{t}_k]]\)

In this structure, S'\text{'} governs the AP which is adjoined to it. The rule of predication which relates the variable to a referring antecedent requires that this antecedent and the AP c-command each other, since the AP is predicated of it, which the S' and the AP do. So, the S' can be interpreted as the constituent giving its value to the variable object of croire. In this case, this variable and the operator in COMP must be
coindexed with it, in order for the relation to be properly effected. This is shown in (59). Note that the coindexing between the S' and the variable holds at S-structure as well, and there is no coindexing between the NP and the variable, as shown here, which is well-formed if coindexing between the head of the construction and the variable is not required until LF, our assumption (cf. Chomsky, 1981):

(60) [Paul aime [[un film] j [difficile à OP1 croire [e],]]_j]\_S1

Concerning the i-within-i condition, it does not matter that the variable is coindexed with a constituent which contains it, so long as it is bound inside this constituent, which it is, by the operator (see chapter 3 for the statement of this problem and its solution, applied to antecedent-contained VP-deletion sentences). And, again, concerning the "improper" relation between the head NP, and the variable, this relation is not realized until LF, where the structure has been changed in such a manner that the NP is no longer the only possible antecedent for the variable.

In theories of ellipsis in which constituents are effectively missing in syntax, as in Napoli (1985a), (1985b), these structures would not contain any empty category in syntax, in the object position of the epistemic verbs. Assuming that the Projection principle could be implemented to take care of such cases, then this would be a viable analysis. However, there are two arguments in favor of the existence of this empty category. The first one is that the tough constructions in question behave like normal tough contractions, in that the position of
the empty category cannot be embedded: 31

(61) *Paul a aimé un film difficile à faire croire
    Paul liked a movie hard to make believe

If the ellipsis analysis is combined with an insertion analysis at LF of the missing material, here the S', then it would be possible to explain this lack of embedding, when the structure is restored. Otherwise, the lack of embedding would have to be explained independently of the operator-variable relation in a non-stipulatary way.

The second argument in favor of the empty category, syntactically defined as a variable is that, if it is true that all variables need Case, then we expect the construction to be impossible when the epistemic predicate is not a Case assigner. That variables need Case can be seen as deriving from the Visibility principle (see Chomsky, 1981), which requires that all chains need Case, or as being a condition on variables themselves, as suggested in Chomsky (1981), or Massam (1985). 32 Anyway, for our purposes, given that the operator itself is not in a Case-marked position, and given the Case Visibility condition, which requires that the chain (Op, e) bears Case, it follows that the variable itself must be in a Case-marked position. Consider (62):

(62) *Paul est allé voir un film facile à être surpris
    'Paul went to see a movie easy to be surprised'

31. In French, tough constructions do not admit any type of embedding.

32. See, however, Borer (1980), for an interesting challenge to this claim, based on the behavior of free relatives in Hebrew.
Note that to be surprised is of the right semantic class: it takes an exclamative complement and it involves a cognizer. So, (62) is not excluded because of ill-chosen lexical items. (62) can be straightforwardly excluded with the postulation of an empty category complement of être surpris. Since this predicate is not a Case-assigner, which is evidenced by the ill-formedness of (63), then the chain (Op, e) violates the Case Visibility condition:

(63) *Je suis surpris la musique
     'I am surprised the music'

There is another possible analysis of the interpretation of S' as the complement of the tough adjective, but we will reject it. This second solution is to obtain the right syntactic structure by insertion of lexical material into the position of the empty S'.

2.4.2 Exclamatory predicates as operators

Our argument will have the following structure. First, we will show that it could be possible to derive a proper interpretation of the missing S', in a sentence containing a tough predicate, like (61) (Paul a aimé un film difficile à croire) by lexical insertion of the implicit sentence governed by croire. Then, we will invoke the same analysis for relative clause constructions, since they are equivalent in the relevant respects. With relative clauses, predicates of various semantic types allow an elliptical sentence as their argument. However, this does not hold when the gap S' is in a subject position. In those cases, only exclamatory predicates allow this gap. We will explain this limitation
by showing that the insertion analysis gives an ill-formed configuration in sentences with a subject gap. Hence, in those cases, only the operator-analysis of the relative clause can work. The confinement of the subject gap to exclamatory predicates will be taken as proof that only exclamatory predicates can behave like operators. Then, returning to the *tough* predicates, since those too are confined to exclamatory predicates in the constructions under consideration, this will show that only the first analysis is at play, and thus that the insertion analysis is flawed. Making a parallel with the relative clauses with subject gaps, which are excluded by the CED (Kayne's ECP), we will suggest that the *tough* constructions are also excluded by the CED. But there might, however, be another explanation for the impossibility of the insertion analysis with *tough* predicates.

We have seen that the modifiers that could behave like operators at LF were APs. Other modifiers which can do so are relative clauses, the predicate of which is a verb which takes an exclamative complement, again, as in (64):

(64)a. Paul aime un livre qui me surprend beaucoup

'Paul likes a book which quite surprises me'

b. Paul aime un livre que je trouve surprenant

'Paul likes book that I find surprising'

Syntactically, these relative clauses are more similar to *tough* adjectives than adjectives like *incroyable*, since they contain a variable, which is identified at LF as having the value of the S' and not the NP. These sentences mean: "Paul likes a book, and it surprises me that he likes this book"; and "Paul likes a book, and I find it
surprising that he likes this book." This means that these constructions should have the same syntactic analysis as the tough constructions. Following the analysis given in the previous section for the tough constructions, the NP is moved to S', the modifier S' extracted from it, and the original S' identifies the variable inside the relative clause, by governing the relative clause, as in the following LF, that of (64)b:

(65) \[ [\text{que}_1 \text{ je trouve } [\text{e}_1 \text{ surprenant}]_j[[\text{un livre } t_j],[\text{Paul aime } t_k]]] \]

Now, note that relative clauses are more lax with respect to the semantic class of their predicates. This predicate can not only take exclamative complements, as in (65)a-b above, but also declarative sentences, as in (66):

(66)a. Paul a bien lu le livre que tu m'avais dit
'Paul did read the book that you told me'

b. Paul n'a pas montré l'enthousiasme qu'il fallait
'Paul did not show the enthusiasm that (it) was necessary'

c. Pierrot est bien allé à l'endroit que tout-le-monde pensait
'Pierrot did go to the place that everybody thought'

d. Paul a vu un film qu'il regrette amèrement, Eraser Head
'Paul saw a movie that he bitterly regrets, Eraser Head'

e. Paul a bien accompagné la femme que Pierre se doutait à l'aéroport
'Paul did take to the airport the woman that Peter suspected'

In all these sentences, the understood argument of the existencial predicate can be the whole clause.

Given this, let us consider the alternative analysis. Instead of determining the original S' as the real argument of the modifier via
predication, one could assume that the gap is filled with the lexical material of its antecedent at LF, by copying, as in the analysis of VP-deletion proposed by Williams (1977). (See also Sag (1976), for a differing conception). In such a case, the modifier would not move out of the NP. The technicality of such a derivation will be explained in more detail in our analysis of antecedent-contained VP-deletion of chapter 3. For the moment, what is to be noted is that such insertion is made possible by the fact that the insertion applies inside a clause which contains an operator. This operator will eventually bind its proper variable at LF', after insertion of the material. In order to avoid an infinite regress, the NP is extracted out of its clause. In other words, the LFs of sentences like (66)a–e are structurally identical (with the difference that the NP is moved by QR here) to the S-structure of their counterpart with full relative clauses. Consider for example (67), the LF and LF' of (66)a, in which LF' is derived from LF by inserting the S' in the position of the variable:

(67) [[le livre que tu m'avais dit [e]_j]_j [Paul a bien lu t_j]]

[[le livre que tu m'avais dit [que j. a lu t_j]_j [J. a bien lu t_j]]

This LF' is well-formed, and it yields the right reading. So, now, it seems that we have two competing LF-treatments of such sentences. One in which the modifier is an operator and the other in which the empty structure gets filled in.34

33. As in May (forth.).

34. Note that the operator in COMP binds the newly inserted NP-variable, inside the relative clause, in (67). See the analysis of
Now, let us note the existence of a minimal pair:35

(68)a. Paul a aimé un livre qui m'a étonné
   'Paul liked a book that surprised me'

b. *Paul a aimé un livre qui m'a plu
   'Paul liked a book that pleased me'

The reason of the ungrammaticality of (68)b must reside in some ill-formedness at LF. Given that (68)a is grammatical, at least one of the two possible LFs, due to the existence of two possible analyses, is well-formed. So, we must determine which structure is ruled out in (68)b and why. Let us consider the LF under the first analysis, where the relative clause is treated like an operator:

(69) [[qui₁ t₁ m'a plu],ₗ[un livre tₗ j][Paul aime tₖ]],₁]

This is a well-formed structure: the relative clause is predicated of the sentence S'₁, which governs it, and the variable in subject position is properly connected to its antecedent, the complementizer qui. So this cannot be the faulty structure. Let us consider the other possible LF (in fact LF', after insertion), the one in which the sentence is filled in:

(70) LF: [[un livre qui [e₁ m'a plu]ₖ][Paul a lu tₖ]],₁]
    LF':[[un livre qui[k[que Paul a lu tₖ]ₖ m'a plu]ₖ[P. a lu tₖ]],₁]

antecedent-contained VP-deletion of chapter 3 for further details.

35. The reading of (68)b which is precluded is the one in which it is the fact that Paul liked the book that pleased me. Otherwise, irrelevantly, the book may be understood as the subject of pleased me.
Let us consider the NP-variable inside the relative clause. This variable is bound by the complementizer *qui*. However, this variable is bound from inside a subject, and this is a violation of Kayne's ECP: the gap is in a constituent, the subject $S'$, which does not percolate up to COMP, so the operator *qui* in COMP is outside the percolation projection of the variable. So, there is nothing to stipulate to exclude such an LF: it is ill-formed according to the ECP. The fact that such an LF is ill-formed indicates that, when the gap corresponds to a subject, there is only one possible LF, when the modifier is treated as an operator.

Now, consider again the fact that subject gaps are possible only with exclamative predicates, whereas non-subject gaps are possible with other types of predicates (see (66)a-e and (68)a-b above). What does this mean? Well, given the result just reached that there is only one well-formed LF with subject gaps, this means that this well-formed LF can be constructed only with exclamative predicates. Now, such LFs are constructed by treating the whole relative clause as an operator. And this is a marked phenomenon, as shown by the limited class of adjectives behaving like *incroyable*. The limitation is exactly reproduced for subject gaps. So, we conclude that what was claimed for *incroyable* extends to relative clauses: only when the predicate of the relative clause is an exclamatory predicate does the relative clause behave like an operator.

Now, the interesting problem is to define the relevant relation between a relative clause and the predicate contained in it, such that, if this predicate is exclamatory, then the relative clause as a whole is
assigned scope: remember that the modifier which is assigned scope is the whole relative clause, and not the predicate inside it. Take a concrete example again:

(71) Paul a aimé [un livre [qui t m'a surpris]]_NP

'Paul liked a book that surprised me'

It is because the predicate inside the relative clause (surprendre) takes an E complement that the relative clause as a whole (qui m'a surpris) can be assigned scope outside of the extracted NP. How is this possible? We have two possible choices. The first one is to identify the properties of the argument which is predicated by the relative clause with the properties of the variable linked to it inside the relative clause. Since this variable is the argument of some predicate, it seems possible to identify this predicate-argument relation with that of the relative clause and the argument that it is predicated of. So, the relation between surprendre and its subject is identified as that between the relative clause and the whole sentence. It is possible that relative clauses in general inherit the properties of the predicate that the variable is an argument of, and in particular, th-marking properties. Relative-clause constructions would thus be syntactic externalizations of th-roles, similar to the morphological processes described and discussed by Williams (1981b). In any case, it could be minimally assumed that the relative clause inherits the selectional properties of the predicate which th-marks the variable inside it. So, if this predicate selects for E, then the relative clause will also select for E. This will straightforwardly define the relative clause as
an exclamatory predicate. Then, the operator behavior of the relative clause would also be straightforward. This view is appealing, but there exists one empirical argument against it.

This empirical argument is based on the fact that the gap corresponding to the sentence cannot be embedded too far, as shown below:

(72)a. *Paul a lu un livre que je crois que tu trouves surprenant
   'Paul read a book that I think you find surprising'

   b. *Paul a lu un livre que je pensais qui t’avait surpris
   'Paul read a book that I thought (that) had surprised you'

Note that embedding in a small clause does not lead to ungrammaticality, as in (64)b, repeated here as (73):

(73) Paul a lu un livre que je trouve surprenant
   'Paul read a book that I find surprising'

Unless we claim that the inheritance of the selectional property of the verb by the relative clause is subject to some local condition, these facts are unexpected. Let us turn to the other alternative.

The other alternative is that it is the operator property which is inherited, automatically, given the definition of the scope of an operator. Predicates like \textit{surprendre} are operators. Now, an operator has the scope which is that of its GC or a maximal projection

\footnote{36. Alternatively, we could also analyze the empty category in subject position as an NP-trace, and not as a variable, making the construction similar to a SC with \(s'\) the predicate, as in chapter 1, section 8, but it is hard to see why it would not be a variable.}
immediately dominating its GC which does not also dominate a subject (cf. (43) and (44), chapter 1). The GC of the VP m'a surpris in (72) is the S immediately dominating it, since it is the first S which contains its governor (INFL) and a subject. So, S or S' should be assigned scope in (72). Hence, if the predicate is treated as an operator, then this operator carries with it the whole relative clause, which is the desired result.

Going back to the impossibility of getting a subject gap in an embedded tensed sentence, as in (72), it is due to the limited scope of the VP operator, which is the embedded S', and not the whole relative clause. Hence, there is no way to obtain a well-formed configuration for predication to apply. For example, the LF of (72)b will be derived from the S-structure with the arrow showing what the scope of the operator surpréndre is:

(74) Paul a lu [un livre [que je crois [qui t'a surpris]]]

In (74), the head of the relative clause is still the NP un livre, so the variable cannot be interpreted as the S', excluding such a reading.

We still have to show that the insertion analysis cannot yield a well-formed result for tough predicates, and may for relative clauses. The insertion analysis yields a bad result in the tough case, because, by inserting the clause, the empty category becomes too deeply embedded, a situation which is forbidden in tough constructions, as in (75):
(75)a. *Paul aime un livre difficile à prouver
   'Paul likes a book hard to prove'

   b. LF: [[un livre [difficile à prouver [e]_j]_j][Paul aime t_j]_j
      LF':[[un livre [difficile à prouver [que Paul aime t_j]]]_j
      [Paul aime t_j]]

As for the relative clauses, one more embedding is incoherent:

(76) ?Paul n'aime pas le livre que tu croyais avoir prouvé
   'Paul does not like the book that you believed yourself to have proven'

   LF':[[le livre que tu croyais avoir prouvé [que Paul aime t_j]]]_j
   [Paul aime t_j]]

2.4.3 Incredible

In English, as was clear from the translations, it is impossible for incredible to behave like an operator when it is associated with a head noun. However, there exist certain sentences in which incredible seems to behave like one, as in (77):

(77) John made an unbelievable mistake

(77) means that John made a certain mistake and that it is unbelievable of John that he made that mistake. Looking at all the constructions in which the AP is allowed, we find that it occurs in expressions, like, for example, make a mistake or run a race. The reason why this adjective is acceptable in such constructions is that the NP itself denotes an action. Given that the NP denotes an action, and that incredible can modify actions, as shown in (78), then the AP is a direct
th-marker in (77), which means that it can be maintained that these APs are not LF-operators in English:

(78)a. This was an incredible decision
   b. John's amazing departure left us voiceless

One other criterion that shows that the AP is directly th-marking the NP in these examples is that it is not necessary for the verb to enter in the meaning of the reason why the action is incredible. For example, (79) is as acceptable as the French sentences:

(79) John found an incredible mistake

In (79), John found a mistake which it is unbelievable that anyone would make. So, the result is that sentence (77) looks like a case of an AP operator, but the reading is obtained directly, as in "John made a mistake which it is unbelievable for anyone to make." So, the NP mistake does not semantically associate with the governing verb, otherwise (79) would mean that it is unbelievable that John found that mistake, which it does not.

So, in (77), the noun does not either associate with the verb, contrary to appearances. The question now is how the NP mistake denotes an action. We can assume that, when nouns occur in expressions with unmarked verbs of existence, like make, as in make a mistake, if the NP a mistake occurs alone, it can be interpreted as the action denoted by the whole expression when this NP occurs in its canonical context. The canonical context of a part of an expression, like N or NP, following Vergnaud (1982), is the verbal context which admits the N or NP, such
that the two form an idiom. So, \textit{make} - is the canonical context of \textit{mistake}. The position of the noun is indicated in the context, too. That the meaning of (77) is related to the meaning of the sentence when the NP is in its canonical context is strengthened by the fact that the degree of acceptability of the sentences in which the adjective enters this construction parallels that of sentences in which the S' is overt and contains this canonical context, as shown below:

(80)a. John laughed at a ridiculous mistake  
    b. John laughed at a ridiculous mistake to make  
    c. *John found a hard mistake  
    d. *John found a hard mistake to make

In chapter 3, we study in more detail constructions similar to these, with an empty category elliptical for the sentence or the VP that it occurs in, inside a modifier.
3.1 Introduction

In the preceding chapter, we investigated a set of modifiers, most of which are semantic modifiers of a constituent other than the NP which accompanies them on the surface. We saw that certain relative clauses contained an "elliptical" variable, whose antecedent is a constituent of the sentence which contains the NP. In the particular examples studied, this constituent was an S'. The proper antecedent was obtained at LF, either by extracting the relative clause to a position such that it became predicated of the right constituent, a possibility specific to exclamative predicates, or by filling the empty S' (the variable) with the content of its antecedent, thus obtaining a normal, non-elliptical relative clause at LF'. In this chapter, we look at deletion inside relative clauses, and by extension, inside comparative clauses, both constructions being a productive domain for deletion of various categories. Deletion can be expressed grammatically either as the presence of an empty category, or as true absence of material, but we will center our attention on deletion as empty categories. We will principally study a particular case of antecedent-contained deletion,
namely, VP-deletion inside a relativized NP contained in the antecedent VP. This construction will exemplify deletion as the presence of an empty category. The problems that will be posed and the results that will be found for these constructions extend to other constructions as well, like deletion of a whole S', instead of a VP, or deletion inside a comparative clause, instead of a relative clause.

3.2 Antecedent-contained VP-deletion

3.2.1 Introduction

(1) John talked to everyone that Peter did

Sentences with antecedent-contained VP-deletion such as (1) raise at least three questions. The first one relates to referential circularity: if the antecedent VP of (1) is compositionally dependent upon the anaphoric element, its interpretation threatens to be circular. These circular constructions may be analyzed as subject to a syntactic condition, rather than a condition on interpretation alone (cf. Halk (1985)). I will disprove this claim here, a task which involves dealing with the supporting evidence that was previously discovered. The second question is to describe the process of

1. I alluded to this problem in chapter 2, section 4, in relation to antecedent-contained deletion of S'. We did not call the empty S' "deletion of S'", but the structure is exactly parallel to what is called VP-deletion, which is in fact an empty VP. So, the term "deletion" will refer to the phenomenon.
interpretation of the missing element. There are two choices: either
the missing VP is directly interpreted as denoting what its antecedent
itself denotes, or it is interpreted off a level beyond LF -- LF' -- at
which it is syntactically represented as a copy of its antecedent.
Then, the denotation of its antecedent does not enter into account, only
its syntactic form. The second alternative can be tested if the copied
VP does not have exactly the same syntactic form as its antecedent. In
certain environments, this can happen, and in fact certain grammatical
conditions apply non-vacuously to the copied VP. whereas they apply
vacuously to the antecedent VP. This favors the second hypothesis.

3.3 VP-deletion as an empty VP

3.3.1 One argument: the condition of no vacuous quantification

One primary question about the missing VP in a sentence like (1),
repeated here, concerns its structure: 2

(1) John talked to everyone that Peter did

Two plausible hypotheses about deletion are either that there is no
material on the surface at all, with possible insertion at LF (cf.

2. Such relative clauses constitute inner islands, in Ross's (1984)
sense:

(1) *John talked to the only person Peter wouldn't

An inner island is an island created by certain operators, in certain
contexts, here the negation, in an amount relative.
Pesetsky's (1982) analysis of Gapping, and here, the analysis of bare XP comparatives in section 5, chapter 4), or that there is an empty category on the surface, following Williams' (1977). Real deletion, as in Sag (1976), seems implausible in the model of Chomsky (1981), because of the form of the grammar. Williams (1977) argues that the deleted VP corresponds to the LF, and not the S-structure, of the antecedent, which means that the level at which deletion applies is LF. However, LF is on the other side of PF from S-structure. Deletion of LF-material would be possible in a theory in which S-structure does not map onto LF by movement, or at least movement of lexical material. For example, it would be possible in van Riemsdijk and Williams' (1981) model, in which only indices of quantifiers are assigned scope, and not the phrases. In such a model, PF could properly have access to LF, and not necessarily S-structure.

There exist arguments in favor of the account of VP-deletion in terms of the presence of an empty category on the surface. We will briefly go back to the possibility that VP-deletion in (1) is true absence of material, and present some arguments of plausibility against it, in the next section. One argument in favor of an empty category on the surface is that the missing VP occurs inside a relative clause, and that relative clauses contain operators in COMP. Since this COMP is present

---

3. That is to say, the S' must have the form of a relative clause, since it has the interpretation of a relative clause. Such interpretations are derived through a predication rule involving an operator-variable relation. So, the interpretation of (1) must also involve an operator-variable relation, hence the operator-variable must be represented.
on the surface, the operator in it must also be present on the surface, given that lexical insertion is done only at D-structure.\(^4\) Note that this operator can be overt, as in sentences like (2):

(2) John parked his car where Mary did [e]

This is the first result about the syntax: COMP must contain an operator, so that the interpretation of the NP can be that of a relative clause. Then, the principle precluding vacuous quantification, which is one of the principles of Full Interpretation applying at LF, requires of all quantifiers that they bind at least one variable. We will extend the principle of Full Interpretation to all levels, and hence claim that this operator has to bind a variable at S-structure too.\(^5\) We also assume that, contrary to tous, which occurs in an adverbial position, hence a licensed position, elements in COMP are not in a licensed position, and hence must not only bind some variable, but must also form a chain with it:

(3) **Licensing of A'-positions**

An element in an A'-position must form a chain with some empty category.

So, if the COMP contains an operator in (1), or (3), this operator should bind a variable in syntax. If VP-deletion is to be expressed as

\[ \text{--------} \]

4. To be precise, insertion of material can be done later than D-structure — as for VP-deletion — but, as an assumption, only by a copying rule, which forces the inserted material to already be somewhere on the surface.

5. This conclusion goes against Napoli's (1985b) analysis, where she allows for vacuous quantification.
the presence of an empty VP in syntax, then the variable that the
operator in COMP must bind is provided: it will be the empty VP, or, if
the VP is itself an expanded empty structure, as in Williams (1977),
following Wasow (1972) the empty NP contained in that empty structure.

The question is to determine what empty category the operator binds:
an empty VP or an empty NP inside an empty structured VP, as shown
below:

(4)a. John talked to [everyone that Op<sub>j</sub> Peter did [VP<sub>e</sub>]<sub>j</sub>]

 binding

 b. John talked to [everyone that Op<sub>j</sub> Peter did [[NP<sub>e</sub>]e<sub>j</sub>]]

 binding

From these two possible structures, the most plausible one will have to
be eliminated, namely the one which duplicates an overt relative clause,
(4)b. Because if (4)b is to have the structure shown, then it means
that what is non-deleted in the sentence is the subject (Peter), the
auxiliary (did), and the NP object (the chain (Op,†)). Such a
postulated structure is thus the one found in sentences like (5):

(5) John saw Peter more often than he did Mary

Now, whatever the correct analysis of the missing verb in (5) is, such
sentences are not acceptable when the overt argument is, not a direct
object, but a prepositional object whose governing preposition is not

6. COMP contains the complementizer that and an empty [+wh] operator, as
in Chomsky (1977) or (1981).
overt. 7 For example, (6) is ill-formed, where the overt argument is the object of *about:*

(6)a. ??Mary talked about Peter more often than she did Bill

b. Mary talked about Peter more often that she did about Bill

If this is the case, that is, if, for some reason, such constructions do not allow the preposition to be silent, then, if the overt object is relativized, we should expect the preposition to have to be overt too, and hence we would expect VP-deletion inside the relative clause to leave an overt preposition. But this is impossible, as shown in (7):

(7) *Mary talked about everyone that Peter did [[e][about [Ne]]]"

Moreover, since silent prepositions are excluded in structures like those of (6)a, we expect a silent preposition to be excluded as well in a relative clause, if it is to have the same structure, but this does not prove to be true, as shown in (8):

(8) Mary talked about everyone that Peter did (silence = [p]

These facts indicate that, if the structural analysis of VP-deletion inside relative clauses (cf. ...everyone that Peter did of (1)) is to be maintained as similar to that of constructions like (5) (John saw

7. This is also the case with Gapping, where prepositional objects must occur accompanied with their preposition. Cf: John went with Mary and Peter ??(with)-Anny. (The grammaticality judgment outside the parenthesis is the judgment obtained if the sentence is without the parenthesized element.)
Peter more often than he did Mary), with the overt argument relativized in (1), then a stipulation has to be made concerning the relativization of prepositional objects, to the effect either that they cannot be relativized at all or that the preposition cannot be overt, to account for (7) and (8), which is quite unjustified. 8

To summarize, if we draw a parallel between (1) and (5), then it means that the empty VP of (1) is structured, the only difference between (5) and (1) being that the object is relativized in (1), and not in (5). However, if the VP is to be structured, then, if it contains a prepositional object, not only should relativization of this object leave an overt preposition behind, but also the sentence should not be acceptable with a silent preposition. However, sentences predicted to be acceptable are not, like (7), whereas those predicted to be unacceptable are acceptable, like (8).

We will take the impossibility for the empty VP to be a structured category in (7) to mean that empty structures are impossible in principle. The other possibility for the missing VP is thus to be a non-structured empty category, so the S-structure of (1) is (4)a, repeated here: 9

8. It could also be that (7) is excluded as a violation of a constraint on P-stranding, preventing it from occurring with an empty V, as suggested to me by Ken Hale, but I have not investigated this possibility.

9. This conclusion is the same when what is missing is a full S', as in (i)-(ii), discussed in chapter 2:

(1) Paul a parlé à la personne que tu m'avais dit  
    'Paul spoke to the person that you told me'
(4)a. John talked to everyone that Op_j Peter did [e]_j

One immediate question is the semantic status of the binding relationship between the operator and the VP variable. The operator is supposed to range over the set denoted by the head NP which is predicated by the relative clause and not over the denotations of VPs. It must be concluded that this anomaly disappears at LF", the level at which the empty VP is represented as a full VP, as a copy of its antecedent. This copying rule yields a structure identical to that of an overt relative clause, as in (9):

(9) John [vp talked to [everyone that Op Peter did [vp talk to [e]]]]

At that level only do we require that the operator bind an NP variable, because it is only to that level that rules of semantic interpretation have access. This means that we interpret the principle of no vacuous quantification in its fullest syntactic sense: at all levels, a quantifier has to bind a variable, no matter what category it is, so long as the meaningful operator-variable relation is restored at LF".

The only syntactic device that must be stipulated in order to allow the operator in COMP to bind a VP variable early in the derivation, and an NP variable at LF", is the change of referential index of that operator. If it does not change indices, then the NP object and the VP will end up coindexed. However, coindexing between a VP and an NP

(ii) Paul a parlé à tous ceux qu'il voulait
    'Paul spoke to everyone he wanted'

The empty S' object of wanted is a non-structured empty category.
signals Predication, in the sense of Williams (1980), where the NP is the external argument of the VP. Since this is not the case in the sentences in question, the NP object and the VP must have different indices.

Given the existence of the operator in COMP and the presence of the empty VP, there exists another a priori possible analysis of the process of interpretation of the empty VP. First, it cannot be that the empty VP is interpreted as denoting what its antecedent denotes, because of the syntactic properties of the constructions: the head of the construction is supposed to determine the value of the variable bound by the operator. So the VP is not free to be directly assigned the value of its antecedent: it has to be assigned this value via predication. This is where our former analysis of relative clauses with exclamative predicates could enter the game again: if the relative clause is extracted from the NP, then it could be assigned the proper antecedent, VP, by adjoining to INFL' (after adjunction of the NP itself to INFL'). Then the antecedent VP and the relative clause would c-command each other, they would thus be in a proper configuration for predication, and the empty VP would be assigned the value of the antecedent VP. This process would amount to relativization of a VP, since the VP would be the head of the relative clause, as shown in (10):

\[(10) \ \text{John} \ [\text{[Op, Peter did} [e]_j, [[\text{everyone} \ t_j, k, \text{talked to} \ t_k, w]], \text{VP}]]\]

This would be a perfectly valid derivation, yielding the right interpretation. However, the facts that we are going to deal with
(examples (20)a-b) show that VP-deletion inside relatives (or comparatives) involve copying the VP. This means that the assignment of a value to the empty VP is not done by LF-relativization of the VP, as suggested by the present hypothesis. So, such relativizations must be ruled out in principle. Remember that we allow LF-relativization of S', in sentences like (64) of chapter 2, like *Paul a aime un livre qui m'a surpris* 'Paul liked a book that surprised me. Relativization of VPs can be excluded by the following principle:

(11)  
\[
\text{LF-relativization} \\
\text{Subjects of predication may only be categories which can be arguments: NP and S' (or S).}
\]

3.3.2 The D-structure of sentences with A'-bound VPs

One other problem for this analysis of VP-deletion (or any deletion expressed as the presence of an empty category) inside a relative clause (or a comparative clause) is the D-structure of the sentences. Consider (1) again and (2), repeated here:

(1) John talked to everyone that Peter did
(2) John parked his car where Mary did

Where does the operator come from, when what is left behind is an empty VP? If such sentences are supposed to derive from movement of the operator into COMP, leaving a VP variable, then two major questions arise. The first one is related to the th-criterion: at D-structure, the wh-word is defined as an argument, but the position that it occurs in is a VP-position, not an argument position. It is thus contrary to
the requirement that all arguments receive a th-role in their
d-structure position, since it cannot receive a th-role in such a
position. In addition, these wh-words are not subcategorized for by
AGR, which itself subcategorizes only for VP (see Zagona (1982); Chomsky
(1981)), so they should not occur in this position at D-structure.
Supposing that such D-structures are in fact allowed, the second
question is the question of the transformation of the category of the
trace of the operator into a VP. To avoid these problems, let us assume
that operators are free to appear in A'-positions at D-structures.
Then, given that they must be licensed at S-structure by entering a
chain with an empty category, it will be the case that, at D-structure,
there is already an empty category which will serve as the variable
bound by the operator. In other words, the operator-variable
configuration in (1)-(2) is base-generated, with the chain between the
two formed at S-structure. What this theory requires is that wh-words
not be defined as r-expressions, so that they do not have to appear in
argument-positions at D-structure. And, if they are not r-expressions,
them they cannot appear in A-positions. This implies that the only
available positions at D-structure for such elements are A'-positions.

One possible problem for such an analysis of the derivation of
(1)-(2) is that one should be able to express the effects of subjacency,
a condition on movement. As (12) presumably shows, VP-deletion inside
relative clauses may not violate subjacency:

(12) *John met everyone that Peter wondered when he could

Since A'-chains must obey Kayne's ECP (which subsumes subjacency), and
since operators must be licensed by entering a chain, the relation
between the operator and the empty VP will have to obey subjacency.\textsuperscript{10}
One of the reasons for claiming that the operator and the VP form a
chain, subject to subjacency, is that subjacency subsumes the ECP, and
that the ECP can explain the distribution of VP-deletion across
languages, discussed in the next section.

To conclude, we analyze such sentences as \textit{John talked to everyone}
that Peter did as involving an A'-chain whose head is an operator in
COMP and whose tail is an empty VP. Let us consider some additional
justification for such a view, before we turn to another problem,
seemingly related to referential circularity.

3.3.3 Three more arguments for VP-deletion as an empty VP

If a deleted VP is in fact an empty VP, it should be the case that
this empty category is subject to the ECP, meaning that it should be
properly governed. Zagona (1982) proposes that VP-deletion is allowed
in English because VP is a properly governed position in this language.
The advantage of the account of VP-deletion in terms of the presence of
an empty category is that it permits to make a prediction. If this
proper-government property is subject to parameterization, this
phenomenon should not occur regularly cross-linguistically. This is in
fact the case. For example VP-deletion is not attested in French.

\textsuperscript{10} Remember that \textit{tous}, which is also an operator, does not have to form
a chain until LF, because its position is licensed, by being an
adverbial position.
Zagora (1982) argues that the node INFL is a proper governor of VP in English, given its verbal properties. If INFL is [+V], it is a lexical category, and hence it is eligible as a proper governor. Evidence for the verbal nature of INFL is provided by the fact that modals, which are verbal elements, can occur under INFL. In French, modals cannot occur under INFL, which suggests that INFL is not verbal, hence that it is not a lexical category, and hence that it is not eligible as a proper governor. According to the ECP as defined by Kayne (1981a), INFL plays the role of the antecedent of the empty VP, where it is coindexed with the VP in English and not in French, in the manner that the complementizer qui may be coindexed with the subject in French. (Note that the fact that INFL is not a coindexed governor in French seems to make the predication that wh movement from object position is impossible, since INFL is not a percolation projection of V.) Then we thus conclude that a matrix V is a percolation projection of an embedded V, following Kayne, and that this percolation can skip the INFL projections. In any case, VP still has to be coindexed with INFL, which must act as an antecedent, not a governor.) On the other hand, if VP-deletion is not to be analyzed in terms of an empty VP, then the difference between English and French remains to be explained.

---

11. One problem with this claim, pointed out to me by Ken Hale, is the question of the position of the auxiliaries être and avoir 'be' and 'have'. These auxiliaries are used for compound tenses, as in (i):

(i) Je n'ai pas encore fini cette these
    'I haven't yet finished this dissertation'

If these are main verbs, taking a Participle Phrase, then they do not have to be in INFL.
The second argument of plausibility for the empty-category analysis is that, as we have seen, it explains how an operator in COMP can obey the principle of no vacuous quantification. If deletion was true deletion, then the operator would violate this principle. One could counterargue that such a principle is relaxed in such sentences, or that the operator is not present in syntax yet, so the principle applies vacuously. A convincing argument in favor of our view must then be to provide a case of true deletion and show that such true deletion is impossible inside relative clauses. Pesetsky (1982) argues for Gapping as being a case of bare constituents in syntax and a number of authors, like Napoli (1983), analyze certain deletions inside comparative clauses as true absence of material in syntax, except for the bare constituent, as in (13)a-b:

(13)a. The phone rang twice and [the door bell] [three times] before I decided to get up
   
   b. Charles sent you a telegram to the same address as [the letter]

Since these display true deletion at S-structure, then if it is right that the operator is present in COMP, subject to the principle of no vacuous quantification, their relative-clause counterparts should be ill-formed, which they are:

(14)a. *Charles sent you a telegram to the address that Elmer his letter
   
   b. *Charles sent you a telegram to the address that the letter

The last argument in favor of the analysis where the operator is in COMP, binding the empty VP has been brought to my attention by Gilbert Rappaport. In English, operators can bind empty categories only, in
environments which do not violate Subjacency, which is to say that there are no resumptive pronouns. So, if the anaphoric VP is overt, then we expect the analogue of (1) to be excluded, which it is, on the same grounds that a sentence with a resumptive pronoun is:

(15) *John talked to everyone that Peter did so

This shown, let us now turn to the LF of sentences like (1).

3.3.4 LF and LF'

In a sentence like (1), John talked to everyone that Peter did [e], if the empty VP is interpreted in situ, then the application of the VP-rule will be infinite, since the content of the empty VP contains itself. This problem can be avoided, if, as implicit in Sag (1976) and Williams (1977) and effectively done in May (to appear), the relativized NP is extracted out of the VP by Quantifier Raising (QR), which then carries with it the empty VP out of its antecedent.\textsuperscript{12} The LF of (1)

\textsuperscript{12} There is another technical way to avoid the infinite regress, which does not involve extraction and allows empty categories to be created in the copying process: the copying rule may be allowed to stop copying at a maximal projection. Then, this maximal projection has to be interpreted, in the new structure. A maximal projection is a projection of a head, so the maximal projection thus copied should be interpreted as a projection of a head, and since it does not dominate any lexical material, an empty head is automatically provided. In other words, if, instead of copying the whole category, say, the book, the copying rule stops at NP, then NP is interpreted as an empty NP category. This device allows for logical forms without QR. Note that the obtention of an empty category in this manner is no different than the obtention of an empty category by the Projection principle: this principle requires that maximal projections (which are the categories selected by predicates) be present in certain positions. It is only because maximal projections are, by definition, projections of some head, that they are in fact empty categories.
looks like (16):

(16) [[[everyone that Peter did [e]]][John talked to \textit{t}_1]]

At LF', the empty VP is represented as a copy of its antecedent, as in (17):

(17) [[[everyone that Peter did [talk to \textit{t}_1]][John talked to \textit{t}_1]]

As an aside, let us note a generalization. What is striking in almost all circular constructions is that they occur practically exclusively in constructions which, in their overt counterparts, involve an empty category-variable, as in relative clauses and comparative clauses. This is because, in order not to be circular, the construction must involve LF extraction of an argument of the VP out of the VP. Such an extraction leaves a trace, and this trace is copied inside the empty VP. This empty category is coindexed with the containing constituent, which threatens to violate the i-within-i condition, if the empty category is not internally bound to this constituent (see section 3.3.6 on this formulation of the condition). So, it has to be bound at that level, and one immediate type of well-formed binding with respect to the th-criterion is binding by an operator. Hence the extracted constituent must contain a potential binder for that variable, a good candidate for it being the operator found in the COMPs of relative clauses and comparative clauses. In

---

13. The only other environment for licit circular constructions, in which no variable is involved, is the environment in which a constituent internally A-binds the anaphoric element, comparable to [[\textit{the man}]\_next to \textit{his}\_dog].
order to see how this holds, let us consider a complex NP which does not involve an operator-variable relation, like the fact that...

(18) *John hates the fact that Mary does [e]

(18) is ruled out under the complement-clause reading, because, at LF', the empty category created by movement of the NP out of the VP is not bound by anything, on the adjoined side, as shown below. Since it is coindexed with the containing constituent, the structure is referentially circular (see section 3.3.6). Also, empty categories have to be bound to be licensed, as we will see in chapter 4, so (18) also violates the licensing condition on empty categories.14

(19) [[the fact that Mary does [hate t₁]]₁[John hates t₁]]

Up to now, we have shown that the derivation of sentences like (1) involves binding of the empty VP by the operator in COMP in syntax, and QR of the relativized NP at LF, followed by the copying of the antecedent VP onto the empty position at LF'. Such derivations are quite unproblematic, the only real stipulation being the possibility for an operator to bind different variables at different levels. Other similar sentences have other interesting properties and we can now turn to these. I will present one unexpected contrast, and the explanation given in Haik (1985). This explanation will then be dismissed in favor of another one, which will use the possibility for a bound pronoun to behave like an operator, namely, to make the constituent that contains

---

14. Thanks to J. Higginbotham for useful comments on this example.
it a quantified phrase.

3.3.5 An unexpected contrast

In the preceding section, we considered sentences in which the NP-variable bound by the operator occurs at LF\(^{14}\), after the VP has been copied into the empty slot. One may wonder what happens when this variable is already present on the surface.\(^{15}\) Is VP-deletion still possible? The answer, interestingly, is, yes and no, as shown below:\(^{16}\)

(20)a. Robert talked to everyone who wanted him to
   b. *Robert talked to everyone who wanted Albert to

(20)a necessarily involves coreference of Robert and him.\(^{17}\) In Haik (1985), this contrast is explained in terms of a variant of the i-within-i condition. There, (20)b is said to violate this condition and (20) not, owing to binding of the empty VP by its subject. To make a long story shorter the analysis is the following (from 3.3.6 to

\________

15. Or when no operator occurs at all in the structure, as in:
   (i) John remembered people surprised that he did
   However, for lack of space, we will not consider these cases.

16. Haj Ross noted the following contrast:
   (i) I told Robert about everyone who wanted me to\(*him to
   This shows that subjects \*him are visible for the coreference effects. See the discussion in 3.4.6.

17. Stress on Albert helps make (20)b a little more comfortable, but my informants still do not accept it. I take this as meaning that (20)b has an ungrammatical status, even with stress.
3.3.6 A previous analysis

Consider the ill-formed (21), borrowed from Brody (1982) and Postal (1972):

(21) *John \[1 \text{ wanted to } [e]_1\]

(21) is a circular construction, to use Brody's (1982) terms. One view about circular constructions is that their exclusion is strictly syntactic, without reference to any consideration about their interpretation, and the other is that circular constructions are excluded for the reason that it is impossible to assign them any interpretation, given the infinite regress that they give rise to. One syntactic treatment of circular constructions is Chomsky's (1981) i-within-i condition, which prohibits structures in which an anaphoric element indexed i occurs inside its antecedent, but it is tempting to see such a condition as a global constraint, predicting the impossibility of assigning an interpretation to constructions of this form. And this would keep intact the view that circular constructions are excluded because of their interpretive properties. However, in the previous analysis, the claim was that a sentence like (20)a (*Robert talked to everyone who wanted Albert to) was excluded as circular at S-structure on strictly syntactic grounds, the LF and the interpretation being perfectly well-formed and acceptable.

As a descriptive generalization, referential circularity is blocked
if, in an XP of the form \([X_{pi} \ldots x_j]\), where \(x_j\) is an empty category or a pronominal to be construed as referentially dependent upon \(X_{pi}\), \((i \text{ and } j \text{ may be equal})\), the anaphoric element \(x\) is bound somehow inside \(X_{pi}\). Referential dependency is annotated in some way in the structure before the interpretive rules apply, indicating what the interpretation is to be, and the principle of referential dependency, applying at S-structure, is the following:

(6) **Principle of Referential Dependency:**

In \([X_{pi} \ldots x_j \ldots]\), if \(x_j\) is to be referentially dependent upon some antecedent \(X_pj\), then \(X_{pi}\) is annotated as referentially dependent upon the antecedent, unless \(x_j\) is bound inside \(X_{pi}\).

Elements which are necessarily dependent are empty categories and lexical anaphors, and those which are optionally dependent are pronouns (and epithets). Circularity is excluded by borrowing Higginbotham’s (1983) condition that referential dependency is transitive but not reflexive:

(22) **Exclude a structure in which a constituent \(X\) is annotated as referentially dependent on itself.**

All types of binding are relevant for the application of the principle of referential dependency (22).\(^{18}\)

---

18. For example, direct binding (cf. the contrast between *a picture of its frame and the man next to his dog*); indirect binding (cf. the contrast between *her brother likes his sister and Bach-Peters sentences*); and A'-binding (cf. the well-formedness of the artist that Ruth prefers).
Now, returning to the antecedent-contained VP-deletion sentences, all of them are potential violations of filter (22), if the anaphoric element is not bound internally to its antecedent. As we saw, for a sentence like (1), the empty VP is internally bound to its antecedent by the operator in COMP, which saves the construction from filter (22):

(1) John [talked to everyone Op₁ Peter did [e]₁]₁

But, as far as the definition of the circularity problem is concerned, it is not clear whether the binding of the empty VP by the operator at S-structure is a saving device from circularity, since this binding relation is independently forced by the Principle of no vacuous quantification. It is thus not clear yet whether the saving device on circular constructions is strictly syntactic. As we saw for the derivation of (1), the construction at LF is saved from the infinite regress by extracting the NP cut of the antecedent VP by QR, and at LF' the structure becomes that of a normal relative clause, which is not circular, owing to A'-binding of the variable and the particular way of interpreting such a relation. So it does not seem that the binding of the empty VP internally to its antecedent by the operator in COMP is required with respect to the circularity problem, and at the same time it is still not clear whether circularity should be checked at S-structure.

But consider (20)b again:

(20)b. *Robert talked to everyone who [e] wanted Albert to [e]

In (20)b, there are two empty categories for one operator. If the empty
category in subject position is not bound by the operator, then it is not bound by anything and thus could only be PRO. However, it occurs inside a governed position, which is not permitted for PRO. Hence it should be the variable bound by who. Then the question of circularity arises with respect to the empty VP, which is free inside its antecedent: the large VP is indicated as dependent on the antecedent of the empty VP, by principle (22), namely itself, a violation of filter (23). If we assume that the circularity conditions are conditions on S-structure representations as well as LF, then (20)b can be ruled out at S-structure.

Note that the strange character of this analysis is that such constructions are saved from the infinite regress at LF, since after QR of the relativized NP, the empty VP is not inside its antecedent, as shown in (23):

(23) [[everyone who [e]₁ wanted Albert to [e]]₁[Robert talked to t₁]]

Representing the empty VP as a copy of its antecedent does not seem to lead to any obvious ill-formedness, as shown in the LF' of this sentence:

(24) [[everyone who [e]₁ wanted Albert to [talk to t₁]]

[Robert talked to t₁]]

Except that one may note the peculiarity of such a sentence, which is that the empty category which appears inside the copied structure can no longer be a variable, because of principle C. Rather, it must be a pronoun, if the structure is to be acceptable at all, since it is
A-bound by the wh trace in subject position in (24). I rejected the possibility of analyzing the ungrammaticality of (24) as being due to the presence of small pro in LF', (that is to say, an empty category which is not in a chain, with a th-role of its own). One could assume that, if English does not have pro on the surface, then it would be expected not to have it at LF'. Alternatively, (24) could also be analyzed as a violation of the condition of identity up to alphabetic variance, proposed by Sag (1976) and Williams (1977) as a condition on VP-deletion: if the antecedent VP contains a variable, and this variable turns into a pronoun in the anaphoric VP, then this condition is violated, since the VPs are not identical. I rejected these two possibilities because of the existence of the well-formed (25), structurally identical to (24), in which the subject of the empty VP is coreferential with the subject of the antecedent VP:

(25) Robert talked to everyone who wanted him to [e]

It seemed that (20)b and (25) (i.e. (20)a) would have to share the same grammatical status, if the ungrammaticality of (20)b was due to the presence of the empty pronominal at LF'. (20)b was thus ruled out by the condition that a construction cannot be marked as circular at S-structure, namely that an antecedent cannot contain an element anaphoric to it which is not internally bound. And the conclusion was thus that certain sentences, which are perfectly acceptable at LF and LF' by escaping from the infinite regress, are nevertheless excluded at S-structure, as violations of a variant of the i-within-i condition. What remained to be explained was (25)'s acceptability.
The minimal difference between (20)a and b is the reference of the subject of the VP. Adopting Williams' (1980) theory of Predication, there exists a formal relation between an external argument of a predicate and the predicate, which is established syntactically as coindexing between the two. My assumption was that this relation may be (but is not necessarily) a binding relation. Moreover, the binding relation between subjects and their predicates in such sentences has all the properties of anaphor binding, as we will see.

One indication that this relation is a binding relation is the following. Supposing that, in parallel to the requirement which holds of indirect binding, if the relation of binding between a predicate and its subject is determined once, it has to hold for all occurrences of the VP, as stated in (26): 19

(26) Condition on bound elements

If some occurrence of X is bound by some occurrence of Y without being in a chain with it, then all occurrences of X must be bound by some occurrence of Y.

X and Y should not be in the same chain, because otherwise this condition would be too strong: it would require, for example, that relativized NPs be bound by an occurrence of the operator in COMP, which binds from COMP the variable coindexed with them, as in

.....[1 the artist that Op₁ Ruth prefers t₁].

binding required by some occurrence of Op₁

19. Thanks to Kyle Johnson and Diane Massam, for conversations that clarified the issues.
Going back to (20)a and b, *Robert talked to everyone who wanted him to and Robert talked to everyone who wanted Albert to, owing to referential circularity, the empty VP must be bound by something inside its antecedent, at S-structure. Its subject is a good candidate as its binder. So, in order for the structure to be well-formed, the subject him has to bind its VP, [e]. Now, condition (26) on bound elements forces all occurrences of VP_i to be bound by some occurrence of NP_j. The other occurrence of VP_i is the antecedent VP. Hence this VP must be bound by NP_j. This forces the two subjects to be coreferential, explaining why (20)a is grammatical and (20)b ungrammatical.

This explained why the two subjects must be coreferential in such constructions. Considering now what type of binding this binding relation can be compared with, where "types of binding" distinguishes between anaphoric binding on the one hand, and pronominal binding on the other, one finds that VP-binding is of the anaphoric type.

3.3.7 VP-Binding: the relation

The characteristics of the binding relation of anaphors is that they do not allow split antecedents (cf. Higginbotham (1983)), whereas the binding of pronouns does. Moreover, the binding of pronouns has to make operative the reference of NPs, as shown by the Disjoint Reference condition discussed in Lasnik (1976) and by some cases of indirect binding, but the binding of anaphors cannot. We consider these cases below.
First, let us consider split antecedents. In (27)a-b, the anaphors cannot be linked to the two separate NPs, whereas the pronoun of (30) can (the notation below in terms of sets of indices is informal and means what is obvious):

(27) a. *John\textsubscript{i} spoke to Mary\textsubscript{j} about each other\textsubscript{i,j}  
b. *John\textsubscript{i} spoke to Mary\textsubscript{j} about themselves\textsubscript{i,j}

(28) John\textsubscript{i} laughed with Mary\textsubscript{j} after they\textsubscript{i,j} left the class

Second, consider examples ruled out by the Disjoint Reference condition and examples where it is impossible to rule anaphors in accordingly:

(29) a. *We\textsubscript{i,j,...} spoke to me\textsubscript{i} about Mary  
b. *We\textsubscript{i,j,...} spoke to myself\textsubscript{i} about Mary

In (29)a, the pronoun is excluded, because it overlaps in reference with we, which binds it in its GC. In (29)b, even though we overlaps in reference with the anaphor, the anaphor cannot be considered bound by it, even though it occurs in the right domain. These two sentences show that inclusion of reference is not sufficient for anaphoric binding, and that pronominal binding must have access to it.

To illustrate how pronominal binding would work for VP-binding, let us consider sentences which show that indirect binding is at the pronominal type, where two indirect binders are allowed to bind an indefinite and a pronoun ccoreferential with the indefinite, even though the indirect binders are not coindexed, but simply overlap in
reference. Consider (30):

(30) Everyone who owns a donkey likes it, but none takes good care of it

In (30), the NP a donkey and the pronouns coreferential with it behave like variables bound by the wide scope NP everyone. In such a case, all occurrences of this variable (and in particular the two pronouns) must be bound by some occurrence of the wide scope NP, by condition (27) on bound elements. In the first clause, the pronoun is bound by the subject NP everyone who owns a donkey. What about the second pronoun? One relevant fact is that the second pronoun may be interpreted as a variable dependent on the assignment of a value for none only if none is anaphoric to everyone, that is to say, if the domain of the variable that it ranges over is the same as the domain that everyone ranges over. This means that the relation of indirect binding is of the pronominal type, because it allows indirect binding to hold when the indirect binders overlap in reference.

With these facts in mind, we can determine whether the relation of VP-binding by its subject is similar to pronominal binding or to anaphoric binding. Consider (31), where them is anaphoric to Tom, Peter and Mary: 20

20. This sentence is not too bad, when the second NP is one of them:
(i) Tom, Peter and Mary talked to everyone who wanted one of them to .

I will leave this unexplained.
(31) *Tom, Peter and Mary talked to everyone who wanted her\some of them to

The fact that (31) is unacceptable, even though the NP some of them or
her denotes a subset of the set denoted by the matrix subject, shows
that the binding relation between a VP and its subject does not allow
inclusion of reference to be a sufficient state of affairs to make the
two VPs properly bound by their subjects. This minimally sets
VP-binding apart from indirect binding, as illustrated by the contrast
between (31) and (32). And more generally, this shows that the
VP-binding relation is not of the pronominal type.

As for constructing examples relevant for testing whether the binding
relation admits of split antecedents, consider right-node raising, as in
the following:21

(32) *Bill thought that he would and Peter that he should
[talk to everyone who wanted them i,j to]

In (32), the antecedent VP has two distinct subjects, he i and he j. So,
if the binding relation that we are testing is to hold between the VP
and its subjects, this relation sees the two subjects as split binders
of the VP. That (32) is unacceptable seems to show that split binding is

------

21. Judgments are delicate, but, if the embedded S' is what is
right-node raised, the sentence is, expectedly, more acceptable:

(1) John tried and Peter even managed [PRO to talk to everyone
who wanted them to]

This is because the anaphoric VP no longer has split subjects, but a
single one, PRO, and it is PRO which has split antecedents.
not allowed there, rendering more complete the parallelism between this binding relation and anaphoric binding.

One last test for the claim that the binding relation is to be identified with anaphoric binding can be the locality condition on this relation. If it obeys principle A of the binding theory, a full set of properties would identify it as anaphoric binding. Note, however, that if evidence shows that only subjects can bind VPs, this could be derived from the binding theory as well as Predication theory. In fact, these theories make exactly the same predictions with respect to the locality condition, since the only NPs which c-command VPs in their GCs are their subjects. So they may both be valid at the same time.

So, we have identified the binding relation between NPs and VPs as an anaphoric relation. Let us now see that the relation does obey principle A of the binding theory, (or does fall under Predication).

3.3.8 VPs as syntactic anaphors

To check whether non-subjects can bind VPs, let us compare sentences in which the NP which is supposed to bind the VP is not a subject in at least one of its occurrences with sentences in which this binder is a subject in both its instances: 22

22. Thanks to Peter Ludlow for a discussion on this question.
(33)a. *Mary$_1$ promised John$_j$ PRO$_i$ to visit everyone who wanted him$_j$ to
b. Mary$_1$ promised John$_j$ PRO$_i$ to visit everyone who wanted her$_i$ to
c. Mary$_1$ forced John$_j$ PRO$_j$ to visit everyone who wanted him$_j$ to
d. *Mary$_1$ forced John$_j$ PRO$_j$ to visit everyone who wanted her$_i$ to

Note that the ill-formed sentences are slightly improved when the
subject of the empty VP is stressed, but still seem to be unacceptable.

In all the acceptable sentences, the binder of the VP is its subject,
not only in the antecedent clause, but also in the anaphoric clause,
even though there are other NPs which c-command the VP. This correlates
with the claim that only subjects can bind VPs. As for the unacceptable
sentences, the only plausible reason why they would be ruled out is
that, of the two potential binders of the VP, namely NP$_1$ and NP$_j$, none
of them is the subject of the VP for both occurrences of the VP, but
only once, and this, apparently, is not enough. Now, these examples
also correlate with the assumption that the binding relation must obey
principle A of the Binding theory, since only their subjects bind VPs in
their GCs. There exists another argument which could support the idea
that the VP is actually bound by its subject, in the circular
constructions under study, in spite of the interference of some other
condition.

3.3.9 No binding by expletives

It is a fact that PRO cannot be an expletive, as shown in (34) (see
Safir (1982); Zubizarreta (1982); Davis (1984); Brody (1984) for
discussion)\footnote{23}

(34) *It seems that oil painting is easy without PRO being obvious
to everyone

This impossibility may indicate that the defect of the binding relation
is to be located on the expletive, rather than on PRO, so suppose that,
because of their non-referential status, expletives cannot bind an
element (unless the element is in the same chain as the expletive). If
this hypothesis is correct, then it can be tested whether subjects are
potential binders of their VPs. In the particular case of an expletive
subject, a circular construction should not be saved by binding of the
VP by it subject, if expletives are not proper binders. Relevant
examples are of the form of (35):

(35) *It turns out that Mary likes every artist who wants it to

Such an argument is not all convincing, because, in general, the empty
VP cannot take as its antecedent a verb which is situated one sentence
higher up. For example, in (36), the empty VP cannot easily be
interpreted as the matrix VP:

(36) John believes that Mary likes every artist who wants him to

Also, expletives may be subjects of empty VPs, so the explanation of the

\footnotetext{23}{(33) contrasts with a sentence in which the expletive is not a pure
expletive, but (possibly) occupies a th-marked position, as with
extraposed sentences:

(i) It amused Mary that I lost may way without really surprising
her that I did}
ill-formedness of (37) cannot reside there:

(37) I don't know if it will turn out that they will show "Tango" again, but I sure would like it to

The assumption of VP-binding is based on one single contrast, (20)a vs (20)b, namely *Robert visited everyone who wanted Albert to vs Robert visited everyone who wanted him to, with the additional support of contrasts like those of (35), but looking at other examples, they all still seem to correlate with the explanation in terms of binding of the VPs, due to circularity.

3.3.10 More facts

In the unacceptible sentences, the empty VP occurs in a relativized NP of the form: everyone who t wanted NP to [e]. We saw that it is because the operator binds an NP variable in syntax that the empty VP has to find another binder. In other words, the ill-formedness of the sentences correlates with the presence of this NP variable in syntax. Does this correlation always hold? Isn't the ungrammaticality of (20)b due, for example, to embedding of the empty VP? If embedding is relevant to the unacceptability of (20)b then a sentence like (38) should be as unacceptable:

(38) ?John visited everyone who Peter wanted to

It seems that (38) is not too bad, when the embedded subject is stressed. However, even when this subject is stressed in (20)b, the sentence sounds unacceptable:
(20)b *Robert visited everyone who wanted Albert to

Whether the embedded sentence is tensed or not, the facts are the same:

(39)a. *Tom talked to everyone who t believed Mary should [e]
   b. Tom talked to everyone who t believed he should [e]

And the empty VP can be embedded, so long as its subject binds it:

(40) Tom visited everyone who Peter thinks he should

Another type of facts was considered a last piece of evidence in favor of the VP-binding account. Given our assumptions, we expect VP-deletion inside a relative clause to be acceptable if the empty VP does not occur inside its antecedent. This is verified in (41):

(41) Everyone who would have preferred Albert to [e] [washed the dishes]

To summarize the paper, the results were the following: antecedent-contained VP deletion are circular constructions, where circularity is excluded at S-structure, which forces the empty VP to be bound by something inside the antecedent VP at S-structure. This empty VP is bound either by the operator in the COMP of the relative clause, or by its subject, when this operator already binds its proper variable. The VP-binding relation by the subject has all the characteristics of an anaphoric relation: it does not allow split antecedents, it does not allow binding by overlap in reference, it does not allow binding by expletives and it obeys condition A. I will henceforth present a different analysis, but all these properties will

- 195 -
have to be explained.

3.4 A new analysis of the strange contrast

3.4.1 Looking for the fault

Suppose that the contrast, repeated here as (42)a-b, is not to be explained in terms of the circularity problem:

(42)a. Robert talked to everyone who wanted him to [e]

b. *Robert talked to everyone who wanted Albert to [e]

It would be a nice result if it was not explained in terms of the circularity problem, since circularity should be a problem of interpretation, and not one of syntax only.24

We will be neutral at first as to the real reason for the coreference facts, the main claim for now being that it is due to the presence at LF of an empty category, which is not an NP-trace nor a variable, in a governed position. This makes it pro-like. As we said earlier, the presence of this pronominal inside the copied VP is a potential

24. In fact, Howard Lasnik pointed out to me that the VP-binding relation was a necessary condition, not a sufficient one, as shown by examples like (i), which are still ungrammatical even if saved by VP-binding at S-structure:

(1) *John [wanted PRO to [e]]

The structure is excluded at LF as a circular construction, because it leads to an infinite regress.
violation of two principles: the first one is that it violates the condition (however it is stated, but see section 3.6 for an account) which precludes empty pronouns on the surface. In that case, (44)b (Robert talked to everyone who wanted Albert to) is ruled out on the same grounds as the overt sentence with an empty pronoun in the second VP: *Robert talked to everyone who wanted Albert to talk to (pro). The second reason for rejecting (44)b is that it can be said to violate Sag's (1976) and Williams' (1977) condition on identity up to alphabetic variance which VP-deletion has to meet.25 This condition says that the antecedent predicate and the anaphoric predicate must be formally identical, save for the names of variables. Sentences like (42)b violate this condition, as we see below, in the LF' of (42)b:

(43)LF \[[\text{everyone who}_{j} t_{j} \text{ wanted Albert to} [\text{e}_{j}]]_{j} [\text{Robert [talked to} t_{j}]]_{i}\]
LF' \[[\text{everyone who}_{j} t_{j} \text{ wanted Albert to} [\text{talk to} t_{j}]]_{i}\]
[\text{Robert [talked to} t_{j}]]_{i}\]

Inside the antecedent VP, the empty category $t_{j}$ is a variable $A'$-bound by everyone.... However, in the left-hand side of $S$, inside the relative clause, this empty category turns into a pronoun. Syntactically speaking, this empty category is a pronoun, since it should not be a pure variable, forming a chain with the wh operator: if it did, the structure would violate strong crossover, with one r-expression (the variable in subject position) c-commanding a coindexed r-expression (the empty category under discussion). So, if the empty

25. That the existence of such sentences are problematic with respect to this condition is hinted at in Partee and Bach (1984).
category is to be acceptable at all in this structure, it should correspond to an empty pronounal. This makes the antecedent VP and the anaphoric VP formally distinct, since one contains a pure variable, and the other a pronounal. So, to start with, we will claim that (44)b is excluded for two possible reasons, one being the impossibility of creating empty pronounals at LF", and the other being that the two VPs are not alphabetic variants. We will then see that our account of the acceptability of (44)b (i.e. when the two subjects are coreferential) is compatible with these two accounts, which means that the two conditions have the same exact empirical coverage, and that they are both empirically adequate. The analyses should also be based on independently motivated conditions. Sag and Williams motivate the condition on alphabetic variance. However, is there any reason to believe that empty pronounals are excluded at LF? I will answer this last question in the affirmative, by showing that this statement explains parasitic-gap constructions. The conclusion for these VP-facts will thus be that, for the moment at least, two plausible conditions, each independently motivated, account for the facts. The technical reason why empty pronounals are excluded will be given in section 3.6. For now, we assume that some condition precludes their existence, which is enough to handle the examples under discussion.

Before proceeding, let us make sure that the newly copied empty category is not defined as something else than a pronounal, and in particular that it cannot be defined as a variable, locally A'-bound from some intervening COMP, so that we are sure that (44)b is excluded because of the existence of an empty pronounal.
3.4.2 The identification of the copied empty category

This other possibility, suggested to me by Juliette levin, is that this empty category is A'-bound by some operator in the COMP of the embedded sentence: COMP him to [visit t₁]. One, small, problem for the idea that an operator binds the trace $t_{\bar{1}}$ in the QRed NP of (43) is that this variable is A-bound by the wh variable in subject position. This is a potential violation of Principle C. Unless, as Chomsky (1981) suggests, variables have to be free only in the domain of their operator, and not necessarily outside of it, in which case the structure would be well-formed:

(44) [everyone who₁ t₁ wanted [Op₁ [him to visit t₁]]]

Apart from this question, it is relevant to ask whether there are structures in which this empty category appears but in which no operator could bind it. If the contrast is still there, then this will indicate that it has nothing to do with the presence of the intermediary COMP, and hence that, presumably, there is no operator in COMP binding the copied empty category. There are two environments in which a COMP cannot host an operator: first, when this COMP position disappears (S'-deletion), as in (45)a, and, second, when this COMP is already filled, as in (45)c:

(45)a. Tom visited everyone who expected [him to [e]]
   b. *Tom visited everyone who expected [Mary to [e]]
   c. Tom visited everyone who told him [where [PRO to [e]]]
d. *Tom visited everyone who told Mary where [PRO to [e]]

These contrasts duplicate those previously discussed. In neither of these sentences is it possible for the LF empty category to be part of a chain with who. So, the empty category appearing inside the relative clause is a pronoun, and the intermediary COMP should not be relevant. We will thus consider that there is no operator in the embedded COMP of sentences like (44)a-b, since the absence of the operator in (45)a-d does not alter the contrast.

Let us call the new analysis, in which the ill-formedness of (44)b above arises because the empty VP happens not to be an alphabetic variant of the antecedent, or because the copying rule creates an empty pronoun, theory B, and the former analysis, in which (44)b is excluded as a violation of the condition on circular constructions, theory A. And let us compare the empirical predictions made by the two, when the two subjects are not coreferential. Theory A predicts that a sentence will be excluded each time the operator in COMP will bind its variable on the surface. As for theory B, a sentence will be excluded each time pro appears, but pro appears each time it cannot be a pure variable, which is either when it is non-subjacent to the wh, or when the wh variable c-commands it, as an effect of principle C (strong

26. Our analysis will be that the sentences become acceptable if pro occurs inside a sentence which can be assigned scope. In chapter 4, we determine that some S's headed by who may be assigned scope. That (45)d is still bad indicates that indirect questions are not movable clauses.

27. We still have to explain why the sentences are not excluded when the subjects are coreferential. For the moment, we limit ourselves to the unacceptability of the sentences, hence to sentences in which the subjects are not coreferential.
crossover). So, both theories make the presence of the wh-trace a crucial factor in the reason of the ungrammaticality of a sentence. For theory A, it is the very presence of this wh-trace which is responsible for the ungrammaticality of the sentence, whereas, for theory B, it is also the fact that this wh-trace c-commands the empty VP. So, to distinguish between the two theories, one has to find a sentence in which the wh-trace is formed, but does not c-command the VP, or perhaps a sentence in which the wh-quantifier does not c-command the empty VP. (46) is a sentence corresponding to the latter description, but, in fact, the two theories have similar acceptability criteria, as we will see: 28

(46) *John talked to everyone whose friends wanted Peter to

The particularity of such a sentence is that, at LF', the copied empty category is neither A'-bound nor A-bound, so it must be a pronominal anyway. By being a pronominal, it is not a variant of the empty variable inside the original VP, so it should be excluded, according to theory B. And it should be excluded, under the second alternative of theory B, because it is an empty pronominal. This is the LF' of (46):

28. The sentence with coreferential subjects is deviant too, which is unexpected:

(i) ?John talked to everyone whose friends wanted him to

This is because the structure corresponds to a weak crossover violation, similarly to a sentence like (ii):

(ii) ?Everyone_i's friends wanted John to talk to him_i
(47) [[[everyone [whose \textsubscript{1} friends] \textsubscript{j \text{t}_j} wanted Peter to [talk to t\textsubscript{j}]]\textsubscript{k}]
\text{[John [talked to t\textsubscript{j}]]\textsubscript{k}}]

This shows that the two theories make exactly the same predictions: if the wh-quantifier does not c-command the empty category, this empty category is a pronominal, so it is excluded by theory B. It is also excluded by theory A, by not having the wh-quantifier bind the empty VP. Note however, that a sentence in which the wh-trace does not c-command the VP is acceptable. Consider (48):

(48) \text{[John kept [every newspaper that Bill read t before Peter did [e]]]}_{\text{NP}}

I will consider the judgment to be on the favorable side. Note that the temporal adjunct is to be understood as inside the NP, and not outside. I will delay the explanation of why (48) is acceptable until the discussion of parasitic gaps. For the moment, let us note that (48) is acceptable on the same grounds as an overt sentence, which involves a parasitic gap, like (49):

(49) [[[every newspaper that Op\textsubscript{1} Bill read t\textsubscript{1} [before Peter did [kept t\textsubscript{1}]]]
\text{[John kept t\textsubscript{1}]]}

It is not certain that a theory like A cannot account for the (relative) acceptability of such a sentence. If it is the case that the wh-trace does not c-command the VP, one can stipulate that the binding of the VP by the operator in COMP is not entirely free, but is restricted to local binding, to the effect that the relation can be set only if no other variable is on the way between the operator and the VP. In the situation
of (50), the S-structure would involve the operator in COMP binding the
wh-trace together with the empty VP, so theory A would also have an
account of the well-formedness of (50), since the VP would be bound.

So, the two theories are as descriptively adequate concerning the
environments in which a sentence threatens to be unacceptable. Let us
now turn to the explanation of the contrast, that is to say, to the
explanation of why sentences become acceptable when the subjects are
coreferential.

3.4.3 Coreferentiality explained

Let us repeat a well-formed sentence:

(50) John\_i talked to everyone who wanted him\_j to \[e\],

Under theory B, the counterpart of this sentence when the subjects are
not coreferential is excluded because of the status of the empty
category (the trace of QR), so we have to find why (50) does not give
rise to a violation of the two alternative conditions of theory B. The
answer is straightforward, given the already motivated principle that
bound pronouns have the behavior of operators. Suppose that we treat
the pronoun him as bound by John, and assign it an operator status.
Being operator-like, it makes the constituent that contains it be
assigned scope. This constituent is its GC (or some maximal projection
dominating it and not dominating a subject), which is the embedded S'.
So, at LF, the embedded clause contained in that NP is assigned scope,
via the syntactic treatment of the bound pronoun as an operator. This
derivation will allow the copied empty category to escape from the
c-command domain of the wh-trace, and to form a chain with the operator
in COMP, making it a variable, and no longer a pronominal, as shown in
the following logical-form derivation:

\[(51) \quad [\text{John}_j[[\text{everyone who}_{k}[\text{him}_j \text{ to } \text{[e]} \text{VP}_i \text{S}'_n[\text{t}_k \text{ wanted } t_n]]]]
\quad \text{NP}_k
\quad [\text{t}_j \text{ talked to } t_k \text{VP}_i]]
\quad [\text{John}_j[[\text{everyone who}_{k}[\text{talk to } t_k \text{VP}_i][t_k \text{ wanted } t_n]]]]
\quad [\text{t}_j \text{ talked to } t_k \text{VP}_i]]\]

Since the copied empty category is not a pronominal, but is a variable,
none of the conditions of theory B are violated: the VPs are alphabetic
variants, which satisfies the first condition. And the empty category
is not an empty pronominal, which satisfies the condition which excludes
empty pronominals. It may be a surprising result that movement of a
constituent out of the domain of a local binder should have such a
drastic effect. This effect is that the LF derivation does not create a
syntactic pronominal, but a variable. This variable is A'-bound by who
in the relative clause and by the whole quantified NP everyone...k,
which is non-distinct for the wh-quantifier, since both are coindexed
quantifiers. The two A'-relations are thus alphabetic variants of each
other, and the sentence is well-formed.\(^{29}\) Treating bound pronouns as if
they were operators is a theoretical innovation, and it should make

---

\(^{29}\) This analysis is contrary to the Bijection principle of Koopman and
Sportiche (1983). Koopman and Sportiche exclude weak crossover on the
grounds that an operator cannot locally bind more than one variable.
Here, I claim that this double A'-binding is what saves the sentence.
certain predictions. In section 3.5, I provide more evidence in favor of this treatment, and, in particular, we will see that some other types of sentences with VP-deletion allow sloppy identity of pronouns, owing to the movement of a whole clause. For the moment, let us consider the predictions made by this analysis, and let us compare this theory, theory B, with theory A, again.

3.4.4 Empty anaphors

One of the strongest results of theory A is that it predicts that the reason for the well-formedness of the sentences under consideration is coreference between the two subjects. No other grammatical function will do. This is because the binding relation is stated as anaphoric binding (or, because VP-binding is a subcase of Predication, that is to say, a relation parasitic on the already established Predication relation). Theory B relies neither on Predication nor on binding, it thus has to explain the two main properties of the well-formed sentences: first, only coreferential subjects save the sentences, and second, the two relations, between each subject and each VP, have the properties characteristic of anaphoric binding and not pronominal binding: the two subjects must be coreferential (overlap in reference does not count), and the empty VP cannot have split subjects. The first task is to explain why only subjects save the sentences, when they are coreferential. In order to do so we have to get one result first, namely, that the empty VP cannot give rise to an empty anaphor at LF'. Consider (52):
This sentence is supposed to mean Tom talks to everyone who talks to himself, but this reading is not available. There are two competing explanations of why this sentence cannot have this interpretation. The first one is that, whatever the true reason of the coreference effects between the two subjects is, it is a general fact that only coreferential subjects make these sentences good. Since John is not coreferential with everyone, the sentence cannot be acceptable. And note that one cannot construct a well-formed example in which the two subjects would be coreferential, since the relativized NP would be bound by the matrix subject, and hence would violate principle C. The second explanation is independent of the coreference requirement. Our theory of chain formation forces governed elements which are bound in their GC to form a chain with their binder. In the LF of (52), the empty category, which is to be interpreted as a reflexive, has not moved to INFL', and, therefore, it cannot be protected from a violation of the th-criterion. In such a theory, anaphors obtained at LF' by copying cannot be allowed, since they cannot move to a non-A-position (of course, under the assumption that LF' is obtained from LF by copying only): an empty anaphor must have a binder, but by the chain-formation condition, it must form a chain with its binder. This violates the th-criterion, since reflexives and their binders must have distinct
th-roles. This second explanation should remain a theoretical speculation, until we find empirical evidence which relies solely on it, and which cannot rely on the fact that the two subjects must be coreferential. But relevant examples which could distinguish between the two explanations are hard to come by, so we will simply assume that the second account is the right one, and that the ungrammaticality of (52) does not have anything to do with the fact that the subjects are not coreferential.

Cormack (1984) discusses an example similar to (52), and notes that it contrasts with a sentence which has the same meaning, as in (53)-(56)

(Cormack's (25)-(26)): 31

(53) *The barber shaves everyone who doesn't
(54) I shaved Bill, because he wouldn't

The particular interest of this contrast is that it indicates that it must be the case that the empty VP goes through a stage in which it is syntactically represented as a copy of its antecedent, and it argues in favor of a rule of accidental coreference. If the analysis to be presented here is right, then the two deleted VPs of (55)-(55) do not,

30. This implies that chains are also formed at LF'. after copying, which is expected, since the th-criterion and the Projection principle should hold at both levels of logical form.

31. Haj Ross notes that (52) is good if the empty VP is embedded under want, as in (i):

(i) The barber shaves everyone who doesn't want to
I will leave this as a problem.
in fact, have the same meaning, in one case the meaning of the VP is that of shave himself, and in the other case that of shave Bill. To solve the problem of the difference between (54) and (54), we must assume a theory which allows for accidental coreference. Let us assume the following:

(55) **Principle of accidental coreference**
A constituent X may have the same referent(s) as a constituent Y.

In that case, X and Y do not have to bear the same referential index. This principle is an interpretive principle, it allows the referents of two constituents to coincide. Then, the difference between (54) and (55) is that the subject of the empty VP is coindexed (marked as coreferential) with the object of the verb in (54), but not in (55), in which case we can assume that he is interpreted as coreferential with Bill according to the principle of accidental coreference. Consider the logical-form derivation of (54):

(56) **LF** 
\[
\text{[[everyone who_j t_j doesn't [e]_j]_j [the barber shaved t_j]]}
\]

 **LF'** 
\[
\text{[[everyone who_j t_j doesn't [shave t_j]_j [the barber shaved t_j]]}
\]

In the LF' above, coindexing between the wh-trace and the trace left by QR is necessary, given all the necessary chains: the QRed NP and its trace must form a chain. hence must be coindexed. Then, this QRed NP is a relativized NP, which then must be coindexed with the relative variable, by construction. By transitivity, the wh-trace is coindexed with the trace of QR, which is the source of the ill-formedness of (54). By contrast, in (55), the pronoun does not have to be coindexed.
with Bill. So, if the deleted VP is shave Bill, the LF' he shaves Bill does not involve coreference between he and Bill. By the accidental coreference principle, it is possible to interpret he as referring to Bill, and under that interpretation, the sentence is satisfied in the situation in which Bill shaves himself. 32

The conclusion of this section is that reflexives in an argument-position cannot appear as a result of the copying rule, owing to the theory of chain-formation developed in chapter 1. We now have to make a stipulation concerning the capacity of a bound pronoun to behave like an operator, before we consider the problem that we started with, namely, why subjects must be coreferential to save the antecedent-contained VP-deletion sentences.

3.4.5 Locality condition between a moved S' and the antecedent-trigger

One thing must be stipulated, in order to get a satisfying explanation of the coreference facts. We saw, in chapters 1 and 2, that certain elements which are interpreted in conjunction with some other elements, like the NP containing same and the "antecedent", that the two

32. Note that reflexives are not ruled out at LF', when they are copied from real reflexives, as in (i):

(i) John washed himself and Bill did not [e]

This is because the copying rule has access to the LF of the antecedent, where the anaphor has already moved. As for (55), it could also be the case that the empty VP is allowed to have the meaning of the intransitive VP, shave, instead of shave Bill, when Bill is the subject, as suggested to me by J. Higginbotham. This would avoid having to introduce the principle of accidental coreference.
are subject to the Same-Scope condition, in order to explain the locality effect which they seem to be subject to. Let us assume that the treatment of a bound pronoun as an operator has the effect of triggering the application of such a locality condition, between the antecedent of the pronoun and the quantified phrase which contains the pronoun. In other words, let us assume that the antecedent of the pronoun and the pronoun are subject to the Same-Scope condition:

\[(57) \text{Condition on bound pronouns as operators} \]
\[
\text{A bound pronoun may be treated as an operator only if it has the same scope as its binder.}
\]

The scope of an operator is the scope of its GC (or some higher constituent), so, technically, this has the implication that the antecedent and the maximal projection which immediately dominates the GC of the pronoun must be adjoined to the same S. This has as a consequence that the maximal distance between a bound pronoun and its binder is one clause, if it it to be treated as an operator: the pronoun cannot occur embedded twice.

So, the result of this section is that, owing to the stipulation above that the Same-Scope condition applies to the quantified phrase determined by a bound pronoun and the binder of the pronoun, bound pronouns and their binders cannot occur more than a clause apart. Now, let us turn to the coreference facts.

3.4.6 Coreference between subjects only

I repeat here a paradigm which was aimed at showing that only
subjects were relevant for the coreference facts. For theory B specifically, these facts show that, even if the pronoun could be bound by NPs with various grammatical functions, the sentence is good only if the binder is the subject. Consider (58):

(58)a. *Mary promised John to visit everyone who wanted him to
b. Mary promised John to visit everyone who wanted her to
c. Mary forced John to visit everyone who wanted him to
d. *Mary forced John to visit everyone who wanted her to

The question is why, given that the pronoun him or her is c-commanded by both NPs Mary and John, in all of these sentences, and thus could function as an operator, some of these sentences are unacceptable. The ungrammatical sentences are those in which the binder of the pronoun is not PRO. All of these examples can be explained with the requirement stated in the previous section, namely the requirement that the pronoun and the binder of the pronoun have the same scope. Effectively, this means that the binder of the pronoun should be in the same minimal S as the minimal S' which contains the pronoun. The sentences are well-formed only if the pronoun is treated as an operator. Now, the only NP which is in the same minimal S as the minimal S' which contains the bound pronoun is PRO. Hence only PRO can be the trigger of the operator behavior of the pronoun, and hence save the sentences at LF'. This result correlates with the facts: only when PRO and the bound pronouns are coindexed are the sentences acceptable.

This is only part of the empirical problems. We still have to deal with sentences in which more than one NP belongs to the minimal S'
which contains the minimal S' containing the bound pronouns. That is to say, sentences where more than one NP occur in the nearest clause dominating the clause of the pronoun. So, consider (59):

(59) John introduced Mary to everyone who wanted her to [e]$_1$$_1$

The particularity of such constructions is that, if we want the pronoun to be bound by another NP than the subject, then this NP has to belong to the VP. In that case, given that the VP gets copied, this NP will be c-commanded by the pronoun subject, a case of binding in one's GC. This creates an anaphoric relation at LF', a situation which the theory does not allow to arise. In other words, the copied Mary is responsible for the inacceptability of (59), since the LF is identical to, overtly, "John introduced Mary to everyone who wanted her to introduce herself (i.e. Mary) to them". The logical form derivation is the following:

(60) LF' [Mary$_j$ [everyone who [[her$_j$ to [e]$_{VPi}$]]$_1$$_1$ [t wanted t$_1$]]$_k$ [John [introduced t$_j$ to t$_k$]$_{VPi}$]]

As we see, at LF', the relative clause has a structure in which the trace of the NP Mary, which is moved by QR in order to make her a bound pronoun, is bound in its GC by the pronoun her. Note that we want to

33. Cf. also, Ross's problematic example, in footnote 16.

34. Remember that the embedded S' in the relative clause is adjoined to its S, in order to make the empty category a pure variable, a possibility allowed by the fact that the bound pronoun is treated like an operator.
coindex Mary with her, in order to treat this pronoun as an operator. Given our theory of chain formation, either the trace of the NP Mary forms a chain with the pronoun, leading to a violation of the un-criterion or it forms a chain with the quantifier Mary, in which case this is a strong crossover violation. 35

To conclude, there is a crucial interaction between the condition that bound pronouns cannot be separated by more than one clause from their antecedent and the fact that an NP cannot end up bound in its GC at LF'. This interaction yields the result that only an NP which is not inside the antecedent VP, but which belongs to the GC of this VP (so that no more than one clause intervenes between this NP and the pronoun, contained in a relativized NP inside the antecedent VP) can be the binder of the bound pronoun.

There is another situation in which an NP belongs to the GC of a VP, without being its subject, namely, when the NP is inside an S-adjunct, or is an S-adjunct itself. In these cases, this NP fulfills the requirement that it is not too far from the pronoun, and the requirement that it is not contained in the VP. However if it is inside an adjunct, unless it is the adjunct itself, this means that it is contained inside a maximal projection, the adjunct, and hence that it does not c-command the VP and everything that the VP contains, and in particular, the pronoun. So this NP cannot act as a binder of the pronoun, unless it is

35. If Mary did not move, as we will see in chapter 4 the structure is a violation of principle C, with Mary coindexed with a c-commanding pronominal.
the adjunct itself. To make these situations concrete, consider (61) first, which is a case in which the binder of the pronoun could be the NP inside the adjunct, were it not for the fact that it does not c-command it:

(61) *John [talked to everyone who wanted her to [e]$_i$]$_i$ in favor of Mary

(61) is bad, even if Mary is outside of the VP, since Mary does not c-command the pronoun and hence cannot bind it. The other relevant case is when this NP is not contained inside an adjunct, but is the adjunct itself. Also, another condition seems to exclude them, which is that VP-deletion in such cases does not seem to allow S-adjuncts. Consider (62), as an illustration:

(62) John [talked to everyone who wanted Mary to [e] then$_j$]$_j$ [yesterday]$_j$

This sentence could be defective for many reasons. One is that an adjunct like then is not acceptable, for some reason, when the deleted VP is an infinitival, as shown in (63)a-b, where (63)a is the counterpart of (62), which should be acceptable, given that the subjects are coreferential:

(63)a. *John [talked to everyone who wanted him to [e] then] yesterday

b. *Mary [went swimming] on Tuesday, and John wanted to [e] then

In these sentences, if the adverbial then is interpreted as an adjunct of want, then it is acceptable. But not when it is interpreted as an adjunct of the empty VP.

So, conspiratory principles and conditions yield the following
descriptive generalization: only coreferential subjects save sentences.

We now have to consider the remaining generalizations of theory A which were captured by stating that the binding relation between a subject and its VP is to be assimilated to anaphoric binding.

3.4.7 Binding of pronoun is similar to anaphoric binding

In this section, we will see that almost all the properties of the relation between a subject and its VP in theory A, which was identified as anaphoric binding, carry over to the properties of the binding relation between an NP and a pronoun bound by it. In other words, the relevant properties which distinguish anaphoric binding from pronominal coreference are also those of quantifier binding. The first property of the VP-binding facts was that the antecedent VP does not allow split antecedents. This was argued for by the use of a sentence in which the antecedent is right-node raised, as in (33), repeated here as (64).

However, the problem now is that such sentences are not conclusive for what we want to test anymore, since the pronoun should be c-commanded by its antecedent, which presumably it is not. Consider (64):

(64) *John thought that he, would and Peter that he should [visit everyone who wanted them, j to [e]]

According to theory B, the pronoun them, if it is a bound pronoun, should allow the structure to be well-formed. However, in order for a pronoun to be bound, it has to be c-commanded by its antecedent. (64) is a case where the VP has been right-node raised, and the status of the position of the pronoun depends on the analysis of right-node raising.
It seems that right-node raising consists in adjunction to the right of a sentence of a constituent which is interpreted as a complement of the clause as well as a complement of a conjunct clause. And an analysis where this complement appears in the form of an empty category in both conjuncts is appropriate. So, the schematic structure of (64) is (65):

(65) [[John thought that he would [e]_k and Peter that he should [e]_k] [visit everyone who wanted them to [e]_k] ]

Note that the extraposed VP must be adjoined to the matrix S, since it must be to the right of the conjunct and this conjunct is a conjunct of the matrix clause. This means that the pronoun them is not c-commanded at S-structure either by the pronouns he_i and he_j, or by the NP Peter. It is c-commanded by the NP John, but this is not enough, it should be bound by both NPs, if it is to be bound by two antecedents. So, the pronoun them cannot act as a bound pronoun in a right-node raising construction of this sort. I do not find any other type of construction which would be a good testing ground for the claim that split antecedents cannot save the sentences, so I will leave this question as still open.

The other property of the binding relation is that it is not saved by overlap in reference: pure coreference is needed, as in illustrated in (66):

36. However, see McCawley (1982), who argues for a representation with discontinuous constituents.
(66) *[Peter and Mary] [talked to everyone who wanted her to [e].]

The reason why this sentence can be excluded, under theory B, is that pronouns interpreted as bound variables must range over the full set of individual values that the antecedent has, by definition. So, if the antecedent is a plural, then the pronoun, if it is to be interpreted as a bound variable, cannot range over only a proper subset of the set denoted by the antecedent. So, as far as this property is concerned, it can be concluded that it is a property that anaphors share with bound pronouns. In turn, this makes theories A and B equivalent for their empirical coverage.

The last relevant property is that an expletive subject of the antecedent VP and of the empty VP cannot save the structure, as shown below:

(67) *It turns out that Mary likes every artist who would like it to

As we noted earlier (see section 3.3.9), this sentence is already marginal because the anaphoric VP has to find its antecedent into the matrix clause (the NP has to move to the matrix clause, violating condition A, which applies to QR). However, it is worse than another sentence which violates this condition but which has the two subjects coreferential, like (68):

(68) ??Peter thinks that Mary likes every artist who would like him to

This contrast is explained according to the plausible assumption given in theory A that expletives cannot be binders (when they are not in the
same chain as the bindee). So, the expletive cannot act as a bound variable, as is plausible, and the structure cannot be saved by treating the embedded S' as a quantified phrase.

This concludes the comparison between the two theories.

3.4.8 VP-deletion in embedded questions

To deal with the contrast in (46)c-d in section 3.4.2, I had to assume that indirect questions do not move at LF, as mentioned in footnote 25, since otherwise VP-deletion inside an indirect question should be acceptable, even if the two subjects are not coreferential:

(69) *Tom visited everyone who told Mary where [PRO to [e]]

We can consider another account of (69), which is that it is excluded as a violation of the crossing constraint, at LF', after lexical insertion of the VP. The crossing constraint rules out A'-relations which cross each other, in linear terms.37 Let us consider the LF' of (69):

37. See Kuno and Robinson (1972), Pesetsky (1982), and chapter 4.
We see that the two wh-relations cross each other, in (70). Sentences in which these two wh-relations would not cross are those in which the relativized NP corresponds to a less close argument of the verb than the question word. However, this means that the non-relativized NP (the NP which corresponds to the questioned argument) inside the relative clause is a close argument of the verb, and hence that its presence is obligatory also in the matrix. But then the VP to be copied would have this argument in addition to the questioned argument, which is one argument too much. Otherwise, relevant sentences should be of the form below, where *someone* should be allowed not to be copied in the anaphoric VP, to avoid having an additional argument:

(71) Tom [introduced someone to [everyone who t told Mary who she could [e]$_1$]$_1$

This sentence would be interpretable as: Tom introduced someone to everyone$_1$ who told Mary who she could introduce to them$_1$. This is not a
possible reading of (71), which means that the empty VP is anaphoric to the whole VP. In fact, (71) cannot have any reading, when [e] is construed as anaphoric to the containing VP.

In this chapter, specifically centered on VP-deletion, we have seen that either the condition that precludes empty pronominals or the condition on alphabetic variance on VP-deletion can explain the impossibility of antecedent-contained VP-deletion, when it gives rise to a pronominal, instead of a variable, at LF'. And, owing to the possibility of assigning scope to sentences by syntactically treating bound pronouns as operators, these same sentences are saved from a violation of the condition, explaining why these sentences are acceptable only when the two subjects are coreferential. Before turning to more deletion constructions, let us study in more detail the consequences of the claim that bound pronouns are like operators in their behavior.

3.5 Sloppy identity allowed by pronouns treated as operators

Concerning the geometry of the tree, the treatment of a bound pronoun as an operator allows a sentence (or, more generally, some category) to be assigned scope. Since a sentence contains formatives in it, it carries with it these formatives, hence allowing them, either to escape from the S-structure c-command domain of some element, or to be, directly, locally bound by some operator in an A'-position, or both, as was the case above. These are the three, and the only three
consequences of this LF movement.

Let us now consider other cases in which a bound pronoun is able to save a structure. A sentence can be saved by allowing an element to become properly A'-bound at LF, owing to the LF movement of the sentence. But let us consider cases in which movement of a clause allows a sloppy pronoun to escape from the c-command domain of an A-binder. By moving, this clause will be considered a \( \lambda \)-expression on its own, whereas if it does not move and stays embedded, the \( \lambda \)-expression comprises some matrix structure in addition, which will be bigger than its antecedent. This movement will save certain sloppy readings for pronouns.

Sag (1976) and Williams (1977) explain an interesting array of data according to one characteristics of the insertion rule (or deletion rule, as proposed by Sag) in place of the empty VP. This characteristics is that, when the antecedent VP contains a variable bound by an operator, if the operator is not contained in the antecedent VP, then the copied VP will contain a variable which, because the operator cannot be copied along, happens to become unbound in the resulting result. For example, Williams or Sag show that it is impossible for VP-deletion to apply to the embedded VP in a tough-construction, but possible to apply to the matrix VP, because tough-constructions contain variables which are bound from the COMP of the embedded sentence. For example, consider

---

38. Such cases, in which some element is inside a clause, too far away to be properly bound if the sentence does not move up, recall wh-islands, from which wh movement cannot escape, except in languages like Italian. For details, see Pesetsky (1984).
(72): 39

(72) *John is easy to please, but Bill is hard to [e]

(72) has the following LF and LF':

(73) LF. J. is easy \(O_p \) PRO to [please \(t_j\)] and B. is hard \(O_p \) PRO to [please \(t_j\)]

LF'. J. is easy \(O_p \) PRO to [please \(t_j\)] and B. is hard \(O_p \) PRO to [please \(t_j\)]

At LF', the copied \(t_j\) is not bound by a proper operator, which should be indexed \(j\). In Sag's or Williams' terms, this structure is excluded because the variable is unbound. Other examples, constructed in the same spirit, show the same thing, like sentences in which the variable which ends up free is a wh-trace as in (74)a, or a trace of QR, as in (74)b (when the quantifier which binds it has scope over the subject of the empty VP): 40

(74)a. *John, who Bill saw and who Bob did too

(Williams, (93))

b. *John has no horse, but Bill does

(Williams, (128))

--------

39. Ann Lobeck pointed out to me that sentences like (72) become acceptable when the embedded VP is negative, as in (i)

(i) John is easy to forget, but Bill is easy not to

I do not know why this should be so. See Lobeck (1984) for an account in terms of phonological conditions on cliticization of to.

40. See Williams' and Sag's work. See also Hirschbuhler (1982), who argues that a quantifier in the antecedent VP may in fact have scope over the subject.
Before we start, it has to be made clear that the domains which must be alphabetic variants of each other are the predicates, with all their arguments, namely, the thematic complex which is subject to deletion VP and its subject. However, we will see that it is better to say that it is the minimal S or S' which contains the VPs, that is to say, the GCs of the VPs or the maximal projection immediately dominating the GCs of the VPs. This is to say that VP deletion can apply whenever the minimal clauses which contain the two VPs are formally identical. The important point is that the domain cannot be bigger than that, explaining the impossibility of deleting the internal VP in a tough-constructions, etc:

(75) Condition on alphabetic variance
VP-deletion only applies to VPs whose GCs, or the maximal projection immediately dominating their GCs, are alphabetic variants.

The examples that show how the prediction works, concerning the pronouns as operators, are those when the variable is a pronoun with sloppy identity.41 Following the concensus on the sloppy reading of pronouns, we will assume that this reading arises when the binder of the pronoun is represented as a quantifier and the pronoun as a variable bound by it.42 Sloppy identity is the phenomenon whereby a pronoun does not have a fixed reference, but varies according as to what its linguistic antecedent is as in (76).

---


42. This traditional analysis is challenged in chapter 4:4.3.2. There, I propose another analysis of the facts of this section, still based on the operator property of bound pronouns.
(76) John saw his friend, and Bill did too

If Bill looked at his own friend, this is the sloppy reading. It is enough to represent the subject of the empty VP as \( \lambda \)-abstracted for the two S's to be alphabetic variants, as in (77):

(77) \[ x = \text{John} \lambda x \text{[saw } x \text{'s friend]}_1 \text{ and } y = \text{Bill} \lambda y [e]_1 \]

\[ x = \text{John} \lambda x \text{[saw } x \text{'s friend]}_1 \text{ and } y = \text{Bill} \lambda y \text{[saw } y \text{'s friend]}_1 \]

Given this, let us look at a sentence in which the relevant variable is a sloppy pronoun and its binder a c-commanding NP which is not the subject of the antecedent VP, as in (78):

(78) John told Susan that Bill likes her and (he told) Mary that Peter doesn't

The string in parenthesis can be omitted, yielding Gapping, in order to get a less crowded sentence. In this sentence, the pronoun her cannot have the sloppy reading. The LF' is the following, with the variable unbound, in the copied VP:

(79) \[ x = \text{Susan} \lambda x \text{(John told } x \text{ that Bill [likes } x\text{]) and } y = \text{Mary} \lambda y \text{(he told } y \text{ that Peter doesn't [like } x\text{])} \]

In Sag's and Williams' terms, this LF is ill-formed because the variable \( x \) in the copied VP [like \( x \)] is not bound by a quantifier, which it should be, if it is a variable. So, (79) is excluded because the sloppy pronoun in the copied VP should be bound by its antecedent, but it is not. This problem arises because its antecedent is not an argument of the minimal S' which contains the deleted VP, so the antecedent is not
inside the antecedent \( \lambda \)-expression.

VP-deletion can apply only if the two VPs belong to \( \lambda \)-expressions which are alphabetic variants of each other. Remember that we said that the \( \lambda \)-expression can be greater than a VP, it can be a whole \( S' \).

Consider (82) again: \textit{John told Suzan that Bill likes her and (he told) Mary that Peter doesn't}, where sloppy identity is impossible. This sentence is excluded because the deleted VP, which contains a variable, occurs in a \( \lambda \)-expression which is greater than the minimal \( S' \) which contains VP, and \( S' \) is the category subject to the condition on alphabetic variance. If the variable happens to occur inside a \( \lambda \)-expression which is just the minimal \( S' \) containing the VP, then it can end up properly bound. In the sentences in question, this can be made possible if the clause that contains the pronoun is defined as a \( \lambda \)-expression on its own. In that case, instead of there being only one \( \lambda \)-expression: \( x \lambda x \) \( (\text{told } x \text{ that NP likes } x) \), there would be two: \( x \lambda x \) \[ (\text{told } x) \) and \( (\text{that NP likes } x) \).

Now, a clause can be assigned scope if it contains a pronoun bound by a \( c \)-commanding NP. Given that the two \( \lambda \)-expressions must be variants of each other, it must be the case that the two clauses move, not one only, hence both clauses must contain a pronoun bound by a matrix NP. If these requirements are met, we expect sloppy identity to be possible, which it is, as shown below:

\[(80) \text{John, told Suzan that he} \_1 \text{likes her and (he told) Mary that he} \_1 \text{doesn't [e]}\]

The LF which allows the sloppy interpretation of the pronouns is the
following, where the embedded clauses are the relevant $\lambda$-expressions, which are identical, as we see:

(81)

Note that, as we said, one pronoun in only one clause is not enough to save the structure, since one of the clauses will be unable to form a small $\lambda$-expression, and this is correct:

(82a. *John$_i$ told Mary that he$_i$ likes her and (he told) Susan that Bill doesn't

b. *John$_i$ told Mary that Bill likes her and (he told) Susan that he$_i$ doesn't

Note that in our examples, the pronoun he subject of the second conjunct is coreferential with John, but, according to our principles, it should not make a difference if this pronoun is not coreferential with John, since this is not the pronoun which triggers scope assignment to the clause. Nevertheless, since John does not c-command the second conjunct, it cannot directly bind the operator he, so the intermediary coindexed pronoun is necessary, as shown in (83):
(83) *John, told Mary that he\textsubscript{i} likes her and Peter told Suzan that 
he\textsubscript{j} doesn't

The reason why this sentence is not acceptable if the subject of told in 
the second conjunct is not co-referential with John is that, if the 
second clause is to be assigned scope, it is because the pronoun inside 
it is bound, which entails c-command at S-structure. So, the pronoun 
should be c-commanded by an occurrence of NP\textsubscript{i} in the second conjunct.

Lastly, the two bound pronouns do not have to be bound by the same 
index, so long as both are treated as bound pronouns. In such a case, 
sentences should be acceptable even if the two subjects are not 
co-referential, so long as both bind a pronoun inside the clause, as in 
(84):

(84) John, told Mary that he\textsubscript{i} likes her and Peter\textsubscript{j} told Suzan 
that he\textsubscript{j} doesn't

This is indeed the case, so the facts fall under our generalizations.

We end this chapter by considering the exclusion in principle of empty 
pronominals.

3.5.1 No governed empty pronominals

In the preceding sections, we found two reasons for the 
ill-formedness of a sentence like: John talked to everyone who wanted 
Peter to, the first one being a violation of the condition of alphabetic 
variance on VP-deletion, and the second one the exclusion of governed 
empty pronominals. In this section, we exclude governed empty
pronominals in principle. Then, in the next chapter, we will devote some attention to parasitic gaps, which are empty pronominals, and hence should be excluded, according to this principle.

In chapter 1, we saw that we do not in principle rule out the existence of empty pronominals, so that PRO and pro are both permissible. According to Brody (to appear), governed empty pronominals are excluded because any governed empty category is a non-head of a chain which violates the th-criterion in case this empty category is pronominal-like, since pronominals must be thematically independent from their antecedents. In the present system, only empty anaphors bound in their GC are excluded this way. So, some principle should exclude pro and still allow PRO. Manzini (1983a) shows that control is determined as a binding condition on PRO, where PRO is taken to be an anaphor. She defines the domain in which anaphors have to be bound as their domain-governing category, which is, either, the GC of the anaphor, or the GC of the minimal maximal projection that dominates the anaphor, when the anaphor does not have a GC, as in (85):

(85) C-domain (definition) (Manzini, (1))
\[ \gamma \text{ is the c-domain of } \alpha \text{ iff } \]
\[ \gamma \text{ is the minimal maximal category dominating } \alpha. \]

(86) Domain-governing category (definition) (Manzini, (120))
\[ \gamma \text{ is a domain-governing category for } \alpha \text{ iff } \]
a. \[ \gamma \text{ is the minimal category with a subject containing the c-domain of } \alpha \text{ (and a governor for the c-domain of } \alpha), \]
b. \[ \gamma \text{ contains a subject accessible to } \alpha. \]

(87) Condition A (M, (70))
An anaphor is bound in its governing category and in its domain-governing category.

Since we do not use the notion of accessible subject, the definition of
domain governing category should be read without clause b. This binding
theory claims that control is a case of binding between an anaphor, PRO, and its antecedent. An alternative view is expressed by Williams (1980),
for whom control consists in the binding of the sentence which contains PRO (and not PRO directly) by the antecedent of PRO. And Borer (1985) presents a view similar to Williams's, where the anaphoric property of the sentence which contains PRO is due, not to the fact that it contains a free position in it, as in Williams's analysis, but to the fact that it contains an anaphoric AGR-element. Manzini's definition is exactly what we need, but we will apply the binding condition on non-variable empty categories, rather than on anaphors, since categories do not have inherent features. First, we repeat the definition of variables.

(88) **Variables** (definition)

\[ X \text{ is a variable iff } X \text{ forms a chain with an operator.} \]

(89) **LF Condition on non-variable empty categories**

Non-variable empty categories must be bound in their domain-governing category.

This condition has the following result: if the empty category should end up with an independent th-role, then it cannot be governed, because if it is governed, then it has a GC, and then it must be bound in it. However, if it is bound in its GC, it must form a chain with its binder, given the definition of chain-links. And this violates the th-criterion. Now, if this empty category should end up with an

43. See also Bresnan (1982) for an alternative theory of control.
independent th-role, and if it is not governed, then it must be bound in its domain-GC, which, in this case, is the maximal projection that dominates it, and not its GC, since it has none, as shown in Manzini (1983). It does not have to form a chain with its binder, since the definition of chain-links only refers to GCs and not to domain-GCs. The result is well-formed, and the empty category is PRO-like. Now, if the empty category should end up with no independent th-role, it should form a chain with some binder, and the definition of chain-links will force it to be bound in its GC. So, empty pronominals are excluded because they should be bound in a domain in which they are forced to be non-heads of chains, a th-criterion violation.

We now turn to parasitic gaps, which are empty pronominals. And, more generally, to variables bound across-the-board.
Chapter 4

Across-the-Board Variables

4.1 Parasitic gaps

4.1.1 Introduction

Parasitic gaps have been brought to general attention by Taraldsen (1981) and Engdahl (1983) and have been further investigated by Chomsky (1982) and (1985b), Kayne (1983), Pesetsky (1982), and others. An account of parasitic gaps will be proposed here, the main lines of which are that parasitic gaps are saved from the impossible status of empty pronominals at LF by being transformed into pure variables, inside a coordinate structure. The characteristic subjacency effects of parasitic gaps inside the adjunct which contains the gap, and the no less characteristic connectedness effects will be respectively explained by the necessity for variable-operator relations to obey subjacency, and the necessity of across-the-board extractions in coordinate structures.

4.1.2 Parasitic gaps are empty pronominals

Engdahl (1983) shows that there are obligatory parasitic gaps, those which cannot be felicitously replaced by overt pronominals, and optional
parasitic gaps, which can be so replaced. The first class is that of gaps which precede the real gap, as in (1)a, the second class is that of gaps which follow it, as in (1)b:¹

(1)a. John is someone that everyone who knows t likes [e]
b. This is the article that John filed t without reading [e]

The first question is to determine the nature of parasitic gaps. Higginbotham (1980) notes an example which is incompatible with the functional determination of empty categories, one in which PRO occurs in a locally A'-bound environment, as in (2):²

(2) Who did PRO doing the dishes annoy t?

In this sentence, PRO is defined as a variable, since it is locally A'-bound by who. However, it is not possible for PRO to be a variable, as shown in (3)

(3) *Who did PRO doing the dishes annoy John?

In a theory in which chain-formation identifies categories, PRO does not

---

¹ She characterizes the obligatory presence of the parasitic gap as the impossibility of having a pronoun instead, and she attributes the latter to weak crossover. Intuitively, this looks right, even though the question remains as to why parasitic gaps themselves do not fall under weak crossover, if these are pronouns. In section 4.1.17, we provide an account of Engdahl's claim. To deal with the reason why a parasitic gap is more acceptable than an overt pronoun, Safir (1984) proposes the parallelism constraint, which forces variables bound by the same quantifier to be either both overt or both empty.

have to be a variable, even if it is locally bound by and subjacent to the operator, so long as there is another possible variable for the operator. So, this example, among others, shows that pronominals may be locally A'-bound without being defined as variables. This is also Brody's (1984) result. For him, the definition of variables does not constitute a way of identifying an empty category.

Wh-chains must respect subjacency in languages like English, so, when more than one empty category are candidates for being in a chain with a wh-operator, only the subjacent one may form a chain with it. If the other empty category is not subjacent to the wh, then it cannot form a chain with it: this is a parasitic gap. Given our definition of variables, parasitic gaps are not variables, even if they are locally A'-bound by an operator, since they are not subjacent to the operator, and since subjacency is a condition on chains headed by operators. They would be pronominals at S-structure, if categories were identified there, and they are pronominals at LF, if LF is not structurally different from S-structure. This is a claim made by Cinque (1984): parasitic gaps are A'-bound pronominals. Now, since parasitic gaps are

3. Two subjacent gaps may form an A'-chain with a unique binder. These are Ross's (1967) across-the-board gaps, as in (1), discussed in this chapter:

(1) What movie did Peter talk about and John see yesterday

4. See Cinque (1984) for the claim that parasitic gaps, and all gaps which are A'-bound by the empty operator (as in tough-constructions) are pronominals. Such a classification allows him to explain why they are subject to the same conditions, which are not always similar to those applying to other empty categories. On the other hand, Pesetsky (1982) argues that parasitic gaps and MTB-gaps are the same objects, adopting the general approach of path theory and connectedness to explain their
empty pronominals and not variables, they should obey condition (90) of chapter 3 at LF, which they do not, if LF is structurally identical to S-structure. Our analysis will be that, at LF, parasitic-gap constructions are structurally identical to coordinate structures displaying an across-the-board (henceforth ATB) extraction, such that the parasitic gap becomes subjacent to the wh-quantifier hence qualifying as a variable. We will see that some of the properties of ATB extractions are shared with parasitic-gap constructions. However, there also are differences between parasitic gaps and the second gap in an ATB extraction, which will be explained by the difference between S-structure or LF ATB extraction. So, we now turn to the derivation of parasitic-gap constructions.

4.1.3 Parasitic-gap constructions at LF

Consider two parasitic-gap sentences again:

---

properties For Huybregts and van Riemsdijk (1984), parasitic gaps and ATB gaps are the same elements only in certain environments, namely when the subordinating conjunction is a coordinator. This leads them to claiming that parasitic gaps in English are not ATB-gaps. My claim is that they are ATB-gaps, but only at LF. Thus, there is no "pure" parasitic gap, contra Cinque (1984).

5. Huybregts and van Riemsdijk (1984) have strong arguments that parasitic gaps are ATB-gaps, in Dutch, and treat the subordinating conjunctions as coordinators on the surface. They claim that the wider distribution of parasitic gaps in English and Scandinavian languages than in Dutch are due to the possibility of P-stranding in these languages. If this is the correct generalization, it would explain why most of the parasitic gaps inside adjuncts are allowed only in these two types of languages, since, according to our assumptions, the LF-derivation of parasitic-gap constructions involves P-stranding.
(4)a. John is someone that everyone who knows [e] wants to go out with [e]
   
   b. This is a book that John stole [e] because he wanted to talk about [e]

As we said in the previous section a parasitic gap is too far from the
wh to be part of a chain with it, and hence is pronominal-like, and has
to obey the condition that it be bound in its domain GC, here its GC,
which is impossible for it to fulfill without violating the
th criterion. Our assumption will be that parasitic-gap sentences give
rise to an ATB extraction at LF. ATB extractions are those which apply
in coordinate structures, where the conjunction is and or but, as in
(5): 6

(5)a. I wonder who John met t and Mary talked to t
   
   b. I wonder who John recognized t but Mary ignored t

Goodall (1984) elegantly argues for the representation of coordinate
structures as the union of phrase-markers in syntax, with a special
phonological interpretation according to which they get linearized at
PF. In syntax, the conjuncts get represented on distinct planes, and
share the nodes which dominate the same lexical items. 7 This three
dimensional representation allows two (or more) conjuncts to occur as

6. For an alternative analysis of parasitic gaps, see Steedman and
Szabolcsi (1985), whose approach uses the combinatorial properties of
functors and arguments, in such a way that the distribution of these
gaps follows from conditions on the possible combinations between these
two kinds of categories.

7. See Williams (1978) for an account in terms of parallel
representations.
two (or more) full structures, even though what is heard often results in "deletions". To give an example, consider (6):

(6) John met

    Mary talked to

(6) is linearized in (7), a right-node raising structure:

(7) John met and Mary talked to Bob

This sentence has the following tree structure:

```
       S
      /\  \\
     NP NP
    /    \\/
   S     S
  /  \
NP   VP
 /  \
John NP
 /  \
V    V
 /  \
VP   NP
      \\
      V
      met
      \\
      talked to Bob
```

In (8), the conjunction and is not represented in the structure, it is implicit in the form of the structure, and added at PF when linearization occurs. Note that the existence of two parallel structures in the representation of single sentences has been proposed by Zubizarreta (1982), Manzini (1983b) and Goodall (1984), to express properties of restructuring, as in causative constructions in Romance.

8. Three dimensional representations have also been argued for in Hellen (1983), to represent different aspects of the meaning of a sentence, especially in opaque contexts. Milner (1978) mentions that such representations are possible in principle.

9. See Postal (1974) and Ross (1967), for a discussion of this phenomenon.
The interpretation of such double structures is different from that of the representation of conjuncts, which remain two independent conjuncts.\textsuperscript{10}

According to the Coordinate Structure Constraint of Ross (1967), it is impossible to extract out of a coordinate structure, with the exception of A\textsc{t}\textsc{b} extractions. This means that, if extraction takes place in one of the conjuncts, then it has to take place in the two conjuncts. Goddall explains the effects of the Coordinate Structure Constraint and its A\textsc{t}\textsc{b}-exceptions as resulting from the principle barring vacuous quantification.\textsuperscript{11} In a coordinate structure each conjunct is equivalent to a separate phrase marker, so, if a wh operator has been fronted to a common node, then the two structures are analyzed with a fronted wh. In that case, because of the condition that quantifiers must bind at least one variable there should at least be one variable in each of the conjuncts. This is a very straightforward and appealing explanation. Such extractions are impossible for the same reason that a sentence like (9) is ruled out:

(9) Who did John see Mary?

In (9), the operator can bind no variable and hence violates the principle excluding vacuous quantification.

\textsuperscript{10} See Baker (1985) for the idea that restructuring is not a case of parallel structures, but a case of incorporation, that is to say, movement of one category onto another.

\textsuperscript{11} See also Williams (1977), for whom this also follows from the definition of a transformation applying to a structure in an A\textsc{t}\textsc{b}-format.
Now, the basic idea about parasitic-gap structures is that the parasitic gap occurs inside an adjunct, which becomes a conjunct at LF, whose governing head (usually a preposition like before, etc) plays the role of the conjunction and in coordinate structures. So, schematically, a parasitic-gap sentence like (4)b will be represented at LF as in (10), in which the matrix clause plays the role of the first conjunct in a coordinate structure and the adjunct clause that of the second conjunct. And the subordinating conjunction becomes a coordinating conjunction, an insubordinating conjunction, in Huybregts and van Riemsdijk's terms:

\[(10) \quad \text{John stole } t \quad \text{(because)} \]
\[\quad \text{...a book that he wanted to talk about } t \]

We have to make more precise what the formal representation is, but is it clear that some results can be obtained from such an assumption. So, let us consider these results first, and then we will look at the exact LF derivation.\(^\text{12}\)

4.1.4 Parasitic gaps obey subjacency

It has been noted in the literature, by Kayne (1983), Contreras (1984), Cowper (1984) and Chomsky (1985b). that parasitic gaps must obey 

\_________________________

\(^{12}\) Kiss (1985) notes that parasitic gaps are subject to a matching requirement in Case with the real gap. This seems to be a requirement on ATB-gaps too, as claimed by Dyla (1984) for Polish. These two separate conditions may be unified in the present analysis, if this Case matching effects has to obtain at LF.
subjacency within a certain domain. These authors express this requirement by making the parasitic gap bound by an operator. The observations in Pesetsky (1982) and Cowper (1984) are interesting in this domain, because they relate the behavior of parasitic gaps to that of ATB-gaps, in environments in which both violate subjacency. They also observe that the subjacency effects disappear when the two gaps are in parallel structures, where the islands have the same form, and that the same holds of ATB-gaps. For example, they become acceptable if they both are inside complex NPs, as in (11):

(11)a. Which book did you accept the fact that they published [e] but reject the notion that the students should buy [e]?
   (Cowper, (13))

   b. A man whom everyone who meets [e] knows someone who likes [e]
   (Chomsky, (79))

   c. A man whom everyone who meets [e] knows someone who writes novels about [e]
   (Pesestky, (206))

That ATB-gaps and parasitic gaps behave the same with respect to subjacency should not surprise us, since this condition applies to these two objects. The subjacency effects are illustrated in (12)a-c:

---

13. The definition of conjuncts adopted here has a similar effect as Cowper's delimitation of the position of the existential quantifier, namely, at the branching point of the branching chain. And this domain is the adjunct which contains the parasitic gap.

14. The fact that subjacency effects disappear when the structures are similar in the two conjuncts must be an effect of the merging of phrase-markers where the nodes responsible for the islandhood of a constituent become intoxicous, according to the analysis to be shortly presented. However, we would have to say that the complex NPs merge, in Chomsky's or Cowper's examples, even though they do not dominate the exact same material which is not a possible assumption for the treatment of pronouns of laziness, as we will see.
The book that John stole because he found someone who would talk about [e]

b. The book that John stole because he did not know how he could afford [e]

c. The book that John stole because he knew who would enjoy reading

When the parasitic gap is inside a subject, subjacency still has to be respected, as shown by (13): 16

(13) *This is the person that [everyone who knows someone who likes [e]] wants to meet t

According to the ATB-extraction theory, the subjacency requirement is checked on the two parallel structures, as is the case in S-structure ATB-extractions. This requirement has to be respected by the parasitic gap, which should turn into a variable, since empty pronominals are not viable. As we will see, when the adjunct-clause becomes a conjunct with the matrix clause, the gap becomes subjacent to the wh operator. 17 This analysis has the advantage of not needing the presence of Contreras (1984) or Chomsky's (1985) empty operator or Cowper's (1984) existential quantifier in the structure, in order to explain the subjacency

15. The term "subjacency" is descriptive. In principle, the ECP as defined by Kayne (1981a) or (1983) is the formal condition on the relation between a wh and its variable, but since subjacency covers well-known phenomena, like the Complex NP constraint or the wh-island condition of Ross (1967), this term is more practical.

16. I am grateful to Kyle Johnson for his comments and advice on such examples.

17. I call the sentence onto which the adjunct is attached a matrix clause for practical reasons, even though it could itself be embedded, and hence not be a pure matrix clause.
requirement inside the adjunct. This is an advantage in so far as the presence of this empty operator violates the doubly-filled COMP filter in adjunct clauses which already contain an operator in COMP, like clauses under before. 18

Let us now look at the details. First, we will deal with adjuncts, and then, with subjects.

4.1.5 Movement of the adjunct

The technical problem is to be able to formally represent the two clauses as parallel structures, out of which the head preposition of the adjunct disappears, or at least does not block the way for subjacency purposes. First, let us consider what the structure of coordinate sentences is, and then we will make a parallel with parasitic-gap sentences. We will follow Goodall (1984) and extend the theory in assuming that we can create coordinate structures at LF. 19 In that case, the LF-representation of parasitic-gap sentences becomes a three-dimensional union of phrase-markers, with an overt conjunction relating the coordinates, which is the subordinating conjunction which

18. See Johnson (1985) for additional problematic cases, especially parasitic gaps inside subjects. The doubly-filled COMP filter eliminates COMP's with more than one element in it. Most presumably, before-clauses contain a moved wh, binding a temporal variable, as suggested by Larson (1984). and originally proposed by Geis (1970). This is a tempting analysis, since before is a comparative preposition. See section 5.

19. See Williams (1981a) for the claim that only alike nodes can be conjoined: the Law of the Coordination of Likes. Also Goodall (1985, p.51 ff).
governs the adjunct on the surface. Given the facts to account for, let us assume the following. By creating a coordinate structure from two constituents which are not coordinates on the surface, one is doing something special, which is not given directly by the tree structure. This thing is to put together and make non-distinct the topmost node of a structure with the topmost node of another structure, even if they dominate distinct terminal nodes. This merging process is limited, by convention, to one node. But, if the nodes already dominate identical material, then these nodes can be made non-distinct automatically. So, the merging may iterate down node by node until it finds a pair of distinct nodes that it makes non-distinct.

(14) **Union of like nodes**

(i) Merge X, where X is a topmost node of a structure A and that of a structure B, and X is the same category in A and B.

(ii) **Iteration of (i)** down the tree is possible if X dominates the same terminal in A and B.

(iii) **Merging a maximal projection** is merging of all the projections of that node, from head to maximal projection.

Also, an important convention is that matrix clauses are implicit S's and not simple S's, that is to say, they have an implicit COMP, which can merge with other COMP's (heads of S's). And lastly, let us assume the following:

(15) **The nature of COMP (convention)**

The COMP of a matrix clause is the same as that of a complement sentence (unless they dominate distinct lexical items, such as different wh operators), but distinct from that of an adjunct clause.

Obviously, this convention should be derived, but we will take it as a
primitive. Following Bencsics (1976) and Klima (1965), as mentioned in Contreras (1984), adjunct clauses are full S's, with a COMP, and are governed by prepositions. So, all adjuncts, like *without-* or *before-*clauses, etc, are to be analyzed as [P S]. What (15) says is that the COMP of S' in [P S] is different in nature from that of a *that-*clause (or of a matrix clause). Here are some illustrations of the effects of this process:

(16)a. \[
\begin{array}{c}
\text{NP} \\
\text{somedr} \\
\text{S} \\
\end{array}
\begin{array}{c}
\text{NP} \\
\text{John} \\
\text{S} \\
\end{array}
\begin{array}{c}
\text{NP} \\
\text{somedr} \\
\text{S} \\
\end{array}
\]  

(16)b. \[
\begin{array}{c}
\text{COMP} \\
\text{who} \\
\text{S} \\
\end{array}
\begin{array}{c}
\text{S} \\
\text{that} \\
\text{S} \\
\end{array}
\begin{array}{c}
\text{COMP} \\
\text{who that} \\
\text{S} \\
\end{array}
\]

I do not know about the result of a union of phrase-markers like the one in (a), but the interesting one is the union in (b), where what is required for the structure to be syntactically well formed is that the elements in COMP (which are now two) find a variable to bind in both conjuncts if these elements are operators. We assume, like Acun, Hornstein and Sportiche (1981), Huang (1982), Chomsky (1981) and others, that the doubly-filled COMP filter is not operative at LF, so the COMP in the (b) structure is not deviant.

Going back to parasitic gaps inside adjuncts, since adjuncts are not governed, given Kayne's definition of g-projections which is based on government, anything inside them will be non-subjacent to anything outside them, whether or not this S' moves and adjoins next to the
matrix COMP, where the operator is. The only way that the parasitic gap can meet the sub-jacency requirement is by having its clause form a union of phrase-markers with the matrix clause at the COMP level. Then, the matrix COMP will be non-distinct from the COMP of the adjunct clause. To obtain this structure, the adjunct-PP moves up to a position such that the $S'$ nodes can be merged, with the PP in between, and the adjunct-clause extracted from it, as in (17):

The reason why the adjunct PP is adjoined to the matrix $S$ is not only one of symmetry, to make the structure similar to coordination, but it also allows us to define conjuncts with respect to the position of the coordinator, as in (18):^{20}

(18) **Conjuncts**

Structures with non-distinct nodes are conjuncts at the level of attachment of the coordinating element.

^{20} This definition will be crucial to derive the ill-formedness of parasitic gaps inside exceptionally Case-marked subjects (section 4.10), and the connectedness effects of parasitic gaps saving other parasitic gaps (next section).
This definition differs from Goodall's view that the level of attachment of the coordinator is not what defines conjuncts. For example, for him, a coordinate NP may be analyzed as participating in an S coordination, as in the following:

(19) John met Suzan and Mary

\[ \text{John met} \quad \text{Suzan} \quad \text{Mary} \]

For us, this sentence has two possible structures: one with a coordinate NP and the other with the coordination attached to S and a bare NP, interpreted as in bare XP comparatives (section 5), as in:21

21. This analysis allows us to understand why the sum-of plural interpretation is impossible with coordinate NPs. One problem for Goodall is that, if these were derivable from a union like the one he postulates, it could not explain why this interpretation cannot have access to embedded coordinate NPs, as in (i):

(i) Mary and Suzan think that Olga likes John and Peter

Goodall allows this sentence to be represented as a union of the two phrase-markers: Mary thinks that Olga likes John and Suzan thinks that Olga likes Peter, which it does not mean (see chapter 1). For us, the embedded coordinate NP can either be an NP coordinate, or a bare XP coordinate, whose scope is limited to one S, as a general requirement: and cannot hang from the matrix S but only from the embedded one due to the distance between the bare NP Peter and its counterpart, John, as in (ii). (See the bare XP comparatives' analysis, section 5.) For lack of time, we will not compare Goodall's theory with the implementations made here, we just have to assume them.
Lastly, we assume that all LF movement has to be justified, so constituents move because they are quantified phrases. Except for the movement of the \( S' \) out of the PP, which is a case of extraposition, we will assume. As we suppose that extraposition is free. So, in the derivation of parasitic-gap constructions, two constituents move: the adjunct and the \( S' \):

(21) Sentences may be extraposed at any level.

However, the PP has to move to \( S \), by scope-assignment. So, this movement has to be justified on the grounds that the PP is quantificational.

Going back to the structure in (17), note that the movement of the clause seems to be movement into a dominating node, of the form

\[ [X_P1 \ldots \tau_1 \ldots] \]

which is impossible, since the trace should be bound by the moved element, and binding does not hold when the anaphoric element is contained inside its antecedent. However, let us assume that binding is possible if the antecedent and the anaphoric element are not on the same plane. In the structure proposed, the adjunct \( S'_{k} \) is not read as containing its own trace \( \tau_{k} \). And in the matrix \( S'_{j} \), the trace of the moved sentence, \( \tau_{k} \) is allowed to be interpreted as bound by its antecedent, because they do not belong to the same plane. This can be
done if c-command is defined so as to hold when \( x \) contains \( y \), so long as they are on different planes:

\[
(22) \quad \text{C-command (Definition) (minimally differing from (58), ch.1)}
\]

\[
x \text{ c-commands } y \quad \text{if and only if} \quad \forall z, \ z \text{ a maximal projection, } \]

\[
z \text{ dominates } x \quad \text{only if it dominates } y.
\]

\[
x \text{ does not contain } y \text{ if they belong to the same plane,}
\]

\[
\text{and } x \text{ may equal } y.
\]

As we see in the schematic structure of the LF of a parasitic-gap sentence, as shown in (17), by merging the S's at the S' node, the COMP of the matrix clause becomes accessible to the adjunct clause (COMP is the head of S', so it merges with the other COMP when S' does). This is a proper representation, so long as, if the COMPs contain operators, these operators find a variable to bind in the two conjuncts. And this is exactly what happens in a parasitic-gap structure: the wh of the matrix clause comes to bind the parasitic gap, as a pure variable, in the second conjunct.

If this is to be the derivation of parasitic-gap sentences, it should have certain consequences, due to the fact that some constituent moves (the adjunct PP, with extrapolation of the clause that it contains). So we now turn to some predictions made by the postulation of the movement of the adjunct at LF.

The adjunct PP cannot adjoin higher than S to S', for example, because the S' extracted from it should be able to c-command its trace, and this S' cannot move higher than the matrix S' since it should merge with it. This entails that the movement of the PP is like QR, so, like QR, it is limited by condition A. This is an important result, and we
are going to see that this property can capture some of Kayne's (1983) insights, termed as connectedness effects.

4.1.6 Connectedness effects

Kayne discovered that a certain locality condition between the real gap and the parasitic gap has to obtain, in order for parasitic gaps to be licensed, which he expresses in his path theory. Kayne defines the ECP as a well-connected relation between a binder and a bindee, where all the nodes between the two are governed, and form a g projection of the bindee. A path is the part of the tree which relates the two elements. He shows that parasitic gaps violate the ECP, because the binder is not properly connected to the gap, owing to the lack of government of one of the nodes (at least) in between the two. But the parasitic gap can be connected to its binder when another path intervenes headed by the same binder. Geometrically, this means that two paths can join, helping the parasitic gap to be connected to its binder.

(23)a. Definitions (Kayne 1984 p.171)

\[ g \text{-projection set } G_\beta \text{ of a category } \beta, \text{ where } \gamma \text{ governs } \beta. \]

a. \[ \forall \pi, \pi = a \text{-projection of } \gamma \rightarrow \pi \in G_\beta \]

b. \[ \beta \in G_\beta \]

and

b'. \[ \delta \text{ dominates } \beta \text{ and } \delta \text{ does not dominate } \gamma \rightarrow \delta \in G_\beta. \]
b. ECP (Kayne, p.171)

Let $\beta_1, \ldots, \beta_j, \beta_{j+1}, \ldots, \beta_n$ be a maximal set of empty categories in a tree $T$ such that $\exists$ a c-commanding $\alpha$, $\forall j, \beta_j$ is locally bound by $\alpha$. Then

a. $\bigcup_{1 \leq j \leq n} G_{\beta_j}$ must constitute a subtree of $T$
and
b. there must exist a $\rho$ such that $\rho \in \bigcup_{1 \leq j \leq n} G_{\beta_j}$ and $\rho$ dominates $\alpha$.

To see how this functions, consider (24)a, which illustrates an unconnected gap, and (24)b, which illustrates how the same gap can be connected via the union of its path with a well-connected path:

(24)a. *The article that John read this book before filing [e]
b. Which article did John file the before reading [e]?

In (24)a, the g-projection of the parasitic gap stops at the adjunct PP, which is not canonically governed. But in (24)b, the node PP is attached to a node which is in a path that connects the wh with the trace. So the path between the parasitic gap and the PP is connected to the path between the wh and its trace, which then makes the parasitic gap connected to the wh.

The lack of connection between a parasitic gap and the wh is due to the lack of government (by a governor up along the path) of one of the nodes projected up from the gap to the wh. The contrast above shows that the real gap is necessary in a parasitic-gap sentence, but is does not prove that the sentence is well formed owing to connectedness. One could say that (24)a is ruled out because of the condition precluding vacuous quantification: if the parasitic gap is not subjacent to the wh, it cannot form a chain with it, and hence the wh is not licensed in the structure. So, a relevant contrast is one in which connectedness does
not hold. In that case, even though the wh will find a subjacent gap to form a chain with, the parasitic gap will be too far from the real gap to be saved, as shown by Kayne.

In our analysis these sentences will be ruled out because the adjunct clause will adjoin lower than the real gap, so, at LF, the clause will be a violation of the Coordinate Structure Constraint. Consider such a sentence:

(25) *Who did you tell a friend of t that Mary came in without noticing [e]?

At LF, the adjunct can adjoin only to the embedded S because its movement is clause-bound, which yields the following:

(26) 

who you told a friend of t that

PRO noticing [e]
This LF displays a coordinate structure that Mary came in + without + noticing [e]. (Recall that conjuncts are defined as the constituents which are at the same level as the coordinating conjunction, here, without.) Since the empty category object of noticing must be related to an operator, in order not to be excluded as an empty pronoun, it has to form a chain with the wh in the matrix COMP. This includes the matrix COMP in the structure of the conjunct. Let us suppose that, when an operator has been included in the structure of one conjunct, it is automatically included in the common structure between the two conjuncts. So, the matrix COMP gets included in the structure of the first conjunct. But then, the first conjunct is without a variable for the operator, and the structure is ruled out as a violation of the condition precluding vacuous quantification (the Coordinate Structure constraint). So, a sentence like (26) is excluded in the same way that a sentence like (27) is, where the conjunction relates two embedded sentences.

(27) *I wonder who you told t that [Mary came in and she noticed [e]]

Note that, if wh movement was successive cyclic, and left traces, we would be able to rule (25) out immediately by looking at the intermediary COMP, the one which forms the union of the two conjuncts, without having to link it up to the matrix COMP because there would be an intermediary trace in the common COMP. However, Kayne's definition of the ECP gets rid of successive cyclic movement (for extraction of properly-governed wh-elements). Nevertheless, it is not clear whether successive cyclic movement may be eliminated, (cf. its crucial use by
Lasnik and Saito (1984) and Johnson (1985)), so it may turn out that we do not need to make our stipulation. For the cases under discussion, the basic idea is that the conjuncts display a violation of the Coordinate Structure Constraint, however the representation of the COMP-to COMP relation is handled.

Let us turn to other connectedness structures. For the connectedness account, assuming that subadjacency is respected inside the adjunct, a parasitic gap will be illicit in structures in which the adjunct-node (the one which blocks government-percolation up) is not attached to a node which participates in government-percolation, like the VP that it is an adjunct of. But it is hard to see what such a node could be, since if this node existed at all, it would presumably be headed by a governor of the adjunct, and hence it would allow government-percolation up, as in (28):

(28) \[
\begin{array}{c}
S \\
/ \\
NP \quad VP \\
/ \\
VP \quad XP \\
/ \\
X \quad Adjunct \\
\end{array}
\]

government

But one situation where connectedness seems to easily account for the facts is when the real gap is the subject, and the adjunct a VP-adjunct, as in the schematic representation:
(29) *Who talked to John after Peter introduced [e] to Mary?

```
S'    
COMP  
INFL'' 
wh1 NP INFL' 
t1 INFL VP

In that case the adjunct is not connected to the subject path, since the node INFL' intervenes. In our terms, this will be explained in terms of Williams's (1978) definition of well-formed factors.\(^{22}\) These cases might be analyzed as violating the anti-c-command requirement between the real gap and the parasitic gap, derived from principle C of the binding theory, which is Goodall's account, where the subject gap is analyzed as c-commanding the other variable in the domain of its operator, a violation of principle C, as stated in Chomsky (1981).\(^{23}\) However, the anti-c-command requirement cannot be our explanation, because the subject does not c-command the other conjunct, in the sentences in question if S is a maximal projection. Consider the configuration:\(^{24}\)

---

22. Or something equivalent to it, in our three-dimensional representation of coordinates, instead of parallel structures. In the text, we will use Williams's account and definition.

23. That is to say, variables must be free in the domain of their operator.

24. The structure is symmetrical, so extraction from either matrix subject is impossible from any conjunct (unless they are both extracted).
We could modify the definition of c-command, in order to allow the subject to c-command the other conjunct, but other facts will crucially rely on the impossibility of an S-adjointed element to c-command the other conjunct, based on the fact that S is a maximal projection. So let us turn to another account of the impossibility of extracting asymmetrically the matrix subject. In his ATB-format representation of coordinates Williams (1978) defines well-formed factors in a manner which excludes these cases. An ATB-format is a representation of two conjuncts as parallel structures. These are factored, that is to say, segmented, with respect to the constituent which is subject to an ATB-transformation, like wh-movement, as in (31), where the relevant brackets are those of the parallel structures, here S:

\[
\text{COMP} \quad \begin{bmatrix} [\text{John saw who}]_S \\
[\text{Mary met who}]_S \end{bmatrix} \quad \text{and} \quad S
\]

Now, a well-formed factor is one in which the left brackets of the conjunct belong to the same factor. In (31), they do. But when the wh

---

25. This is how we account for the unavailability of pronouns of laziness in adjuncts.
corresponds to a matrix subject and a non-matrix subject, the left
brackets of the conjuncts do not belong to the same factor, as shown
below:

(32) \[ \text{* COMP} \begin{array}{c}
[\text{who}]_S \\
[\text{left}]_S \\
\text{and} \\
[\text{Mary saw}]_S
\end{array} \]

This is what is responsible for the ungrammaticality of \textit{I wonder who
left} and \textit{Mary saw}. We will thus adopt this explanation for the
corresponding parasitic gap sentences, such as (29). Descriptively, the
condition on well-formed factors can be expressed as the impossibility
of having one of the conjuncts only starting with an empty category.

Now, the definition of connectedness has a very interesting effect,
which is basically, that gaps save gaps. This means that a parasitic
gap can be embedded inside islands, so long as there is a path
connecting each island to the operator above. Does this serial effect
carry over to our analysis? The question is to determine how a
multiply-non-subjacent gap is saved by other gaps. The answer is that
this serial effect in Kayne's theory is transposed in our theory as an
iterative formation of conjuncts, each containing a gap, hence forming a
multiple ATB-extraction representation. Let us take a particular
example:

(33) \textit{What person did Tom visit [because he liked [e] after
talking to [e]]}

As indicated, the lower adjunct clause is understood as a modifier of
the lower VP. According to Kayne's theory, the second parasitic gap is saved by the first one, itself saved by the real gap, as shown below:

(34)

The circled nodes are the nodes at which the junction between two paths occurs. This is to be contrasted with a non-connected parasitic-gap structure, like:

(35)

According to our theory, connected gaps will be ruled in because the most embedded clause will be able to be extracted from a clause which is itself able to be extracted, all of them conjoined with the matrix sentence. In the connectedness theory, the intermediary gap is understood as licensing for the deeply embedded gap, whereas in our
account, the presence of this gap is interpreted as forced by the Coordinate Structure Constraint, and its ATB-extraction exceptions: if extraction occurs inside the most embedded clause (which will end up being the third conjunct), then all conjuncts must contain a gap, and in particular the middle clause (represented as the second conjunct).

Then, the two former parasitic gaps become pure variables at LF:

\[(36)\]

By joining the S's of the adjunct clauses with the matrix S', the two parasitic gaps become subjacent to the operator in the COMP of the matrix clause. Also, given that the COMP becomes part of the structure of each conjunct separately, each conjunct has to contain a variable, which it does. So, the condition on vacuous quantification derives the connectedness effect of such constructions.

Let us now turn to the other consequence of the claim that the adjunct moves at LF, in order to make the parasitic gap a proper variable.
4.1.7 Movable clauses

We have implicitly assumed that movement of the adjunct was unproblematic. However, movement at LF should be constrained in some way. Following common assumptions (cf. Higginbotham (1983), May (forth.)), QR may apply only to quantified phrases, through the operator which they contain, implicit or overt. In the particular case of the adjunct, these constituents can move if they contain an operator, like a bound pronoun in the S' that they contain or an operator in COMP. We thus expect a parasitic gap in an adjunct to be bad if the sentence is not movable, if it cannot be treated as a quantifier. Let us first determine what clauses can move, apart from those which contain a bound pronoun.

Comparative clauses are one type of adjunct clauses which can move, as shown below (an example of Chomsky and Lasnik (1977) adapted from Bresnan (1977), cited in Engdahl (1983):

(37) The books that Mary read t as often as Bill read [e]

This is because comparative constructions are headed by an operator, which is assigned scope at LF. Given the definition of quantified phrases as the GC or some category dominating that GC of some operator, and given that the comparative operator is the specifier of the Adverb Phrase, this Adverb Phrase is a quantified phrase. So, the LF of (37) is (38), in which the parasitic gap forms a chain with the wh-quantifier:
(38) the books [that Oₗ[ [Mary read tᵢ tₖ] [as often as tⱼ] ] ]

[Bill read [e]ₗ tₖ] j,₁

Note that certain prepositions are semantically comparative, like the
temporal prepositions before and after. A parasitic gap inside adjuncts
headed by these prepositions behaves like one in an overt comparative
clause, which is expected, since the treatment of such prepositions
should be the same as that of overt comparative operators, as argued by
Geis (1970): the preposition is assigned scope over the sentence which
is compared:

(39) Every newspaper that John kept t after Bill read [e]
    was out-of-date

Other clauses, headed by operators, behave like quantifiers, and can
be assigned scope because of the operator which heads them, like
sentences headed by when or if:

(40)a. The draft that Albert glanced at t when Robert got rid of [e]
    was full of contradictions

b. Tom was cooking the stew that Richard would eat t only if the
cat had a little of [e]
Now, in certain classes of adjunct-clauses, the subject of the adjunct must be coreferential with the subject of the matrix clause, a fact which is even more obvious when the adjunct clause has an overt subject. In that case control does not automatically make these subjects coreferential, so the potential choice for the reference of the subject of the adjunct renders more obvious the absence of choice. This is explained by the hypothesis that the LF derivation of parasitic-gap sentences forces the adjunct to be assigned scope, and hence forces it to be quantificational. And bound pronouns allow sentences to be quantificational. Note the contrast: 26

(41)a. This is the book that John stole [e] because he wanted to talk about [e]

b. This is the book that John stole because his best friend wanted to talk about

Because-clauses are not headed by operators, so the presence of a bound pronoun is necessary to license a parasitic gap.

As for without-clauses, they are not headed by operators either. Given that such adjuncts contain a controlled PRO, this PRO will act as the operator which will license the presence of a parasitic gap. Without-clauses cannot be tested very well with respect to the coreferentiality requirement, because of obligatory control, which will always make an operator available in the adjunct -- PRO -- unless they take a gerund with a lexical NP, which is already a little awkward.

26. If his is bound by John in (41)b. this will make the pronoun an operator, whose scope is that of its GC, namely, the NP. and not the S'. This is why the presence of the bound pronoun does not help in (41)b.
stylistically. Still, consider the sentence:

(42) ??What book did John read t without Peter talking about [e]

So, the coreferentiality requirement seems to constitute empirical evidence in favor of the treatment of the adjunct-clause as quantificational. We can explain why adjunct clauses headed by because require coreference between the subjects, and not those headed by before: because is not an operator and the adjunct itself does not contain an operator so it cannot move to form a conjunct with the matrix S', leaving the parasitic gap an impossible empty pronounal at LF.

Note that we have said that, in order for an adjunct which is not already quantificational to be assigned scope, the adjunct which contains the parasitic gap should contain a bound pronoun. The parasitic gap itself is a pronounal, so it should be able to make the sentence that contains it behave like a quantifier, and it should not need the presence of another bound pronoun. The reason that the parasitic gap cannot be the bound pronoun which makes the sentence quantificational is that this bound pronoun becomes a pure variable at LF, as the tail of an A'-chain, and hence can no longer act as an operator at that point, since operators should always head chains, as an assumption. So, the moved sentence would end up in a quantifier position without containing an operator, an improper representation.

27. If sentences with because and without are not so bad, even when the subjects are not coreferential, it is presumably because the adjunct PP hangs from S already, and hence does not have to move.
To conclude, we have seen that the adjunct-clauses that allow parasitic gaps are all quantificational. Now, the sentence above (every newspaper that John kept before Bill read...) recalls the same sentence presented in section 3.4.2:(49), which gives rise to a parasitic gap at LF, which we now investigate.

4.1.8 LF parasitic gaps

Parasitic gaps can occur on the surface but they also may occur at LF', after some copying rule applies, as in VP-deletion sentences. Given that the condition on empty pronouns is meant to apply at logical form, we expect parasitic gaps not to be able to occur in positions in which they car-ot end up as pure variables. That is to say, we expect certain contrasts to arise, depending on whether or not the clause containing the parasitic gap has been able to form a conjunct with another clause. Consider (43)((49) of 3.4.2):

(43) John [kept [every newspaper that Bill read t before Peter did [ e]j]]j

This sentence is grammatical, owing to the union of the adjunct S' with the relative clause, and the extraction of the NP out of the VP, in order to avoid circularity, as seen in the following logical form derivation.  

28 Also, the empty VP in this sentence gets represented as

28. At LF', the comparative clause headed by before contains the temporal variable t. This variable is A'-bound by an operator in the COMP of the clause Complement of before, following Chomsky's (1977) analysis of comparative clauses. For clarity, I haven't represented this operator in this LF. That these are to be represented as
materially identical to the matrix VP:

\[(44)\]

Like overt parasitic gaps, LF-parasitic gaps force the sentence which contains them to be assigned scope, and hence to be quantificational. So, they require coreference in because-clauses if these hang from VP and not S, as illustrated in the slight contrast below:

\[(45)a. \text{What team did John, bet on because Peter, thought he should?}\]

b. What team did John, bet on because he, thought he should

It seems that \((45)a\) can hardly be interpreted as: what team did John bet on because Peter thought he should bet on it.\(^{29}\)

29. I have called the gaps reconstructed inside the empty VP LF-parasitic gaps for the reason that, if the sentence was overt, they would be parasitic gaps on the surface.
4.1.9 Parasitic gaps in LF-conjuncts

The necessity for a parasitic gap to become a variable at LF forces the adjuncts which contain the parasitic gaps to become conjuncts at LF. We are going to see that, for the derivation of parasitic gaps, the conjuncts cannot be complements of the sentence that they merge with. They may only be pure coordinates or adjuncts, or subjects, but not complements. However, we will also see later that this restriction on the kind of sentences which may merge with the matrix clause does not extend to the derivation of other cases of LF conjuncts, as in the derivation of sentences with sloppy identity of pronominals, and free relatives (of sections 2 and 4). Anticipating the study of these cases, the distinction between parasitic-gap sentences and the other sentences seems to be that the operator which ATB-binds the relevant variables has been moved on the surface in parasitic-gap sentences and not in the others. In any case, let us wonder why such a discrimination against complements should exist at all.

Semantically, a conjunct is a proposition which is interpreted as related to the rest of the sentence it occurs in with the connective and. Adjuncts, similarly to subjects, are interpreted this way, via predicating with the INFL of their clause, unlike objects, which combine with a predicate to form a complex predicate.30 Given that

syntactic predication and complementation give rise to different ways of interpretation. conjuncts, subjects and adjuncts form a natural class of constituents, as opposed to the class of complements. So, let us assume the following:

(46) **Condition on LF-conjuncts**

If a single \( A' \)-chain in a main clause becomes a double \( A' \) chain, the double chain occurs inside semantic conjuncts.

Two semantic conjuncts are defined as clauses which are related to each other with the logical connective **and** or **or**. Since predication provides the connective of coordination, **and**, adjuncts and subjects are semantic conjuncts with the clause with which the relation of predication holds. Remark that adjunct-clauses are in a predication relation with the clause that they are attached to only by extension, because, in general, they are governed by a preposition, and it is really the PP which is predicated of the clause they are attached to. We will consider, however, that the adjunct clause itself is defined as a semantic conjunct. Usually, the preposition adds some semantic information about the coordination between the two, but the two clauses do constitute semantic conjuncts.\(^{31}\)

Returning to the purpose of (46), in parasitic-gap constructions, the \( A' \)-chain formed by the real gap becomes a double chain, when the parasitic gap merges with it. So, it falls under (46), and can occur only in semantic conjuncts. Let us now turn to the facts. One type of sentences is indirectly stated as being unable to felicitously contain

\(^{31}\) See Larson (1983) for the semantics of adjuncts.
parasitic gaps, namely, complements.

We have to find a complement sentence which contains a parasitic gap. In (47), the sentence is a complement and it contains a gap:

(47) Who did John tell t that Mary would call [e]

This sentence is not too bad, but the gap in brackets is subjacent to the wh-quantifier, so it is not a real parasitic gap. This sentence simply displays a double chain at S-structure. So, we have to find a non-subjacent gap inside a complement which is in the same clause as the wh-trace, owing to the connectedness effect discussed in section 4.1.6. Such sentences are relative clauses and comparative clauses, which are islands:

(48)a. *Tom read the books that John discussed t with everyone who had talked about [e]

   b. ??Tom read the book that John discussed t with everyone who
       Mary had talked to t about [e]

(49)a. *I met the person that John gave more books to t than Op
       Mary gave [e] to t

   b. ??Who did you show a picture of t to more people than you talked to t
       about [e]?

The reason why these sentences are excluded is that the island-category which contains the gap is a complement and not an adjunct. Hence, the LF representation of such sentences cannot turn the S's contained in these complements into adjuncts. There is thus no way to get rid of the

32. Condition C still has to hold. Since (47) is acceptable it means that the first gap does not c-command the second gap.
distance between the wh operator and the parasitic gap with respect to subjacency (Kayne's ECP). Note that the ungrammaticality of (49)a-c does not really arise from the fact that the relative clause or the comparative clause cannot be assigned scope. In fact, both occur in quantificational constituents, a relative clause headed by every, and a comparative clause, so both can be assigned scope. What cannot happen is for the NP and S nodes of the relative or comparative clause to be eliminated from the way in between the wh operator and the gap. Given that it is impossible to transform these sentences into conjuncts, the parasitic gaps will remain inside islands. The structure is like (50), where, in terms of traditional subjacency, NP and S are bounding nodes, and in Kayne's terms, NP is not a g-projection of S':

(50)a. *the books that Op_1....[NP everyone [S_1...[e_1]]]
b. *the person that Op_1...[NP more books [S_1...[e_1]]]

Let me argue against two possible alternative analyses of the ill-formedness of (48)-(49). First, it may be that (48)-(49) are bad because of the anti-c-command requirement between the real gap and the parasitic gap. Even if true, this would not be an exhaustive

33. (48)a is also excluded. in our account, for the reason which excludes parasitic gaps inside relativized NPs which are not subjects, like (i):

(i) *This is a person that everyone to whom John gives a present to [e] likes t

The reason is that the variable in the relative clause must be a subject, if the relativized NP is a subject, and may not be a subject, if the relativized NP is not a subject. See next section. The fact that (48)a is better than (i) is a problem.
explanation, because when the real gap is slightly embedded, the result is the same even though it can no longer be said to c-command the indirect object: 34

(51)a. *I met the person that John introduced a friend of t to everyone who wanted to talk to [e]

b. *Tom read the books that John discussed a chapter of t with everyone who had bought [e]

The second alternative account of (48)-(49)'s ungrammaticality is that it is a crossing-constraint violation, or a violation of the Path-Containment condition, in Pesetsky's (1982) terms: 35

(52) Crossing Constraint
   Two A'-dependencies cannot cross.
   or: Path Containment Condition (Pesetsky (1982): (94)p.309)
      If two paths overlap, one must contain the other.

The LF of (49)a is the following:

--------

34. See Johnson (1985) for additional arguments that the object does not c-command the adjunct in the previous sentences.

35. See Kuno and Robinson (1972) for the proposal of the crossing constraint.
We see that the line between the direct object and the COMP is contained in the line between the indirect object and the quantifier phrase that binds it, both in the adjoined phrase, and in the original S. So there is no way to claim that the crossing constraint is responsible for (49)'s unacceptability.\(^{36}\)

Now, let us turn to parasitic gaps inside subjects. The interesting property of subjects is that they are predicated of the VP. Hence, they can be represented as conjuncts, respecting the condition on semantic

\[\text{36. But the other order does violate the crossing constraint, as shown in the simplified structure (without LF movement of more people...):}\]

\[(i) *I read the book that Op}^{i} \text{ John gave } t_{i} \text{ to more people than Op}^{j} \text{ Mary gave } [e]_{j} \text{ to } t_{j}\]
4.1.10 Parasitic gaps inside subjects

Consider (54):

(54) Tom is someone who [everyone who knows [e]] likes to talk to t

If we follow the syntactic analysis given for the adjunct cases, the relativized NP should move to a position such that the matrix S' and the relative clause can join. Cowper (1984) and Stowell (1985) have noticed that parasitic gaps inside subjects were worse when the NP was not quantificational. This NP has to be non-specific if it is an indefinite, or it must be headed by quantifiers like every.\(^{37}\) This is because the NP has to move, and hence has to be quantificational.

One thing remains to be clarified. The relativized NP has to adjoin to S', in order for the COMP of the relative clause and the COMP of the matrix clause to form a non-distinct node. We see below that this movement, which crosses S and S' in one swoop, is allowed with respect to the ECP as formulated in Kayne (1981), without having to go through COMP, since the trace of this movement becomes bound by the wh already in COMP (that of the relative clause), in the resulting structure.

\(^{37}\) That is to say, the NP has to be used as a description, and not as a referring expression. In the first case, anything which fits the description can make the sentence true, but in the second case the description is used to point at a particular individual, so not any individual satisfying the description will make the sentence true, as used then.
Then, the relative clause is extracted from the NP, that is to say, is
extraposed, in the position such that it combines with the matrix S'. 38
So, we assume the LF representation of (54) to be (55):

(55)

```
\[ S' \]
\[ \left\{\begin{array}{c}
NP_j \\
NP_k \end{array}\right. \]
\[ S' \]
\[ \left\{\begin{array}{c}
COMP \\
S \\
S_k \end{array}\right. \]
\[ \left\{\begin{array}{c}
\text{everyone} j \text{who}_j \text{who}_k \text{NP}_j \text{VP}_j \text{NP}_k \text{VP}_k \\
t_j V j \text{NP}_j \text{VP}_j \text{NP}_k \text{VP}_k \\
\text{likes } t_i \text{ knows } t_i \end{array}\right. \]
```

Note that, in this LF, the common COMP now dominates two operators, so
these two operators must bind the two variables inside each conjunct,
which they do: in both conjuncts, [t_j knows t_i] and [t_j likes t_i], there
are two variables, each one ABT-bound by one of the operators in COMP,
and each one forming an A'-chain which obeys the ECP. So, this is a
proper representation. That one of the S's is interpreted as a relative
clause restricting everyone is indicated by the trace of the relative
clause. And even if such movement did not leave traces, as a direct
interpretation of the Projection principle makes possible, then the
predication relation between the head NP and the relative clause is
annotated on the surface, as in Williams (1980), so the clause can be
recognized as a relative clause. So, all the necessary information can
be read from this structure, and all the gaps are subjacent to their

38. I am talking of these two movements as if they were ordered, but
they are not.
operators, the desired result.

In (55), the relative clause and the original S' are interpreted as conjunct clauses, and this obeys condition (46) on semantic conjuncts, because the moved NP is in a predication relation with the INFL' of the sentence that it is extracted from. Note that this analysis forces the NP which contains the parasitic gap to be the subject of the VP of the S' that the relative clause combines with, which is verified when it is inside an island, as expected, (cf. (56)a) but also when it occurs inside an NP from which it could be extracted without violating any extraction-conditions, as in (56)b.39 For example:

(56)a. *John is someone that everyone who knows someone who likes [e] annoys t

    b. *John is someone that a picture of everyone who likes [e]
       surprises t

(56)b is unacceptable even though it is possible for the relativized NP to be extracted from the picture-NP, as shown below:

39. Such an example is predicted to be acceptable for Cowper (1984), who provides the existential quantifier at a point subjacent to the parasitic gap, and also for Chomsky (1985a), for whom the non-subjacent node (here the relative clause) has to be subjacent to the wh operator.
4.1.11 The relativized NP must correspond to a subject

We have just seen that, when the parasitic gap is inside a subject which is a relativized NP, the relative clause and the matrix clause become conjuncts. Now, given that the operator of the relativized NP becomes an ATB-binder, binding both its trace in subject position (obtained by QR) and the wh-variable (of the relative clause), one of these two variables will be a subject, namely the trace of the NP, since this NP is originally in a subject position. However, in a coordinate structure, as given by Williams's (1977) definition of factors, it is not possible for an operator to asymmetrically bind a matrix subject; descriptively speaking, it is impossible for only one conjunct to start with an empty category. This thus predicts that, if a subject relativized NP is forced to form a conjunct, then this NP will have to correspond to a subject also inside the relative clause. This is verified, as in Chomsky's (1985a) important example:

40

40. I am grateful to Kyle for helpful discussion on this example and its theoretical significance.
(58) *He's a man that [any present they'll give to [e]] will please t

(58) is ill-formed at LF, because the conjuncts have an ATB-gap
(corresponding to a matrix subject in one conjunct and to an object in
the other).41

(59) [any present

[that] \_Op

[they'll give \_to [e]]]

[t \_pleases [e]]

4.1.11.1 A difference between overt parasitic gaps and LF parasitic gaps

This section gives a theory-internal reason for making the
impossibility of asymmetrically extracting from a subject a consequence

41 Johnson (1985) discovers a contrast for which our theory may provide
some kind of explanation. Safir (1984) gives (i) as well-formed, but
since there is a sharp contrast between it and (ii), we will consider
that it is on the ungrammatical side:

(i) Who did you tell friends of [e] PRO to compromise with t
(ii)*Who did you tell friends of [e] that someone will compromise with t

The reason for this contrast could be Tense, but, rather, it must be the
fact that the NP which contains the non-subjacent gap indirectly
corresponds to a subject, by being the controller of PRO. As evidence,
consider (iii):

(iii) *Who did you promise PRO friends of [e] to compromise with

The facts are the same if, instead of PRO, the embedded sentence
contains an overt pronoun or if the NP which contains the parasitic gap
is an island by itself, as in (iv)-(v):

(iv) Who did you tell friends of [e] that they should compromise with?
(v) ?Who did you tell everyone who likes [e] PRO to compromise with?

These interesting facts are left for further research, but it is clear
what the directive line should be.
of the definition of a well-formed factor, rather than a consequence of the anti-c-command requirement. In our study of antecedent-contained VP-deletion, we saw that a certain coreference requirement was due to the creation at LF of an empty pronoun. And in order to avoid this faulty status, the empty category had to find itself subjacent to a wh.

Consider again such an example:

(60) John talked to everyone who wanted him to

The LF of (60) must involve movement of the clause [him to [e]] to INFL", in such a manner that the inserted NP variable forms a chain with the wh quantifier, as in (61):

(61) [John [everyone who [him to [talk to x]]] [t₁ wanted t₂]

[tₖ talked to x]]

At LF, the parasitic-gap structure becomes similar to that of (63) below, which, if our analysis in general is correct, must be well-formed, since the sentence is grammatical: 42

(62) *Who t met you before you wanted to talk to [e]

As we said earlier, the LF of (62) involves assigning scope to the adjunct PP, which is a movable constituent, since it is headed by an

---

42. Note that the subject does not c-command the moved S, because its S is a maximal projection. So, the empty category inside the S does not violate principle C.
operator, and the S' extraposes and merges with the matrix clause, to become a conjunct, as in:

(63)

\[
\begin{array}{c}
S' \\
\text{COMP} \\
who_i \\
\text{S} \\
\text{PP}_k \\
\text{NP} \\
\text{VP} \\
\text{t}_j \\
\text{before} \\
\text{t}_i \\
\text{VP} \\
\text{t}_k \\
\text{V} \\
\text{NP} \\
\text{met} \\
\text{you} \\
\text{S} \\
\text{NP} \\
\text{VP} \\
\text{FRO} \\
\text{V} \\
\text{NP} \\
to\text{talk}\to\text{t}_i
\end{array}
\]

Let us now turn to another important example, which shows that connectedness does not always rescue parasitic gaps, and which speaks in favor of an analysis in terms of conjoined structures, as we have done.

4.1.12 The real gap and the adjunct have the same GC

Connectedness predicts that an embedded real gap will save a parasitic gap which is inside a higher adjunct, but this is not true, a fact which has been previously discovered independently in Johnson (1985):

(64)??I [wonder who John remarked that the problem annoyed t before criticizing [e]]

Descriptively speaking, it seems to be the case that the real gap and the adjunct containing the parasitic gap must have the same GC. In some
preceding cases, when the real gap was situated higher up than the
adjunct, the sentence was ill-formed because the LF conjuncts violated
the Coordinate Structure Constraint -- or because the parasitic gap
failed to be connected to the wh when the saving path was too far up, in
Kayne's terms. In the example under consideration, the adjunct is
situated higher up in the tree than the real gap, and this is what makes
the sentence marginal. However, this is not excluded by Kayne's
connectedness account, since the parasitic gap is properly connected via
the path of the real gap. This shows that, if connectedness is a
necessary condition on gaps, it is not a sufficient one:

(65)

\[ S \]
  \[ I \]
  \[ VP \]
    \[ wonder \]
      \[ S' \]
        \[ who \]
          \[ S \]
            \[ VP \]
              \[ remarked \]
                \[ that \]
                  \[ S \]
                    \[ P \]
                      \[ PP \]
                        \[ that \]
                          \[ the problem \]
                            \[ annoyed \]
                              \[ NP \]
                                \[ criticizing \]
                                  \[ t_i \]
                                    \[ NP \]
                                      \[ t_i \]

The impossibility of embedding the real gap is strikingly similar to
overt ATB-extractions, in which an embedded first gap is bad if the
conjunction is attached to the matrix clause:

(66) This is the man who Mary remarked that the boss was going to
      hire \( t \) and Peter criticized \( t \)
Now note that the gap may be embedded in the first conjunct, so long as the gap is also embedded in the second conjunct, as shown in the pure coordinate structure:

(67) I wonder who Mary thinks that John will invite t and Peter believes that he will leave out t

This effect is reflected in the following parasitic-gap construction, in which it is now possible to have the adjunct in the matrix clause, so long as the parasitic gap is embedded, like the real gap:

(68) This is the person that Mary thought that John would invite t after she realized that he might leave out [e]

So, considering that parasitic-gap constructions do involve a coordinate structure at LF, it seems to be the case that, in a coordinate structure, the matrix clause cannot contain a gap which has a greater degree of embedding than a gap in the second conjunct, as stated below:

(69) Condition on embedding in double chains
A gap in a first conjunct may not have a degree of embedding greater than that of a gap in a secondary conjunct.

The degree of embedding of x is the number of S nodes which dominate x. What is called a first conjunct is the first conjunct in an S-structure coordination, and the conjunct formed by a matrix clause, in an LF coordination. The secondary conjunct is the next conjunct, in an S-structure coordination, and the moved conjunct, in an LF coordination. 43 There exists another argument in favor of this

43. Note, however that this convention does not seem to apply to subjects, in either direction: the distance between the wh and the gap
correlation, presented in the next section.

4.1.13 LF-conjuncts are S-structure coordinates or predicates

In this section, we are going to see that there is another type of sentence which suggests that the coordinate-structure approach is the right one. In the preceding sections, we saw that there must be a close relationship between the phrase-markers which get merged at LF. they must be semantic coordinates.

Let us grant this, and consider the following sentence

(70)a. "This is someone that I believe everyone who knows [e] to like t"

This is to be compared with the tensed counterpart:

(70)b. "This is someone that I believe that everyone who knows [e] likes t"

As was said earlier, parasitic gaps are acceptable inside subjects, because subjects are in a predication relation with the VP (possibly with INFL') of their sentence. So, if the subject needs to become a conjunct at LF, like when it contains a parasitic gap, then it can merge with the sentence of which it is the subject. This is the only possible inside a subject may be longer than that between it and the gap in the matrix clause, and vice versa, without having any effect on the acceptability of the sentences:

(i) John is someone who everyone who thinks that someone should visit [e] likes t
(ii) John is someone who everyone who likes [e] thinks that someone should visit t

This will be left as an open question.
sentence that it can merge with, since merging with other constituents would imply that the subject is in a predication relation with them, which it cannot be, since there is a one-to-one correspondence between subjects and predicates. Now, the quantified NP, in (70)a above, is in an Exceptional Case Marking environment, which means that it is governed by the matrix verb. Now, since this NP contains a parasitic gap, it has to be assigned scope in order to merge with its predicate. However, the scope of a quantified phrase is determined as its Governing Category, as in Schein (1983). This means that, if the subject is to move at all, it moves to the matrix clause, not the embedded one, given the definition of scope. Consider the structure, when the NP is extracted, before landing on the matrix S'-adjoined position:

(71)

```
Op...

S'  
|   |
|   |
NP  S
|   |
NPj S' NP VP
|   |
ev.COMP S I V S
|   |
who NP VP believe NP VP
|   |
tj V NP tj V NP
|   |
knows tj to like tj
```

The matrix S'-adjoined node is the site at which the moved NP would merge with the sentence. But if the NP is attached to the matrix S', it will form a union with the matrix S' node, in which case it must be interpreted as a semantic conjunct with the matrix S', which it cannot: the predication relation associates it with the embedded S', not the
matrix one.

So, the relative unacceptability of the sentence in which the subject is governed by the matrix verb is due to a conflict between the scope assigned to it by general rules of scope-assignment, and the impossibility for it to form a common phrase-marker with the sentence that it has scope over, given the original relation of predication, which is with the embedded S. Contrasting this example with one in which the scope of the quantifier is not greater than the S' of which it is the subject, as in (70)b, we find a difference, which can be explained in our general account of parasitic gaps.

A connectedness account of parasitic gaps would also have to implement its theory to explain these facts, given that they are independent from connectedness per se: the path of the real gap saves the parasitic gap inside a subject, whether or not this subject is governed from a matrix verb or from the INFL of its own clause. Let me stress that the possibly relative acceptability of (70)a is not a sign of its being a grammatical sentence. If (70)a is relatively only slightly deviant, it is because there exists another LF for the sentence, one in which the relativized NP is assigned embedded scope. It is marginally possible for quantifiers to be assigned narrow scope in Exceptional-Case-marking environments, as shown here:

(72) John believes nobody to have left

It is marginally possible to interpret this sentence as meaning that John believes that nobody left (vs there is no person such that John
believes of that person that he\she left). Theoretically, we have banned the possibility of the narrow scope of quantifiers in such positions. but to the extent that they are possible in effect, this is what will explain the marginal acceptability of the parasitic-gap sentence: if the quantifier is able to be assigned narrow scope, then it may form a conjunct with the embedded clause without violating the condition on semantic conjuncts, saving the parasitic gap.

Since there is much less choice with small clauses, we expect parasitic gaps inside subjects of small clauses to be worse than exceptionally governed subjects, but in fact the contrast is less sharp:

(73)a. This is someone that friends of [e] are angry at t
    b. This is someone that I consider friends of [e] angry at t

This may be because the NP friends of [e] does not constitute an island for the parasitic gap. But if it is an island, then the two sentences get worse:

(74)a. *This is someone that everyone who knows [e] is angry at t
    b. *This is someone that I consider everyone who knows [e] angry at t

So, I will leave these last pairs as problematic. 44

----- ---

44. Chomsky pointed out to me that a complex NP embedded under expect may contain a parasitic gap, as in (i):

(i) John is someone that I expect everyone who knows [e] to like t

I would assume that this is because expect does not necessarily attract quantifiers to the matrix clause. Cf. the greater acceptability of narrow scope of everyone in:

- 283 -
4.1.14 Two differences between parasitic gaps and ATB-gaps

4.1.14.1 Reconstruction in parasitic gaps and ATB-gaps

Syntactically speaking, parasitic gaps and the second gap in an ATB-extraction are the same object at LF: variables forming a chain with a wh-operator, but they are different objects on the surface: an ATB-gap is a variable and a parasitic gap is not. Given this difference, there may exist other distinctions between the two, which, hopefully, could be subsumed under this one.

One difference between a parasitic gap and the second gap of an ATB-extraction is that reconstruction is not possible in the latter, but possible in the former. The phenomenon under consideration now is one called connectedness by Cinque (1983), whereby a moved constituent seems to behave with respect to certain conditions as if it was in the position of its trace. For example, if the moved constituent contains a reflexive, the reflexive is checked with respect to the binding theory as if the constituent that contains it was still in its original position. Consider (75)a-b:

(i) John expects everyone to understand this theory

45. Similar observations have been noted in an unpublished (1983) manuscript, in Hellan (1984), Cinque (1983) and Chomsky (1985b). Thanks to Barry Schein for suggesting sloppy identity, which is a more compelling test for reconstruction than the mere binding of anaphors.
(75)a. I wonder which picture of himself John likes t and Mary hates t

b. *I wonder which picture of himself Mary likes t and John hates t

(75)a is grammatical, which means that himself is properly bound in the sentence. It is not in a position in which it can directly be bound, so it must be bound by being connected to one of the gaps, at least. Given that only one of the gaps leads to a well-formed reconstructed binding relation namely, the first gap. This indicates, first, that reconstruction is possible in the first gap of an ATB-extraction, and second, that reconstruction is not necessary in the second gap of an ATB-extraction. Applying the same reasoning, the ungrammaticality of example (75)b shows that reconstruction is obligatory in the first gap of an ATB-extraction, and that, if reconstruction was possible at all in the second gap, it is not enough to make the sentence good. Given that we are more concerned with the behavior of the second gap than the first one, the binding of reflexives is not very telling, since reconstruction in the first gap is necessary and sufficient, making the acceptability of the reflexive depend solely on the first gap. We must use another test, the sloppy reading of the reflexive. Consider (76):

46. Or, more generally, the sloppy reading of a pronominal, since elements which give rise to sloppy readings are bound pronominals or anaphors. As with anaphors, this binding requirement forces them to be connected, in Cinque's sense, to the gap. See Ross (1967), Williams (1977), Sag (1976) and Reinhart (1983), for the interpretation of sloppy pronouns. (But see section 3 for a new account of sloppy identity.) Along these lines, the behavior of pronouns bound by quantifiers like every... are a good test, also, since they require syntactic binding.
(76) I wonder which picture of himself Peter likes and John hates.

The reflexive can be interpreted as sloppy, that is to say, as bound by the subject Peter in the first conjunct and the subject John in the second conjunct. The fact that the reflexive can be interpreted as bound by John shows that reconstruction may apply in the gap of the second conjunct. Since otherwise, this interpretation would be impossible. Since binding implies c-command, by being connected to the second gap, John c-commands the reflexive. (I will be more precise as to the form that reconstruction should take. in order to make notions like c-command work, when c-command does not directly hold in the structure.) Now consider reconstruction inside a parasitic gap:

(77) I wonder what picture of himself John looked at before Peter destroyed [e].

In (77), it is impossible for the reflexive to be interpreted as sloppy, that is to say, as bound by Peter in the adjunct. This shows that reconstruction cannot apply inside a parasitic gap. Otherwise, reconstruction would make the reflexive in a position such that Peter properly binds it, as was the case with ATB-gaps.47

At this point, the theoretical questions are numerous. First, what

47 Note that there is a difference between a parasitic gap and a VP-gap, since VP-deletion may lead to a sloppy reading of the anaphoric element: as in (i):

(i) Which of his pictures did John look at when Peter did [e]? This is discussed in section 4.2.6.
is reconstruction, second, why do first gaps in ATB-extractions or parasitic-gap sentences require reconstruction and not the second gaps, and third, why cannot parasitic gaps allow reconstruction? For reasons of organization, we will answer the first question here. We then will have nothing to say about the second question. And as for the third question, we will postpone it until the section in which we investigate the phenomenon of sloppy identity (section (4.3)).

So, let us consider the technical question of reconstruction.48 First, in van Riemsdijk and Williams's (1983) NP-structure model, connectedness phenomena can be accounted for by requiring that reflexives be bound at NP-structure. NP-structure is a postulated level in between wh-movement and NP-movement. However, Chomsky (1981) notes that certain connectedness effects hold between constituents which cannot be related by reconstruction, as in cleft sentences like (78):49

(78) [PRO to be 18 years old] is what everyone wants most

But there is a more compelling question than the binding of the anaphor, as we said, which is that this anaphor may have a sloppy interpretation. Given that this interpretation is obtained from LF, in a structure in which it is represented as bound by the subject NP, this

48. See Hellan (1984) for a theory in which reconstruction for anaphors, consists in ruling them in at each level at which they are properly bound. And Kuno (forth.) brought to my attention by K. Johnson; and Johnson (1985).

49. Such a construction is examined at length, with respect to the connectedness effect, in Barss (1984). See Higgins (1976) for a syntactic and semantic study of cleft and pseudo-cleft constructions.
means that, at LF necessarily, the reflexive must be c-commanded by the two subjects, hence forcing a representation in which the anaphor is connected to the gap at that particular level. If one wants to maintain that connectedness signifies applying the relevant rules at the level at which the NP is in its pre-wh-movement position, then one is lead to saying that the sloppy interpretation is obtained from a level before wh-movement applies, but this is contrary to the concept of the form of the grammar, where interpretation takes place from LF. (Moreover, we will see that the sloppy reading is obtained from a very special representation, which does not usually occur on the surface, hence making difficult any attempt to obtain sloppy reading from the structure that sentences have on the surface.) One could then presume that some global notation could make the interpretation sensitive to the pre-wh-movement structure, but we should avoid global notations. This means that the two gaps must be related to the wh-constituent at LF, the level from which sloppy identity is obtained. And if reconstruction cannot be viewed as insertion of lexical material back into a trace, given Chomsky's (1981) and Barss's (1984) argument, then we can adopt Higginbotham's (1980) or Barss's view that binding of anaphors (and variables as well, one should add) is defined with respect to a chain, rather than with respect to a single formative of the chain. Barss

50. To be more accurate, since we adopt the theory in which anaphors move at LF, the LF representation of the NP which PRO picture of himself is one in which the reflexive adjoins to N' and is bound by PRO, subject of the NP. (See Chomsky, 1985a.) Then, the proper interpretation is obtained by control.

51. As for the connectedness effects, Szabolcsi (1985) interestingly shows that not all A'-dependencies allow them, and that all those which
(1984) shows that connectedness effects may be rendered by utilizing the presence of traces of movement directly relevant in the definition of principle A, the condition on anaphors.52 I will give his definition, and then will transfer it into that of c-command, in order to allow the phenomenon of reconstruction to encompass the behavior of all anaphoric elements.53

(79) **Principle A (Barss, (22))**

If A is an anaphor, A must be linked to one and only one phrase B, where:

(i) The Container of $A = a_1, a_i \in C$, C an $A'$-chain

$$(a_1, \ldots, a_i, \ldots, a_n)$$

(ii) B c-commands $a_i$ in the governing category for $a_i$.

Let us change this and express the effects of reconstruction in the definition of c-command. Then, some additional statements will be added, to define the domain in which an element bound by reconstruction is in fact bound.

<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>None of these do not have the form of the insertion of a variable inside the domain of an operator, which is logically improper.</td>
</tr>
</tbody>
</table>

52. Higginbotham (1980) formalizes variable-chains such that any relation to a variable is a relation to any member of the variable-chain. These chains are not the usual A or A'-chains but they allow binding of a pronominal by a quantifier through the variable-chain that this quantifier binds.

53. This definition mentions A'-chains only, but a certain number of facts show that reconstruction is also sensitive to NP traces. Like when the trace of en seems to be bound by reconstruction:

(1) "Plusieurs t_1, en, ont été retrouvés t_j"  
'Many of-it have been found'
C-command by reconstruction (definition)

$X$ c-commands $Y$ if $X$ c-commands $Y$ or if $X$ c-commands $Z$, $Z$ a
formative of an A'-chain that contains $Y$.

If $Y$ in this definition contains $W$,
(i) $W$ is c-commands by $X$ in case $Y$ is c-commands by $X$ and
(ii) the domain in which $W$ is c-commands by $X$ is: the
category $Y$ + the domain which dominates $Y$.

(We will not discuss (ii) in the text. It is meant to define the domain
in which an element is bound, if it is contained inside a constituent
which is itself "reconstructed", that is to say, evaluated with respect
to c-command with respect to some potential binder. When talking about
sloppy identity, we will use pronouns, and maybe anaphors, but we will
not talk about the binding domain, implicitly admitting that the
pronouns can be shown to be free in their GC and the anaphors to be
bound in it.) Given this definition, if $Y$ heads a double chain, and
hence forms a chain with two gaps, and if the two gaps themselves are
c-commands by distinct quantifiers, then the head of the chain is bound
by the two quantifiers at the same time, by reconstruction. Moreover,
if the head $Y$ contains a variable, as in which PRO picture of himself,
where PRO plays the role of the variable, then this variable is also,
simultaneously, c-commands by the two different quantifiers, and hence
bound by them, via the two gaps, even though it materially is a single
formative and is not directly in the scope of the quantifiers, under the
previous, former, definition of c-command.

Now, returning to the difference with respect to reconstruction
between the second gap of an ATB-extraction and a parasitic gap, let us
consider what the difference could be, under the consequence of the
statement of reconstruction as a property of formatives of chains and our assumption that well-formed chains must have their links obey subjacency: the ATB-gap is subjacent to the wh, hence if falls under reconstruction, but the parasitic gap is not, hence it does not fall under reconstruction. However, this is not as simple as it seems. Parasitic gaps are not subjacent on the surface, but they are, at LF. This means that reconstruction, which is here expressed as an extended notion of c-command, should not apply at LF, when the parasitic gap forms a chain with the wh-quantifier. However, sloppy readings are derived from LF-representations, not S-structures, as proved by the fact that VP-deletion, which does not get filled with lexical material before LF', gives rise to sloppy readings. So, reconstruction should apply at LF. Hence a contradiction. We will return to this problem in section 4.3.8.54

The other question is why reconstruction is obligatory in the first gap of a conjunct, and not obligatory in the second. Evidence for this is repeated below:

(81)a. I wonder what picture of himself John looked at and Mary destroyed

b. *I wonder what picture of himself Mary looked at and John destroyed

I have no answer to this question. As we saw earlier in section 11.12.

54. Pending a better solution one could use a global constraint, that would prevent a former non-subjacent gap from allowing reconstruction. Or one could order LF-derivations, in a manner such that reconstruction precedes the formation of the chain of the former parasitic gap.
concerning the fact that the depth of embedding of the first gap dictates that of the second gap, it seems that the first conjunct has precedence over the second conjunct, in the sense that all of its formatives must be relevant to the well-formedness of the sentence, whereas this is relaxed for the second conjunct.

4.1.15 Parasitic gaps cannot be subjects

The second difference between parasitic gaps and the second gap in an ATB-extraction is that parasitic gaps cannot be subjects except in small clauses, but ATB-gaps may, as shown below.55

(82) a. I wonder who John met t and thought t was intelligent
   b. *John is someone who Steve liked t even before thinking [e]
      was intelligent
   c. *John is someone who Steve liked t even before believing [e]
      to be intelligent
   d. John is someone who Steve liked t even before finding [e]
      intelligent

According to Cinque (1984), parasitic gaps are locally A'-bound empty pronominals. Now, locally A'-bound pronominals are resumptive pronouns, and resumptive pronouns are excluded in subject positions (of sentences)

55. This characteristics discovered by Taraldsen (1981), is discussed at length in Cinque (1984). Postal (1972) discovered the difference between infinitivals and small clauses in tough-constructions, and Barry Schein the fact that the gap in a tensed clause is as unacceptable with a complementizer as without, in tough-constructions. Cinque's analysis unifies these facts.
which are adjacent to operators in COMP. The behavior of parasitic gaps is thus subsumed under the behavior of resumptive pronouns at large. We cannot adopt this account, since categories are not identified at S-structure, so we must leave it as an open question.

Johnson (1985) notes examples which can also be explained by the can on empty pronominals in subject position:

(83)a. *Who t met you before [e] talked to you
   c. *Who did you meet t before [e] talked to you

Kayne (1983) can explain these by connectedness, since connectedness subsumes the ECP, and these are ECP violations on the side of the parasitic gap. However, for us, the ECP applies at LF only, when the parasitic gap forms a chain with the wh operator. And at LF, it respects the ECP, since it becomes a normal subject gap bound from the nearest COMP. At S-structure, it just needs to be bound by the operator, so the ECP does not apply then (see principle (85)). Still, the ECP cannot handle (82)c, indicating that something else is happening here.

4.1.16 Licensing parasitic gaps

In our account, it is not clear why parasitic gaps are licensed by quantifiers overtly moved, as has been observed in the literature, and illustrated in the contrast below:

----------

(84)a. What combination did you try to before Ramuald discovered [e]?
    b. *Who tried what combination before you discovered [e]?

Following Chomsky (1982), let us assume that empty categories have to be licensed at S-structure.57 If A'-binding licenses an empty pronominal at S-structure, then this would explain why wh movement is obligatory:58

(85) Licensing of empty categories
    At S-structure, governed empty categories must be licensed by binding.

This almost concludes our discussion of parasitic gaps. Let us simply mention now that one fact that this analysis can capture is the contrast between parasitic gaps and pronouns with respect to weak crossover.

4.1 17 No weak crossover effects with parasitic gaps

Consider (86)a and b:

(86)a. *Which student did your attempt to talk to him scare t to death?

(Engdahl, 1983: (58))

b. Which student did any attempt to talk to [e] scare t to death?

57. On a different line, see Johnson (1985), who deduces the licensing of the parasitic gap from Case theory.

58. A'-binding does not necessarily mean forming a chain with. So parasitic gaps are A'-bound, but they are not variables.
The reason of the contrast lies on the nature of weak crossover. Weak crossover is the phenomenon where an anaphoric element cannot be interpreted as bound by a quantifier, which itself means that it cannot be assigned the value attributed to another element in an A-position, its antecedent. The main type of sentence that it is supposed to account for is the following:

(87) *Who did his brother help t?

The question is why cannot the pronoun be assigned the value of the variable object of help. Certain authors, like Koopman and Sportiche (1983) or Safir (1984) claim that weak crossover is due to an illicit local binding between an operator in an A'-position and two anaphoric elements. If weak crossover does not say anything about the relation between elements in A'-positions and anaphoric elements, but rather is preoccupied only with elements in A-positions (as in Reinhart (1977)), then the reason why parasitic gaps do not give rise to weak crossover effects is straightforward. At LF, the parasitic gap becomes an ATB-gap. ATB-gaps, like all wh-variables, are assigned values directly, by affixing one to the operator. They are not dependent on elements in A-positions. Hence they are not subject to weak crossover. This account, naturally, locates weak crossover at LF. At S-structure, categories are not identified, so they cannot be subject to conditions on their reference. To make this account complete, one could define

59. Others claim that local binding of a pronominal by an operator is what is responsible for weak crossover, like Milner (1982) and Sells (1984).
weak crossover as the requirement for an anaphoric element which is to be interpreted as a variable bound by a quantifier to be A-bound by the variable that this quantifier binds.\textsuperscript{60} We now turn to pronouns of laziness, and the phenomenon of sloppy identity at large.

4.2 Pronouns of laziness

4.2.1 The similarity between parasitic gaps and pronouns of laziness

We have just seen that parasitic gaps are ATB-gaps at LF. The purpose of this section is to argue that pronouns of laziness are similar to parasitic gaps in all respects but one. First, it will be shown that pronouns of laziness can only occur inside conjuncts, and second, that they are quantifier-bound in an across-the-board manner at LF. However, contrary to parasitic gaps, which may occur in LF-conjuncts, pronouns of laziness can only occur inside overt conjuncts, and not LF-conjuncts (that is to say, adjuncts). And the connectedness effects so characteristic of parasitic gaps are not duplicated for pronouns of laziness. These differences will be explained according to the difference between wh-movement and QR: one is to COMP, limited by

---

\textsuperscript{60} This is a translation of Reinhart's (1977) or Haik's (1984) account. Empirical problems arise, like the following, pointed by May (1977) and (forth.):

(i)a. Everyone's brother helped her
   b. Somebody in every\underline{city} despises it

See Gueron (1984) for a solution for (i) at least, where the definition of scope of a quantifier allows everyone to have her in its scope.
subjacency, and the other is to INFL", limited by principle A. We will also see that, in order to account for the well-known "paycheck"-sentence, we will use Goodall's representation of coordinates as unions of phrase markers in the same way that dealt with parasitic gaps inside subjects.

First, let us characterize pronouns of laziness. These pronouns were introduced in linguistic theory by Karttunen (1969), in the so-called paycheck-sentence:

(88) The man who gave his paycheck to his wife was wiser than the man who gave it to his mistress

The special property of the pronoun it consists in referring to the second man's paycheck, as if it stood for the expression his paycheck, where his would be a sloppy pronoun. In the literature, one often comes across pronouns of laziness which have as their antecedent an NP, like the paycheck-NP, but the antecedent can actually be any other type of constituent, so long as this type of constituent can have an anaphoric function, like a pronominal, an anaphor or an epithet. If the antecedent contains a pronominal interpreted as a variable (yielding the sloppy reading), then the pronoun related to this antecedent is a pronoun of laziness. Consider the following:

(89)a. The man who is convinced that someone reads his mail is more nervous that the man who does not believe it

b. The man who [had paid his taxes] was more upset than the man who had not done it\so

c. The man who sleeps in his office is happier than the man who dozes there
The antecedent of the pronoun is an S' in (89)a, a VP in (89)b, and a PP in (89)c. For reasons of space and time, I will not give examples of all these types in the discussion, but will confine the matter to pronouns having NPs as their antecedents. Then, in section 3.1. I will show that all sloppy identity should be treated the way sloppy identity is treated with NP-pronouns of laziness. That is, sloppy identity "is read" inside anaphoric elements, and it is these anaphoric elements which are special: they become ATB-bound variables. So, up to now, we have the following informal characterization:

(90) **Pronouns of laziness (characterization)**
    A pronoun of laziness is an anaphoric element which is interpreted as if it stood for an expression containing a sloppy anaphoric element.

Let us consider the first property of pronouns of laziness, namely, the fact they can only occur in conjuncts, and not in adjuncts or complements. This is an empirical claim, which can be tested immediately:

61. When the antecedent is a VP, the pronoun can be an NP, it or it can be a VP so. When the pronoun is an NP, the sentence corresponds to sentences of the form of (i), where the NP ranges over actions, which may be denoted by VPs:

(i) What did John do t?
(ii) John did something terrible today

There is a difference of use between the pronoun it and the pronoun so as pronouns of laziness, which is that it can be used only if the action denoted by the VP is expected to be performed by everybody. It does not yield very acceptable sentences otherwise:

(iii)??Those who skipped their footnotes were wiser than those who did not do it
The distribution of pronouns of laziness

Pronouns of laziness occur inside overt conjuncts only.

The original paycheck-sentence looks like a plain counterexample, since the pronoun occurs inside a complement, but the parallelism between the constituents which contain the pronoun of laziness and its antecedent will provide the explanation for their grammaticality. We will leave the original paycheck-sentence aside for a moment, and all sentences similar to it, that is to say, those in which the antecedent and the pronoun occur inside structurally similar complex categories like (92) below:

(92) Playing his violin was more fun for Peter than tuning it
for John

Sentences like these will require representations similar to those involving parasitic gaps inside subjects, and we will account for them after the main results have been obtained (section 4.2.8). First, consider pronouns of laziness inside complements:

(93)a. *John told his brother a story that was funnier than what Robert had told him (Robert's brother)

b. *John told his brother more stories than Robert told him (Robert's brother)

c. *Nancy reminded her brother that Lucy was convinced that he would be drafted (Lucy's brother)

d. *Linda did not want to leave her hometown, because Elsa

62. Extraposed sentences should be tested, in that respect.

63. (93)a is also excluded because of subjacency, since the pronoun is embedded in two islands. This is why (93)b is added, where the pronoun is embedded in only one island.
said that she missed it (Elsa's hometown)

Now, consider sentences in which the pronoun occurs inside an adjunct containing a movable clause. These were acceptable constituents for parasitic gaps, as we saw, but they are not for pronouns of laziness:

(94)a. *John told us his life-story before anyone could remember it

b. *Bob thinks his exam was a disaster but Tom would never imagine that it could be

And lastly, contrary to parasitic gaps, pronouns of laziness are unacceptable inside relative clauses, when no structural parallelism obtains between this NP and the constituent which contains the antecedent:

(95)a. *The man who fixed his radio told John about it

(John's radio)

b. *The man who had fixed it told John about his radio

(the man's radio)

We will see that this is a locality effect.

As we said, pronouns of laziness are somewhat acceptable inside overt coordinate structures, and here are some examples: 64

64. An interesting fact: the more generic the sentence, the better the lazy interpretation. Also, inalienable possessions or part-whole relationships sound better than alienable possessions. In general, the sentences are improved when it is pragmatically presupposed that there is a one-to-one correspondence between the thing possessed and the possessor. Hence paychecks and glasses are more felicitous than dollar-bills or suitcases, etc... These observations indicate that the notion of natural relations (like inalienable possessions) play a role in the acceptability of the sentences, and perhaps are at the root of it. See Kempson (1984) for potential extensions of these remarks.
(96)a. John reads his newspaper in the morning and Peter glanced at it in the afternoon

b. Mary has to put her glasses on even to answer the phone, and Jane would not wear them to see the blackboard

c. Bob thinks his exam was a disaster and Tom is certain that it isn't

Let us now turn to the subjacency effects.

If these pronouns are parallel to parasitic gaps, they should not be allowed inside islands. Since we have just seen that they are not acceptable inside adjuncts, we have to test overt conjuncts with an island in the second conjunct. Consider the following sentences:

(97)a. *John told us his life-story and Bill told us what it was (Bill's life-story) too

b. *Bob thinks his exam was a disaster, and Tom is sure of the fact that it isn't (Tom's exam)

c. *I was reminding John of his friend, whereas Peter was wondering when he would see him again (Peter's friend)

Now, let us look at the LF derivation of these sentences. First, the claim is that pronouns of laziness are special, they are not like other pronominals. If we express some of the special effects of their referential properties as deriving from the fact that they stand for expressions which contain pronouns which themselves must be interpreted as bound variables, then we can explain why they behave like variables themselves, bound by their antecedent, similarly to sloppy pronouns: their semantic content somehow comprises a variable, and they behave
like variables themselves, a straightforward entailment. 65 That is to say, these pronouns must be c-commanded by the NP which would bind the sloppy pronoun, if pronouns of laziness were represented as full expressions at LF. To illustrate how close this is to the right approach, consider a schematic illustration of this binding phenomenon, where the subjects are $\lambda$-abstracted, and the pronoun of laziness represented as a full NP:

(98)a. John like his radio loud, and Peter likes it low

LF: John $\lambda x$ (x likes x's radio loud) and Peter $\lambda y$ (y likes y's radio low)

b. *John likes his radio loud, and it distracts Peter, low or loud

(98)b is excluded, because the pronoun it is not c-commanded by Peter at S-structure, a requirement on bound pronouns which seems to be stronger in the case of sloppy pronouns and pronouns of laziness than in the case of quantifiers like every N'. 66 So, the above contrast seems to be

65. See Cooper (1979), Hausser (1979), Reinhart (1983), Haik (1984b) for the interpretation of these pronouns.

66. Similarly for "donkey"-pronouns:

(i) *Everyone$_j$ owns a donkey$_i$, but it$_j$ bit them$_j$

This is to be compared to the well-formed:

(ii) Everyone$_j$ owns a donkey$_i$, and they$_j$ like it$_i$

If the indefinite and the pronoun coindexed with it are interpreted as variables bound by everyone, as in Heim (1982), Kemp (1980) and Haik (1984a), then (i) is a case of weak crossover, parallel to (98)b. See Higginbotham (1983) for discussion. "Donkey"-sentences, discussed by Geach (1962), are those in which a pronoun linked to an indefinite denotes the value of a function which is set by assigning scope of a quantifier over the indefinite, as in (i):
strong confirmation of the hypothesis that a pronoun of laziness must be bound by the NP which would bind the sloppy pronoun, if they were represented as full NPs at LF.

4.2.2 Analysis of pronouns of laziness

The LF derivation that we will propose will mainly get the result that the pronoun of laziness stands for the full antecedent-expression, without having to actually replace it by this expression, as a direct consequence of the definition of reconstruction. The antecedent NP (which we will also call the paycheck NP) contains a pronoun to be interpreted as a bound variable (the sloppy pronoun), so it can be treated as a quantified phrase, like all constituents containing bound pronouns. By assigning it scope, it will not only bind its own trace but also the pronoun of laziness, across-the-board. This will form a double chain, with the trace of the paycheck-NP on one side, and the pronoun of laziness on the other. Now, given the definition of reconstruction, the paycheck-NP is understood as reconstructed in both conjuncts, hence his is bound by two NPs. And the pronoun his of the paycheck-NP is bound by the two antecedents, each in its own conjunct. Note that is it acceptable for a variable not to be an empty category, so long as its binder forms a chain with at least one empty category.

(1) Everyone who owns a donkey beats it

67. This is not all. The bound reading of his in the paycheck-NP has to be obtained in some way. We will assume that the binders of his are treated as quantifiers, but this is only a temporary step. In section 4.3.2, we give a full account of sloppy identity, explaining why the binders of his must have the same grammatical function.
So, the derivation is the following: the paycheck-NP, which is the antecedent of the pronoun of laziness, is assigned scope, to be able to A'-bind this pronoun and to form a chain with it. And, in order for the pronoun inside the paycheck-NP to be bound by its two binders, hence yielding the sloppy reading, these binders are also assigned scope. Consider (99) and its LF. Note that this movement contradicts our general view that movement is due to the assignment of scope to some operator. And NPs like John do not contain operators. So their movement is not justified. I will later give a representation of bound pronouns such that their antecedents do not have to count as quantifiers necessarily, which will get rid of the stipulated movement of the binders Peter and John. For convenience, we assume for the moment that they are quantificational.

The sentence under discussion is an S-coordination, so the two Ss are joined on their topmost node, S (the maximal projection of INFL). Now, three quantifiers are adjoined to this S. As a convention, adjunction nodes do not count as distinct from the original node. So, it is possible for the union of the two Ss to merge all the S-adjunction nodes:

68. I repeat that this will be changed, in section 4.3.2.
69. See May (forth.), Chomsky (1985b) for finer distinctions.
(99) John likes his coffee with milk and Peter drinks it black

In such a representation, the common node is the topmost one, that is to say S and the adjunction-nodes S, and the head of S, INFL. The NP his coffee locally binds its trace and the pronoun. And it is subjacent to both of them, so it forms a chain with both. This NP contains the pronoun his, treated as a bound pronoun. This pronoun must be bound by a quantifier, and in fact it is bound by two quantifiers. These are John in the first conjunct, and Peter in the second conjunct.

To close this section, so far, this analysis does not say that the treatment of pronouns of laziness should not entail binding by the antecedent of the sloppy pronoun. On the contrary, it forces this to happen. However, its aim is to claim that the pronoun-of-laziness reading is a reading which can arise only in structures in which the antecedent of the pronoun of laziness (the paycheck-NP) forms an A'-chain with it. Without this chain, the interpretation in question is not derivable. This is thus a more restrictive analysis than what the

70. In fact, the relation should obey condition A, since the NP is QRed, and not wh-moved. More on this in the text.
common assumptions about these pronouns have claimed. We have seen that
the syntactic domain of these pronouns is quite restricted, not
unsurprisingly. We now wonder why the formation of the A'-chain is a
necessary ingredient in the derivation of the interpretation of these
pronouns.

4.2.3 Why pronouns of laziness form an A'-chain

It is interesting to note that, in the general case, a pronoun may be
interpreted as bound by a quantifier, even though the quantifier is
situated higher up in the structure, as in (100):

(100) Every girl thought that Mary said that someone would hope that
she would appear

So, if pronouns of laziness were merely required to be bound by the men,
in the paycheck-sentence, then we would expect the distance between the
pronoun of laziness and these two NPs (denoting the men) to be freer
than it is, but we will see in the next section that it is not. The
fact that it is not indicates that pronouns of laziness are subject to
another syntactic condition than mere binding by the men. Our
suggestion is that they also are interpreted as pure variables, in the
same A'-chain as the antecedent NP, the paycheck-NP. This requirement
has to be justified.

If one assumes that the interpretation of a pronoun of laziness is
done through some replacement analysis, where the pronoun is replaced by
its antecedent, then the justification is somewhat easy. What would
make this chain-formation obligatory could either be the fact that
replacement-rules are not permitted by Universal Grammar, or, less generally, cannot erase lexical material. But insertion-rules are allowed in the grammar, given the antecedent-contained VP-deletion sentences of the form of John talked to everyone who wanted him to, which were shown to be arguments in favor of such rules (the coreference fact was explained as indirectly due to the presence of an empty pronominal after replacement of the empty VP by the antecedent VP). So, one would have to assume that, since pronouns of laziness are overt lexical items, they cannot be replaced by their antecedent. For its stipulatory nature, we will reject this reason. As we will claim, this chain-formation is necessary because there is no other way that the pronoun his can be sloppily bound. Before this, let us turn to one alternative. Cooper's (1979) theory of pronominals, where pronouns denote the value of a function, and this function is interpreted with respect to the context of utterance.

Again, one syntactic problem with this hypothesis is the limitation on the distribution of these pronouns. The fact that these pronouns are more limited than bound pronouns, as in "donkey"-sentences, or simple bound pronouns, is unexplained a priori, since there is no other syntactic condition on the latter pronouns than being in the scope of the NP which binds them. Obviously, this does not show that the hypothesis that these pronouns are functions is false, it just means that it is incomplete. In the present analysis of pronouns of laziness, I have merely shown that the distribution of such pronouns can be explained if they receive a certain syntactic treatment. However, the claim is deeper than this. There still remains to exclude in principle
the possibility of interpreting pronouns of laziness as functions, in the way we interpret pronouns in "donkey"-sentences. It is a particularity of [+definite] pronouns in general to be interpreted as pointing at objects salient in the discourse, as in Heim (1983), Kamp (1980) and Reinhart (1983). These objects can become salient by the use of linguistic expressions. Now, the main difference between an NP like his coffee, where his refers to John, and an NP like a donkey, when this NP is in the scope of a plural quantifier is that the former introduces John's coffee and only John's coffee in the discourse, whereas the latter introduces all the donkeys which are in relation with the wide scope quantifier, as a consequence of the expression of plurality taking scope over the indefinite. That is to say, the situation in which every person is associated with a donkey is described by the truth-conditions of the sentence which contains the two quantifiers, in a donkey-sentence.

So, consider the difference between a pronoun in a donkey-sentence, and a pronoun of laziness:

(101)a. Everyone who owns a donkey₁ beats it₁

b. John likes his coffee with milk and Peter likes it black

In the donkey-sentence, the existence of a function relating people and donkeys is set, by assigning scope over the indefinite to the universal quantifier. However, what the second sentence says is that John is related to his coffee, but no relation is set between people and cups of
coffee. Since there is no relation set between people and cups of coffee, the use of it as "related" to the NP his coffee, when his has the value of John, can only yield the coreferential reading: John's coffee.

Now, another case when a function is set linguistically, not using scope over indefinites, is when an expression of plurality binds a pronoun, as in the following:

(102) Everyone should like his coffee with milk, but it appears that Peter likes it black

The pronoun it looks like a pronoun of laziness, but it in fact is a donkey-pronoun: it denotes the value of the function 'coffee-of applied to Peter. The use of a function is allowed here, because it is introduced formally in the first clause, where everyone binds his, and hence his coffee. By the interpretation of the first proposition, the universe of discourse contains people that everyone ranges over and their cups of coffee.

So, the conclusion is that a pronoun of laziness has an antecedent, and that this antecedent does not necessarily provide a function in the

71. To go back to footnote 63, this is where the genericity of the sentences could come into play. If, for pragmatic reasons, certain objects are easily associated with every individual separately, then the use of an NP like a paycheck-NP could, by the simple mention of the N' paycheck, introduce the relation paycheck-of in the discourse. In that case, the interpretation of a pronoun would be similar to that of a donkey-pronoun.

72. This is evidenced by the fact that Peter must be one of the individuals that the quantifier everyone ranges over.
discourse. So these pronouns cannot themselves denote the value of such (inexistent) functions. Let us now consider the locality effects that pronouns of laziness are subject to. We will see in section 4.3.2 that sloppy identity can be obtained only from coordinate structures, given a particular formal representation of bound pronouns in general.

4.2.4 Pronouns of laziness and locality effects

I repeat that, for the moment, we deal with pronouns of laziness inside simple constituents in conjuncts, and not with pronouns of laziness inside relativized NPs as in the original "paycheck"-sentence. In that case, our account of pronouns of laziness makes a prediction. Since pronouns of laziness are ATB variables in a double chain obtained by QR, the distance between them and the paycheck-NP should obey principle A. This implies that, contrary to ATB-extraction, or parasitic gaps, where the second gap may occur deeply embedded, as shown in (103)a-b, a pronoun of laziness cannot occur further down than the GC of its binder, the paycheck-NP, as shown in (103)c or d:

(103)a. I wonder who John met and Mary thought that she would invite t
   b. Who did John meet without thinking nobody would invite [e]?
   c. *Fred saw his doctor and Mary managed to convince Bob
to pay him a visit (him=Bob's doctor)
   d. *Fred drank his coffee, and I think that Bob just made it
      (Bob's coffee)

The LF of (103)d is (104), with the QRred NP his coffee, too far from the pronoun him, with which it should form an $A'$-chain obeying condition A:
This is why the sloppy reading of the pronoun *him* is not available.

Note that the reason why this LF is ill-formed is not that the pronoun *his* is not properly bound by NP, because in fact, NP could bind it by reconstruction, if the NP *his coffee* formed a chain with *his*.

Another condition—A effect is that a pronoun of laziness looks as if it should occur in the same clause as the binder of the pronoun *his* of the paycheck-NP, as shown in (105):

(105) *John is satisfied with his movie, and Bob suspects that the critics will hate it (it=Bob's movie)*

Again, at LF, the pronoun *it* is too far from its antecedent *his movie*, since it is embedded inside an intermediary clause.
4.2.5 Pronouns of laziness in adjuncts or subjects

Condition A also explains one of the main differences between parasitic gaps and pronouns of laziness, which is that pronouns of laziness cannot occur inside adjuncts or inside subjects acting as GCs for the pronoun. Consider the case of an adjunct:

(106) *John bought his car after Bob got it fixed

The LF of (106) is (107). Remember that I assume that matrix clauses have an implicit COMP position, which can merge with that of the adjunct clause ((17), section 4.1.5). Also, the two indices on the S are the indices of the two S-conjuncts:

(107)

I have not represented the subjects as moved by QR, since the problem resides in the position of his car with respect to it. Given that S is a GC, and that the quantifier his car does not bind it in the S that contains this pronoun, no chain can be formed between the two, so the interpretation of the pronoun as standing for the expression his car,
where his would be bound by Bob, is not available. This is because the
NP his car adjoins to S, and because merger takes place at S', not S,
since the topmost node of the adjunct is an S', not an S. Pronouns
which are not part of such a chain are interpreted either as bound or as
coreferential, which forces them to have the same referent as their
antecedent, but they do not have this third possibility of being
interpreted as with their antecedent reconstructed in them. This
explains why the only conjuncts which accept pronouns of laziness are
those which conjoin Ss. Adjuncts make S'-conjuncts, which do not allow
an external binder to bind inside them while respecting condition A.
Note that this crucially relies on the assumption that the COMP of a
matrix clause is inherently different from that of an adjunct clause.
Otherwise, the COMP would be considered non-distinct, when forming the
union of phrase-markers, and merging could go one more node down.

This claim seems to predict that S'-coordinates do not accept
pronouns of laziness. This is not verified in (108)a or b:

(108)a. ?I think that John started his book and that Peter finished it
b. ?I wonder when Peter lost his wallet and when John found it

But this is verified in (109), where the COMPs are filled by distinct

---

73. This crucially assumes that all sentential adjuncts are S's and not
Ss, an assumption that follows from X-bar theory: any unmarked sentence
is realized as an S'. See also Emonds (1976) and Klima (1965).
Exceptional Case-Marking environments, in which a predicate happens to
govern an S and not an S' may be analyzed either as deriving from
S'-deletion, as in Chomsky (1981), or from the subcategorization
properties of the governing predicate, as suggested in Chomsky (1985a)
and Massam (1985), if S may be subcategorized for.
operators:

(109) *I wonder when Peter lost his wallet and where John found it

And, if the complementizer is non-deletable, i.e. when it is not
governed by the matrix verb, as claimed by Stowell (1981) and Guercn
(1981), the sloppy reading is still available:

(110) ?Paul regrets that Peter lost his job and that John keeps it

There is no way that the GC of an element inside S could be extended to
S', if we do not want to contradict ourselves, given the
ungrammaticality of (111), where the head NP should bind the empty
category in its GC (cf. chapter 1 last section).

(111) *Paul a sa soeur que Pierre amuse
    'Paul has his sister being amused by Pierre'

Now, note that all these sentences are overt coordinates. So, the
explanation will lie in PF (Phonological Form), and the conventions on
the phonological interpretation of conjuncts. We will assume that (110)
is represented as an S-coordination, and not S', as it seems to be:

(112) John lost his job

    Paul regrets that

    Peter keeps it

Goodall seems to say that this structure is unambiguously linearized
this way.
(113) Paul regrets that John lost his job and Peter keeps it

However, it is possible to consider that PF-linearization can phonologically interpret twice certain constituents which belong to the common structure of the coordinates, if they are adjacent to the non-common parallel structures. Such a convention permits the linearization of (112) as, either (114)a or b, in addition to (113):

(114)a. Paul regrets that John lost his job and that Peter keeps it

b. Paul regrets that John lost his job and Paul regrets that Peter keeps it

(The other possibility: Paul regrets that John lost his job and Paul that Peter keeps it is excluded, given the condition that the repeated string Paul must be adjacent to the parallel structures.) So, with this convention, (108)a-b and (110) above may in fact display an S-coordination in syntax, with the complementizer repeated in PF in the second conjunct. And if it is an S-coordination, this solve the problem of the distance between the QRed antecedent and the pronoun. This account explains why the pronoun cannot have the sloppy reading in (109). Since the COMPs dominate distinct lexical items, the conjuncts cannot be S-coordinates, they must be S'-coordinates. Also, note that, in the case of adjuncts, it will not be possible to make them S-coordinates since they are S', and since the COMP of an adjunct is not considered similar to that of that-clauses or matrix clauses.

Let us now consider pronouns of laziness inside relative clauses, as in (115)a-b (repeated from (95)a-b):
(115)a. *The man who fixed his radio told John about it
   b. *The man who fixed it told John about his radio

Since the pronoun, or its antecedent, occurs inside an island, the
distance between the two cannot respect principle A. Moving the
relativized NP to S' and merging the COMPs, as was done for parasitic
gaps, cannot yield a better result than with adjuncts, since the Ss will
still not merge.74

When the structure of the relative clause is identical to that of the
matrix clause, a pronoun of laziness inside the relative clause should
be able to have its antecedent in the matrix clause (or vice versa),
since the two Ss are allowed to merge, along with the COMPs, as in
(116). Consider:

(116) The man who gave her paycheck to Mary also gave it to Suzan
      (Suzan's paycheck)

The LF of (116) involves moving the relativized NP to S', and extracting
the relative clause out of it, to merge it with the matrix S', S and VP
since these are the same. However, (116) is not very acceptable. The
low acceptability of the sloppy reading is due to weak crossover effects
between the binders of the sloppy pronoun and the sloppy pronoun, as in
a parallel sentence with a true quantifier binding the pronoun ?Peter
gave her paycheck to every woman.

74. Pronouns of laziness are a general phenomenon. Island effects
should also be perceptible with indefinite pronouns, but it is not clear
whether there is a contrast there:

   (i) John has a picture of his brother and Peter got one too
   (ii) John has a picture that his brother likes and Peter got one too
4.2.6 No connectedness effects

Pronouns of laziness, contrary to parasitic gaps, are not rescued by connectedness. This is because, again, they would be contained inside an $S'$, which makes the antecedent NP too far for them. Consider (117):

(117) John is satisfied with his movie and Bob$_i$ laughed because he$_i$
hates it

This sentence is excluded at LF, similarly to when the adjunct is not embedded, because the antecedent his movie cannot form a chain with it, even though the embedded clause contains a bound pronoun and may be assigned scope:

(118) [his movie [John is satisfied with t$_j$] and [Bob laughed because t$_j$
[he$_g$, he hates it]$_j$]]

This concludes the first part. Let us note that there exist two other possible analyses.

4.2.7 Two alternatives

As we said earlier, pronouns of laziness could be analyzed as denoting the value of a function taking as its argument the men and as its value the men's paychecks (in the paycheck-sentence), but this is semantically incorrect, and syntactically insufficient.

The other alternative is a syntactic treatment close in spirit to the one that we have adopted in the text: the structure could be
base-generated with an empty operator binding from an adjoined position the two NPs at the same time: the paycheck-NP and the pronoun of laziness. The double chain would be headed by a base-generated operator Chomsky-adjointed to S, in the manner of Huang's (1984) analysis of pro in languages in which pro may occur in all argument positions. According to Huang, this pronoun is bound by an operator which licenses it. So, the S-structure of a sentence like (119) would schematically be the one below:

(119) John likes his coffee with milk, and Peter drinks it black

\[
\begin{array}{c}
S \\
\downarrow \\
Op \\
\downarrow \\
S \\
\downarrow \\
NP \quad VP \\
\downarrow \\
NP \quad VP \\
\downarrow \\
\ldots \text{his coffee} \ldots \quad \ldots \text{it} \ldots \\
\end{array}
\]

The semantic rules would interpret such a structure with all elements in the chains as identical to each other, so *it* would be interpreted as identical to *his coffee*.

This analysis is less appealing than the one in which the antecedent of the pronoun of laziness is treated like the quantifier which forms the chain with the pronoun because it has to stipulate the existence of such base-generated A'-chains.\(^5\) We now consider the classical

\[^5\text{However, the base-generation of the empty operator is not to be rejected, following Bowers' (1982) ideas, it could account for the distribution of epithets, like the fool, in an example pointed out to me by Haj Ross and en, as studied by Carden and Lamiroy (1984):}

(1) My brother, *ids* in every project that the idiot, thinks will make him rich
4.2.8 Paycheck-sentences

The classical paycheck-sentence cannot be accounted for by our treatment so far, since the antecedent NP is in an island, which makes it too far from the pronoun. The desired LF is one in which the antecedent NP c-commands both its trace and the pronoun, so it has to be extracted from the subject of the sentence that it occurs in, as in (121):

(120) The man who gave his paycheck to his wife was wiser than the man who gave it to his mistress

(121) [[[his paycheck][the man who gave it to this wife] was wiser than the man who gave it to his mistress]]

In (121), the relation between the paycheck-NP and its trace or the pronoun of laziness violates both condition A and subjacency, since this NP is extracted from the relative clause. Similar examples, in which the pronoun and the antecedent belong to subjects, accept pronouns of

(11) Pierre, ne sait pas que Marie connaît quelqu’un qui en i pense du bien
   'Pierre does not know that Marie knows someone who thinks well of-him'

Carden and Lamiroy have shown that en was acceptable in domains non-subjacent from its antecedent, and the same kind of generalization seems to hold between epithets and their antecedents (the facts are not so clear, though). If this was a correct empirical generalization, then the licit domain of occurrence of epithets or en could be delimited by making them A'-bound by the empty operator. Like all variables, they would have to be free in the domain of their operator, only. See Bowers (1982).

paycheck-sentences.
laziness, as in (122):

(122) [PRO running his experiments] is more hazardous for John than [PRO ruining them] for Peter

Let us consider the acceptability of pronouns inside island subjects, as exemplified in the original paycheck-sentence. The problem is to reconcile the account given earlier of pronouns of laziness with the fact that the general condition on chains is not respected in the LF of the paycheck-sentence. Similarly to parasitic gaps inside subjects, the solution resides in the fact that the structures in which the pronoun and the antecedent occur are identical structures which can form conjuncts at LF. In paycheck-sentences, the comparative clause is a bare XP comparative, where the comparative preposition governs a bare NP, interpreted as the subject of a predicate identical to that of the matrix clause: is x-wise. The derivation of such sentences involves the creation of an LF coordination, where the matrix clauses sustains the second conjunct of the comparative construction. And since the two subjects have parallel structures, they can also merge, at least down to the S level. Then, the moved paycheck-NP will be able to form a licit double chain with its trace and the pronoun of laziness, as in the following LF (the two relative clauses are represented as merged down to the VP level):
In such a representation, not only both NPs are fully represented, and hence both fully interpretable, but also the elliptical comparative clause, which becomes represented materially by being matched with the matrix clause, as in the representation of conjuncts in general: the interpretive rules see this structure as the sum of two full conjuncts, and not as one sentence only. (See also section 4.5.)

To summarize, our analysis explains why a paycheck-NP can bind from outside its clause: in coordinate structures, most of the nodes which form island become superimposed by the nodes of the other structures, making it unnecessary for the NP to actually be extracted from its clause. So, since islands neutralize each other by forming conjuncts, we expect them to act as real islands if there is no counterpart in the other conjunct. This is true, as shown in some minimal pairs below:

(123)a. *The man who gave his paycheck to his wife was wiser than \[ \text{the person \[}_s \text{ who knows the man who gave it to his mistress]}. \]

b. ?The person who likes the man who gave his paycheck to his wife is more relax than the person who likes the man who gave it to his mistress.
c. ??[Talking to [someone who does not like his ideas]] annoys Mary more than anyone who is proud of them

d. Talking to someone who does not like his ideas annoys Mary more than talking to someone who is proud of them

The bracketed constituents indicate the structure which is part of one conjunct and not the other, hence constituting an island. In the counterparts, these two islands disappear by being matched with the nodes of the other conjunct.76

Let us now consider whether the condition on semantic conjuncts holds for pronouns of laziness.

4.2.9 Pronouns of laziness and the condition on semantic conjuncts

Consider (124):

(124) [The man who gave his paycneck to his wife]₁ likes [the man who gave it to his mistress]₂

One first question is to identify the two conjuncts. The only possible one is the coordination between the subject and the matrix clause, but they do not provide parallel structures. The other two parallel

76. One problem is that it is sufficient that the structure be the same in the two conjuncts, without it having to be the case that they dominate the same terminals, for the unacceptability of (123)a and c to reduce. For example, (123)c constrasts with (i) as well:

(i) Talking to someone who does not like his ideas annoys Mary more than disagreeing with someone who is proud of them

These cases are similar to Pesetsky's (1982) or Cowper's (1984), where subjacency effects are reduced in identical structures, without it having to be the case that the lexical items be the same, section 4.1.
structures are formed by the NPs. But it is impossible to merge the NPs without violating the th-criterion, since the object NP will have to move onto the subject position, in order to merge with the subject NP. So, one other possibility is that the relative clauses themselves merge. Let us assume that extraposition extracts the relative clauses out of the NPs, and that they merge, as in (125):

(125)

```
[Diagram of tree structure showing the merging of relative clauses]
```

Note that, since we merge only the relative clauses, it is not clear whether this violates the condition that LF-conjuncts must be semantic conjuncts. Since these clauses have nothing to do with each other, it is possible that they are neutral with respect to it. Other examples are the following, which are to be compared with the parasitic-gap examples (49) of section 4.1.9:

(126)a. *John sold his book to more people than Peter gave it to

b. *John sold his book to the man who sold it to Peter

( the man's book )

If these are bad, this shows that the condition on semantic conjuncts also holds with pronouns of laziness. This would be a nice result,
since it would show that there must be a one-to-one correspondence between the form and the interpretation of LF's with conjuncts.\footnote{But VP deletion sentences (section 4.3.2) and free relatives (section 4.4) seem to violate this condition. So, it is debatable whether there is a one-to-one correspondence between the form and the interpretation of conjuncts.}

Before we turn to another topic, let us consider an alternative approach to that adopted in the text, which makes use of the concept of indirect binding.

4.2.10 A indirect-binding approach to pronouns of laziness

One of the properties of indirect binding is that respective scope between two quantifiers may be defined as a binding relation, whereby the wide scope quantifier binds the narrow scope quantifier. This binding relation is reflected semantically by treating the bound NP as a function of the wide scope quantifier, that is to say, as a complex variable of the form $f(x)$, where $x$ is the variable bound by the wide scope quantifier. For example, in (127), if the underlined NP is interpreted as in the scope of every woman, then it may be interpreted as a function of every woman, and it is syntactically bound by it, and the same thing holds of (127)b, if the pronoun is interpreted as bound by the universally quantified NP:

(127)a. Every woman met someone

b. Every woman met her former classmates

This binding relation between a wide scope NP and the NPs that it has in
its scope explains the distribution of pronouns coreferential with the dependent NP: if the dependent NP is interpreted as a function of the wide scope NP, it has to occur in its scope, since it semantically contains an occurrence of the variable bound by the wide scope quantifier. So, if a pronoun is coreferential with that NP, then the pronoun also has to occur within the domain of binding of the wide scope NP, by the condition on variables: all variables must occur in the domain of scope of the quantifiers which are supposed to bind them. Such a binding relation describes the possible environments for crossing coreference with definite descriptions and quantifiers (cf. Haïk (1984b)).

Returning to the paycheck-sentence, we note that the pronoun contained in the paycheck-NP is to be interpreted as a bound pronoun, and that the binder of this pronoun is the relativized NP. By having the pronoun his of his paycheck in its scope, and, because this pronoun is contained in the paycheck-NP, the relativized NP also has the paycheck-NP in its semantic scope: the value of the paycheck depends on that of the man. This means that, in a paycheck-sentence, the paycheck-NP is always dependent on the NP which binds its internal pronoun. This entails that indirect binding can always be set between the paycheck-NP and the binder of its internal pronoun. Setting the scope relation between two NPs as a binding relation could serve to help the movement of the paycheck-NP out of its clause, by having the paycheck-NP locally bound in its GC by the indirect binder, in the way that "illicit movement" is saved for Chomsky (1985a) or Lasnik (1985). This would contradict what we said about movement, namely, that it
cannot be saved by intervening binders, if the binders are not part of
the chain, since we set condition A on chain-links, but let us consider
what would happen in a paycheck-sentence, if condition A held only of
the traces, and not of chain-links:

(128) [The man who \( t_1 \) gave [his paycheck] to his wife] \(_i\) is wiser than
[the man who \( t_j \) gave it to his mistress] \(_j\)

As we said, the NPs denoting the men may be the indirect binders of the
NPs denoting the paychecks, since this reflects how the sentence is to
be interpreted: with the paychecks a function of the men. In that case,
it may be possible for the wide scope NPs to bind the trace of the
paycheck-NP and the pronoun of laziness at LF. Consider the LF of (128),
again (the sentence no longer needs to be represented as a union of
conjuncts):

(129)

\[
S \rightarrow \ NP_i \rightarrow S \rightarrow NP \rightarrow VP \rightarrow PP \rightarrow S
\]

his paycheck

\[
NP_j \rightarrow S' \rightarrow V \rightarrow AP \rightarrow P \rightarrow NP
\]

the man COMP S was wiser \( t_1 \) than \( \NP_k \rightarrow S' \)

\[
\text{who NP VP,}
\]

\[
\text{who NP VP,}
\]

gave \( t_j \) ...

gave it \( t_k \) ...

We see that, even if the NP his paycheck has been extracted far from its
original position, the two anaphoric elements indexed \( i \) are in fact
bound in their GC by some indirect binder: \( t_i \) is bound in its GC by the wh trace \( t_j \), and, symmetrically, \( t_i \) is bound in its GC by the other wh-trace. This account, in terms of indirect binding, could work in a theory like Chomsky's (1981) or (1985a), where improper movement can be salvaged by intervening binders. This would account for why the pronoun of laziness and its antecedent must belong to the same minimal clause as the wh-traces, in the original paycheck-sentence.

But one set of facts would still remain to be explained, which is that the two indirect binders must belong to the same minimal clause. See, for example, the island effects, as in (123)c, above: talking to someone who does not like his ideas annoys Mary more than anyone who is proud of them. I will leave this solution unexplored here, even though it looks like a potential answer in a theory which does not have binding principle A as constitutive of chain-links.

4.2.11 Pronouns of laziness as resumptive pronouns

The analysis of pronouns of laziness identifies them as LF resumptive pronouns. Then, if, following Cinque (1984), the impossibility of parasitic gaps in subject positions is due to their status of resumptive pronouns, and if the condition that prevents this situation holds at LF

\[ (1) \text{They think that Mary said that they should visit each other} \]

\[ (1) \text{cannot mean: Peter thinks that Mary said that he should visit John, and John thinks that Mary said that he should visit Peter.} \]
as well as S-structure, then pronouns of laziness should also be unacceptable in such positions. In order to test this, we have to make sure that the pronoun of laziness ends up in the GC of its antecedent at LF, because condition A should hold between the two. So, since a subject position is already in an embedded sentence with respect to the antecedent, the only subject position which obeys condition A is the subject position of an infinitival. Consider (130):

(130)a. John believes his brother to be intelligent and Peter believes him to be handsome

b. John likes his brother more than Peter believes him to be a good fellow

These sentences are good, with the pronouns interpreted as referring to Peter's brother. This shows either of two things. First, this could show that the condition which rules out resumptive pronouns in subject position is an S-structure condition. Or it could show that this condition applies at LF, but that it applies to non-subjacent A'-bound pronouns. This latter possibility is disconfirmed by the data cited in Cinque (1984), where subjacent gaps were excluded in such positions. Hence this condition must be an S-structure one.

4.2.12 Coordination across discourse

I have claimed that the interpretation of pronouns of laziness goes through a syntactic stage at which their antecedent have to take scope

79. Or a small clause, irrelevantly for the test here, since resumptive pronouns are acceptable anyway in such subject positions.
over two conjuncts. The conditions that the relation between this moved
quantifier and the conjuncts are subjects to are syntactic conditions of
sentence-grammar, however these conjuncts may belong to two different
sentences of the discourse, as in (131):

(131) Speaker A: Does John like his job?
   Speaker B: I don't know, but Peter hates it.
   (Peter's job)

This shows that the rule which derives the LF of a clause containing a
pronoun of laziness must have access to the discourse, and that it is
not limited to the sentence.

Williams (1977) argues that rules of Discourse Grammar make the LF of
a sentence have access to the LF of sentences previously uttered.
Following his idea, but not distinguishing between discourse grammar and
sentence grammar, let us assume that sentences previously uttered are
represented as coordinates ordered with respect to each other. That is
to say, if sentence A has just been uttered, and B is uttered, then the
representation of A followed by B is that of two coordinates, dominated
by the node E (for Expression), postulated by Banfield (e.g. Banfield
(1973):

```
   E
  /|
 / |
A B
```

One fact about coordination is that two independent clauses can be put
together without adding any nodes on top of either of the sentences.
And in case they are merged, they share at least their topmost node.
This means that something which could be bound from outside the sentence
if it was not in a coordinate structure could still be bound from outside, in a coordinate structure.

So, a discourse like (131) above can be represented as:

(132)

\[
\begin{array}{c}
E \\
\text{S'1} & \text{S'2} & \text{but} & \text{S'3}
\end{array}
\]

S'1 and S'3 should be represented as conjuncts, so we have to allow forming the union of two conjuncts which are not adjacent, like S'1 and S'3. Since coordination takes place with conjuncts which do not necessarily belong to the same plane, S'1 may be coordinated with S'3. Then, their topmost nodes can be merged, S. In that case, the paycheck-NP may be adjoined to S, and bind its trace and the paycheck-NP in its GC.

The paycheck NPs are only one particular instance of sloppy interpretation. We now turn to this phenomenon in VP-deletion sentences, which is often referred to in the literature by sloppy identity.

4.3 Sloppy identity

4.3.1 The condition on alphabetic variance

According to Sag (1976) and Williams (1977), VP-deletion can take place only if the anaphoric VP is an alphabetic variant of the antecedent VP. This condition is meant to obtain the result that
pronouns which have a sloppy interpretation are bound by the "same NP" in the antecedent clause and in the anaphoric clause, that is to say, an NP with the same Grammatical Function (henceforth, GF) in the two clauses. To illustrate this, consider (133):

(133) Tom told Bill about his book, and Peter did too

Considering the possible sloppy interpretations of his, if (133) is to be interpreted as with Tom the antecedent of the pronoun, then its antecedent must be Peter in the second conjunct, and it cannot be Bill. And if the pronoun is interpreted as bound by Bill in the first conjunct, then it has to be bound by Bill in the second conjunct, and cannot be bound by Peter. In other words, the binder of the pronoun has to have the same GF in both conjuncts.

Eventually, our aim will be to dispense with the condition of alphabetic variance, and provide an alternative analysis of sloppy identity to that of Sag (1976), Williams (1977) and further studies. The point of this section is that the condition that the binder of the sloppy pronoun should have the same GF in both conjuncts also holds for sloppy identity in the case of NP-pronouns of laziness, as illustrated in the following:

(134)a. *Peter told his psychiatrist about his illusions, and Mary asked Bill about them (Bill's illusions)

b. *Peter told his psychiatrist about his illusions, and Bill told his doctor about them (Bill's illusions)

c. *Mary asked Peter about his illusions, and Suzan inquired about them to Bill
In (134)a, the binder of the pronoun is the subject Peter in the first conjunct, and the indirect object in the second conjunct. And this sentence is worse than one in which the antecedent has the same GF in the two conjuncts, as in (134)b, where this antecedent is a subject, and (134)c, where it is an indirect object.

This shows that the condition on alphabetic variance should hold in all cases of sloppy identity, that is to say, sloppy identity induced by pronominal VPs and by pronominal NPs as well. And its domain of application is not only the domain of the "identical material" (the VP for VP-deletion, and the NP in a paycheck-sentence), but also the whole clause which contains the NP or the VP. So, both in VP-deletion sentences and sentences with pronouns of laziness, the sloppy pronoun has to occur in a λ-expression which is alphabetically identical to that which contains its antecedent.

Here, we will argue that sloppy identity in VP-deletion sentences is to be analyzed along the same lines as sloppy identity in paycheck-sentences. So, VP-deletion also gives rise to an LF with conjuncts. We will see that the λ-expression which is relevant in the application of the condition on alphabetic variance, in the commonly assumed theory of sloppy identity in VP-deletion sentences, corresponds to what we define as a conjunct. Describing the facts in mixed theoretical terms, something like the following holds:

(135) Condition on alphabetic variance
Conjuncts must have the same operator-variable structure, when the operator is QRed.
This merely describes the facts, however, it generalizes over VP-deletion and paycheck-sentences. It is also meant not to force wh gaps to have to have the same GF in the two conjuncts, so (136) is well-formed (Goodall, 1984 (154c')):

(136) ?That is the man who I think Mary likes _ but they think _ kissed Jane

Note that this condition not only implies that the binders of the sloppy pronoun have the same GF, as shown above, but is also implies that the moved paycheck-NP must have the same GF in both conjuncts, which it does. (That the binders must have the same GF will partly follow from the definition of bound pronouns, as we will shortly see.)

(137) a. ?A man who talks about his illusions to his doctor is wiser than a man who talks about them to his psychiatrist

(138) b. *A man who talks about his illusions to his doctor is wiser than a man who talks to them about his doctor

In our discussion of pronouns of laziness, we have not shown how the sloppy pronoun (his, in his paycheck) is represented formally. All we said is that it is interpreted as a bound variable, and that its binders were the moved NPs inside the two conjuncts. And if we do not use \( \lambda \)-abstraction to be able to represent the binder of the pronoun as a quantifier, then how can the pronoun behave like a bound variable, that

80. See Hirshbuhler (1983), who proposes that a quantifier inside a VP may bind across-the-board, in VP-deletion sentences, hence binding one position in the antecedent VP and one position in the anaphoric VP:

(1) A Canadian flag is hanging from every window and an American flag is too
is to say, how can it be assigned different values?

4.3.2 The representation of bound pronouns

The aim of this section is to get rid of the necessity of representing a clause as a function of an NP, by \( \lambda \)-abstraction of this NP. This is because \( \lambda \)-abstraction has only one main role in grammar, which is to be able to obtain sloppy readings for pronouns. However, the representation of a sentence as a \( \lambda \)-expression has the power of entailing that there exists a function between the referent of the antecedent of the pronoun and the referent of the pronoun, but this is wrong. When someone says: Diane is packing her books, this person is not setting a function between people and their books.

Let us first give a formal representation of bound pronouns, and then we will see that sloppy identity follows straightforwardly from this representation, given the double chain-hypothesis. We will use Vergnaud's (1982:chapter2) formal definition of addresses in a tree. A tree can be described as a set of nodes, each of which is identified as having a unique address, notated as a number attached on the node, and these are randomly assigned, so long as they are different for each node. Now, pronouns which have a linguistic antecedent can be analyzed as elements with a set of instructions which indicate what the relation is between it and the antecedent. Let us assume that the basic interpretation of a pronoun is the identity function, "is identical to \( (x_i) \)", where \( x_i \) is a category \( x \) whose address is \( i \). So, coreference or
binding is obtained from a representation of the pronoun as the identity function applied to a category with a certain address. So, if she is interpreted as coreferential with NPi, and NPi refers to Mary, then she is interpreted as identical to that NP, and hence refers to Mary. And if everything binds a pronoun it, then it is interpreted as identical to the element which has the address occupied by everything (or its trace, left by application of OR), which is a variable y assigned values. So, the pronoun is interpreted as identical to "y", and is assigned the same values as its antecedent.

Now, the interesting property of coordinate structures is that, often, they have symmetrical structures, in the sense that their shape is the same. For example, if one contains a double-object construction, the other one does too; if one contains a SC construction, the other one does too. etc. We can express this structural resemblance by saying that the nodes of one conjunct have the same address as the nodes in the second conjunct. In his work, Vergnaud (1982) gives the condition that two nodes may not have the same address. However, it is possible to relax this condition in the case of conjuncts, given that the assignment of addresses is intended by Vergnaud to apply within the domain of single sentences. If the assignment of addresses is an algorithm whose domain is the sentence only, then, since coordination puts two sentences together, it is possible that the nodes of one conjunct have the same address as the nodes of the conjunct. Let us assume that this happens under special circumstances, namely, when the nodes can be defined as being in exactly the same environment, where the environment takes into consideration the whole sentence from the node into consideration up to
the maximal dominating node. So, let us assume the following:

(139) **Symmetrical structures**
Two (or more) structure(s) are symmetrical if each node of the first structure has the same address as a node in the other structure(s).

(140) (i) Two nodes have the same address only if they are in the same structural environment.
(ii) The structural environment of a category A is the whole structure that dominates A, and the nodes which are sisters to S.

Note that (140)i does not entail that constituents with the same GF (grammatical function) may always have the same address, because not only their local environment must be the same, but also the complete structure which dominates them, by (ii). So, for example, a matrix subject may not have the same address as an embedded subject, even though they both are in the environment sister to an INFL', since one occurs in an embedded clause and the other does not. However, in coordinate structures, two matrix subjects may have the same address. Moreover, two matrix subjects may have the same address, even if their VPs have different shapes. This is because VPs are sisters to (not counting INFL') to their subjects, which is sufficient, given (ii).\(^81\)

Condition (140) is important. It will explain (section 3.8) the possibility of sloppy readings in ATB-gaps and their impossibility in parasitic gaps. Now, if two nodes have the same address, they are non-distinct with respect to their address. Consider, for example,

\[^81\] As for defining internal arguments of a predicate as occurring in the same environment, which implies for them to have the same GF, more work needs to be done than my sketchy presentation. See Chomsky (1981), Marantz (1984), Baker (1985), Massam (1985) and Perlmutter (1983).
(141):

(141) John sees his psychiatrist every fortnight, and Peter visits his doctor sometimes

We have the choice of representing the two conjuncts either as symmetrical structures (where nodes have the same addresses) or as independent structures. The representation with symmetrical structures is the one which interests us, the following:

(142)

Now, in order to get the sloppy reading of pronouns, as in paycheck-sentences, the pronoun his should be interpreted as a bound variable, that is to say, as identical to an element with a certain address. And, in symmetrical structures, symmetrical elements have non-distinct addresses, which entails that, if the pronoun is bound by one of them, it is bound by all of them. Consider again a sentence with a pronoun of laziness:

(143) John likes his coffee with milk and Peter drinks it black

The LF, with the paycheck-NP moved, is the following:
In the structure above, the pronoun his can be interpreted as bound by NP3, in which case it is represented as i(NP4), where i stand for the identity function. Importantly, note that NP4 does not have to be extracted from the sentence, in order to bind the pronoun, since it c-commands the trace of this pronoun, and hence binds it by reconstruction. Now, NP4 is represented twice in the structure, as John and as Peter, so the value of his is unambiguously John in the first conjunct, and Peter in the second conjunct.

This technical analysis allows us to explain why sloppy identity requires the binders to have the same GF in both conjuncts. In fact, they must have the same address. And this condition is due to the way we define bound pronouns in general: they mention addresses.

Returning to the stipulation that coordinates seem to relax Vergnaud’s condition that nodes may not have the same address, one may wonder whether structures are allowed to merge whenever they are symmetrical, with their nodes having the same address, even if they do not dominate the same lexical material. This would solve Pesetsky’s (1982) and Cowper’s (1984) examples, where subjacency is allowed to be violated in symmetrical structures, even if the nodes do not dominate.
the same material (section 4.1.4). However, we cannot allow for such a possibility for the following reason. We have excluded pronouns of laziness inside adjuncts on the grounds that the paycheck-NP cannot c-command the pronoun, due to the fact that the merging of the two conjuncts can take place only with S', and may not go down to S.

(145) *John drank his coffee before Peter made it (Peter's coffee)

\[
\begin{array}{c}
S' (j) \\
S \\
NP_i \\
| \\
\frac{\text{his coffee}}{S} \\
| \\
\frac{\text{S}}{PP} \\
| \\
\frac{\text{NP}}{VP} \text{ before } t_j \\
| \\
\frac{\text{Peter}}{V} \\
| \\
\frac{\text{NP}}{NP} \\
| \\
\frac{\text{drank }}{t_i} \\
\end{array}
\]

In (145), the moved NP does not bind it in its GC. It would if the S nodes could merge, as in the following structure:

(146)

\[
\begin{array}{c}
S' (j) \\
S \\
NP_i \\
| \\
\frac{\text{his coffee}}{S} \\
| \\
\frac{\text{S}}{PP} \\
| \\
\frac{\text{NP}}{VP \text{ before } t_j} \\
| \\
\frac{\text{John}}{V} \\
| \\
\frac{\text{NP}}{NP} \\
| \\
\frac{\text{drank } t_i}{t_i} \\
\end{array}
\]

This means that we are forced to saying that the conjuncts formed with
adjuncts, like those above, cannot merge more than the topmost node, as we have generally assumed.

To summarize, sloppy identity is obtained in and only in double-chains headed by a constituent which contains the sloppy pronoun, as in the following:

(147)  
\[
\text{[his ... ] antecedent} \\
\text{sloppy \ pronoun of laziness} \\
\text{of pro of \ trace} \\
\text{pronoun of laziness}
\]

1. There must be a chain between the paycheck-NP and the pronoun of laziness, because the sloppy reading of the pronoun can only be obtained by reconstruction (into the position of the two A'TB-variables), and reconstruction is defined on formatives of chains only.

2. Paycheck-NPs must head two chains, because binding of a single pronoun by two or more NPs is possible only if the NPs have the same address (given that no function has been set by the relevant linguistic elements). And non-distinct addresses are possible only in conjuncts.

Let us now turn to the analysis of sloppy identity in VP-deletion sentences.

4.3.3 Sloppy identity and VP-deletion

Empty VPs are anaphoric elements. As was the case with pronouns of laziness, if the antecedent contains a pronoun, the sloppy reading can be obtained by a double A'-chain, headed by the antecedent VP. Consider
the following sentence:

(148) John likes his coffee black, and Peter does too

The sloppy reading of his is obtained from a representation in which it is interpreted as the value of the identity function applied to the element which has the address of the subject. In order to get it bound in the two conjuncts at the same time, these conjuncts must be represented as symmetrical structures. By convention, let us assume that, since the equivalent of the paycheck-NP is a VP, it is the VP which is assigned scope and heads a double A'-chain. So, the LF of (148) is (149):

(149)

\[
\begin{array}{c}
S1 \\
\text{likes his coffee black} \\
\text{NP4 INFL'5 and NP4 INFL'5} \\
\text{John INFL6 VP7} \\
\text{Peter INFL6 VP7} \\
t_i \text{ does } t_i \text{ too}
\end{array}
\]

In (149), the pronoun his inside the moved VP can be represented as \(i(NP4)\), in which case, it is interpreted as identical to John in the first conjunct, and to Peter in the second conjunct, being bound by them by reconstruction. Of course, binding still requires c-command at LF, to get the right interpretation, so John and Peter must c-command the variable.\(^82\)

\(^82\) Cf. the ungrammaticality of (i), where Peter and Bill do not c-command the subject position:

(i) His radio annoys Peter and it bothers Bill too
Let us now determine the syntactic domains in which sloppy identity is possible in VP-deletion sentences.

4.3.4 The syntactic domain of sloppy identity

The central hypothesis about the interpretation of the VP, in VP-deletion sentences with sloppy identity, is that the VP is assigned scope by wh-movement, and not by QR. So, the distribution of sloppy identity in VP-deletion sentences is more similar to that of parasitic gaps than pronouns of laziness. The reason why it is like wh-movement may be that this is the LF counterpart of VP-fronting, which is a case of wh-movement:

(150)a. John wanted to clean his desk and clean his desk he did [e]

b. John wanted to clean his desk and clean his desk I think he did [e]

c. *John wanted to clean his desk and clean his desk I can remember when he did [e]

(150)b shows that the movement can be long-distance, and (150)c shows that it is limited by subjacency. So, given that predicates can move to COMP in syntax, they are allowed to do so at LF (presumably because they bear [+wh] features), which they do in the derivation of VP-deletion with sloppy identity. As a parenthesis, note that, as we said above,
since sloppy readings are obtained when the pronoun is represented as a bound pronoun, the binder of the pronoun must c-command it, by reconstruction or directly. So, the binder must c-command the VP that contains the sloppy pronoun, by reconstruction or directly.

Sloppy identity crucially depends on the possibility for the moved VP to form a chain with its trace and with the empty VP. This can be done if the moved VP is subjacent to them. If it cannot be subjacent directly, the clause that contains its trace and the clause that contains the empty VP have to be made conjuncts at LF. We may wonder whether conjuncts can be formed with non-semantic conjuncts, in this case, and the answer is yes. Consider the following contrast, where VP-deletion is acceptable inside complements, as illustrated in (151)a, but not parasitic gaps, as in (151)b:

(151)a. John introduced his brother to [everyone Peter did [e]]

b. This is the man that, John introduced t₁ to [everyone that, Peter introduced [e] to t₂]

(151)a is interpretable as: "John introduced his brother to everyone that Peter introduced his (own) brother to", but (151)b is unacceptable. These facts are the main reason for stating the condition on semantic conjuncts as applying to chains already formed at S-structure in the main clause. In the particular sentence under discussion, (151)a, the VP forms a chain with two positions, but at LF only.

We will also generally assume the following:
(152) Convention on Predication
If a predicate has the same address in two (or more) conjuncts, its subject does too.

Below is the LF of (151)a:

(153) 

In (153), the VP in COMP trivially has the same address in the two conjuncts, since it occurs in the non-distinct node. By convention (152) above, its subjects may have the same address. The pronoun inside it is interpreted as the identity-function of the subject, which, in this case, has the same address in the two conjuncts. So, the pronoun gets simultaneously bound by the two subjects, by reconstruction. Each subject c-commands a variable which forms a chain with the constituent which contains the pronoun. As for the well-formedness of (153), the operator Op$_j$ of the relative clause binds the wh variable inside the VP which is in COMP, and which belongs to the common structure. So, the relation between Op$_j$ and some variable respects the coordinate structure constraint. And the VP in COMP binds a variable inside each of the conjuncts, so it also respects the condition.

To conclude, without the movement of the VP to a position such that it has the same address in the two conjuncts, sloppy identity would not
be possible, since the VP comes from an embedded sentence, subject to
Vergnaud's condition that, in a single sentence, nodes may not have the
same address. However, forming a conjunct at LF is the formation of two
sentences, so the structures may be symmetrical. Given that the VP
shares the same address in the two conjuncts, the subjects' addresses
may be rewritten as the same address, by convention (152). Let us now
consider other types of structures.

4.3.5 LF VP-movement is wh-movement

One reason for claiming that the movement of the VP is like
wh-movement is that the empty VP can be embedded, as in the following:

(154) Someone who [broke his arm] thought that the doctor said
that Bill did too [e]

At LF, the VP *broke his arm* moves to the COMP of the relative clause,
which itself merges with the COMP of the matrix clause, in the manner of
(153) above, and the VP in COMP is subjacent to it its trace, and to the
empty VP. This is to be compared with a sentence where the empty VP is
inside an island, as in (155):

(156) ??Someone who broke his arm thought that the doctor did not
remember when Peter did too

This is also to be compared with a sentence with a pronoun of laziness,
which does not admit embedding because the relation between the
paycheck-NP and the pronoun is subject to condition A:
(157) *John drank his coffee, and I think that Peter just made it

Another relevant type of structure is when the empty VP belongs to an adjunct. In that case, sloppy identity is possible:

(158)a. John broke his arm when Peter did [e]
   b. Lisa finished her drawing without thinking Mary would [e]
   c. (?) Elsa drank her beer because Betsy did [e]

Let us consider the LF of (158)a, treating the when-clause as a relative clause headed by when. As with (153) above, the relative clause is extraposed and merged at S' with the matrix clause, and VP is moved to COMP:

(159)

As for the LF of (159), it involves merging the matrix S' with the adjunct S', after movement of the adjunct to S' and extraction of the adjunct S_k.

(157)c is only slightly worse than the other sentences, but it should be unacceptable, since the adjunct is not quantificational. This

---

84. We will see that when-clauses may be free relatives and that (matching) free relatives have basically the same structure as normal relatives.
indicates that the because-clause already hangs from S, and hence does not have to move, before the S' extrapolates.

Let us now turn to the example with sloppy identity which was given as one type of evidence that bound pronouns behave like operators, and let us see if our former account, in terms of the condition on alphabetic variance may be transposed in the new analysis.

4.3.6 The sloppy binder and the empty VP must have the same GC

As we saw in chapter 3:3.3.5, the binders of sloppy pronouns may not occur higher up than the clause whose VP is the "deleted" VP, as shown in (160) (79) of section 3.3.5:

(160) *John told Mary that he likes her (=Mary) and (he told) Suzan that Peter doesn’t [e]
      (=like suzan)

Following Sag (1976) and Williams (1977), such an example was explained by the condition on bound elements, which required of the copied variable in like x that it be bound in the second VP, which it is not:

(161) λx x= Mary (John told x that he [likes x]) and λy y= Suzan
      (John told y that Peter [likes x])

The variable x is unbound in the second VP, because it is copied from an antecedent VP which does not form a λ-expression by itself, so the λ-abstracted element cannot be copied along with it, and the variable gets copied without its proper binder. So, as we see, the unacceptability of sentences like (161) are due to the fact that the λ-expression is too far from the binder of the pronoun represented as a
bound variable. As evidence for the claim that (161) is excluded on locality grounds, we compared it with the well-formed (162):

(162) John told Mary that he likes her and he told Suzan that he doesn't (like Suzan)

The well-formedness of (162) is accounted for by the LF-movement of the embedded clause that he doesn't [e], owing to the presence of a bound pronoun, he. This allows this clause to be defined as a $\lambda$-expression, which is an alphabetic variant of the matrix one, which allows the pronoun to be copied along with its binder, the $\lambda$-abstracted NP, Suzan.

Now, we have to give an account of the above contrast which is not based on the notion of alphabetic variance.85 Let us consider (159), first, and let us pose the problem of its unacceptability. Given our assumptions, the VP should be able to move to COMP and bind its trace and the empty VP, as in (163):

(163) [[like her], [John told Mary that he t, and he told Suzan that Peter doesn't [e]t]]

Given that the structure is a coordinate structure, it has to have an extraction site in both conjuncts, which it does. Moreover, both A'-chains respect subadjacency. So, there does not seem to be anything wrong with it. However, let us make a correlation with this LF and certain overt counterparts. Johnson (1985) notes that extraction of

85. Many thanks to Kyle Johnson, who gave a neat solution to this problem, allowing us to keep the notion that VP-movement is wh-movement at LF.
adjuncts from double object-constructions are unacceptable long-distance. And he suggests that this may be what happens in these cases: VPs cannot be felicitously extracted long-distance from double-object constructions:

(164)a. John thought he should like coffee, and like coffee you said that he does [e]

b.??You thought John should like coffee, and like coffee you persuaded me that he does [e]

This is a less striking contrast than the contrast between the possibility of a sloppy reading and its unavailability in (159), which is unexplained, since they should reflect each other. Also, Kyle Johnson informs me that extraction is less unacceptable with tell NP, so our example with this verb should not be too deviant, but it is. These problems put aside, let us consider the solution of the contrast between the sentence with a non-coreferential subject and that with a coreferential subject. Suppose that, for some reason, presumably ECP, long-distance wh-movement of non-arguments is impossible, in double-object constructions. Then, the VP can move only to the COMP of its clause. In that case, it will be unable to c-command the empty VP in the second conjunct and hence to form a chain with it, as shown in the following tree (this tree represents the sentence with Gapping, with a union of the matrix phrase-markers, in order to show that Gapping has no influence to lessen the distance):

(165) John told Mary [that[like her]₁[he t₁]] and [he told Suzan [that[Peter does [e]₁]]]
As we see, the VP adjoined to the embedded S is contained in the maximal projection S', which does not dominate the other conjunct, so the VP does not c-command the empty VP inside the conjunct.

Now, when the embedded clause contains a bound pronoun, it can be moved to its own GC. And if both embedded clauses contain a bound pronoun, then they are both able to move to their own GC. In that case, the fact that these sentences occur in adjuncts becomes relevant, because the VP moved to COMP will be able to c-command and form a chain with its trace and also with the empty VP in the moved clause of the second conjunct. And moreover, the moved clauses can also form conjuncts. Consider the LF of (162), where what has to be done is to merge the S' node of the quantifier clause (the topmost S is already merged on the surface, so LF merging is allowed to go one node down, here the S' of the moved clause):
To conclude, we see that we can account for sloppy identity in VP-deletion sentences along the lines of the interpretation of pronouns of laziness, without having to represent sentences as $\lambda$-expressions. This is an advantage in so far as it does not say anything special about sloppy identity, except that nodes may have the same address in conjunctions, and that subjects have the same address when their predicates do. The phenomenon is allowed just in case two structures are symmetrical, and the VP can form a chain with two positions at the same time.

Let us now consider a sentence where sloppy identity is impossible to obtain, a fact which is attributed to a violation of the crossing constraint in Halk (1985).

4.3.7 Sloppy identity and crossing

VP-deletion inside relative clauses (or comparative clauses) is an interesting domain for the study of $A'$-relationships formed at LF', since one $A'$-chain appears at that level, the chain between the operator and the variable that it binds. They are good testing grounds for the interaction between $A'$-chains with respect to crossing effects. The
facts will show that the relation between a sloppy pronoun and its binder interacts with the A-chain created by the wh in COMP and the variable it binds. In (167) the judgment indicated is the one where the second-person pronoun should turn into a first-person pronoun inside the deleted VP that is to say, where the pronoun is to have the sloppy reading:

(167)a. ?You can show your books to all the persons I did
   b. *You can show to your brother all the books I did

Why is there a difference of acceptability between these two sentences? In a theory which admits the representation of sentences as λ-expressions, the contrast can be explained in terms of the crossing constraint, applying between the wh operator and its trace, and the pronoun of laziness and its binder. Consider the LFs of (167)a-b:

(168)a. [[[all the persons][Op₁[I λx (x did show x's books to tᵢ)]]
   [you λy (y can show y's books to tᵢ)]]
   b. [[[all the books][Op₁[I λx (x did show tᵢ to x's brother)]]
   [you λy (y can show tᵢ to y's brother)]]

If lines relate the operators to the sloppy pronoun on the one hand, and to the wh variable on the other hand, these lines cross in (168)b and not in (168)a, as shown below:
The explanation for such a fact could simply be what it seems to be, namely, an ill-formed crossing configuration between two A'-chains. The problem is that it is not clear why bound pronouns like the sloppy ones should create a visible relation for the crossing effects. Other bound pronouns do not create such effects, as shown in Pesetsky (1982), and illustrated below: 86

(170)a. Who did every girl give her books to?
   b. What did every girl give to her brother?

There is no noticeable difference of acceptability between (170)a and b, even though their LF is similar to that of (168)a-b, as shown below:

(171)a. who_1 did [[every girl]_x [x give x's books to t_i]]
   b. what_1 did [[every girl]_x [x give t_i to x's brother]]

Bound pronouns and sloppy pronouns should thus not be treated on a par totally.

One other relevant fact for the analysis of VP-deletion is that

86. However, see May (forth.) for the claim that crossing effects are sensitive to the relations formed by bound pronouns.
quantifiers like everyone do not induce crossing effects, in the general case, as shown below:

(172) What did you give to everyone?

The LF of (172) is (173), in which the two A'-relations cross each other, without precluding the reading:

(173) \([\text{what}]_i[\text{everyone}]_j[\text{you gave } t_i \text{ to } t_j]]\]

So, quantifiers moved by QR do not create crossing effects (but see May (to appear) for disagreement).

In addition, let us raise the question of the behavior of anaphors with respect to crossing. Anaphors obligatorily induce sloppy readings (except when they are not direct arguments of verbs: see Bouchard (1982)), it is thus interesting to wonder whether they display crossing violations, like sloppy pronouns, or whether they are exempt from them. The surprising fact is that they are exempt from them, which indicates that the formal representation of anaphors must be different from that of pronouns with sloppy readings:

(174)a. You can introduce yourself to all the people I did
   b. You can explain to yourself all the books I did

If anaphors induced crossing effects, (174)b should be ruled out along with (166)b, which minimally differ from each other with respect to the pronoun\anaphor distinction. The absence of crossing effects with anaphors is explained if they are not represented as bound variables, but rather if they stay anaphors in the copied VP. Since anaphors must
obey the binding theory, the fact that they are interpreted as
coreferential with the new subject derives from the fact that the
binding theory will not permit it otherwise.\textsuperscript{87}

Now, in the theory which does not represent sentences as
$\lambda$-expressions, the contrast will still be explained in terms of the
crossing constraint, if we consider the crossing constraint a condition
on paths, as in Pesetsky (1982): the path between the sloppy pronoun and
the subject crosses the path between the wh-moved VP and its trace, as
in (175):

\begin{equation}
\text{(175)}
\end{equation}

Moreover, we see that the paths do not cross in the LF of the other
sentence:

87. Subjects are not the only possible antecedents. An anaphor could be
bound by, say, a subject in the antecedent VP and an indirect object in
the copied VP, which is wrong, as discussed in Williams (1977), and
here. Such a problem can be solved in Chomsky's (1985a) theory, which
is adopted here, in which anaphors are LF-clitics. Given the position
to which they move, they get antecedents which are unambiguously
determined: moving to INFL, only the subject, and moving to VP, only the
indirect object. This conclusion was also arrived at by Andy Baras
(personal communication).
This account also depends on the particular property of sloppy pronouns to create crossing effects.

4.3.8 Why no sloppy reading inside a parasitic gap

We noted, in section 4.14.1, that reconstruction, for which sloppy identity is a good test, is not possible inside parasitic gaps, but is possible in ATB-gaps. Our assumptions account for this difference. Given that addresses are assigned to nodes as soon as D-structure, and given that only conjuncts may be assigned the same address, it follows that nodes in conjuncts formed only at LF cannot have the same addresses. Since parasitic gaps occur in LF-conjuncts, they will be unable to find themselves in symmetrical structures with the real gap, hence the absence of sloppy readings associated with the position occupied by a parasitic gap.

Now, a parasitic gap does not allow reconstruction, but the sentence's counterpart with VP-deletion accepts it:
(177)a. *Which picture of himself did John look at before Peter threw out [e]?
   (Peter's picture of himself)

   b. Which picture of himself did John look at before did [e]
   (look at a picture of himself)

This is because VP-deletion leads to a rewriting of the address of the subject, by convention (152), owing to the fact that the new address of the VP, after LF movement, is the same in the two LF-conjuncts.

4.3.8.1 Comparative Deletion

Williams (1977) reports an observation of Bach, Bresnan and Wasow (1974), where comparative clauses allow sloppy identity with VP-deletion, but not with Comparative deletion:

(178)a. John likes more of his children than Bill does

   b. John likes more of his children than Bill likes

I assume that, in Comparative-deletion sentences, the variable is the determiner of the NP, not the whole NP, in a way similar to comparative subdeletions, except that the N is not overtly realized. The subdeletion cases show that operators may take as their variable the determiner of the NP, as in (179):

(179) John likes more of his children that Bill likes [[e] of his friends

88. See Bresnan (1973) for a study of comparative clauses and various types of deletions in them.
Another type of evidence comes from French, where subdeletion is possible as in (180)a, but comparative deletion is not, as shown in (180)b:

(180)a. Paul aime plus de films que Pierre n'aime [[e] de livres]  
'Paul likes more books than Pierre likes books'

b. *Paul aime plus de films que Pierre n'aime  
'Paul likes more books than Pierre likes"

It is likely that the ungrammaticality of (180) is the same as the ungrammaticality of (181):

(181) *Paul aime beaucoup  
'Paul likes many'

In other words, it is impossible to leave an empty N in French. Moreover, both ill-formed sentences can be rescued if the N is an overt pronominal, the clitic en:

(182)a. Paul aime plus de films que Pierre n'en aime  

b. Paul en aime beaucoup

In Belletti and Rizzi (1981) and Haik (1982), the ungrammaticality of (180) is assumed to be due to illicit government on the empty N, which is analyzed as a PRO. I suppose that the English well-formed cases show that the option of having a category N which dominates nothing at all is possible in English. So, the particularity of comparative deletion is that the operator in COMP forms a chain with the determiner of the NP. The chain being between more and the determiner, it does not include the pronoun his contained in the N' of his children, so reconstruction does not help put back this N' in the position of the variable, hence
preventing the sloppy interpretation of his. 89

Lastly, we note that sloppy identity is not possible with VP-deletion itself, when the antecedent VP is moved overtly. In that case, the empty VP acts like a parasitic gap:

(183) John wanted to talk to his son, and talk to his son he did t before Peter did [e]

It is impossible to interpreted the adjunct as: before Peter talked to his (own) son. To show that this is due to the movement of the antecedent VP, consider the counterpart of (183), with the antecedent in situ:

(184) John talked to his son before Peter did [e]

In (184), the adjunct may be interpreted with the sloppy reading. This could be explained if overt VP-fronting is characterized as topicalization, that is to say, movement to the node TOP (Chomsky (1977). Since adjuncts do not contain TOP nodes, the topmost node of the adjunct cannot merge with the topmost node of the matrix clause, since these are different. Merging is a sine qua non condition on sloppy identity, so (182) is explained.

89. Such an example is treated as a case of unbound variable in the copied structure, in Williams (1977), because the λ-abstraced binder is too far from the copied material to be itself copied, similarly to tough-constructions.
4.4 Free relatives

4.4.1 LF union of phrase-markers

In this section, I will sketch an analysis of free relatives, based on the formal work done for the analysis of parasitic gaps, and sloppy identity, which is the formation of union of phrase-markers at LF. A striking fact about free relatives is that they are much more natural when the verb of the relative clause is the same as that of the matrix clause. Practically all the examples seen in the literature contain this repetition:

(185)a. I'll go wherever you go
   b. He'll like whoever you like
   c. They'll run however fast she runs

This is not obligatory, but it is a revealing clue of what the syntax of these relatives is:

(186) Bob will read whatever Suzie gets

There has been much debate about the structure of free relatives. Overtly, free relatives all have the form of a wh-operator followed by a clause without a complementizer, as in (187):

(187) John will buy whatever Peter makes

Certain authors, like Grimshaw (1977) and Bresnan and Grimshaw (1978)
argue for a structure in which the free relative is like a relative clause, its head being the wh-phrase, followed by a bare S. without a COMP position, like (188):

(188) $$[[\text{whatever}]_{\text{NPI}}[S \text{ Peter makes } t_1]]$$

Others, like Groos and van Riemsdijk (1979), argue for a structure in which the wh-phrase in is COMP, and the relative construction headed by an empty head, of the category of the wh-word, like (189): 90

(189) $$[[e]_1[S, [\text{whatever}]_{\text{NPI}}[\text{Peter makes } t_1]]$$

Here, we will adopt the view that free relatives are headed by the lexical wh-phrase, but that the clause that follows it is a full S', with a COMP. So, they look like normal relatives.

Let us note that the hypothesis that free relatives have the structure of relative clauses is meant to avoid having something special to say about the matching phenomena, pointed out in many places. 91 The matching properties of free relatives is that the wh-phrase behaves as if its was subcategorized for by the predicate of which it is an argument. This follows straightforwardly from the hypothesis in which they have the structure of relative clauses. However, certain free relatives, in certain languages, display non-matching phenomena, as in the following (cited in Levin, (1982), Harbert (1983) and Süß (1984)):

90. See also Hirshbuhler (1978).

(190) A qui has parlat esta malalt
to who (you) have spoken is sick
'(The one) to whom you have spoken is sick'

In (190), the subject is, overtly, a PP, instead of an NP. Altogether, it is probable that the non-matching properties of free relatives come from their having a different structure than the one postulated here for the matching free relatives. In this work, we will only consider matching free relatives.

Our hypothesis is that free relatives have the structure and the interpretation of true relative clauses, but their characteristics is that, instead of being in the COMP position, the wh-operator occupies the head position of the relative construction. Given that, at LF, the structure should not contain two wh-operators, the whole construction may contain only one such operator in the whole derivation (no insertion or deletion of material is allowed, except insertion by copying rules at LF'). This entails that there cannot be another wh-operator in the structure, and in particular in COMP. And, by the Projection Principle, an empty category occurs in some position inside the relative clause. So, the structure of a free relative must be something like (191):

(191) [[whoever]_{NP_i} [\_S, COMP [\_S you meet t_i]]]

Given that the structure is that of a relative clause, the interpretive rules read this NP by the predication rule, which attributes the property described by the relative clause to the referent of the head NP. Since this is done via predication, the trace inside the relative
clause cannot form a chain with the head NP. Since it has an independent th-role, it is pronominal like, which is impossible at LF. Also, like parasitic gaps, this empty category has to be licensed at S-structure. So, in fact, COMP must contain an A'-binder at S-structure. So, free relatives are like relative clauses in all respects, except that the feature [+wh] is on the head, instead of being on the operator in COMP. This means that free relatives are defective, if nothing happens at LF to save them.

At LF, the structure must become that of a well-behaved relative clause, with the proper operator-variable relation restored, that is to say, with the wh-operator in COMP, and no wh operator in the head position. The LF derivation will make the operator, head of the relative clause, transmit its wh-feature to the operator in COMP, by moving into it, and leaving its feature behind. Technically, this can be done only if the matrix clause, of which the free relative is an argument, sustains the relative clause in the way that the matrix clause sustains an adjunct, in parasitic-gap constructions. 92

Consider (191) again:

(190) I'll like [whoever₁[COMP [you meet t₁]]]

COMP in (191) has to be filled by a wh-element indexed ₁. This element cannot be the NP whoever, since it would create a movement downward.

92. This analysis has been influenced by Tim Stowell's suggestion (personal communication) that the movement of a category can provide an empty category in a structure.
which is impossible in principle (see Chomsky (1985a)). It cannot be
the empty category \( t_i \), since it does not bear the feature [+wh]. \(^{93}\) So, the only way to make the movement of \textit{whoever} upward, and at the same
time move it to the COMP below, is to merge this COMP with the matrix
COMP, in the now familiar manner: the whole constituent \( \text{NP}_i \) moves inside
the empty element in COMP, which is coindexed with it, and then adjoins
to this \( S' \), leaving its [+wh] feature on the empty category in COMP,
and forms a union of phrase-markers between its COMP and the COMP of the
matrix \( S' \), as in (192), where the relative clause is extraposed from \( \text{NP}_i \)
and merged with the matrix COMP:

(192) S-structure: I'll like [\texttt{whoever}_i[COMP [you meet \( t_i \)])]]

\[
\text{LF: } & [\text{you meet } t_i]]] \\
& [\text{I'll like } t_i] \\
& [S, \text{whoever}_i[COMP[e]]] \\
& [S', (k)] \\
& S' \\
& \text{NP}_i \\
& \text{whoever} \ t_k \\
& \text{COMP} \\
& \text{S} \\
& \text{S} \\
& t_i \\
& \text{NP} \\
& \text{VP} \\
& \text{NP} \\
& \text{VP} \\
& \text{you} \ V \\
& \text{NP} \\
& \text{I} \ V \\
& \text{NP} \\
& \text{meet} \ t_i \\
& \text{like} \ t_i
\]

\(^{93}\) If it contained the feature [+wh], it would violate the condition
which makes wh-movement obligatorily in relative clauses, in English.
This condition is mysterious; see Lasnik and Saito (1985) for discussion
of wh movement in questions.
Note that the usual restrictions on the distance between the empty operator and the variable that it binds do not apply here:

(193)a. John will meet whoever you think Mary will like _

b. John will buy whatever they say one can fix _ easily

This is to be compared with:

(194) *This car is easy to say that one can fix _ easily

This shows that the restriction on embedding must be an S-structure condition since if the empty operator is created at LF, it can escape from it.

Also, note that the presence of the element in COMP in the whole derivation allows us to explain why the empty category is not subject to the ECP, contrary to parasitic gaps in such positions: 94

(195)a. John will hire whoever he thinks _ will understand his handwriting

b. *The person that John hired t without thinking _ would understand his handwriting

One outstanding problem for this analysis is that (195a) violates the definition of well-formed factors, since one of the conjuncts starts with one of the variables. This is to be compared with a coordination of Ss, as in (196)b:

94. See Kayne (1983) for an ECP account and also Cinque (1984) for an alternative explanation, in terms of the impossibility of resumptive pronouns in subject positions in general.
(196)a. John will follow whoever _ enters the room  

   b. *I wonder who John follows _ and _ likes him

This means that there is more to say about this condition.

Lastly, we note that merging is possible even if the two sentences are not semantic conjuncts, contrary to parasitic-gap constructions. This follows our characterization that the condition on semantic conjuncts holds only if the matrix clause contains one branch of the double chain at S-structure. As for the relative clause, it contains one branch of this chain, but the condition applies to the branch in the main clause only, not the second conjunct.

Now, the reason for deriving free relatives the way we have done is that, first, it could explain why free relatives sound better when the verbs are repeated. It is a particularity of unions of phrase-markers that they give rise to this preference.95 For example, coordinates sound more natural when their constituent structure is the same, parasitic gaps sound better when the real gap and the parasitic gap have the same type of th-role, and pronouns of laziness are preferred in comparable environments too:

(197)a. I know who he likes _ and Mary hates _  

   b.(?)I know who he likes _ and Mary met -  

   c. This is a donkey that everyone who knows - likes -  

   d.(?)This is a donkey that everyone who knows - wants to kick -

95. See Goodall (1984).
e. The man who gave his paycheck to his wife is wiser than the man who gave it to his mistress

f.(?) The man who gave his paycheck to his wife is wiser than the man who described it to the taxman

This preference is quite subtle, and maybe it is an illusion. However, there exists a striking difference of behavior between free relatives headed by -ever- phrases, and those where -ever is missing: with those lacking -ever, this preference becomes a requirement, and this is what we now attempt to explain.

4.4.2 When -ever is missing

Consider the minimal pair:

(198)a. I'll like whoever you meet
   b.*?I'll like who you meet
   c. I'll like who you like

Let us repeat one of our assumptions about LF movement: all movement has to be justified, in that it must be due to the assignment of scope to an operator.96 In addition, successive movement is not possible.97 So, if a constituent has to make two types of movement, it will only be because it contains two operators, each of which is the trigger of movement. We also assume that, if a lexical item implicitly contains more than one operator, these operators may split at LF, so they may make the lexical

item move twice:

(199) **Condition on LF movement**
Operators are assigned scope once and only once.

Now, concerning the derivation of free relatives, we have seen that they necessarily involve a double movement of the NP: up to COMP and up to an S'-adjoined position, as in (200), repeated from (191). The movement into COMP is the only derivational difference between free relatives and, say, parasitic gaps inside subjects.

Both movements are necessary. The movement to COMP is required in order for the wh-operator to leave its [+wh] behind, and the extraction of the NP out of COMP, in order for the relative clause and the matrix clause to merge at the S'-node, so that the relative clause eventually contains the wh-operator in the merged COMP. For example, let us see what the structure would be if the second movement did not take place:
In (200), the COMP of the free relative is still without a wh-operator in it. It is only if the COMP of the matrix clause (which contains NP_{1}) merges with this COMP that it will contain it. 98

Given these assumptions, let us ask what allows the two movements in the case of -ever-phrases. For the movement into COMP, the wh-feature allows it, as is usual with wh movement. And for the extraction of this NP onto the position adjoined to S', we can assume that it is the universal quantifier expressed by ever. This accounts for the possibility of deriving a proper LF for the common matching free relatives. 99 Now, free relatives headed by wh-operators which lack the universal quantifier should not be able to be extracted out of COMP.

98. As for the first movement, if the NP does not move into COMP at all, then this COMP will not contain its wh-feature. The empty category inside the S' will not be able to be defined as a proper variable, but it will stay a pronominal, and hence the structure will be ruled out for containing a governed empty pronominal.

99. In case the head NP is what or when, the VPs no longer have to be the same. This is because they have universal force by themselves, contrary to who. Thanks to Ken Hale for a discussion on this.
because they do not have ever to help them out. So, they are expected to be bad, which they are, as shown in (197)b: *I'll like who you meet.

But they are not always bad, as shown by (197)c, namely, when the verb is exactly the same in the two clauses: I'll like who you like. Why not? In fact, the exact formulation of the similarity between the matrix clause and the relative clause is that they have the same VP, not the same V. As we saw in the derivation of sloppy identity in VP-deletion sentences, it is possible for VP to move into COMP at LF. We assume that this is because VPs may bear [+wh] features. So, in case the VP of the free relative is [+wh], this VP will act as the operator allowing movement of the free relative into COMP. In COMP, the head NP who will transmit its [+wh] feature to the empty category there, and, by being a wh-quantifier itself, who will be able to be assigned scope over S', and will be extracted from COMP.100 What about the VP? The VP is treated as a wh-quantifier, assigned scope to COMP, and it stays there, since operators are assigned scope once and only once. However, if there is a VP in the merged COMP, as in all cases of conjuncts with an operator in the common COMP, this operator must bind a variable in both conjuncts, as an effect of Ross's ATB-exceptions to the Coordinate Structure Constraint. Consider the structure up to this point of the discussion:

100. This presumes that the movement of who out of COMP is possible, simply because it bears an operator feature.
In (201), the free relative has been moved to $S'$, and the relative clause forms a union of phrase-markers with the matrix $S'$ at the COMP node. This is not well-formed for the reason that, if the VP in COMP binds a proper variable inside the relative clause, this VP does not find a variable to bind in the matrix $S'$, which already has a filled VP. In order to get a proper representation, with a variable in both conjuncts, the VP of the matrix $S'$ also has to move to COMP. In that case, it will merge with the moved VP of the relative clause, leaving only one VP in COMP, binding in an ATB-manner the two empty VP-categories inside the two conjuncts, as in the following:

Of course, it is because the two VPs are identical that they can merge. If they dominated different lexical material, they could not, leading to
a violation of the Coordinate Structure constraint.

4.4.3 Condition on scope assignment is a condition on representation

Note that the requirement that each operator is assigned scope once and only once should be stated as a well-formedness condition on LF-representation rather than a condition on movement, for the following reason: free relatives are well-formed when they contain an elliptical element, instead of being fully overt, as in (203) (Bresnan and Grimshaw (1978): (63a), citing Jespersen (1927)):

(203) Eat what you please\like

Bresnan and Grimshaw do not recognize this example as one of a true free relative, because they notice that embedding is impossible, contrary to true free relatives:

(204) *Eat what you seem to please

vs. Eat what you want to eat

As Jespersen notes, the verb like "has no reference to [the person]'s personal feelings". This is because these verbs take an elliptical sentence, which is interpreted as identical to the matrix one: (203) is interpreted as: "Eat what you (please)\like to eat". So, (203) is indeed a case of a true free relative, with an empty S', interpreted as coreferential with the S' that contains it, in a way equivalent to VP-deletion inside relative clauses: here, instead of a missing VP,
there is a full S'. That these have not been recognized as involving an elliptical S' is due to the fact that, in general, elliptical S's are impossible in English, and that the overt form with please is unacceptable, as in the following:

(205)a. *John likes the person you said
   b. *John likes whoever you said
   c. *John likes who he pleases to like

I do not know why these are marginal in English, and why they are possible with like, and please. However, they are acceptable in French, as we saw in chapter 2, and as seen here, with free relatives:

(206)a. Paul aime qui tu m'as dit
   'Paul likes who you said'
   b. Paul a rencontré qui tu crois
   'Paul met who you believe'

So, let us consider these cases, assuming that they contain an empty S' as complement of the V inside the relative clause. This S' is then represented at LF via a copying rule identical to the VP-copying rule of Williams (1977). Such cases show that the condition that the operators be assigned their scope once and only once is a condition on representation, because, on the surface, they lack the VP which is

101. I have no answer as to why embedding is impossible for these sentences. The sentences which I claim are the equivalent in French are possible:

(1) Paul a bien rencontré qui il me semble que tu lui avais dit
    'Paul did meet who it seems to me that you told her (that he had met)'

---
responsible for the movement of the free relative into COMP. This VP appears at LF' only, after copying of the S' onto the position of the empty S', as in the following LF':

(207)

```
NP
what S'k

S'1

S'1,k

COMP

S

S

V NP (you) t_j you V S'1

eat t_i

please COMP S

NP VP

(you) t_j
```

In this LF', S'1 was an empty category and it is a copy of the matrix S'. It contains the copied VP, an empty category, so this empty VP can be properly bound from the common COMP. As for the representation, each operator is in its proper place. And given that the VP in COMP is interpreted as belonging to the relativized NP (as well as to the matrix S'), it can be interpreted as the operator which licenses the movement of this NP into COMP, represented there as the trace of the NP, t_i.

Let us now consider an argument suggesting that the structure of free relatives is the one postulated here.

4.4.4 Extraction from extraposed sentences

We have studied VP-deletion inside relative and comparative clauses, and we have seen that the operator in COMP may correspond to an argument
of the empty VP, or empty S' (cf. chapter 3):

(208)a. John visited [every person that you did [e]_VP]_NP
    b. Paul a vu [le film que tu avais dit [e]_S']_NP
       'Paul saw the movie that you said'

One may wonder what would happen if the relativized element was extracted from an extraposed sentence, instead of a complement sentence. The facts go in two directions:

(209)a. *Paul a vu le film que ça l'arrangeait
       'Paul saw the movie that it was convenient for him (to see)'
    b. Paul est allé où ça l'arrangeait
       'Paul went where it was convenient for him (to go)'
    c. Paul a rendu visite à qui ça l'arrangeait
       'Paul visited who it was convenient for him (to visit)'

First of all, we have to make sure that the relative clause does contain an empty category S', interpreted as coreferential with the containing S'. This is ensured in (209)a by the fact that COMP must contain an operator on the surface and that this operator must bind a variable, since this is a relative clause, and since lexical insertion is not possible later. Since variables cannot be overt, it has to be the case that the structure contains an empty extraposed S', playing this role. Now, the reason why (209) is excluded is the same as when normal wh-movement occurs from that position, as in (210):

(210) *Qu'est-ce que ça t'arrange [e]?
      'What is it convenient for you?'

The reason why it is impossible to question an extraposed sentence is that, given that variables must be empty categories, the only possible
variable in (210) is $t_i$. Now, we assume that variables corresponding to arguments must bear Case themselves. Given the ungrammaticality of extraposition of pure NPs, as in *it is clear the theorem, extraposed positions are not in a position assigned Case. So, the empty $S'$ cannot be a variable, making the structure ill-formed.

So, now, if (208) is ruled out, why is (208)b acceptable? The claim here is that (208)b is a case of a free relative, and that, as shown in (208)c, free relatives are acceptable in such environments. Let us consider how this is possible. First, note that (208) is OK because it contains the operator où. A sentence synonymous with (208)b, with a relative que is as bad as (208)a, which shows that the contrast between the two is due to the operator and not the argument- or adjunct-status of the relativized element:

(211) *Paul est allé à l'endroit que ça l'arrangeait
'Paul went to a place that it was convenient for him (to go to)'

102. Pronouns of laziness show that variables do not have to be empty, but it must be the case that the operator forms a chain with at least one empty category. Since there is no possible double chain in (210), the operator can only bind an empty category.

103. We differ here from the view that variables need Case because of the Visibility condition, since the coindexed subject could form a chain with the extraposed $S'$, which would then be included in a Case-marked chain, and would be OK with respect to the Visibility condition. (The Visibility condition of Chomsky (1981) requires of NPs that they belong to a Case-marked chain, it does not say anything about the particular elements of the chain individually.) However, it is not clear whether the expletive in fact does form a chain with the extraposed $S'$, a complex issue, as shown by Pollock's (1985) work, and I have assumed (chapter 1:17.1, (79)) that it does not. See also Davis (1984), Safir (1983), Zubizarreta (1982), for the relation between expletives and the arguments linked to them.
The generalization is that, with certain wh-operators, like où 'where', and quand 'when', relativization is possible out of an extraposed empty sentence. And the same is true of free relatives, as in (208)c, for example.

To make a greater generalization, let us claim that the où- and quand-clauses are free relatives too. Then, the S-structure of all these sentences is that of a wh-operator in the head position of the relative construction, with a non-operator in COMP, and with an empty S' in the extraposed position. Given that COMP does not contain an operator, it does not have to bind a variable, i.e. a Case-marked empty category. However, it has to be licensed in the structure, like all elements in A'-positions, so it must form a chain with an empty category, and the empty sentence, as an empty category, also has to be licensed in the structure, simply by being bound (like parasitic gaps), which it is, if the category in COMP forms a chain with it. But this does not entail that the empty S' has to be Case-marked, since the binder is not an operator. Then, at LF, the derivation of the free relative is like that of (191) above, where the relative clause forms a union of phrase-markers with the matrix S', and where the empty extraposed S' gets represented as a copy of its antecedent. In other words, the free relative becomes a well-formed relative clause at LF and LF'.

The aim of looking at these examples was to show that the assumption that COMP contains something, but that this thing is not an operator has some empirical support. Note that, as is the case with true
relative clauses, it is impossible to obtain Gapping in a free relative:

(212) *John gives his books to whoever Mary her papers

This is because COMP is filled by an empty element, and this element has to be licensed by forming a chain with an empty category. As we mentioned in chapter 3, Gapping does not involve empty categories, either in Pesetsky's (1982) analysis, or in Goodall's (1984). Bare XPs are also impossible, making the same point:

(213)a. *John reads whatever Mary
    b. *John gives his books to whoever his papers

4.4.5 Extraposed S's and bound pronoun as operators

All the predicates that accept this extraposition are verbs of psychological attitude, and tough predicates, which contain an implicit (or explicit) experiencer, as in the following:

(214)a. Paul partira quand ça lui chantera
    'Paul will leave when he'll please (to leave)'
    b. Marie fait la connaissance de qui ça lui plaît
    'Marie comes to know who it pleases her' (to know)'
    c. Pierre grimpe où c'est le plus facile
    'Pierre climbs up where it's the easiest (for him to climb)'

And they are impossible when the predicate does not involve an experiencer:

(215) *Pierre est amoureux de qui c'est clair
    'Pierre is in love with whom it's clear (that he is in love)'
Suppose that tough-predicates are represented with an implicit argument with the th-role [experiencer], such that this thematic position, in the th-grid of the predicate, is able to be semantically visible. All the predicates which accept relativization from an extraposed clause are those for which the predicate takes such an argument. Importantly, this argument is always understood, or overtly represented, as coreferential with the matrix subject, as indicated in the translations.

If relativization from extraposed sentences is possible only when the experiencer is coreferential with the matrix clause, this indicates that it is because the experiencer is treated as a bound pronoun, and makes the sentence that it occurs in behave like a quantified phrase. So, this bound pronoun is probably one of the two necessary operators which make NPᵢ move in the derivation of free relatives. The question is then, why is the VP not able to be the relevant operator in extraposed sentences, as in (216)b, contrary to complement sentences, as in (216)a? (Remember from the preceding section that this VP is represented at LF as a copy of its antecedent, and that, at LF', it is seen as the quantifier that makes the movement of the NP possible.)

(216)a. Paul est parti quand tu as dit [e]
   'Paul left when you said'

    b. Paul est parti quand ça l'arrangeait [e]
       'Paul left when it was convenient for him'

In extraposition configurations, S' is not the real argument of the predicate, this argument being the subject, which is a referential
demonstrative pronoun. We can assume that this is the reason why something inside this non-argument sentence cannot play a role in the semantics of the whole sentence. In other words, the quantificational property of the VP inside the extraposed $S'$ are obliterated by the fact that it is contained inside a constituent which may not interact semantically with the rest of the sentence. If this is right, then $NP_i$ needs another operator, to allow it to move twice, and this operator is the bound pronoun, explicit or implicit, in the relative clause. In the same way that the movement of the VP to COMP is enough to carry with it the NP that dominates it, we have to assume that the movement of the quantificational clause may carry with it the relativized NP. Also, given that $S'$s do not move by wh-movement, we have to assume that the movement of $NP_i$ to COMP is due to the [wh] feature of $NP_i$, and the movement out of COMP due to the quantificational clause. This yields the following LF, where $S'_k$, the relative clause which contains the bound pronoun is also extraposed from the head NP and merges with the COMP of the matrix clause. The quantificational nature of this clause is given by the presence of the pronoun (a clitic), which is bound by Paul. And since it is related to $NP_i$ via its trace, it is interpreted as the quantifier that makes $NP_i$ move out of COMP:

104. See Pollock (1985), and Zubizarreta (1982) for strong arguments for this view.

105. This does not mean that $S'$ is not interpreted at all, but rather, that it is external to influences inside the sentence.
In (217), the wh in the common COMP must bind a variable in both conjuncts. It binds the trace of $NP_i$ in the matrix clause, but it does not bind a proper element inside the relative clause. This can be rectified by copying the matrix clause $S'_j$ in the position of the empty $S'$, as usual. Then the operator will bind a proper variable.

Let us now consider another fact which may support the analysis of free relatives as involving a union of phrase-markers.

4 4.6 Free relatives from PPs

Consider (218)a.

(218)a. Paul est parti avec qui tu es parti
'Paul left with whom you left'

The free relative is the NP complement of the preposition *avec* with, but this preposition is not stranded inside the relative clause, as it should be. In Bresnan and Grimshaw's (1978) analysis, the gap in a free relative is analyzed as obtained by deletion under control of an
identical antecedent. In the PP case above, they analyze the free relative as headed by the PP avec qui, which controls the deletion of the corresponding PP inside the free relative clause. They relate the possibility of control of a PP in free relatives to the general property of French that PP can have 'pro-forms.

(219)a. Pierre en a mangé la moitié of-it 'Pierre has eaten half of-it'

b. Marie lui a parlé to him 'Marie spoke to him'

Under our analysis it is possible to keep the idea that free relatives are always NPs, as in usual relativization, and that the French cases are obtained via the union of phrase-markers in COMP, from which the VP merges, and then is interpreted as the VP of the matrix clause together with that of the relative clause:

(218)b.

So, this means that a piece of the VP in the free relative clause may be elliptical (an empty category), like the PP headed by avec, since its content is supplied by merging the V of the matrix clause with it.
The difference between French and English still remains to be explained. It seems that the best explanation should rely on the different way these languages allow empty PPs. English accepts pied-piping in wh-constructions, so it accepts PP variables. However, following Bresnan and Grimshaw's (1978) idea, English does not have PP-clitics, a fact which itself could reflect the inexistence of non-variable empty PPs. Given that the empty category in the free relative clause is not a variable on the surface, because the A'-binder is not a [+wh]-operator yet, the empty category must be a non-variable empty PP, but these elements do not exist in English. So, these constructions are barred in English. Since French accepts non-variable empty PPs, it allows such free relatives.

As a last piece of information, free relatives in French do not have the ever equivalent of the English ones. So, they all are headed by bare wh-operator's, which makes the requirement that the VFs be the same in the matrix clause and the embedded clause a permanent restriction on matching free relatives. Hirshbuhler (1978) notes certain free relatives where the V can be different, in the matrix clause and in the free relative clause, as in (220):

(220) J'ai parlé à qui tu avais fait allusion

My judgment is that this sentence can only be interpreted with the verbal complex faire allusion à taking an elliptical object, coreferential with the matrix clause. That is to say, the sentence is elliptical for. "J'ai parlé à qui tu avais fait allusion que j'avais parlé". In that case, the derivation of (220) is similar to those seen
above. So, it seems that certain dialectal differences arise with respect to the matching effects. 106

We now consider bare XP comparatives.

4.5 Bare XP comparatives

4.5.1 Bare XP analysis

Consider (221).

(221)a. John told as many funny stories as Mary

b. John met Mary earlier than Bill

In these sentences, the comparative clause is elliptical, there being overtly only an NP. (As a notation, the underlined NPs are those which are understood as having the same Grammatical Function (GF).) We will mostly be concerned with examples where the corresponding NPs are not subjects, and will leave aside the treatment of sentences like (221)a. Our main concern will be to show that these constructions are bare XPs on the surface, and that the complement of the comparative preposition becomes a full sentence at LF. Given the formalism used in this chapter.

106. Non matching free relatives occur in infinitival relatives, as in (i)-(ii) (cf. Levin (1983)):

(i) J'ai trouvé à qui parler
    'I found who to talk to'
(ii) J'ai de quoi écrire
    I have of-what to write 'I have what to write with'
the missing sentence will be restored by forming a union of phrase-markers with the matrix clause. This device will have the result that insertion rules have, without having to really copy any material. So, on the surface, the complement of the preposition is a bare XP, but it becomes a full S' at LF.

In this section, we see that the complement of the comparative preposition as or than must be a bare NP (or bare XP in general).

Hankamer (1973), cited in Hoeksema (1983) gives one strong argument in favor of the bare XP analysis, which is that the XP can be extracted but not in the corresponding overt form (when the bare NP corresponds to a subject):

(222)a. Who did John tell as many funny stories as -?

   b. *Who did John tell as many funny stories as he did\told stories?

(222)b illustrates the fact that extraction of a subject out of a comparative clause is ill-formed, but not in (222)a

One second argument, proposed by Hoeksema, is that an anaphor is possible in place of the NP in (222), but not in the corresponding overt structure:

(223)a. No man is stronger than himself\*him

   b. No man is stronger than *himself\he is

These two tests are harder to make use of when the XP does not correspond to the subject of the comparative clause, but Bresnan's (1976) condition on Analyzability, which determines what constituents
are eligible for wh-movement can be a good test, since S's are not eligible, and NPs are. Consider (224):

(224)a. *At the same time as what article did John read your paper?
   b. *At the same time as he read what article did John read your paper?

This contrast indicates that the complement of the preposition is in fact a bare NP.

There exists an argument of plausibility for our analysis, based on the existence of temporal comparative prepositions, like before and after. It is worth noting that these prepositions also take bare XPs:

(225)a. John saw Mary before Peter
       (ambiguous)
   b. John went to London before New York

If it can be shown that the complement of these prepositions is a bare XP which becomes an S' at LF, then it will mean that this phenomenon of ellipsis is allowed by the grammar, hence rendering plausible the analysis of all these comparatives as taking bare XPs on the surface and full S's at LF.

So, consider the prepositions while and during. These are comparative, in that they are comparisons on the times of two events, and they say that the times are the same. They differ minimally in their subcategorization frames and are in fact in complementary distribution with respect to subcategorization: while takes an S' and only an S', and during takes an NP and only an NP. Otherwise, they are synonymous. Since before and after are able to take what looks like a
bare XP, these two prepositions should also be able to take bare XPs.

But this is not so

(226)a *John solved a math puzzle while a chess problem

b. *John solved a chess problem during a math puzzle

If the complement of the preposition is an S', then (226)a should be acceptable, since it would fit the subcategorization frame of the P. So (226)a shows that this argument is a bare NP. However, if this is true, (226)b should be grammatical. The fact that it is not shows that the assumption that the complement becomes an S' at LF is right: if the subcategorization requirements of a predicate have to be met at every level, as required by the Projection Principle, they have to be met at LF too, explaining the ungrammaticality of (226)b. Also, even if the NP of (226)a turns into an S' at LF, this does not take away the violation of D-and S-structure. So, because these prepositions are restricted to only one category of complement, they cannot be acceptable in bare XP constructions, in contrast with before and after, which take both types of categories, and are acceptable in bare XP constructions.

To conclude. we will consider, without questioning it anymore, that the XP is a bare XP, in the comparative constructions under study. We will also assume that the comparative PP is inside the comparative constituent, and not outside it. So, a sentence like John read this

---

107. See Napoli (1985a) and (1985b) for arguments and discussion. In general, Napoli opts for a theory of deletions where no element is present on the surface. We differ on this point, since sometimes deletion should be a case of empty categories, as shown in preceding sections.
book at the same time as this article has the following S-structure.

(227) John read this book at [NP the same time [PP as [NP this article]]]

4.5.2 The analysis

The problem created by the bare XP analysis is that bare XPs are a violation of the th-criterion, since these are constituents which do not receive a th-role, or, if they are PPs, which do not enter any thematic structure. Let us assume that certain positions are licensed on the surface, to the effect that they may and must contain a constituent there. Such positions are those governed by prepositions which are non-th-role assigners, like comparative prepositions. The class of these prepositions is that of prepositions which can act as coordinators in LF-coordinate structures, like the comparative preposition without, etc. However, since these prepositions are coordinators, they must put together two conjuncts. Since the first conjunct is a matrix clause, their complement must be a clause too, in order to be the second conjunct. So, the missing S' must be integrated in the structure at LF.

In his account of Gapping, Pesetsky (1982) allows the integration of the bare XP inside the structure. For him, Gapping is a case of two bare XP constituents, instead of the one of the comparative construction, as in (228):

108. See McCawley (forth.) for the claim that comparative prepositions behave in certain respects like coordinators.
(228) John saw Mary and Peter Suzan

In (228), the coordinating element governs two bare NPs Peter and Suzan. Now, he notes that, in such coordinate contexts, the counterparts are contrasted with respect to each other, which indicates that they are focussed. It is the assignment of focus scope which permits to solve the problem of the integration of the two arguments inside the missing structure, where the missing structure is copied onto the position governed by and. By assigning focus scope to the antecedent NPs, John and Mary, the antecedent sentence becomes one with unbound variables. By allowing the two bare constituents to bind these two copied variables, these NPs become part of the structure. So, consider the LF proposed by this analysis:

(229) [John [Mary [[e] saw [e]]_Si and [Peter [Suzan [[e] saw [e]]_Si]]]]

Our analysis will differ minimally, in that, instead of copying the missing sentence, it will read this sentence as parasitic on the matrix-clause structure. We will nevertheless use the idea that the counterparts are focussed. One other thing has to be added, which is that comparative clauses must contain an operator in COMP, as was shown by Chomsky (1977). So, since the comparative preposition must govern a comparative clause at LF, this clause has to contain an operator in its COMP. Since this operator is not in COMP on the surface, the comparative constituent has to move into COMP. There, it will be interpreted as the operator. From this constituent, the bare XP is extracted and adjoined to S', and is interpreted as focussed. Also, in
order for the $S'$ to be interpreted as a union of phrase-markers, the
counterpart of the bare XP also moves, (as in Pesetsky's analysis), in a
position dominated by the common $S'$-adjointed node with the extracted
bare XP.

Consider (230) and the first steps of the LF:

(230) John met Mary before Peter

Two things go wrong, in (230). The first one is that the comparative PP
binds two variables, one in each conjunct, which should receive
different values, so we assume that this coindexing does not entail
coreference when the operator is a comparative operator.

The second problem is double. First, $NP_1$ binds its trace, and does
not bind the trace in the position object of meet. And the preposition
before does not govern an $S'$, but an NP. Following Pesetsky (1982), we
assume that free indexing applies at all levels, LF included. At this
level, the variable is allowed to bear two indices, one bound by each
focussed NP. Moreover, given that the presence of traces is forced by
the Projection Principle, the trace of $NP_1$, governed by before, is not
necessary, if \( \text{NP}_i \) binds the empty category, object of \textit{meet}, as in the following:

\[
\begin{array}{c}
\text{NP}_i \\
\text{Peter} \\
\text{COMP} \\
\text{PP}_1 \\
\text{before} \\
\text{VP} \\
\text{met} \end{array}
\quad \begin{array}{c}
\text{S'} \\
\text{S' k} \\
\text{S} \\
\text{VP} \\
\text{V NP t_1} \\
\text{t_1, j} \\
\text{NP}_j \\
\text{Mary} \\
\end{array}
\]

In (230), all the \( A' \)-chains are well-formed, and the structure is interpreted as the union of the phrase markers: \([\text{PP}_1 \text{John met Mary t_1}]\) + \([\text{before}_1 \text{John met Suzan t_1}]\), where \( \text{PP}_1 \) is read as an operator binding a variable ranging over times.

This is the technical analysis of such constructions.\(^{109}\)

Let us consider whether there exist empirical evidence in its favor.

---

\(^{109}\) The bare XP and its counterparts occur adjoined to COMP, so their relation with their variables should be subject to subadjacency only, but it also seems to be subject to condition \( A \), which we cannot explain. The counterpart and the comparative constituent have to belong to the same GC, as shown in (i) and (ii). This could be accounted for if the two postulated movements were subject to condition \( A \):

(i) *John told Mary that he would [visit Peter before Suzan]
(ii) *John [thought that Mary saw this movie before that one]

(i) cannot have the meaning where before moves to the matrix clause. And (ii) cannot have the reading in which before is in the matrix clause, and the bare NP corresponds to an embedded argument, since the bare XP would bind the empty category from a matrix \( S' \)-adjoined position.
4.5.3 The bare XP is extracted from the comparative PP

Consider (231)a-b:

(231)a. How angrily did John enter the kitchen?
   b. *How angry did John enter the kitchen?

This contrast shows that an adjunct predicate cannot be wh-moved on the surface. Let us assume that the impossibility of extracting the AP is due to an illicit structure with respect to predication. If the AP how angry is predicated of the subject, then the predicate and its subject should c-command each other, following ideas expressed in Williams (1980). However, given that S is a maximal projection, the subject does not c-command COMP. So, predication cannot hold when the predicate is moved to COMP. This account predicts that LF movement of the AP should be as bad as its overt movement, given that predication has to hold at LF. This is verified in (232):

(232) *John entered the kitchen as angry as the living-room

This is to be compared with (231), where predication does not have to hold between the Adverb Phrase and the subject, as also shown by (233)'s acceptability:

110. It is not clear whether Williams (1980) would require the predicate to c-command the subject. Williams allows traces of predicates to count. However, in his discussion, the predicate is a primary predicate. Secondary predicates presumably do not leave traces, so they must c-command their subject in the position they move to.
John entered the kitchen as angrily as the living-room

This contrast indicates that the comparative constituent moves into COMP at LF. 111

4.5.4 Crossing effects

Consider the following sentence.

(234)a. *John gave his shirt to the same person as his boots

(234) is a crossing violation at LF between the comparative constituent and its trace, and the focussed NP and its variable as shown below:

(235)

111. Note that when the bare XP corresponds to the subject, the sentence seems to become grammatical:

(i) John entered the kitchen as angry as Tom

This may due to the fact that (i) is ambiguous between an interpretation in which Tom also entered the kitchen angry, the elliptical reading, and an interpretation in which Tom is said to be angry, but it does not imply that Tom entered the kitchen or that he did so angry.
The other order is well-formed structurally and the sentence is grammatical:

\[(236)\] John gave to Mary the same books as to Peter

The contrast illustrated here shows that the focussed analysis of the bare XP together with the syntactic representation of the missing sentence (which contains the comparative operator-variable relation), however it is realized, is on the right track.

4.5.5 Exceptions to the crossing violations

The only types of complements which give rise to crossing effects are close complements of the verb. In case the complement is not subcategorized, sentences are acceptable, as shown in (237), which is parallel to (237) above but is grammatical:

112. For some reason, Heavy-NP Shift of the direct object seems to be a necessary state of affairs, probably to make the bare XP peripheral.

113. An insertion analysis of the ellipsed sentence yields the same configuration of quantifications, and is thus also capable to explain the facts.
(237) John read this book more often than this article.

We can assume either of two things. First, that non-subcategorized complements do not give rise to crossing violations and second that non-subcategorized PPs may head comparative constructions, leaving PP variables, instead of NP-variables inside stranded PPs.

Evidence for the second hypothesis comes from the following contrasts:

(238)a. John bought his book in the same store as Peter sold his
   b. *John bought his book in the store that Peter sold his
   c. *John gave his book to the same person as Peter sold his

(238)a versus c shows that it is possible for a non-overt PP to occur in a comparative clause only if it is non-subcategorized. And (238)a versus b shows that only comparative clauses, as opposed to relative clauses, accept these non-overt PPs. What this suggests is that, in comparative constructions, the PP can be the head, binding a PP variable. In that case, as Pesetsky (1982) shows, A'-dependencies whose tail is the PP start at the VP node. Since the VP node is also the terminal point of the A'-dependency of the object, the two will not cross.
4.6 Reanalysis as a union of phrase-markers on a single plane

4.6.1 Introduction

In this section **tough**-constructions will be examined in the light of an extension of the notion of union of phrase-markers. The union of phrase-markers, as we have used it in this chapter, allows two syntagmatic trees to coexist, each one on its plane. Here, we will investigate the possibility of forming the union of phrase-markers, such that the resulting phrase-marker is represented on only one plane. We will see that there is a minimal difference between the merging of nodes such that the result is two conjuncts or the merging of nodes such that the result is on one plane only. Then, we will apply this formalism to account for **tough**-constructions. Given our hypotheses of chapter 1, which lead to the conclusion that the th-criterion does not hold at S-structure, only two stipulations will have to be made, namely, that the sentential argument of **tough** is an S and not an S', and that

![Diagram of phrase structure](image-url)
tough-predicates are operators. This section will be brief, its main interest being the suggestion that these constructions can receive an account in these terms.

4 6.2 Union of phrase-markers on a single plane

Let us consider again the form of the union of the phrase-markers below, at the X point:

(240) \[
\begin{array}{c}
X \\
Y \quad Z \\
\end{array} & \begin{array}{c}
X \\
V \quad W \\
\end{array} \quad \rightarrow \quad \begin{array}{c}
X \\
Y \quad Z \quad V \quad W \\
\end{array} \\
\text{plane 1} & \text{plane 2}
\end{array}
\]

In the general case, in the resulting structure, the two merged phrase-markers will have to belong to different planes, because if they were represented on a single plane, they would be interpreted as one single structure. And licensing principles will ensure that such duplication may not arise, for the same reason that these structures cannot be base-generated: they do not fall under the Projection principle. To take an example, consider the junction of two S's. If the resulting structure was a single constituent, it would be an S dominating two subjects and two VPs (the head of S would merge, by the convention that heads merge when their projections do, so there would only be only one head, which is well-formed, by X-bar theory.) Such a structure would be excluded by the fact that INFL, the head of S, takes only one complement, VP. So, one of the VPs would not be licensed,
ruling the structure out.

Now, suppose that we do want to obtain a well-formed single structure from the union of two phrase-markers on a topmost node. Then, by convention, the union of two nodes of category X will be obtained by forming a branch between the two nodes X, as in (241):

(241) \[\begin{array}{c}
X \\
Y Z \\
V W \\
\end{array} \rightarrow \begin{array}{c}
X \\
Y Z \\
V W \\
\end{array}\]

Then, licensing principles will apply on the resulting structure, but, a priori, it is not necessarily excluded, since merging does not put either structure under the influence of the head of the other structure. So, technically, the difference between the union of two nodes on different planes or on a single plane is reflected as the difference between superimposing the nodes or relating them with a branch.

(242) **Union of phrase-markers on a single plane (definition)**

X forms a union on a single plane with Y only if X and Y are of the same category and X and Y are related by a branch.

Note that it impossible to form the union of more than one node, since the structure would be an ill-formed tree:

(243) \[\begin{array}{c}
X \\
Y Z \\
\end{array} \rightarrow \begin{array}{c}
X \\
Y Z \\
Y W \\
\end{array}\]

Let us now analyze tough-constructions.
4.6.3 Tough-constructions

Consider (244):

(244) John is easy to please

The traditional problem of tough-constructions is that the syntactic structure cannot be a simple projection of the thematic structure, given conditions on A-movement, i.e. movement to an Argument position. In our theory, it is possible to analyze the S-structure of (244) as deriving from movement of John to the subject position, since the th-criterion does not hold at S-structure, so long as the derivation does not violate Sportiche's (1983) Isomorphy principle (cf. chapter 1), which prevents an argument from changing its thematic role in the derivation: thematic structure is invariant. The licensing principle on empty categories, which forces them to be bound, will analyze the empty category as bound by John at S-structure. But they do not have to form a chain, so the relation does not have to obey condition A.

We assume that D-structure is the following, where the S after easy is not its object, but is in an adjunction position, and receives the external th role of easy:

(245) [e] is [[easy] \text{AP} [\text{PRO to please John}]] \text{AP}

The subject position is non-thematic, so any NP can move in it, so long

114. For an overview of the properties of these constructions, see Browning (1984).
as the th-criterion is satisfied at LF.

The problem is that, at LF, there should be a chain between John and the empty category, but this chain violates condition A, since the GC of the empty category is the extraposed S, which does not contain John. We will assume, first, that the sentential argument of tough is an S, not an S'. So, the S-structure of (247) is (S is short for INFL):

\[
(246)
\begin{array}{c}
\text{S} \\
\text{NP} \quad \text{INFL'} \\
\text{John} \quad \text{INFL} \quad \text{VP} \\
\text{V} \quad \text{AP} \\
\text{is} \quad \text{AP} \quad \text{S} \\
\text{easy} \quad \text{PRO} \quad \text{INFL'} \\
\text{INFL} \quad \text{VP} \\
\text{to} \quad \text{V} \quad \text{NP} \\
\text{please e}
\end{array}
\]

In short, our analysis of tough-movement is that the extraposed sentence, [PRO to please e] will become the main clause at LF, by forming a union of phrase-markers with it. And the matrix itself will "disappear" in the process, in that there will be only one S' left -- thematically, the following. [\text{sg}, PRO to please John] And this sentence will be in a configuration of predication with the tough-predicate. Let us consider the details.

For the LF derivation of (246), the second assumption is that tough is an operator. Being an operator, it can be assigned scope over its
argument, carrying with it one of the projections which dominates it, here, INFL'. Since easy is an operator, and since it takes a sentential complement, it is possible to move it to a position such that it is in a proper syntactic relation with its argument at LF. This argument is the whole sentence, which comprises the subject NP John, so it should be assigned scope over the matrix clause. Now, we are going to look at the structure that each step of the derivation leads to, which is only a way of speaking, since the derivation does not really go through stages in between two levels, but it helps to see what the ultimate LF structure will be. So, the first step is to move the INFL' to a position such that the AP is in a predication relation with the matrix S. Note that it moves to S', not S (C is COMP):

(247)  

```
          S'       
        / \     
       INFL'  S'   
      /   \   / \ 
     INFL  VP  C INFL''
    / \   / \  / \
   AP  NP  tj  
  /  \   /  \  /   \
 AP  S  John
```

easy PRO VP
to please e

In (247), the AP is in a configuration such that, by transmission through the predicable verb be, it can be interpreted as assigning its thematic role to the S' that it adjoins to; INFL' c-command its subject, S', and vice versa. Note that the dominating adjunction node does not count as the category which is the argument of the predicate, but rather, the original node, which does not dominate it. This is expected, in
adjunction structures: the dominating node does not count, and the relevant relation is always between the original node and the category adjoined to it (as in, e.g. Small Clauses).

The second step is the extraposition of the S contained in the AP. As we have assumed in this dissertation, categories in an adjoined position, such as relative clauses or comparative clauses, may be extraposed at LF. So, suppose that S gets extraposed, and forms a union of phrase markers on the same plane with the other S. Since they have the same category label, they may do so. This yields an additional branch, as in (248):

(248)

Now, in (248), the two movements have been represented as leaving traces. But they can be eliminated. Consider what happens if they are.
In terms of the structural change, the extraposed $S$, PRO to please $e$, has been moved to a position such that it forms a segment with a node of the same category: this is an adjunction configuration. And these two nodes are directly dominated by $S'$. So, in fact, the structure becomes a full $S'$ predicated by the AP easy through the INFL INFL is easy. And the NP John, initially the subject of the moved clause, becomes represented as adjoined to the new $S$ on its right. Given free indexing, the empty category may be coindexed with it, making it a QRed NP. And, from an A-position its position becomes an A'-position. This does not matter, so long as it may form an A'-chain with the empty category. It does since this chain respects condition A. Note that the fact that quantifiers must bind at least one empty variable entails that the position inside the sentence complement of the AP must be empty.°°

We should claim that the relation between John and the empty category obeys condition A, as shown by the ill-formedness of (250):

115. Thanks to Ken Hale, who warned me against some unwanted possibilities.
(250) John is easy to convince Mary that she will be found of

The exceptions to this generalization involve infinitivals, and we may assume that some property of QR makes infinitivals transparent, in English.

(251) John is easy to forget to criticize

Going back to the main lines, the only change from D-structure to LF, in terms of the thematic structure of the sentence, is that the argument of tough is the S [PRO to please John] at D-structure, and the S' [COMP [PRO to please John]] at LF. We may assume that S and S' are both manifestations of the same argument, a clause, and that the presence or absence of the S' node does not induce a change in the category of the constituent. That is to say, S and S' have to be considered projections of the same syntactic category, even if both are projections of different heads, INFL and COMP. This dilemma can be solved if we adopt Aoun's (1981) view that these are the two formatives of one discontinuous constituent. And the theory of discontinuous constituents should have as a consequence that projections of discontinuous heads may form different maximal projections but constitute the same category.

Note that the same problem arises in S'-deletion phenomena: if the predicate believe th-marks an S' at D-structure, and this S' turns into an S in the derivation, then the predicate th-marks an S at later levels. This is not a Projection Principle violation if S and S' are
considered non-distinct categories for selection.

Returning to our stipulation that tough-predicates take an S and not an S' at D-structure, the other choice, S', yields an ill-formed result, as shown in the following LF:

\[(252) \]

```
S'
  \ /
 INFL' S'
  \ /
 INFL VP C S S'
  \ /
 is AP NP C S
  \ /
 easy John PRO INFL'
  \ /
 INFL VP
     \ /
 to please e
```

In (252), merging takes place with the topmost nodes, that is to say, with S'. Then, the NP John is too far from the empty category to form a chain respecting condition A. The sentence is thus excluded as a violation of the requirement that an element in an S-adjoined position forms a chain with a variable respecting condition A. Note also that merging is allowed only with one node, since one can add only one branch in the tree, so there is no way to go down the tree and lessen the distance between John and the empty category, as was possible, for example, in paycheck-sentences.
4.6.4 Consequences

One consequence of our view of S-structure as being oblivious of the th-criterion and the Projection principle, and of the particular analysis of tough-constructions, is that the empty subject of the tough-sentence may be PRO-like all along. The fact that it is inside an S, at S-structure, and hence is governed by the matrix predicate, is not a problem, since the binding theory (chain-formation theory) applies at LF only. So PRO may be governed at S-structure. One consequence of the change of S into S' at LF, is that PRO becomes protected from government at LF, the desired result.

One other result is that the generalization that NP-traces occur in complementary distribution to A'-bound empty categories with respect to Case-marked positions can be kept unchanged: all variables occur in Case marked positions (maybe questionable), and all NP-traces occur in non-Case marked positions.\textsuperscript{116} One problematic question about tough constructions is to integrate the NP John inside a thematic chain, without making the chain an A-chain, if this generalization is to keep intact. A reanalysis account like that proposed in Chomsky (1981) has to deal with this problem, as pointed out to me by Rita Manzini.

Another consequence of this analysis is that it does not require the presence of an empty operator, to license the empty category at

\textsuperscript{116} But this contradicts our analysis of S'-s as predicates of SCs, where an A-chain is allowed to be formed with a Case-marked empty category. Thanks to Diane Massam for mentioning this question to me.
S-structure, since the subject can do it, by binding it (any binding licenses empty categories).

As for the fact that the moved NP cannot correspond to the subject, the question will remain only partially answered. Following Clark's (1984) claim about gerundival constructions such as (253), we will assume that tough-constructions are control structures.

(253) Dinette told me that Jespersen's book deserves reprinting [e]

In the tough case, the implicit argument is understood as a controller. This means that there must be an argument to control. The fact that it must be PRO, and that it cannot be an object, should be accounted for, but I have no answer to this. Another unanswered question is posed in the next section.

4.6.5 Other empty-operator constructions

Up to now, we have studied two types of constructions, parasitic-gap and tough-constructions, both of which require the presence of an empty operator in Chomsky's (1985a) theory. And in both cases, we have dispensed with it, trading it for the union of phrase-markers, either on different planes, or on the same plane. If this exchange works for two types of constructions, a desirable result for simplicity would be that all analyses requiring the empty operator are interchangeable with the union-of-phrase-markers analysis. Hence, extent-adjectival constructions, as in (254), should involve a union of phrase-markers:
(254) John is too stubborn to talk to [e]

But since it is not clear how to deal with such sentences without the empty operator, this will be left open.

4.6.6 Clitic-movement in tough-constructions in French and Italian

One advantage of the analysis proposed here is that it explains how the relation between John and the empty category may respect condition A, without at the same time allowing it for any other category. In Italian, where clitics may climb to the matrix verb in restructuring sentences, as has been shown and discussed by Rizzi (1978), Burzio (1981), Manzini (1983), Baker (1985), and others, it is a puzzling fact that clitics cannot climb in tough-sentences, as pointed out to me by Luigi Rizzi:

(255) *Il libro gli fu difficile da offrire
   'The book to-him was hard to offer'

One must allow a distinction to be drawn between the subject and the clitic, with respect to their relation with their traces. At LF, the clitic is attached to the verb which moves to S', so it still is too far away from its trace, contrary to the subject, which gets included inside the S that contains its trace, as shown below:
Despite its explanatory character for this type of clitic movement, our analysis has nothing as yet to say about the fact that clitics cannot move to the embedded verb either, which Luigi Rizzi also pointed out to me. I will give the example in French, but the facts are the same in Italian:

(257) *Ce livre a été facile à y obtenir
'This book was easy to there get'

Since the clitic-chain does not change by the extraposition of the sentence, it is impossible to state this ungrammaticality in terms of the form of this chain alone. I will not provide an answer here, but let us note that there exists a difference between subcategorized and non-subcategorized clitics. The former are acceptable, and not the latter. Compare (257) with (258):

(258)a. Ce livre a été facile à lui offrir
'This book was easy to to-him offer'

This shows that more needs to be said about tough-constructions, but
this does not falsify our general approach.

In conclusion, this chapter has shown that a certain number of constructions fall under an analysis where conjuncts may be formed at LF, with constituents like adjuncts and relative clauses. We have come to the result that LF structures may differ drastically from S-structures, in situations in which these structural changes appear to be necessary, as in tough-constructions, or constructions involving operator adjectives, or parasitic gap constructions.
Bibliography


Schein, B. (1984): "Reference to Events and Quantification," ms. MIT.


