ON THE CONTENT OF EMPTY CATEGORIES

(Vol. I)

by

Denis Bouchard

B.A., Université de Montréal
(1977)

M.A., Université de Montréal
(1979)

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

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 1982

© Denis Bouchard 1982

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 Denis Bouchard
Department of Linguistics and Philosophy
July 29, 1982

Certified by
Noam Chomsky
Thesis Supervisor

Accepted by
Samuel Jay Keyser
Chairman, Departmental Graduate Committee

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
DEC 1 1982
Hum.

ON THE CONTENT OF EMPTY CATEGORIES

by

DENIS BOUCHARD

Submitted to the Department of Linguistics and Philosophy on August 1, 1982 in partial fulfillment of the requirements for the Degree of Doctor of Philosophy

ABSTRACT

We assume that some principle such as the Projection Principle of Chomsky's Lectures of Government and Binding is basic in the grammar because of its contribution to the descriptive part of the grammar since it simplifies the mapping between the different levels of the grammar, this simplified mapping also having a contribution to explaining language acquisition. Such a principle depends crucially on empty categories: if determining the properties of ECs requires specific statements in the grammar, this weakens the import of the Projection principle, but if the properties of ECs follow from principles independently motivated in the grammar, then the Projection principle receives significant support.

The goal of this thesis is to defend this use of ECs and to strengthen the theoretical choice of such abstract elements by showing that no specific statement ever has to refer to ECs in the grammar. We develop a modular grammar where the different components apply to categories in general, and it is the interaction of the systems of the grammar that allows some categories to be empty in some cases. Thus this gives crucial support to a model of grammar where ECs play a role since the properties of ECs depend on systems of rules and principles that are independently motivated for categories in general.

Assuming a model of grammar such as the one proposed in Chomsky and Lasnik (1977), where S-structure is mapped into two levels that do not interact directly with one another, Phonological Form and Logical Form, we propose a complementary approach to ECs: an EC is a category that has no feature that is "visible" in PF, although it has the features relevant for interpretation in LF, so that an EC is simultaneously "empty" in PF but "full" in LF. Taking the category NP as a case-example, a general principle

of Lexicalization applies in PF which states that a N is "pronounced" if it has features and not pronounced if it has no features. In LF, the principle of Denotability requires that an NP have a R-index for proper interpretation to take place, and the principle of Agreement governs the need of proper features of person, number and gender.

A study of lexical anaphors shows that the notion of Binding, which makes crucial use of government, plays a central role in determining what an anaphor is. This notion of Binding extends to ECs, which are shown to be either a Bound anaphor or a pronoun.

In Chapters 3, 4, and 5, we show that the interaction of these principles and of the notion of Binding allows us to derive some of the statements of Chomsky's Lectures on Government and Binding that referred specifically to ECs, all the principles and rules that we propose being stated on categories in general. Thus the ECP, the analysis of Pro Drop in both configurational and nonconfigurational languages, and the theory of Control can all be derived from principles that apply to NPs in general, this giving crucial support to the Projection Principle and the use of ECs.

Thesis Supervisor: Noam Chomsky

Title: Institute Professor

ACKNOWLEDGEMENTS

I would like to thank all of those who helped me in so many ways in pulling this through. First, to my advisor, Noam Chomsky, many thanks for his detailed discussions of every point and for generally showing me what it is like to be a linguist. Thanks also to the other members of my thesis committee: Paul Kiparsky, for sharing with me part of his marvelous scholarly knowledge and for his indepth views of the field: Paul is the one who would often say: "one must take the theory seriously"; that is what I have tried to do; Ken Hale also helped me to have a glimpse of what it means to know something about language. I also had the pleasure to have Morris Halle as an ever-reminding voice that one must always doubt and question. Thanks also to Joan Bresnan for her example of a fine syntactician.

A special salute also to the guys at Université de Montréal: Raj Singh, who was the first one to show me what linguistics is all about and who got me to take a closer look at a field for which someone could have such a passion; Yves-Charles Morin and John Reighard whose strong commitment to accuracy and honesty have had a strong influence.

Life at MIT has been a great intellectual adventure, and the pleasure that I had is due to the fine personal and intellectual qualities of the students and visiting scholars that were around: J Aoun, H. Borer, N. Fabb, I. Haik, J. Huang, L. Levin, R. Manzini, M. Montalbetti, C. Neidle, D. Pesetsky, B. Poser, P. Pranka, M. Rappaport, H. vanRiemsdijk, S. Rothstein, K. Safir, M. Saito, J. Simpson, D. Sporticke, T. Stowell, D. Steriade, E. Wehrli. Special thanks to Lisa Travis who had the job to type the thing. In Anne Rochette and Dough Pulleyblank (and family) I have made new friends with which to share a bit on "mal du pays" once in a while. During my stay at MIT, I have benefitted from the very generous support of the conseil de recherches en sciences humaines du Canada and la direction générale de l'enseighement supérieur du Québec (Fonds FCAC). My parents have been a source of encouragement all along this period of three years.

I could never have gone through this adventure and remained sane had it not been for Catherine's constant support and reminder that life still existed out there: thank you Cat, and I hope I can sort of get back now. Finally, Merci Sébastien: on sait tous les deux que ça n'a pas toujours

été facile, mais j'ai fait çà en très grande partie pour toi:
je te dédis ma thèse et mes trois années de recherches.

In a minute there's time
For decisions and revisions which a minute will reverse.

T.S. Eliot

TABLE OF CONTENTS

ABSTRACT 2

ACKNOWLEDGEMENTS 4

CHAPTER 1: THEORETICAL OUTLINE 8
Footnotes: Chapter 1 40

CHAPTER 2: DETERMINING THE PROPERTIES OF THE EMPTY CATEGORY 47
2.1 The issue 47
2.2 What is an empty category? 50
2.2.1 The content of an empty category 50
2.2.2 F-features and Domain D 53
2.2.3 The principle of Denotability and lexicap NPs 59
2.2.3.1 Names 59
2.2.3.2 Pronouns 59
2.2.3.3 Anaphors 60
2.2.3.4 Summary 89
2.2.4 The principle of Denotability and [NPe] 89
2.2.4.1 Assignment of R-indices 90
2.2.4.2 F-features and the EC 94
2.2.4.3 Consequences of the complementary approach to ECs 98
2.2.4.4 Summary 110
2.2.5 The partitioning of the EC 111
2.3 Some consequences with respect to coreference 115
2.4 Summary 167
Footnotes: Chapter 2 170

CHAPTER 3: THE EMPTY CATEGORY PRINCIPLE 192
3.1 General Comments 192
3.2 The NP-trace relation 195
3.3 The WH-phrase trace relation 211
3.3.1 General comments 211
3.3.2 WH-trace and Case 218
3.3.2.1 Hebrew Free Relatives 220
3.3.2.2 French Relative Clauses ...224
3.3.2.3 Problems 240
3.3.2.4 The Visibility Hypothesis .257
3.3.3 Successive cyclicity and subjacency 273
3.4 ECP and LF Movement 310
Appendix: The Avoid Pronoun Principle and the Elsewhere Principle 318
Footnotes: Chapter 3 327

CHAPTER 4: PRO DROP LANGUAGES 358
4.1 General Comments 358
4.2 Pro Drop in configurational language 360
4.3 Pro Drop in non-configurational languages ..392
Footnotes: Chapter 4 402

CHAPTER 5: CONTROL THEORY 411
5.1 General Comments 411
5.2 PRO as an anaphor 416
5.2.1 The distribution of PRO: Case or government? 432
5.2.1.1 Government but no case 432
5.2.1.2 Case is assigned, but there is no government 438
5.2.1.2.1 Gerunds 439
5.2.1.2.2 Case assignment with PRO 451
5.2.2 Other aspects of local control 461
5.3 PRO as a pronoun 471
5.4 On differences between trace and PRO 481
5.5 Concluding remarks 483
5.6 Appendix: control in French 487
Footnotes: Chapter 5 496

BIBLIOGRAPHY 506

CHAPTER 1: THEORETICAL OUTLINE

The basic assumption of generative linguistics has been from the start that the purpose of linguistic theory is to understand the nature of the language faculty and to explain the acquisition of language. So, faced with the diversity of existing (or even possible) grammars, the theory must be constructed in such a way that it will be compatible with this wide range of often apparently poorly related grammars on the surface. But at the same time, the theory must be sufficiently constrained and restricted in the options it permits to explain the fact that language is acquired on the basis of impoverished stimuli and despite the unavailability of direct negative evidence. This problem of constructing a descriptively adequate grammar while maintaining explanatory adequacy with respect to language learning has led to the assumption that the language faculty is best characterized as a biological faculty, a mental organ of some sort (cf. Chomsky (1955) and the work that followed, especially Chomsky (1975)). This mental organ is often referred to as Universal Grammar (UG). UG is assumed to have inherent properties of its own. In order to meet both conditions of adequacy, the theory of UG is postulated as highly structured so that UG can narrowly restrict the class of possible grammars that a child can infer on the basis of limited and defective data. So the theory of UG is assumed to be "based on a number of fundamental principles that sharply restrict the class of attainable grammars and narrowly constrain their form,

but with parameters that have to be fixed by experience" (Chomsky (1981a) p. 3-4).

If the theory of UG is sufficiently rich in structure, then fixing values for the parameters embedded in UG will allow for a great diversity of languages, since the closely knit structure of UG will have for effect that the choice of one value for a parameter will possibly have repercussions in several components of the grammar. This approach to UG is meant to solve the problem of descriptive adequacy. Yet at the same time, it can solve the problem of explanatory adequacy since it allows the learner, presented with limited evidence, in fact evidence just sufficient to fix the parameters of UG, to determine a grammar that may be very intricate.

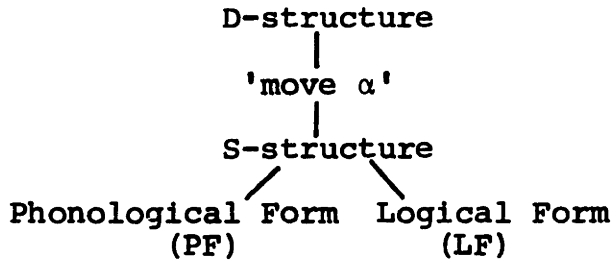
Informally speaking, we can imagine UG as a grid with which the child is equipped, and that his task is to pigeon-hole the linguistic data to which he is exposed, thus determining the values of the parameters, with an intricate intertwining net of consequences on the grammar of the language which the child will eventually end up with. Such a theory of acquisition is therefore not one of testing of hypotheses, but rather one of fixing parameters.

The approach to learnability sketched above implies that a certain basis, common to all language learners, i.e. UG, will have to be very general, and therefore that UG will contain rather abstract principles and operations if it is to account

at the same time for language variation. This approach also implies that the structure of UG will be highly modular. One framework which seems to us to meet these conditions and which we will therefore adopt in its broad outlines is the Government and Binding (GB) framework, as sketched in Chomsky (1980, 1981a, 1981b) and related work, which came out of the Extended Standard Theory (Chomsky 1973, 1975, 1976; Chomsky and Lasnik 1977, and related literature).

In the GB framework, a core grammar is determined by setting the values of the parameters of UG. Core grammar is assumed to have a structure as in (1), following Chomsky and Lasnik (1977).

(1)



As it is clear from the diagram in (1), the essential claim of such a model is that S-structure representations feed into two components, PF and LF, and that these components do not interact with each other.[1] What takes place in one of these two components is "invisible" to what takes place in the other, so that an operation taking place in LF cannot have any effect on rules applying in PF, nor can an operation in PF have any effect on operations in LF.

The D-structure component in (1) can be factored out in two subcomponents: the lexicon and the categorial component. The

lexicon provides specifications about the abstract morphophonological structure of each lexical item, its syntactic features (i.e. categorial and contextual features), and also thematic and selectional specifications for its complements. The categorial component contains phrase structure rules which meet some variety of the $\bar{X}$-system (see Chomsky 1970; Jackendoff 1977; Stowell 1981a; and others). D-structures are generated by inserting lexical items into the structures generated by the phrase structure rules. D-structure is the level at which thematic and subcategorization requirements for interpretation are met locally: that is, there is a one-to-one correlation holding between referential expressions and thematic roles, and between syntactic categories and subcategorization frames.

The transformational component consists essentially of the rule move α which maps D-structures into S-structures. This rule leaves traces, that is, empty categories coindexed with their antecedents. The rule move α may also operate in the PF and LF components. In the former, it performs stylistic movements, that is, movements that have no bearing on the interpretation of the sentence but only on the surface order of some constituents for example. In the latter, it is responsible for QR, WH-Raising and Focus interpretation, rules that have no effect on the surface order of constituents but which modify the structure of a sentence for interpretive purposes (see May 1977; Chomsky 1981a; Aoun, Hornstein, Sportiche 1981; and others).

This is a sketch of the subcomponents of the rule system. In addition to the rule system, it is assumed in GB that there are subsystems of principles that interact with each other and with the rule system. In fact, in recent years, the focus of linguistic research has shifted from the study of the rule system to the study of the subsystems of principles. This shift is similar to the change of perspective brought about by Ross (1967) to the approach to transformational rules: instead of having complex transformations which often showed redundancies in the statement of their structural descriptions, their structural changes and the conditions which were stated on their application, Ross proposed to have simpler transformations with more general conditions that applied to all transformations. Another similar search for the properties that underly successful grammars is the case of Conditions on Transformations (1973) where Chomsky pushes further this quest of the abstract properties that make successful grammars work. The final outcome of this change of perspective came about in Chomsky and Lasnik (1977) where all transformational rules of movement were reduced to move α, the other interacting systems of the grammar excluding all the cases of overgeneration.

In the GB framework, the subsystems of principles are the following:

(2) (i) bounding theory

(ii) government theory

(iii) θ-theory

(iv) binding theory

(v) Case theory

(vi) control theory

Let us have a quick look at what these notions convey, keeping in mind that these notions will be made more precise and in some cases will be modified as we go along in the next chapters.

The bounding theory states locality conditions on certain processes: a case-example is Subjacency, which constrains the application of move α (or its output structure if it is stated not as a condition on rules but as a condition on structures).[2]

The notion of government is central in GB. The core notion of government is meant to express the relation that holds between a head and its complements: it has a clear thematic content. The operative notion of government, however, involves structural configurations which generalize the core notion. Consider the examples in (3).

(3) a. $[_{VP}\ V\ [_{NP_1}\ NP_2\ \ \bar{N}\ ]]$

b. $[_{VP}\ V\ [_{\bar{S}}\ COMP\ [_{S}\ NP_1\ \ INFL\ \ VP\ ]]]$

c. $[_{VP}\ V\ [_{AP}\ A\ of\ NP_1\ ]]$

d. $[_{PP}\ P\ [_{NP_1}\ NP_2\ \ \bar{N}\ ]]$

In (a), we want V to govern NP_1, but we do not want V to govern NP_2. Similarly in (b), we do not want V to govern NP_1, and in (c), V must not govern NP_1. In (d), the P must govern NP_1, but not NP_2. The notion to capture is that the daughter node α of

a node $\underline{\gamma}$ governs all the nodes under $\underline{\gamma}$ except if a maximal expansion intervenes between $\underline{\alpha}$ and the governed node. So the core notion of government has the following properties: in the structure $[_{\underline{\gamma}} \ldots \underline{\beta} \ldots \underline{\alpha} \ldots \underline{\beta} \ldots]$, if $\underline{\alpha}$ governs $\underline{\beta}$,

1° $\underline{\alpha} = X^{o}$

2° $\underline{\alpha}$ is an immediate constituent of $\underline{\gamma}$

3° $\underline{\gamma} = \overline{\underline{\alpha}}$

4° $\underline{\beta}$ is an immediate constituent of $\underline{\gamma}$

Note that property 3° is derivable from 1° and 2°: if $\underline{\alpha}$ is an X^{o}, and if $\underline{\alpha}$ is an immediate constituent of $\underline{\gamma}$, then $\underline{\gamma}$ is necessarily $\overline{\underline{\alpha}}$. So 3° does not have to be stated in the definition of the core notion of government. Furthermore, if the core notion of government is meant to express subcategorization relations, then it is not necessary to mention 1° in the core notion of government: since subcategorization is a lexical property, only X^{o} elements can be governors as far as the core notion of government is concerned. So properties 1° and 3° do not have to be incorporated in the formal definition of the core notion of government. Therefore, the core notion of government can be stated as in (4).

(4) Government[3]

In the structure $[_{\underline{\gamma}} \ldots \underline{\beta} \ldots \underline{\alpha} \ldots \underline{\beta} \ldots]$,

$\underline{\alpha}$ governs $\underline{\beta}$ if and only if

(i) $\underline{\alpha}$ is an immediate constituent of $\underline{\gamma}$

(ii) where $\underline{\psi}$ is a maximal projection, if $\underline{\psi}$ dominates $\underline{\beta}$ then $\underline{\psi}$ dominates $\underline{\alpha}$.

Maximal expansions are NP, PP, AP and $\bar{S}$. Note that we do not consider VP to be a maximal expansion. The consequences of this choice will be explored in the following chapters.

θ-theory is concerned with the assignment of thematic roles (θ-roles) such as agent-of-action, theme, etc. (see Gruber (1965), Jackendoff (1972)). The correlation between the assignment of a θ-role to certain positions and the referential expressions which fill these positions must be one-to-one. This is captured by the θ-criterion which can be informally stated as in (5).

(5) θ-criterion

A. Each θ-position is assigned an argument.

B. Each argument is assigned a θ-role.

(For some discussion of these properties, see Freidin 1978; Borer 1981; Bresnan 1982; and Chomsky 1981a).

Binding theory is concerned with relations of anaphors, pronouns, names and variables to possible antecedents. The following binding conditions are given in Chomsky (1981a):

(6) Binding conditions:

A. An anaphor is A-bound in its governing category.

B. A pronominal is A-free in its governing category.

C. An R-expression is free everywhere.

The definitions of the notions bound and governing category of Chomsky (1981a) are given in (7) and (8) respectively.[4]

(7) α is X-bound by β iff α and β are coindexed, β c-commands α and β is in an X-position. (X=A, $\bar{A}$)

(8) $\underline{\beta}$ is a governing category for $\underline{\alpha}$ iff $\underline{\beta}$ is the minimal category containing $\underline{\alpha}$, a governor of $\underline{\alpha}$, and a SUBJECT accessible to $\underline{\alpha}$ (where SUBJECT is "the most prominent nominal element" in an expansion, including AGR. Accessible means informally that linking of $\underline{\alpha}$ and the SUBJECT must not violate well-formedness conditions like the i-within-i condition for example).

The notions of anaphor, pronominal and variable will be made more precise in Chapter 2.

Case theory is concerned with the assignment of abstract Case and its morphological realization. The Case Filter, informally stated in (9), sanctions the appearance of lexical NPs.[4]

(9) *N, if N is lexical and has no Case.

As with government, the core notion of Case is a simple one, namely that lexical NPs have to bear Case. But when the notion is made more precise by specifying the structural configurations where Case can be assigned and how it can be used in the grammar, then the operative notion of Case is extended to cover a much wider range of facts as we will see further on.

Control theory is concerned with the potential for reference of the abstract pronominal PRO, which is the subject of infinitival clauses in the GB framework. As we will see in Chapter 5, control theory deals primarily with the recoverability of the content of PRO.

The notion of government is pervasive throughout the GB mo-

del of grammar. It is directly relevant for subcategorization, θ-marking and Case assignment. Government is also relevant for binding theory since it enters in the definition of governing category, to which crucial reference is made in the binding conditions. Attempts have been made to derive at least part of the bounding condition of Subjacency by the use of the Empty Category Principle, once again crucially involving the notion of government (see Kayne 1981a).

Another element that plays a crucial role in GB is the Projection Principle of Chomsky (1981a), stated informally in (10).

(10) Lexical requirements must be met at every level.

These lexical requirements include subcategorization frames and θ-role assignment. The levels in (10) are those where these lexical requirements are relevant, namely D-structure, S-structure and LF. It follows from (10) that at D-structure, there must be a one-to-one correlation between lexical requirements and single elements, since no linking mechanisms have applied yet at D-structure. On the other hand, when S-structure is reached, a network of links has been established by move α. The lexical requirements can be satisfied in two ways at S-structure: either the single element is still in a position where it can satisfy the lexical requirements, or it is linked to a trace which is in such a position, that is, the position from which the single element originated in D-structure before being displaced by move α.

These links which are created by move α are called Chains. Chains allow all interpretation to be done at S-structure, since

the thematic and subcategorization information provided by the D-structure configurations are preserved at S-structure by means of the links to the traces left by move α in the D-structure positions. Consider the example in (11).

(11) a. Cato saw Doug.

b. $[_{NP}e]$ was seen Doug (by Cato)

c. $Doug_i$ was seen t_i (by Cato).

In (11a), the verb saw subcategorizes for an NP complement to which it assigns a θ-role in post-verbal position, presumably the θ-role of patient. When the verb appears in its past participial form seen as in (11b), we can assume that its subcategorization frame and θ-assignment properties have not changed. So the Projection Principle is met straightforwardly in (11b). If we assume that (11b) is the D-structure from which (11c) is derived by move α, we now have an S-structure with a Chain which consists of $Doug_i$ and its coindexed trace. It is this Chain which allows the Projection Principle to be satisfied since the Chain consists of a trace in the θ-position following seen and the argument Doug, which can fulfill the lexical requirements of seen.

The Projection Principle was introduced by Chomsky (1981a) to overcome the problem of an unwanted redundancy between the rules of the categorial component and the lexicon. In a grammar of the sort outlined in Chomsky (1965), for example, the lexicon contained information about the type of complements that a verb can take. So, for example, it would be specified in the

lexicon that a verb like persuade takes an NP complement followed by an $\bar{S}$ complement. But on the other hand, the fact that a verb can take an NP and an $\bar{S}$ complement was also specified in the phrase structure rule of VP expansion. So in such a model of grammar, the categorial component was in fact a specification of redundancy rules of the lexicon, i.e. an explicit formulation of the class of subcategorization frames that are found in the lexicon.[5] It is clear that someone learning English will have to discover in some way what the subcategorization features of a verb are: this is part of learning its meaning. With the introduction of the Projection Principle, and given that the subcategorization features have to be learned anyway, this means that the basic properties of the syntactic structures in which a particular verb appears can be determined by the Projection Principle and need not be represented independently in the grammar. This has interesting consequences on aspects of the acquisition of language, as pointed out by Chomsky:

> A person who knows the word persuade (hence knows its lexical properties, specifically, its subcategorization features) can at once assign an appropriate LF-representation and S- and D-structure when the word is heard in an utterance, or in producing the word, and will recognize the sentence to be deviant if other properties of the utterance conflict with this assignment. Hence languages satisfying the projection principle in their basic design have obvious advantages with respect to acquisition and use.
>
> (Chomsky, 1981a, p.31)

The Projection Principle has for effect to draw out of the rules of the categorial component most of their impact: apart from order, all the rest of the information that they provide

for syntactic structures can be extracted from the lexicon, given the Projection Principle. But even specifications about order can be reduced drastically, as shown in Stowell (1981a), since many restrictions on the order of constituents can be derived by independently motivated restrictions on Case assignment. Stowell shows that if one adopts a requirement of strict-adjacency on Case assignment (as proposed in Chomsky 1980), the order of NPs with respect to other complements in a VP for example can be accounted for without having to specify it in a phrase structure rule as in (12).

(12) VP → V NP PP

If the order of NP and PP were reversed, then the NP would not be strictly adjacent to V, and so would not be Case-marked. In the case of $\bar{S}$ complements as in (13),

(13) VP → V PP $\bar{S}$

Stowell assumes that tensed $\bar{S}$ complements bear a Case assigning feature (i.e. the NP subject of a tensed $\bar{S}$ is assigned Case by INFL). He proposes the Case Resistance Principle in (14).

(14) Case may not be assigned to a category bearing a Case-assigning feature.

Therefore, a $\bar{S}$ complement cannot be adjacent to V when Case is assigned or else it would violate (14).[6]

If an approach such as Stowell's turns out to be satisfactory, then specifications about order in the categorial component can be radically reduced: one could only be required to speci-

fy if the head of a phrase is in initial or final position. Assuming this hypothesis to be workable, then the categorial component might be reduced to general specifications of the $\bar{X}$-type as in (15) (from Stowell 1981a).

(15) A. A phrase is labelled according to the category of its head.

B. The head is specified as final or initial.

C. Specifiers appear at the $\bar{\bar{X}}$ level, complements at the $\bar{X}$ level.

Note that if we assume, as it seems quite plausible, that specifiers are not subcategorized for but that complements are, then the specification in (15C) can be derived if subcategorization is dependent on government, which we assume, and if the notion of government is the one presented in (4) above. By this definition of government, a lexical head can only govern nodes that are under $\bar{X}$, not $\bar{\bar{X}}$. So the notion of government in (4) allows a further simplification of the categorial component.

This general approach to attempt to derive the properties of the categorial component from other independently motivated properties of the grammar is a step in the right direction. There is one problem, however, with respect to this in Chomsky (1981a) in that Chomsky keeps a categorial rule of expansion of S: this rule is given in (16).

(16) S → NP INFL VP

Chomsky (1981a) gives this rule as a principle of UG. This is unfortunate because it weakens the conceptually attractive ap-

proach to the categorial component presented above. In fact, once one such rule is admitted in the grammar, then there is little motivation not to admit more rules of the same type. Stating (16) as a principle of UG is no real answer to this problem. Note furthermore that the three nodes to the right of S in (16) are obligatory. This is a violation of basic $\bar{X}$-theory, where normally only the head of a phrase is obligatory, since labelling of the phrasal nodes depends on the syntactic category of the head. So in (16), one must stipulate what element is the head of S (namely INFL for Chomsky).

The reason why Chomsky (1981a) keeps (16) is because of the obligatoriness of the subject of a sentence. Thus he wants the infinitive complement in (17) to have a PRO subject.

(17) The barbarians tried $[_S$ PRO to destroy the city $]$

But the fact that destroy has a subject in (17) is already derivable from the Projection Principle and the θ-criterion since the VP formed by destroy assigns a θ-role to the subject. Chomsky says that this approach to derive (16) cannot work since the nominal derived from destroy should be subject to the same conditions, and yet it does not require a subject, as we see in (18).

(18) $[_{NP}$ the destruction of Rome$]$

So Chomsky says that (16) must be stipulated in the grammar in order to account for the difference in obligatoriness of subject between S and NP. However, it is interesting to see how we get a phrase like (18). Compare the sentences in (19) and the phras-

es in (20).

(19) a. The barbarians destroyed Rome.
b. Rome was destroyed by the barbarians.
c. Rome was destroyed.

(20) a. the barbarians' destruction of Rome
b. Rome's destruction by the barbarians
c. Rome's destruction

The (b) phrases are related to the (a) phrases since they also have Rome in object position at D-structure, this NP then being moved into subject position by move α. The by-phrase is base-generated as such. The (c) phrases are passives without the by-phrase: the agent is simply unspecified. This is exactly the property of the phrase in (18). Now if we compare (18) and (20c), it is clear that these phrases are closely related. They have the same underlying structure (21).

(21) $[_{NP}$ destruction $[_{NP}$ Rome $]]$

In (21), Rome has no Case assigned to it, so that something has to take place for Rome to fulfill the Case requirements. In (20c) move α has applied to Rome and it has moved it into the subject position where it gets Genitive Case. In (18), of-insertion provides the Case for Rome.

If we now look at the sentential equivalent in (19c), we see that Rome has been moved into the subject position in order to be assigned Case. The reason why Rome appears in the subject position in (19c) is not because a sentence has an obligatory subject position that must be filled; the NP is moved into subject position in order to get Case, just like it is in the NP

in (20c). And the NP must be moved in subject position in S because, whereas of-insertion can save Rome from the Case Filter (or its equivalent) in the NP construction, there is no such rule that applies after a past participle. The only position where the NP Rome can get Case is in the subject position: thus the apparent obligatoriness of the subject position in S. But what is in fact obligatory in this case is for the object NP to get Case in some way or other, and the movement to the subject position is the only one available in S.

Another argument presented in favor of (16) has to do with non-arguments: if there is no obligatory NP position in S, says Chomsky, why should there be place fillers like it? So although the obligatoriness of arguments in subject position can be derived from the Projection Principle and the θ-criterion, no such explanation can account for expletive elements.

It must be observed, however, that expletive it falls into two categories for Chomsky (1981a): a really semantically empty element as in (22), and a quasi-argument with a quasi-θ-role as in (23).

(22) It seems that John will come.

(23) It often snows here.

In the case of (23), we could in fact derive the obligatoriness of it in subject position from the θ-theory like other arguments if it has a quasi-θ-role. As for (22), Chomsky himself provides some arguments for its necessity elsewhere in Chomsky (1981a) when he says that the expletive element is necessary to cosuper-

script the $\bar{S}$ complement with a Case marked element so that it can be "visible" for θ-role assignment, since he assumes a visibility condition that applies at LF which is stated informally in (24).

(24) A Function Chain can be θ-marked in LF if it has Case or is headed by PRO.

(This visibility condition allows a derivation of the Case Filter on the LF side of the grammar. See Chapter 2 for some comments).

So in (22), the presence of it might be required to transmit Case to the post-verbal $\bar{S}$ with which it is cosuperscripted.

This explanation is not tenable for us, however, since in Chapter 2 we will reject the visibility condition in (24) in favor of another principle. How then are we to account for the obligatoriness of the subject in S without adopting (16), since we do not want to weaken the results that we have arrived at with respect to the properties of the categorial component? It would be unfortunate to have to drop such a strong hypothesis about the categorial component. As pointed out by Chomsky about methodology in linguistics, "it has often proven to be a wise move to persist in maintaining principles of UG that had significant explanatory power in some domain, even when they were faced with what appeared to be counter-evidence" (Chomsky 1981a, p.281).

Adopting (16) has damaging effects on the head constraint of $\bar{X}$ theory, and on the whole approach to the redundancy between lexical entries and phrase structure rules presented above. There is a possibility that we could find an elegant solution to

this problem by assuming an analysis of the tense of infinitives similar to the one presented in Stowell (1981a,b). (This solution is also implicit in Marantz (1981)). Stowell assumes that both tensed and infinitival sentences have the feature [+tense], but that they differ in that tensed clauses are also marked as [±past], as we can see in (25).

(25) a. tensed clause: (+N, -V, +tense, ± past)

b. to-infinitive: (+N, -V, +tense)

He proposes that tense be considered as a type of operator which ranges over the whole sentence. The fact that infinitives are neither present nor past, although they have the feature [+tense], "has the effect of specifying that the time-frame of the infinitival clause is unrealized with respect to the tense of the matrix in which it appears" (Stowell 1981b, p.2).[7]

Now suppose that the tense operator can only range over a full proposition: this would automatically account for the fact that the subject is obligatory since a VP is not a full proposition. We would be assuming, therefore, that a VP always forms a sentence with a subject because a predicate always forms a proposition with an external argument slot. If the VP does not assign a θ-role to the subject, as in (22), this would mean that the predicate that the verb organizes does not assign a semantic role: it is a defective predicate, but it still has an external argument slot, and therefore has the syntactic structure of a proposition. Assuming such a solution to be workable, then we would be provided with an extremely reduced categorial component,

most of the specifications that were formally put in this component now being derivable from independently motivated principles of grammar. The only specifications to be made in the categorial component would be those given in (15), repeated here as (26).[8]

(26) A. A phrase is labelled according to the category of its head.

B. The head is specified as final or initial.

C. Specifiers appear at the $\bar{\bar{X}}$ level, complements at the $\bar{X}$ level.

The expansion of S would now obey these minimal requirements of $\bar{X}$-theory since the subject would now be optional, just like any specifier, its apparent obligatoriness depending on other components of the grammar.

(27) S → (NP) INFL VP

Therefore, we will assume that the subject of a S is obligatory but that this obligatoriness can be derived from some other principle(s) of the grammar, which is the direction that the theory tells us to take. Exactly how the obligatoriness of the subject of S can be derived will not be relevant for our discussion, so we will leave this topic open for further research.

As for the obligatoriness of both INFL and VP, we could solve this problem by assuming that V is the head of the sentence, and that affixation of INFL to the verb takes place in the lexicon: there would be no INFL node within the syntax. There are arguments for this position in lexical phonology (see Pesetsky

1979, Kiparsky 1982, and others), so we will assume it to be workable.

Another consequence of the Projection Principle, combined with the θ-criterion, is that movement of an argument is only possible into a non-θ-position. Since all θ-positions have to be filled in D-structure, a moved NP will always originate in a θ-position (at least for its first movement) (assuming that expletive elements like it and idiom phrases have some kind of quasi-θ-role assigned to them as in Chomsky (1981a)). Therefore moving into a θ-position would create a Chain where an argument would be assigned two θ-roles, one from its D-structure position and one from its S-structure position (or any intermediate position). If one assumes that expletive elements and idiom phrases do not receive a θ-role in their D-structure position, then moving such an element into a θ-marked position would still be impossible since a θ-role would then be assigned in the position to which it moved to a non-argument.

The fact that lexical requirements have to be met at D-structure means that an NP could not be generated in a non-θ-position and then moved to a θ-position afterwards. So movement is strictly to non-θ-positions. A non-θ-position can be either the external argument position of a predicate that is not assigned a θ-role by this predicate (i.e. the subject position of the predicates formed by raising verbs like seem, or of passivized verbs like was seen), or a non-argument position in D-structure (COMP for example).

We have seen that a phrase displaced by moveα is coindexed with its trace. One way to implement this mechanism of indexation is to say that move α creates an index when it moves an element and that the index is given to all the elements in the Chain that is created. Another mechanism is proposed in Chomsky (1981b) and we will adopt it because it has interesting consequences with respect to parasitic gaps (see Chomsky 1981b), and inversion in Pro Drop languages (see Chapter 4). Chomsky's proposal is to assume that all NP positions are freely indexed at D-structure, but that this indexing procedure is restricted to A-positions, that is, positions where an argument can appear at D-structure (i.e. subject position, object position, etc., but not COMP or adjoined positions).[9]

This indexing procedure can be at least partially derived or motivated by considering a proposal that Stowell (1981a) makes about how θ-roles are assigned to complements. For example, consider a verbal complement: Stowell proposes that a θ-role be assigned to an argument by taking the index of this argument and inserting it in a θ-grid that would be part of a verb's entry. This could be extended to subcategorization frames and Case-features, since the unmarked case is for all of these three relations between V and complement to be dependent on government from the V. So if an index is required at D-structure for this identification of the complement in the verb's grid, it can be assumed that such indices will be provided to all positions where they are necessary, that is, A-positions. The case of phrases

base-generated in $\bar{A}$-positions (see fn. 9) and of the subject is a bit more tricky. Since the subject is assigned its θ-role compositionally by the VP (see Chomsky 1981a, Marantz 1981, for discussion), its index does not enter into a grid of any sort. We could say, however, that it is required for agreement purposes. We will come back to this issue in Chapter 4 when we look at the Pro Drop phenomenon.

Coming back to how the indexing procedure functions, if, for example, a phrase is moved to an $\bar{A}$-position like COMP as in (28), the index of the A-position is carried along by the moved phrase.

(28) $[_{\bar{S}}\ [_{COMP}\ Who_i]\ [_S\ did\ John_k\ see\ \underline{t}_i]]$

If the phrase is moved from an A-position to another A-position as in (29),

(29) $Malka_i$ was seen $\underline{t}_i$ by $Lori_k$

the landing position is one that already has an index since it is A-position and all A-positions get an index at D-structure. We could either say that move α changes the index of the landing position in such cases as in the derivation in (30), or that the index assigned to the landing position in D-structure prior to movement has to be the same as the one of the position of origin, this being allowed by the free indexing procedure as in (31).

(30) D.S. $[_{NP}\ e_i]$ was seen $Malka_j$ by $Lori_k$

S.S. $Malka_j$ was seen $\underline{t}_j$ by $Lori_k$

(31) D.S. $[_{NP}\ e_i]$ was seen Malka_i by Lori_k

S.S. Malka_i was seen $\underline{t}_i$ by Lori_k

The first procedure apparently runs into problems since a phrase could be moved into a position that is already in another Chain for example, since move α can change indices. But such derivations would be ruled out by the Projection Principle and the locality conditions on Chains since the Chain that is moved into is now broken up. One could also say that it is not the positions but the elements inserted into these positions that bear the indices, so that the subject position prior to movement in (30) would not have an index.

The second procedure is interesting because in this case it is assumed that the positions ultimately related by move α have to be coindexed from the start at D-structure: so with a slight reinterpretation of what a D-structure configuration must be, move α could be considered as an interpretive rule. As pointed out by Chomsky (1981a), it is quite difficult in fact to establish that there is a difference between move α as a syntactic rule of movement and as an interpretive rule: in fact, it is not clear that the question is of any real interest.[10]

We have seen, and we will find again later on, that the Projection Principle has far reaching consequences. It is crucially dependent on trace theory. It precludes any model of grammar, like the On Binding model for example, that incorporates "structure-building" rules, or any theory that has a rule of Raising-to-Object, if subcategorization entails θ-marking. Furthermore,

it forces one to adopt a structure for infinitives as in (32) rather than (33).

(32) Catherine$_i$ tried [$_S$ [$_{NP}$ e$_i$] to call Jacqueline]

(33) Catherine tried [$_{VP}$ to call Jacqueline]

It also has for consequence that Pro Drop languages must have an empty category in the subject position of S as in (34).

(34) [$_{NP}$ e] mangia le mele.

This means that the Projection Principle crucially depends on the existence of empty categories and their presence in the structures at the levels at which it is relevant. In fact, the central concern of GB is to determine the positions in which different manifestations of NPs can appear. The Projection Principle, in conjunction with other components like the base and the θ-criterion, determines where an NP must be present. This NP can be manifested as a lexical NP: a fully realized referential expression or a lexical anaphor or a lexical pronoun; or it can be manifested as an empty category: PRO, t (NP trace), variables (WH traces, traces of move α in LF), or pro (the "missing subject" in Pro Drop languages). In order to determine what manifestation of NP is possible in any given position where its presence is required by the Projection Principle, GB assumes several subsystems, as we have seen above, and each of these predicts a certain distribution of nominal elements in a certain domain. Thus, the lexicon, Case theory, and the binding theory interact with each other in several ways to determine the type of NP that

can appear in any given position. Case theory will determine whether the NP can be lexical or not. And the three subsystems will interact to determine whether the NP can be a name, a variable, an anaphor or a pronominal. Chomsky (1981a) has noted furthermore that since the different types of empty categories cover the possible partitioning of $[_{NP}\ e]$ without overlapping, this suggests that there is only one empty category as far as internal properties are concerned (like the fact that it has person, number and gender), and that the different types of empty categories are determined on functional grounds only (like the fact that the empty category is governed or not, Case marked or not, A- or $\bar{A}$-bound).

So we see that there is strong theory internal motivation to have ECs. Given ECs, one can formulate very general principles in the grammar that apply to lexical NPs and to ECs, so that the derivation of sentences with non-overt elements is greatly simplified and is ultimately virtually reduced to derivation of a sentence with overt elements. There is also empirical motivation to posit entities such as ECs. One such case has to do with reflexives and their relation to an antecedent. A minimal condition on this relation is that the reflexive and the antecedent must agree in features of person, number and gender. There are also structural restrictions on this relation: the antecedent must be structurally "close enough" to the reflexive in some sense, as we can see in (35) (we return to the exact formulation of this structural relation in 2.2.3.3).

(35) a. The girl said [the boy shaved himself]
b. *The boy said [the girl shaved himself]

But this structural relation seems to be reversed in (36): the NP which is structurally the farthest from the reflexive seems to be functioning as the antecedent in (36).

(36) a. Which boy did the girl say shaved himself?
b. *Which girl did the boy say shaved himself?

The solution to this problem is to posit an EC, the trace of the WH-phrase, in the subject position of the embedded clause. This is the now traditional argument of the long dependencies: the relation between the WH-phrase and the reflexive holds however distant the WH-phrase is from its original D-structure position.

Once the obligatoriness of the presence of an NP and its possible type have been determined, the interpretation of the sentence is quite straightforward if the NP is lexical. If the NP is an empty category, however, its content must be determined for interpretation to take place. So, some kind of recoverability process must operate. For example, the content of PRO will be determined by the theory of control: so the content of PRO can be the content of an antecedent that PRO is linked to by control, or PRO can be arbitrary in interpretation. For NP traces and variables, the Empty Category Principle (ECP) will function as a recoverability mechanism. The ECP is used extensively in GB.

(37) <u>The Empty Category Principle</u>[11]
$[_{\alpha}\ e]$ must be properly governed.

(38) Proper Government

$\underline{\alpha}$ properly governs $\underline{\beta}$ if and only if $\underline{\alpha}$ governs β, and

(i) $\underline{\alpha}$ is lexical ($=X^{O}$), or

(ii) $\underline{\alpha}$ is coindexed with $\underline{\beta}$

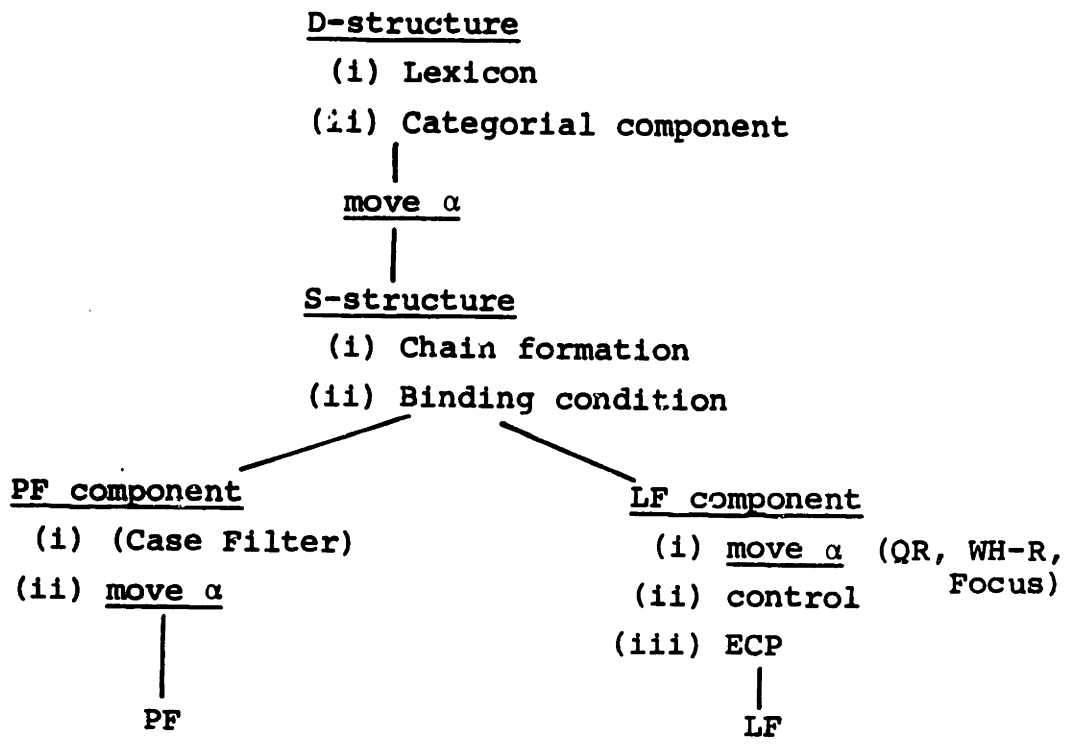
Kayne (1981a) has noted that the ECP is some sort of recoverability mechanism since, informally, it has the following content:

> An empty category must have an antecedent; the antecedent may itself govern the empty category; if not, the empty category must, through its governor, be "closely connected" to the antecedent. (Kayne 1981a, p. 103)

As for the fourth type of empty category, the subject of Pro Drop languages pro, its recoverability depends in some way on the richness of the inflection of the verb with which it agrees. This will be made more precise in Chapter 4, but for now we will simply refer to it as the Pro Drop condition.

To recapitulate what we have seen so far in this sketch of GB, the picture is one of a highly modular grammar, where components act at different levels. The description of the structure of core grammar that was given in (1) can now be made a bit more precise as in (39).

(39)



(There is no ordering implied by the listings at each level.)

We have put the Case Filter in parentheses since it is ultimately derived from the visibility condition in Chomsky (1981a), this condition applying at LF. Case is presumably assigned (or checked) at D-structure or S-structure. If Subjacency is a condition on rules, then it presumably holds at S-structure and at LF. On the other hand, if it is a condition on structures, then it might be sufficient for bounding theory to apply at LF only (it could then at least partly be derived from the ECP according to Kayne 1981a). To (39) must be added the conditions on well-formedness which are the Projection Principle and the θ-criterion.

We do not intend this summary to be a comprehensive introduction to GB. We only intend to provide an overview of the theoretical considerations in which is embedded the investigation in

the following chapters. As we have often pointed out, the various subsystems and notions will be studied in more detail as the discussion unfolds.

The point that we do want to emphasize is the central part that empty categories play in GB. The goal of this dissertation is to defend this use of empty categories and to strengthen the theoretical choice of such abstract elements by showing that no statement ever has to refer specifically to empty categories in the grammar: existence of such elements depends on the interaction of systems of the grammar which are independently needed for lexical categories. Thus this gives crucial support to a model of grammar where empty categories play a role.

There are three things that must be determined about an empty category, as we have seen: its presence, its type, and its content. In GB, the presence of an NP can be determined by the interaction of the Projection Principle and the θ-criterion; and the type of an NP can be determined by the interaction of the lexicon, Case theory and binding theory. All of these subsystems apply to categories, whether they are lexical or not, just like move α for example moves categories, whether they are lexical or not. But determining the content of NP in GB is not done in a unified fashion: lexical NPs have inherent content, while the content of empty categories is dependent on an antecedent and must be recovered by mechanisms that apply strictly to empty categories. Furthermore, there are three such mechanisms: ECP, control theory, and the Pro Drop condition. It is clear that by their very nature, empty categories cannot have their content de-

termined in the same fashion that the content of names is determined for example. However, this does not mean that their content has to be determined in a manner that is different from lexical NPs: it could be determined in a fashion similar to the way that the content of lexical anaphors or lexical pronouns is determined.

Recall the theoretical breakthrough that Conditions on Transformations brought about: it led to a three-fold unification of linguistic theory. First, Chomsky (1973) showed that movement transformations do not differ with respect to boundedness: so an apparently unbounded transformation like WH-movement was reinterpreted as an iteration of a more local, bounded rule. The second unifying proposal was to reduce all apparently unbounded transformations to the single rule of WH-movement (a further step was taken in Chomsky and Lasnik (1977) in reducing all movement rules to the single rule of move α). The third kind of unification was trace theory. Trace theory is a unifying idea because it brings the theory of movement rules closer to the theory of bound anaphora: the only movements affecting NPs that are allowed are those with an output where the antecedent-trace relation/ is a possible antecedent-anaphor relation. This considerably reduces the class of possible movements and it is of a low cost in the grammar since conditions on anaphora are independently needed for lexical anaphors.[12]

However, this unification of trace theory is greatly weakened if additional stipulations are added in the grammar that restate in some other form conditions on the relation between an

empty category and its antecedent. The goal of this thesis is to eliminate this weakness by showing that the conditions which govern the relation between a lexical anaphor or pronoun and its antecedent, thus determining the content of the lexical anaphor or pronoun, carry over to all empty categories so that no specific statement ever has to be made about empty categories.

In Chapter 2, we will propose a model of grammar where no statement ever refers specifically to the empty category: its distribution and type will be determined by the same subsystems that determine the distribution and type of all NPs, whether they are lexical or not; and the content of the empty category will be determined by the same subsystems which determine the dependency of content of lexical anaphors and pronouns. Since this model of grammar has only general statements about NPs and yet predicts the possibility of having an empty category, with its properties and distribution, it provides support to the claim that a theoretical entity such as an empty category exists.

In the subsequent chapters, we will show how the data accounted for by the three recoverability subsystems of GB can be accounted for in the present model and what further insights the analysis provides: Chapter 3 will deal with ECP, Chapter 4 with the Pro Drop phenomenon, and Chapter 5 with facts of control.

FOOTNOTES: CHAPTER 1

1. In (1), LF and PF denote levels of representation. In the literature, these terms are often used not only to refer to these levels, but also to the set of rules which map S-structure representations onto LF and PF. We will also use these terms in this fashion, hoping that the context will avoid confusions.

2. It is very likely that part of Subjacency is derivable from the ECP, as suggested by Kayne (1981a), since both of them are locality conditions on binding if Subjacency is interpreted as a condition on structures. The fact that "$\overline{S}$ deletion" allows to circumvent barriers of both government and Subjacency is also an indication that they are probably related. We will return to this topic in Chapter 3.

3. Note that with regard to (3c), we want to say that A governs its complement <u>of NP</u> (where <u>of</u> might presumably be inserted only in PF, see 3.2). This means that in (i), we still want the adjective to govern its complement.

(i) John is very proud of his story.

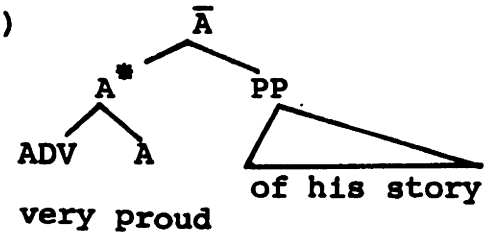
But in (i), the modifier <u>very</u> and the adjective <u>proud</u> seem to form a constituent of which the complement <u>of his story</u> is not a part, as can be seen by the fact that just this constituent can be extracted without the complement.

(ii) How proud is John <u>t</u> of his story?

What this suggests is that there is a small constituent as in (iii), call it A^{*}, which is similar to Williams' "small VP"

(see also fn. 5).

(iii)



This seems to call for a modification of the notion of government given in (4) since according to clause (i) of (4), proud does not govern its complement in (iii) because of the presence of very. But since there is only a small class of specifiers that can appear under A^*, or more generally X^*, we will consider that the core notion of government still holds in these cases, so that elements under X^* are closely related so that they form a unit that can govern complements for example. If we did not allow this for the core notion of government which is used for subcategorization, then we would have to consider that the second property of the core notion of government given above (i.e. clause (i) of (4)) must be modified so that it would now be as in (iv).

(iv) In the structure $[_{\gamma} \ldots \beta \ldots \alpha \ldots \beta \ldots]$, α governs β if and only if

1° $\alpha = X^{o}$

2° γ is a projection of α

3° β is an immediate constituent of γ

Note that the third basic property, that $\gamma = \bar{\alpha}$, would not hold under this new modification. This description of the notion of government runs into the problem that subjects would now be

governed in structures like (v).

(v) a. $[_{V}$max NP $[_{\overline{V}}$ V ...]]

b. $[_{NP}$ NP $[_{\overline{N}}$ N ...]]

But subjects are not part of the subcategorization frame of a V or N (see Marantz 1981, Burzio 1981, Chomsky 1981a for arguments). Thus (iv) cannot be a description of the properties of the core notion of government and something to the effect that X^{*} is a low-level unit that allows government by the X has to be assumed. Note that an extended notion of government with properties like 2° in (iv) is proposed by Aoun and Sportiche, so that this extended notion must incorporate a specification to the effect that only a narrower notion of government is relevant for subcategorization (see Chapter 2, fn. 24, and also the discussion on gerunds in 5.2.1.2.1.)

4. In Chomsky (1981a), the Case Filter is given as in (i).

(i) *NP, if NP is lexical and has no Case.

The reason for this change is that gerunds must be assigned Case like any other NP. But in Chomsky (1981a), it is assumed that gerunds have the structure in (ii).

(ii) $[_{NP}$ NP's VP]

Since gerunds do not have a N head in such an analysis, and yet still require Case marking, the Case Filter is given as in (i). Note that this is the only motivation for the change in the filter. However, we will see in Chapter 5 that there are reasons to believe that gerunds do have a nominal head and that the Case Filter can be stated as in (9) (although we will ultimately de-

rive the Case Filter from some other principle of the grammar).

5. As noted by Chomsky in class lectures at MIT in the winter of 1982, phrase structure rules are not really redundancy rules. A redundancy rule has the following form: if A has F, then it has F'. Phrase structure rules are not of this form: they do not allow us to extract anything out of the lexicon and simplify it. So phrase structure rules have the status of generalizations rather than that of redundancy rules.

6. See Stowell (1981a) for the details of this analysis. The condition of strict-adjacency on Case-assignment could be explained if we assumed, as in Bouchard (1979), that labelling of syntactic nodes is done by lexical insertion, and that structural relations are strictly binary. For example, two adjacent elements in a string can only be bracketed as in (i) or (ii).

(i) $[_{A}\ A\ B\ ]$

(ii) $[_{B}\ A\ B\ ]$

This has for consequence that three adjacent nodes could never be all three immediate daughters of a same higher node since three adjacent elements can only be bracketed as in (iii).

(iii)

a. $[_{\bar{\bar{A}}}\ [_{\bar{A}}\ A\ B]\ C]$	a'.	$[_{\bar{\bar{B}}}\ A\ [_{\bar{B}}\ B\ C]]$
b. $[_{\bar{\bar{B}}}\ [_{\bar{B}}\ A\ B]\ C]$	b'.	$[_{\bar{\bar{C}}}\ A\ [_{\bar{C}}\ B\ C]]$
c. $[_{\bar{C}}\ [_{\bar{A}}\ A\ B]\ C]$	c'.	$[_{\bar{A}}\ A\ [_{\bar{B}}\ B\ C]]$
d. $[_{\bar{C}}\ [_{\bar{B}}\ A\ B]\ C]$	d'.	$[_{\bar{A}}\ A\ [_{\bar{C}}\ B\ C]]$

In (iii), A and C could never be immediate daughters of the same node, so that A could never govern C according to the definition

of government given in (4) above; therefore, A could never assign Case to C. In fact, no lexical node in the position of A could govern more than one position. This can be related to Marantz's (1981) proposal that a semantic role assigner can only assign one semantic role. So for example, in (iv),

(iv) Elmer gave a porcupine to Hortense.

the verb gave assigns a semantic role to a porcupine, the preposition to assigns a semantic role to Hortense, and the predicate gave a porcupine to Hortense assigns a semantic role to Elmer.

This approach to node-labelling also has far reaching consequences on bar-levels and government which I will not pursue here. See Farmer (1980) and Nash (1980) for some discussion of related topics. This type of node-labelling is also akin to Williams' (1980) notion of small verb phrase for the English double NP construction.

7. Stowell notes that this interpretation is similar to Bresnan's (1972) observation that an infinitival complement describes "something hypothetical or unrealized".

8. We have seen that (26C) can be derived if the notion of government given in (4) is used for subcategorization. As for (26A), it amounts to saying that lexical insertion is responsible for node-labelling. If we adopted the labelling procedure briefly described in footnote 3, therefore, (26A) could also be accounted for. So (26B) might be the only specification to be made in the categorial component. And even then, a part of the possibilities for the combinations of Subject, Verb and Object

for example, would be highly marked in a node-labelling of the type presented in footnote 3, as we can see in (i).

(i) S [V O] *V S O
S [O V] [O V] S
[V O] S *O S V

This seems to fit with the studies that have been done on this kind of language typology (cf. Greenberg (1963)). However, we will see in Chapter 4 that an additional stipulation will have to be made to account for the difference between so called non-configurational languages and configurational languages.

9. Chomsky's proposal is in fact to index A-positions at S-structure. The proposal made here to index these positions at D-structure keeps the important distinction that Chomsky wants to make between indexation of A- and $\bar{A}$-positions, and it allows an elegant account of indexing by move α as we will see directly. The distinction that Chomsky wants to make is between A-position indexing and indexing of a phrase base-generated in COMP and/as in (i).

(i) ?Who$_i$ do you think that Mary would prefer that he$_i$ stay home?

Such instances of the resumptive pronoun strategy, where a phrase in COMP binds a pronoun in the sentence, indicate that there must be an indexing procedure for $\bar{A}$-positions too. However, $\bar{A}$-position indices are only relevant at LF: so we will say that these indices are only assigned or visible at LF, to differentiate them from A-position indices. So, depending on whether they have a resumptive pronoun strategy or not, languages will differ as to whether they allow such indexing of an $\bar{A}$-position at LF (see

Chomsky 1981b).

10. However, the relation between an $\bar{A}$-position like COMP and an A-position in the sentence seems to favor a syntactic movement analysis of move α since, in the treatment of parasitic gaps for example, the difference in indexing between a phrase moved to COMP and one base-generated in COMP is crucial: the former is indexed at D-structure, while the latter's index is only relevant at LF, and it is only in the case that the phrase is moved to COMP that it can license a parasitic gap.

11. We give here one version of the ECP. There are several different versions that have been proposed in the literature. This version will suffice for our purposes since we will see shortly that such stipulations as the ECP must be eliminated from the grammar on principled conceptual grounds.

12. Note also that the fact that trace theory allowed all interpretation to be done at surface structure was a precursor to the Projection Principle.

CHAPTER 2: DETERMINING THE PROPERTIES OF THE EMPTY CATEGORY

2.1 The issue.

In this chapter, we will propose some modifications of the GB framework in order to satisfy the conditions that must be met so that the null hypothesis about empty categories can be made: the distribution, type and content of $[_{NP}\underline{e}]$ must be fully determined by conditions and principles that apply to the category NP, without discriminating as far as lexical content is concerned. This, we hope, will strengthen the claim that a theoretical entity like $[_{NP}\ \underline{e}]$ exists.

The study of empty categories is particularly fascinating. The properties of empty categories can hardly be determined by observation of overt data; therefore, it is reasonable to assume that they reflect deeper principles of UG, so that empty categories are "windows" into the nature of the human language faculty. This might be the reason why the study of empty categories has proven to be such an important probe for determining properties of syntactic and semantic representations.

Once one postulates the existence of empty categories with statements that refer to them like the ECP, control theory or the Pro Drop condition, this does not resolve the problem of empty categories, it poses the problem: one must then explain why the phenomenon exists with the properties it has. Chomsky (1981b) takes up this question with respect to parasitic gaps: once one has a statement about the conditions under which a parasitic gap may appear in a sentence, the problem is to explain why the phe-

nomenon of parasitic gaps exists with certain given properties. Chomsky's answer is the following:

> It is highly likely that the answer is that the phenomenon exists with the properties it has because of other properties of UG and the particular grammar that are quite independent of parasitic gaps; it is, again, most unlikely that particular grammars have rules governing parasitic gaps or that UG has some specific principles bearing on this phenomenon, which is a particularly interesting one for just this reason. While this seems a very plausible assumption, it amounts to quite a strong claim, as noted earlier. It means, for example, that all of the quite intricate properties of parasitic gaps must be reducible to general principles of UG, given rules and structures of the particular grammar that are established on other grounds; and it also means that if languages appear to differ with respect to the existence or properties of parasitic gaps, these differences must be completely explained on the basis of other structural differences among the languages in question. The task for the linguist, then, is to show how independent properties of a particular grammar interact to yield the distribution and interpretation of parasitic gaps. (Chomsky, 1981b, p. 43g-44)

Our intention in this thesis is to apply this approach to <u>all</u> gaps. This is an important test for the theory. As we have seen, the validity of the Projection Principle crucially depends on whether the other interacting components of the grammar can provide an independently motivated account of the properties and distribution of different manifestations of the empty category. If these components do not provide such an account in any reasonable fashion, then we could doubt the validity of the Projection Principle and of the whole approach of using empty categories (the two being closely linked in an optimal theory), since this account would have to be supplemented by stipulations referring specifically to empty categories. But on the other hand, if these components do provide such an account, then the Projection

Principle receives significant support. We have seen that, in GB, at least three stipulations have to be made strictly about empty categories, so that they are not independently motivated: the ECP, control theory, and the Pro Drop condition. The task therefore is to eliminate these stipulations from the grammar. This could be done because one sees some redundancy between these stipulationsand other components of the grammar: it would therefore be good methodology to try to derive them from these other components. But the reason why we want to eliminate them is more principled: they must be eliminated because they weaken the whole approach given the null hypothesis about ECs. Furthermore, by removing these stipulations from the grammar and deriving the properties of empty categories from principles and conditions that apply to NPs in general, one reduces the class of possible grammars since no special kind of (abstract) element is allowed to be introduced in the grammar anymore, but the existence of such an element must be independently motivated. (Recall the advantage of trace theory which reduces the possible transformations to those which have for output configurations that are like those of independently motivated antecedent-anaphor relations).

If the properties of gaps are determined by components that operate on NPs in general, then it is most likely that gaps should share properties with lexical NPs. Off-hand, this seems to be the case. So, for example, a gap can bear a θ-role; it seems to function like an anaphor or a pronominal in some cases; and it enters into agreement with elements that agree with lexi-

cal NPs, like lexical anaphors as in (1), attributive adjectives in (2) and tensed verbs in (3).

(1) a. Anne seems [t to have hurt herself/*himself]
b. Lisa tried [PRO to hide herself/*himself behind the door]

(2) a. Sébastien semble [t être très content/*-te de son coup]
b. Pierrot a encore essayé [PRO de ne pas sembler surpris/*-se]

(3) Which men did you say [t [t were/*was at the door]]

These shared properties have to do with what the type of an NP is (pronominal, anaphor, variable) and what its content is (for example, agreement features). The need for an NP in a given position we will consider to be determined by the Projection Principle and the θ-criterion without further discussion. Let us now consider how the content and the type of an empty category can be determined from the stipulations independently motivated to account for properties of lexical NPs.

2.2 What is an empty category?

2.2.1 The content of an empty category.

A strong hypothesis to hold with respect to the content of empty categories is that these internal properties of gaps are a subset of the set of properties of lexical NPs, and that this subset contains the minimal properties required for an NP to be an argument: the properties necessary and sufficient for an NP to be an argument would be the properties of a NP gap.

In order to determine what these minimal properties are, it is interesting to look at what the three GB stipulations on reco-

verability of the content of empty categories have in common. First, consider control theory. Control theory establishes a link between PRO and an antecedent, or it assigns (or checks possibly) the referential index arbitrary to PRO. So it seems that PRO must have a referential index, either by being indentified as coreferent with an antededent, or by getting the specific index arbitrary. On the other hand, the Pro Drop condition requires that the verbal inflection be "rich enough", that is, that it possess certain grammatical features like person, number, and gender that can be assigned to the pro subject, pro and AGR forming a sort of discontinuous element.

As for ECP, it requires that the empty category be either coindexed with an antecedent or governed by a lexical element. If one adopts Stowell's (1981a) idea of a verbal grid in which the index of the complement is inserted for θ-role assignment and if, as is quite likely, one extends this to all lexical heads that assign θ-roles (and/or assign Case and have subcategorization frames as we have seen in Chapter 1), then the disparity in the notion of proper government is eliminated and what emerges is that the empty category must be governed by an element with which it is coindexed. We return to this topic in 2.2.4.3. So once again, just like in control situations, an index seems to be crucial for an empty category. Note that this index is also needed for pro since we will want to say that it enters into coreference or disjoint reference relations with other NPs in a sentence. In this case, we could either say that pro

has an index of its own, or that it gets one from the AGR on the verb.[1]

So a referential index (henceforth R-index) seems to be an important feature of the internal properties of an empty category. As for the grammatical features of person, number, and gender (henceforth F-features), they also seem to be relevant for all four manifestations of the empty category in GB. In Chomsky (1981a), PRO has these features intrinsically, hence its pronominal status. According to Chomsky, NP trace and variable also have these features since they also enter into agreement with other elements: these features are said to be "left behind" by move α. And we have seen the dependency of pro on AGR for these features. What we see, therefore, is that a R-index and F-features seem to be part of the crucial content of a gap. This seems to be intuitively right: determining the reference of the empty category is certainly required if it is to be an argument, and the agreement facts suggest that F-features are also part of the content of an empty category.

We noted at the beginning of this section that we were looking for the minimal properties required for an NP to be an argument. We expect therefore that these properties will be relevant at LF, where the properties of arguments are ultimately necessary for proper interpretation to take place. It is interesting to note with respect to this that, in GB, the stipulations for recoverability of the content of three out of four of the manifestations of the empty category are stated to apply at LF.

Thus, the ECP and control theory apply at LF. It is not clear where the Pro Drop condition applies, however. In Chomsky (1981b), where the manifestation of the empty category as pro is introduced, the indexing relation between AGR and pro is established at D-structure; but this is not the same kind of indexing as the one that takes place between the other three empty categories and their antecedents. It is a co-superscripting, and co-superscripting is not relevant for the binding theory, for example. It seems that the minimal properties which are necessary to be an argument could be relevant only at LF, so it might turn out that these properties are recovered only at LF. But this raises the question of the status of the F-features as part of these minimal properties for argumenthood. Could it be that F-features are relevant only at LF, or is it simply that we were wrong in assuming that F-features are part of these minimal properties, this being only an apparent phenomenon? If the former is true, then this raises questions about how and where the theory of agreement applies. If the latter is true, then one might want to explain this accidental relation of F-features and argumenthood. In order to answer these questions about R-indices and F-features, we will explore what exactly would be the use of these elements in LF.

2.2.2 F-features and Domain D

Consider the domain D of individuals associated with arguments at LF as values of variables, denotata of names, etc., which is presented in Chomsky (1981a). This is the domain that

contains "mental objects" to which language can refer. Thus domain D contains mental objects which are the mental representations of real world objects like chairs and linguists, and birthday cards. But domain D also contains objects which have no real world equivalents, like flaws in arguments, unicorns and the like. Domain D stands at the interface between real world and "linguistic world". This domain D even contains objects which are not really mental objects, since they are not possibly conceivable, like square circles, for example.

When we use the term 'referential expression', therefore, we mean to refer to objects in domain D, not in the real world, since we clearly want flaws and unicorns and square circles to be referential expressions as far as linguistic rules and principles are concerned. This is because NPs like these behave like truly referential expressions with respect to grammatical processes like establishing coreference with a pronoun or taking a θ-role. So, for example, we want to account for the grammaticality judgements in (4) in the same way that we account for those in (5) (where underlined NPs are meant to be read as coreferential).

(4) a. John believes that a dog will chase its shadow.
b. *John believes that it will chase a dog's shadow.

(5) a. John believes that a unicorn will chase its shadow.
b. *John believes that it will chase a unicorn's shadow.

Note that this step in the process of interpretation is not to be confused with what might be called "real semantics", that is, the study of the relation between language or

> language use and the world. Rather, it should be considered to be in effect an extension of syntax, the construction of another level of mental representation beyond LF, a level at which arguments of LF are paired with entities of mental representation, this further level then entering into "real semantic interpretation".
>
> (Chomsky 1981a, p. 324)

Assuming, therefore, that there is such a domain D, what can be said about the individuals in it? First, we can say that the individuals in domain D can be divided into subsets of individuals that can or cannot have a property P predicated of them. For example, each of the following predicates subdivides domain D into two sets of individuals: the individuals that can fill the subject position of the predicate, and those that cannot.

(6) a. x saw Mary.
b. x saw that Mary was a smart girl.
c. x is turning green.

But this tells us little that can be useful in grammar about the individuals in domain D. A more interesting approach would be to see if the individuals in domain D have features that are linguistically relevant. For example, these individuals seem to be marked for selectional features: the examples in (7) and (8) show that these individuals can be grouped into subsets of domain D depending on whether they have the feature [±human].

(7) a. Who hit Bill?
b. For which x, x a person, x hit Bill

A: a. John hit Bill.
b. *The rock hit Bill.

(8) a. What is on the table:
b. For which x, x an object, x is on the table

A: a. The book is on the table.
b.??Bill is on the table.

However, since selectional features can be attributed to individuals in the real world, the above facts tell us little about the existence of a domain D which would be an interface between the real world and the linguistic world. The interesting question, therefore, is whether strictly grammatical features can define subsets of individuals in domain D. The answer is yes. Consider (9) and (10).

(9) a. Laquelle préfères-tu?
b. Which x, such that x is feminine in French, you prefer x

(10) a. Which ones do you prefer?
b. Which x, x a set of individuals, you prefer x
c. Which x, such that x is plural in English, you prefer x

In (9), the individuals in domain D are divided into two subsets, depending on their gender in French: the answer to the question can be anything that is feminine in French. Similarly, in (10), under reading (10c), the individuals in domain D are divided into two subsets, depending on their number: the answer to the question can be anything that has the intrinsic feature plural in English (e.g. trousers, scissors, etc.). The reading in (10b) is the case where a plural in the real world is a proper answer, i.e. more than one individual.

The same observation can be made about pronouns as in (11) and (12).

(11) a. Cette table est belle et elle irait bien dans notre salon.

b. These scissors are sharp because they are new.

(12) a. Elle est belle (la table, la lune,la fille d'à côté...)

b. They are new (the scissors, the trousers, ...)

In (11), we have instances of pronouns which are coreferential with NPs expressed in the sentence. In (12), the pronouns are deictic, that is, pronouns that are considered to be pragmatically controlled. What we see is that even the deictic pronouns are sensitive to some aspects of linguistic form: deictic pronouns can only refer to mental objects which have matching F-features, even if these F-features are not recoverable from the syntactic context.

This observation is not new. Hust and Brame (1976) for example, showed that whether a pronoun is deictic or not, it must agree, and it must do it in a consistent way. By this, they mean that in the case of a noun like ship, for example, both it and she can be used to refer to the same entity that ship refers to. Yet, in a sentence like (13), it is not natural for it and she to refer to the same entity.

(13) She is not as fast as it once was.

Similarly, in German, where we may refer to an automobile as das Auto and der Wagen, the same judgement holds as to the coreferentiality of es and er in (14).

(14) Es war schneller als er früher war.

This type of deviance is due to the fact that coreference is de-

termined on linguistic objects in domain D; which have grammatical features, and that coreferential NPs, i.e. NPs with the same R-index, must agree in F-features, as it will be made explicit shortly.[2]

What this discussion on gender and number suggests is that individuals in domain D have grammatical features, so that domain D is really an interface between real world and linguistic world. Because of this, we can assume that there exists a general principle which governs the pairing of NPs with entities of mental representation, i.e. the individuals in domain D, and that this principle crucially makes stipulations about F-features. Furthermore, we must also assume that some device is used to indicate reference: so we will assume that NPs bear referential indices (R-indices) to account for facts of coreference and disjoint reference. We will call this principle which governs the pairing of NPs with individuals in domain D the principle of Denotability. We will assume that the principle of Denotability operates on the mapping of NPs in LF onto individuals in domain D. It can be stated as in (15).

(15) Principle of Denotability I

An NP will denote an object in domain D iff that NP has an R-index and that NP has F-features (= person, number, gender).[3]

This principle can have quite far reaching eonsequences on the analysis of empty categories since it requires that an empty category that is in an argument position bear both an R-index and F-features at the level of LF. This then will be the content of

the empty category that must be recovered at LF.

2.2.3. The principle of Denotability and lexical NPs.

Before looking at the effects of the principle of Denotability on empty categories, let us look at how it deals with lexical NPs, since we want to respect the general conceptual approach presented above and make sure that all the principles and rules that we propose apply to the category NP in general.

2.2.3.1. Names.

First consider the case of a name. This is the most straightforward case since we will make the reasonable assumption that the lexical entry of a name contains F-features as part of the abstract morphophonological structure of this item, and that names are assigned an R-index when they are inserted in D-structure.[4]

2.2.3.2. Pronouns.

Lexical pronouns will be like names as far as their F-features are concerned. However, the R-index of a pronoun is different from that of a name in that the R-index of a name is fixed in some sense, whereas a pronoun can freely pick any index at all, subject to coreference (or disjoint reference) conditioning which operates independently from this free choice of R-index.[5] Apart from the coreference conditioning, to which we will return in 2.3., there is also an agreement conditioning on pronouns, whether they are pragmatically or syntactically controlled as we have seen in (9)-(14) above. This can be stated as in (16).

(16) Agreement
$\underline{\alpha}$ assigns (redundantly) its F-features to $\underline{\beta}$ if $\underline{\alpha}$ and $\underline{\beta}$ have the same R-index.

The introduction of redundantly in parentheses allows the rule to be a feature checking rule or a feature assigning rule, depending on whether redundantly is taken into consideration or not, respectively. This will be made clearer as we go along.

2.2.3.3. Anaphors.

The last instance of a lexical NP to consider is that of a lexical anaphor, i.e. a reflexive or a reciprocal. Lexical anahors are like names and pronouns in that they have F-features specified in their lexical entry as part of their abstact morphophonological structure, and they are like pronouns in that there is an agreement requirement on their F-features. We have seen in (16) that agreement takes place when an item's R-index is the same as the R-index of another item. The fact that lexical anaphors must agree with some element is expected since, by definition, anaphors are referentially dependent elements. We have seen that pronouns are also referentially dependent in some sense since, although they can freely pick any index (subject to coreference and agreement restrictions), they do not have a specific R-index of their own like names do. But there is a difference between pronouns and anaphors in that the dependency of the anaphor is much stronger and also more local. The dependency is much stronger in that the anaphor must have a syntactic antecedent from which it can pick its reference, whereas the pronoun's reference can possibly be only pragmatically controlled. This can be seen in the contrast between (17) and (18).

(17) a. $John_i$ hurt $himself_i$.
b. *Himself came for dinner.

(18) a. $John_i$ saw that he_i was going to be late.
b. He came for dinner.

To express this difference between lexical anaphors and pronouns, we will postulate that lexical anaphors do not have an R-index in their lexical entry, so that an R-index must be assigned to the lexical anaphor by its antecedent in order to satisfy the principle of Denotability. This lack of an R-index would capture the intuitive idea that anaphors are referentially dependent. The fact that this dependency is very strong, in that the reference of the anaphor must be syntactically controlled, suggests that the assignment of an R-index to an anaphor is determined by a rule, hence is grammatically determined. This rule, or possibly a component interacting with it, will also express the locality requirements on the antecedent-anaphor relation that are violated in (19).

(19) *Martin said that himself would come alone.

The fact that the R-index of an anaphor is assigned by a rule would also account for the fact that the antecedent of an anaphor is unique, whereas a pronoun can be coreferential with more than one NP.

(19') a. $John_i$ told $Mary_j$ that $they_{i+j}$ should leave early.
b. *$John_i$ believes $Mary_j$ to like only $themselves_{i+j}$.

Let us call the linking between antecedent and anaphor gram-

matical binding. We turn now to the task of determining what this grammatical binding is, and where it takes place in the grammar.

The first thing to determine is when we are indeed dealing with such an anaphoric relation. It is often assumed that this can be easily done since anaphors have morphological features that identify them as such. Thus, for reflexives for example, English has the suffix -self attached to a pronoun to indicate that this is an instance of an anaphor. Similarly, Dutch has the suffix -zelf, and French has the suffix -même. Furthermore, in Romance languages, some clitic forms can be inherently reflexive, like French se or Italian si. It is also generally assumed that pronouns and anaphors are in complementary distribution, in the sense that pronouns enter obligatorily into disjoint reference where anaphors may appear. Thus, in (20b), John and him are obligatorily disjoint in reference.

(20) a. John_i shaved himself_i.
b. John shaved him.

But this complementary distribution is blatantly violated if the distinction is a purely morphological one. This can be seen in (21)-(22), where both pronominal and anaphoric forms are acceptable, and in (23)-(24), where pronouns and anaphors seem to appear in the same syntactic context, although in this case pronominal and anaphoric forms are not interchangeable with one another.

(21) a. John heard stories about him/himself
b. John says that a picture of him/himself was hanging on the wall.
c. They read each other's/their books.

(22) Victor est content de lui/lui-même.

(23) a. John always keeps his wits about him/*himself.
b. John is always talking about himself/*him.

(24) a. Victor a toute l'équipe avec lui/*lui-même.
b. Victor bavarde avec lui-même/*lui.

One early attempt to account for these facts was presented in Chomsky (1965). There, Chomsky proposed to account for the contrast between the sentences in (25a) and (25b) by assigning them different structures, as in (26a) and (26b), respectively.

(25) a. I kept it near me.
b. I aimed it at myself.

(26) a. I kept it $[_S$ it is near me$]$
b. I aimed it $[_{PP}$ at myself$]$

Chomsky (1965) assumes a rule of reflexization that applies only sentence internally, so that I and me are not both immediately dominated by the same S node, and reflexivization does not apply in (26a). But even such a powerful grammar could not account for sentences like those in (27), which Chomsky (1965) points out as problematic.

(27) a. I pushed it away from me.
b. I drew it toward me.

Here, deep structures with S complements are difficult to motivate in Chomsky's (1965) framework, and yet reflexivization is

is impossible.

The attempt in Chomsky (1980) to account for these facts is very similar to the one in Chomsky (1965). Thus a sentence like (28a) is given the underlying structure (28b), with to be deletion applying subsequently.

(28) a. John considers Mary angry at him.
b. John considers $[_S$ Mary to be angry at him]

So him in (28) is in an opaque domain, the domain of the subject Mary, and so it can have John as its antecedent since John is not in this domain. But this runs into the problem of the non-synonymy of pairs like (29) and (30), which was pointed out in Chomsky (1970).

(29) John considers Mary silly.

(30) John considers Mary to be silly.

In Chomsky (1981a), a new attempt is made at solving this problem. First Chomsky postulates the binding conditions in (1-(6)), repeated here as (31), to account for different possibilities of coreference.

(31) A. An anaphor is A-bound in its governing category.
B. A pronominal is A-free in its governing category.
C. An R-expression is A-free everywhere.

Governing category is defined in (32).

(32) β is a governing category for α iff β is the minimal category containing α, a governor of α, and a SUBJECT accessible to α.

So, in a sentence like (25b), an anaphoric relation can be established between I and myself because, although myself is governed by at, the PP is not its governing category because the PP at myself has no accessible SUBJECT. But then to account for sentences like the one in (33a), one is forced to postulate that PPs can have PRO subjects, as suggested by Manzini (1980), and that the D-structure of the PP is similar to the one of an AP like in (34).

(33) a. John saw a snake near him.
b. John saw a snake [PRO near him].

(34) a. John considers Mary [PRO angry at him]
b. John left the room [PRO angry]

So the PP in (33) is a governing category, and him is therefore free in its governing category in accord with the binding condition B. This solution is very similar to the one presented in Chomsky (1965). And Chomsky has the same qualms about this solution that he had in Chomsky (1965): The idea seems quite inappropriate in other cases. For example, compare the sentences in (35).

(35) a. John turned the child against him/*himself.
b. John turned the argument against *him/himself.

In order to account for this difference by a structural difference between the two sentences, one would have to claim that turn takes a clausal complement in (35a) but not in (35b). This is not very appealing, and furthermore, there is no independent LF form John-turn-clause.

Chomsky (1981a) also points out that Manzini (1980) shows on the basis comparative evidence that the plausibility of this structural solution is illusory. Manzini compares the Italian analogues of (34a) and (34b) in (36a) and (36b) respectively.

(36) a. Gianni considera Maria arrabbiata con lui/*se/*se stesso.
b. Gianni vide un serpente vicino a lui/se/*se stesso.

Italian has two reflexive elements, se and se stesso, which correspond roughly to the English X-self. Manzini points out that se stesso can appear only in a subclass of the environments where se can appear, and that the additional restrictions on the distribution of se stesso do not relate to the theory of binding. She suggests that the English X-self is more closely related to se stesso than to se, and that the same irrelevant reasons which prohibit the occurrence of se stesso in (36b) are responsible for the impossibility of X-self in (33).

Chomsky's conclusion from this set of data is that a unified treatment of (33) and (34) is not warranted and that the apparent similarities between these sentences "may be an artifact based on peculiarities of English reflexivization. It is highly doubtful, then, that such facts as (33) should be used as the basis for any structural argument" (Chomsky, 1981a, p.291).

This conclusion seems justified since constructions of the type $[_{pp}$ P NP] with respect to reflexivization (considered not as a rule but as a phenomenon here) show all the typical properties of marked constructions, and therefore should not be included in the core analysis of reflexivization. Thus, judgements of acceptability about these constructions tend to waver, as it

has often been noted. For example, Helke (1970) noted that there is variability from speaker to speaker in these constructions, and furthermore, that the same speaker will sometimes have a different treatment of the same structure depending on the lexical items.

> In particular cases, some speakers allow either a pronoun or a reflexive, others allow only a pronoun, and yet others only a reflexive. It is doubtful, however, that those for whom a particular pair of examples ... is questionable do not have other examples in the dialects they speak in which a pronoun and a reflexive may alternate. To account for the diversity of dialects on this point, one might choose ad hoc markings in the lexical entry for each of these expressions indicating whether an anaphoric pronoun or a reflexive would be considered acceptable. Such ad hoc markings are quite appropriate in this case, since they reflect the absence of a systematic variation from one dialect to another. (Helke 1970, p.4)

There is also variation across languages that does not seem to be explainable on the basis of structural arguments. We have seen differences between English and Italian above; other examples of such contrasts can be found, for example, between English and Dutch, as in (37), and between English and French, as in (38).

(37) (from Koster 1978b)

a. Tom believed that the paper had been written by Ann and himself.

b. *Tom geloofde dat de paper was geschreven door Ann en zichzelf.

(38) a. Victor est content de lui/lui-même.

b. Victor is proud of *him/himself.

Another standard feature of marked constructions is the susceptibility to lexical and non-grammatical complexity factors.

An illustration of this for the construction at hand is given in Zribi-Hertz (1981). She shows that there seems to be a significant generalization to make about when a pronoun or a reflexive will be used in a $[_{pp}$ P NP] construction in French. She shows that there seems to be a correlation between the degree of possibility of coreference that the predicate formed by the V allows between the subject and a $[_{pp}$ P NP] complement, and the choice of pronoun or reflexive. We summarize her proposal in (39), with respective examples in (40).

(39) a. obligatory coreference: lui *lui-même
b. possible coreference: lui lui-même
c. improbable coreference: *lui lui-même

(40) a. Victor a toute l'équipe avec lui/*lui-même.
b. Victor est content de lui/lui-même.
c. Victor bavarde avec *lui/lui-même.

The generalization seems to be that the more the coreference is forced by the predicate, the less there is the need of an obligatory bound reflexive form, since a pronoun will have to be coreferent anyhow for pragmatic reasons. And conversely, the less coreference is probable, the more the reflexive form is needed, being forced by pragmatic reasons to get this improbable reading.

Note that these, or similar, pragmatic tendencies seem to hold in English also: (21a) would be an instance of possible coreference, hence the optionality of choice between a pronominal and an anaphoric form; (23a) would be an instance of obligatory

coreference, hence the obligatoriness of the pronominal form; and (23b) would be an instance of improbable coreference, hence the obligatoriness of the anaphoric form.

We would normally expect pragmatic factors to be identical from one language to another. The slight differences between French and English in the distribution of reflexives and pronouns could be due to additional factors of the following kind: 1° syntactic factors (reanalysis, etc.); 2° lexical factors: the fact that lexical items are rarely fully equivalent cognates; possibility of other unknown factors interacting.

Another example of this type of markedness factor is presented in Koster (1978a). Koster compares the Dutch sentences in (41) and in (42).

(41) Peter$_i$ zag Mary naar zich$_i$ toe komen
Peter saw Mary to himself (part) come
(Peter saw Mary come towards him.)

(42) *Peter$_i$ zag Mary zich$_i$ wassen
Peter saw Mary himself wash
(Peter saw Mary wash himself.)

Koster proposes to account for this contrast by adding condition (iii) to his Locality Principle given in (43).

(43) Locality Principle (Koster 1978a)[6]

No rule involves $\underline{\alpha}_{i+1}$, $\underline{\gamma}$ (where $\underline{\alpha}$ c-commands $\underline{\gamma}$) in

$\cdots \underline{\alpha}_{i+1} \cdots \underline{\alpha}_i \cdots \underline{\gamma} \cdots \underline{\alpha}_i \cdots \underline{\alpha}_{i+1} \cdots$ $(i \geq 1)$

unless (i) $\underline{\alpha}_{i+1}$ (or $\underline{\gamma}$) is more prominent than $\underline{\alpha}_i$

(ii) $\underline{\alpha}_i$ is coindexed with a phrase properly containing $\underline{\gamma}$.

(iii) $\underline{\alpha}_i$ and a reflexive $\underline{\gamma}$ are (linked to) co-arguments of an irreflexive predicate $\underline{\psi}$ that

does not include a transitive verb.

So again we see that lexical and pragmatic factors have some influence on reflexivization in $[_{PP}$ P NP] constructions. Note also, by the way, that one could not simplify the complexity of the factors involved in cross-language differences by saying that the distribution of the Dutch reflexive is more restricted than that of the English reflexive because of the contrast in (37) where a reflexive form is possible in English but not in Dutch, since just the opposite is true in (41): the English equivalent to (41) with a reflexive is ungrammatical.

So we see that constructions of the type $[_{PP}$ P NP], with respect to reflexivization, exhibit all the standard features of marked constructions: there is variation in judgements of acceptability, there are differences across languages, and there is susceptibility to lexical and non-grammatical complexity factors. This is implicitly recognized by Chomsky (1981a) and explicitly by Koster (1978a, b), and others.[7]

There are two possible attitudes in the face of these data depending on how one construes opacity: the first one is to assume that the pronominal forms are marked here since they occur in the domains of reflexive constructions; the second one is to assume that the reflexive forms are marked here since they occur in the domains of pronominal constructions. Chomsky (1981a) assumes the former: for him, PP constructions like those discussed here are constructions where reflexives should generally be possible, but some additional conditions, which do not relate to

the binding theory, block the possibility of having these reflexives and this is why pronouns appear in positions where anaphors are expected. However, a closer look at the facts suggests that the second attitude is to be preferred, as we will show immediately.

We will assume that reflexives are not to be defined on a strict morphological basis, but rather on the basis of a specific type of structural relation with an antecedent in addition to the morphological basis. The domain in which this type of relation can be established will be very strictly limited as we will see immediately, and elements which will have the morphological structure of reflexives but which will not fall under these strict limitations in their relation to an antecedent will be considered as deviant forms which are in fact pronouns. So elements with the morphological form of a reflexive will not be automatically determined to be typologically classified as anaphors: the type anaphor will be functionally determined by specific properties. As we will see, there are four of these properties: 1° obligatoriness of the antecedent; 2° a one-to-one relation between antecedent and anaphor; 3° locality of the relation; 4° a structural condition on the relation. This way of dealing with anaphors on a functional basis rather than a morphological basis is similar to Chomsky's (1981a) approach to ECs. ECs are all of a similar morphological form, and this can presumably give a general basis for their distribution: so for example, some ECs can appear where no Case is assigned, but this is not possible for

lexical NPs. But one wants to say more about ECs: it seems that different types of ECs are at hand, i.e. anaphors, pronouns and that it is the distribution by type which is interesting. This distribution by type cannot be determined morphologically for the EC: it must be functionally determined. Our claim is that the same is true of reflexive forms: although they form a coherent morphological class, they do not form a coherent typological class, i.e. they are not all anaphors. Let us try to determine what is this restricted domain in which an anaphoric relation can be established between two noun phrases.

One structural relation which must always be an antecedent-anaphor relation is the one between subject and object, if they are to be coreferential.

(44) a. John shaved himself.
b. *John shaved him.

(45) a. John's hitting himself on the foot was not accidental.
b. *John's hitting him on the foot was not accidental.

This appears to be true cross-linguistically, and we will consider it to be the core case of an anaphoric structural relation.

The only other structural positions where a lexical NP, hence a lexical anaphor, can occur are subject position and in a PP as object of the P. This is because these are the only other positions where Case assignment is possible, and the requirement of Case for lexical NPs proscribes all other positions, however this requirement will be stated in the grammar. We have already

seen that in the case of the PP construction, it seems that the choice between a proximate pronoun or a reflexive is quite erratic and depends on factors that are not of the type that one would want to incorporate in a formal binding theory. However, there are two subclasses of PP constructions that seem to fall into the core case of the anaphoric structural relation. One is when the P has been reanalyzed with preceding elements. But first, consider the other one which is when the PP is the result of the insertion of a dummy Case marking preposition like of or de before the object of a non-Case assigner, like a noun or an adjective, as (46) and (47).

(46) a. A man's destruction of himself/*him is always sad to see.

b. Ronald is proud of himself/*him.

(47) a. Sa peur de lui-même/*lui l'a perdu.

b. Victor est content de lui/lui-même.

b'. Victor est égal à lui-même/*lui.

The example in (47b) seems to contradict this claim, but we return to it shortly. Consider first the cases where the generalization holds. The explanation for the fact that these constructions behave like the core case of the subject/object relation as in (44)-(45) could be that, at the level where this relation is relevant, the two constructions are identical: the reflexive would not be in a PP at that level. One way in which this could be possible is by assuming that of or de insertion in these cases only takes place at PF, for example. As for the problem of accounting for (47b), it could be solved by assuming that (at

least some) prepositions inserted before adjective complements in French are not inserted in PF but at D-structure. There are independent reasons to believe that dummy Case marking prepositions in French do not all behave uniformly and that this depends on lexical conditions: for example, there are the well-known facts about cliticization of à complements. The PP complements with à can either cliticize as a third person dative clitic lui/leur, or as the clitic y.

(48) a. Tim a donné son adresse à Anne.
b. Tim lui/*y a donné son adresse.

(49) a. Eric a pensé à son chat.
b. Eric y/*lui a pensé.

We will see in Chapter 3 when we deal with French relative clauses that there are reasons to believe that the à of penser is a dummy Case marker, just like the dative à in (48) has been analyzed as a dummy Case marker (see Rouveret & Vergnaud 1980, for example). But to account for the contrast between (48) and (49), we will have to assume some difference between the two manners in which à is introduced. We will assume that the à of penser, contrary to the dative à, is not inserted at PF, but is inserted at D-structure. It is most likely that this independently motivated distinction which is responsible for the contrast between (48) and (49) is also relevant in (47b): the NP lui/lui-même is in a PP throughout the derivation in (47b) assuming that dummy de can also be inserted at different levels like à. This difference in the level of insertion of the dummy

Case marking preposition has effects that are visible elsewhere in the grammar. For example, the clitic y is a pro-preposition since it can also be a locative clitic as in (50).

(50) a. Je vais à Montréal demain.
b. J'y vais demain.

On the other hand, lui, although it is marked for dative when it is cliticized, can also be a stressed (but not necessarily emphatic) pronoun as in (51) when cliticization is impossible, so that it seems to be more nominal than prepositional in nature.

(51) a. Je rends ces lettres à vous et à lui (Vigny).
b. Un paquet à lui adressé
c. C'est lui qui me l'a dit.
d. On n'admire que lui.
e. Racine était contre lui.
f. Cela lui fera plaisir, à lui.
g. Paul et lui sont de bons amis.

Vergnaud (1974) also provides two tests that show that some indirect objects in French are NPs rather than PPs. The first test shows that while objects of prepositions can be conjoined, this is not the case for objects of dative prepositions, which indicates that dative à is more like a clitic on the noun than a proposition that creates a PP.

(52) a. *Ils ont parlé à Marie et le directeur.
b. Ils se sont assis sur la table et les chaises.
c. Ils ont pensé à Marie et le directeur.

The second test has to do with the fact that conjoined PPs

cannot be heads of relative clauses, whereas conjoined NPs can be.

(53) a. Il a parlé à l'homme et à la femme qui se sont rencontrés hier.
b. Il a vu l'homme et la femme qui se sont rencontrés hier.
c. *Il a compté sur l'homme et sur la femme qui se connaissaient.
d. *Il a pensé à l'homme et à la femme qui se connaissaient.

So considering that the clitic corresponding to the à-phrase of verbs like penser has a prepositional nature, whereas the clitic corresponding to the dative à-phrase has a pronominal nature, plus the facts presented in (52)-(53), we expect the à of penser to be present throughout the derivation, whereas the dative à is probably inserted late in the derivation possibly only in PF. The prediction with respect to reflexivization, therefore, is that it should be subject to erratic variations with the penser-à type, but that the dative à type should be a core case of anaphoric structural relation, so that a reflexive should be obligatory in this case. This prediction is borne out as we can see by the contrast between the sentences in (54) and (55).[8]

(54) Dative à
a. Pierre s'est donné des coups.
b. *Pierre a donné des coups à lui.
c. *Pierre lui a donné des coups.

(55) Non-dative à
a. Pierre pense à lui/*lui-même avant de penser aux autres.
b. Pierre s'intéresse seulement à lui/lui-même.

In the constructions with a dative à-phrase like in (54), the reflexive is obligatory, whereas in the non-dative à-phrase constructions like (55), the choice between a reflexive and a pronominal form is subject to conditions like those put forth by Zribi-Hertz (1981). So if we now reconsider the sentence in (47b) given here as (56),

(56) Victor est content de lui-même/lui.

we see that the erratic choice of NP form can be explained by the fact that the dummy Case marker de here is present throughout the derivation so that the NP is in a PP throughout the derivation, hence not in a core anaphoric relation with the subject Victor.

What we have seen so far is that reflexives are obligatory in the underlined positions in(57).

(57) Obligatory reflexives:

a. $NP_i\ [_{VP}\ V\ \underline{NP_i}]$

b. $NP_i\ [_{XP}\ X\ [_{PP}\ P*\ \underline{NP_i}]]$ (i) where P*= dummy P Case assigner in NPs, APs, dative in French.

Another construction that seems to be part of the core anaphoric structures is a construction where Reanalysis has applied, (cf. Hornstein & Weinberg (1981), Reimsdijk (1978), Rothstein (1981)). Recall that sentences like (58), anaphoric forms are obligatory.

(58) a. They spoke to each other/themselves/*them.

b. They are sorry for each other/themselves/*them.

There might be a functional explanation for why Reanalysis is obligatory here. Since what one wants is a coreferential relation inside an S, and since the core reflexive relation can be established after Reanalysis, it might force Reanalysis to apply, thus allowing a contrast between coreferential and non-coreferential relations in the sentence. If Reanalysis did not apply, this contrast would be lost since a pronoun would be ambiguous between a coreferential and a non-coreferential interpretation.[9]

Since anaphoric forms are obligatory in constructions where reanalysis has applied, we will add these to the core cases given in (57).

(59) (57) ... (ii) where P* is reanalyzed with other material.

There are other facts that suggest that the reflexive forms that show up in constructions which are not the core reflexive constructions as described in (59) are more like reflexive forms functioning as pronouns than like true anaphors. For example, some of these "false anaphors" can appear in constructions where they do not have a syntactic antecedent, as in (60a), and "false anaphors" can also appear in constructions where the usual c-command relation between anaphor and antecedent does not hold as in (60b).

(60) a. Faith in yourself is important in this job. (Fiengo 1977)

a'. This is a picture of myself/me which was taken years ago. (Koster 1978b)

a''. Physicists like herself are rare. (Fiengo 1977)

b. This is a book by John about himself.

b'. Pictures of himself amused him.

"False anaphors" also enter into pronominal constructions like coordination: in these cases, there is variation across languages, as expected.

(61) a. Tom believed that the paper had been written by Ann and himself.

b. *Tom geloofde dat de paper was geschreven door Ann en zichzelf.

"False anaphors" can also have more than one antecedent.

(62) $John_i$ showed $Mary_j$ pictures of $themselves_{i+j}$/$them_{i+j}$.

Similar facts are also presented in Helke (1970). For instance, Helke shows that "false reflexives" differ from "true reflexives" in that they can have multiple antecedents, as in sentences like the rich girl showed her husband a picture of themselves, and that they can violate the usual linear order (i.e. hierarchical relation) of true reflexives, as in sentences like that the picture of himself in the paper is ugly enrages John. So the fact that reflexives form a heterogeneous class is well established (see Helke 1970, Koster 1978a,b, among others). Furthermore, Ken Hale informs us (personal communication) that many languages have only the core cases of reflexives, and have no pronominal reflexives, so that in such languages, there is a one-to-one correspondance between the morphological form reflexive and the type anaphor.

All of this suggests that "false anaphors" behave more like

pronouns than like true anaphors.[10]

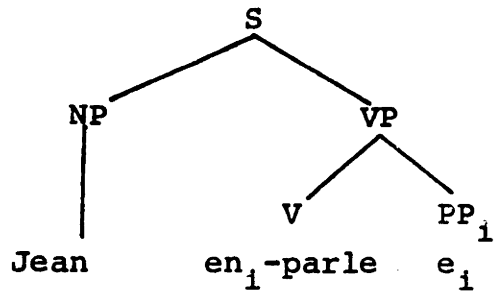
There is one more fact that must be accounted for about the possibility of a reflexive relation between a subject and the complements of a phrase: although pronouns inside PPs can be coreferential with the subject, this is never possible for a clitic PP (or any non-reflexive clitic).[11]

(63) a. Jean a parlé de lui.
a'.*Jean en parle.
b. Jean pense à lui (avant de penser aux autres).
b'.*Jean y pense.

Chomsky (1981a) suggests that, in languages that allow cliticization, it might be the case that it is clitics rather than full pronouns which fall under condition B of the binding theory. This would explain the contrasts shown in (63). However, when Chomsky (1981a) later takes into account the facts about pronouns and reflexives in PPs, he concludes that any structural argument seems to be doubtful.

In our analysis, reflexives are obligatory when the relation between the subject and the bound position is direct, in the sense that there is no structural PP node intervening at the relevant level (cf. (59)). The corollary of this is that pronouns coreferential with the subject are barred from such positions where direct binding can be established. Consider the S-structure of (63a'), given in (64).

(64)



In (64), the relation between Jean and PP_i is direct, in the same sense described above, since Jean binds the PP itself. So Jean and PP_i cannot be coreferential. The same conclusion holds if one considers that it is en rather than PP_i that is bound here. Note that one could not appeal to a condition that would block coindexing of NPs and PPs, since such a relation is possible in configurations where the generalization in (59) is not contradicted.

(65) Jean$_i$ aime Marie$_j$ et il$_i$ en$_j$ parle souvent.

So the fact that clitics cannot be coreferential with the subject falls into the generalization in (59). There is no difference in the operation of the binding theory between languages that allow cliticization and languages that do not: the binding theory applies in the same way to the core cases described in (59), and PPs are opaque in both types of languages. The only difference is that clitics are always in a direct relation with the subject since no PP node intervenes between the two. So clitics are not like English pronouns in general, but like English pronouns in object position.

So far, we have looked at two of the three positions where a lexical NP can appear and the possibility of having anaphors

in such positions: the object position of a verb, and the object position of a preposition. There is a third position where a lexical anaphor could potentially appear as far as Case requirements are concerned: in some subject positions. We now turn to the task of determining which, if any, of the subject positions are core positions for anaphoric elements, i.e. positions where we can find only truly anaphoric phrases, and no pronouns. Consider the following phrases.

(66) a. <u>They</u> read [$_{NP}$ <u>their</u> books]
b. <u>They</u> read [$_{NP}$ each other's books]

In (66), we see that both pronouns and reciprocals can appear in the subject position of a NP. Reflexives are not possible in such position, but it might be due to morphological factors: genitive Case might be incompatible with reflexive morphology. One could make an argument in favor of the view that these positions are in fact accessible for reflexivization but that this morphological constraint forces the use of a pronoun, thus supporting an analysis where most of the positions that we have discussed so far are in fact possible positions for anaphors and that pronouns are marked cases. This attitude is the opposite of the one that we are taking, as we have mentioned earlier. This would explain why <u>each other</u> is possible in (66b). But it would be at a loss to explain why the accidental clash between reflexive and genitive morphology is pervasive in languages. Surely one would not want to make the <u>ad hoc</u> claim that such a condition on reflexive and genitive morphology is universal.

On the other hand, if we say that the subject position of an NP is opaque, then we must explain the possibility of having each other in (66b). Intuitively, we say that each other is an anaphor because it must be linked to another element in the sentence. So let us define an anaphor very generally as in (67).

(67) Anaphor = an element requiring an antecedent.

What we must then determine is what creates this dependency of anaphors on an antecedent. We have already proposed that reflexives must have an antecedent in order to get an R-index to satisfy the principle of Denotability. We will assume that reciprocals must have an antecedent for a different reason: along the lines of an analysis by A. Belletti (personal communication), we will assume that some element of a reciprocal must be moved by an LF rule of each movement to the antecedent for the reciprocal to be properly interpreted at LF. We already know that each is subject to such a rule in a sentence like (68).

(68) The diplomats were assigned one interpreter each.

Furthermore, there are languages like French where the reciprocal can actually be separated into its two elements which then occupy two θ-positions, as in (69b).

(69) a. Ils doutent les uns des autres.
b. Les uns doutent des autres. (reciprocal or distributive)

If a reciprocal is linked to its antecedent by an LF rule of each movement, then we have an explanation for the fact that

each other is possible in an opaque position in (66b): this position might be opaque for grammatical binding, which is the process that relates an anaphor to its antecedent (and which we will make more precise shortly); but each movement is a rule that is clause bound, so that it can link they and each other in (66b) since they are in the same S. Assuming this analysis to be on the right track, we consider that the subject of an NP is not a position accessible by grammatical binding, i.e. not a possible position for an anaphoric reflexive form.[12]

Another reason to distinguish between the way in which a reflexive and the way in which a reciprocal is linked to its antecedent is that reciprocals can appear in some PPs where reflexives are impossible.

(70) a. *They saw snakes near themselves.
b. They saw snakes near them.
c. They saw snakes near each other.

The facts in (70) also show that an analysis of the distribution of anaphors and pronouns in such constructions cannot be derived from the binding theory by simply positing a PRO in the structure which creates an opaque domain. So, for example, in such an analysis, the contrast between (70b) and (70c) could be due to the presence of a PRO in the former but not in the latter, as in (71a) and (71b) respectively.

(71) a. They saw snakes $[_{PP}$ PRO near them]
b. They saw snakes $[_{PP}$ near each other]

But then there is no explanation as for why this option of not inserting a PRO in the structure is not available for a reflexive as in (70a). So again, it seems that the distribution of reflexives and pronouns in PPs cannot be determined on structural grounds (except in cases of reanalysis and dummy Ps, where a reflexive form is obligatory: but then there is no PP at the relevant level as we saw above).

Note that given an analysis as in Chomsky (1981a) where the distribution of PRO is dependent on government, and given the notion of government of Aoun and Sportiche adopted by Chomsky (1981a), PRO would not be possible in (71a) anyhow for theory internal reasons: the PRO would be governed by the P. This would also hold for <u>picture</u> phrases, where the N would govern the subject position. Furthermore, that EC could not be a trace since movement is generally impossible from that position, and there would be no antecedent available for the trace. Finally, the EC could not be a pro, the pure pronominal EC of Chomsky (1981b), since the distribution of this element seems to be dependent on features of INFL in that analysis, and such features are not available for pro in these positions.

Consider the sentences in (72).

(72) a. <u>They</u> thought [$_{\bar{S}}$ that <u>they</u>/*<u>themselves</u>/*<u>each other</u> would come]

b. <u>They</u> want [$_{\bar{S}}$ for <u>each other</u>/?<u>themselves</u>/*<u>them</u> to win]

b'.?*<u>They</u> would be happy [$_{\bar{S}}$ for <u>each other</u>/<u>themselves</u> to win]

c. They believe [$_S$ each other/themselves/*them to be intelligent]

In (72a), we see that the subject of a tensed clause is not accessible by grammatical binding.[13]

In (72b) and (72b'), we see that for subjects in infinitives are generally not accessible by grammatical binding, except for the want-type verbs which are marginally better. Some speakers accept sentences like (72b'), others, like Chomsky, for example, consider them "marginal at best". So we will consider these not to be core anaphoric structures. (We will return to the analysis of want in Chapters 3 and 5.)

Finally, in (72b) and (72b'), we have an instance of an Exceptional Case Marking (ECM) construction. In this type of construction, we will assume with Chomsky (1980, 1981a) that Case is assigned to the embedded subject by the matrix verb because "$\overline{S}$-deletion" has taken place, so that the S-structure of (72c) is (73).

(73) They believe [$_S$ each other/themselves to be intelligent]

The subject position of the infinitive in (73) is a position where a reflexive is obligatory if coreference is intended, as we can see in (74).

(74) *They believe [$_S$ them to be intelligent]

So we will consider the subject position in an ECM construction to be a core anaphoric position. Note that this is the only Case-marked subject position to be a core anaphoric

position: the subject of an NP (cf. (66)), the subject of a tensed clause (cf. (72a)) and the subject of a for infinitive (cf. (72b,b')) are all non-accessible by grammatical binding.

We are now in a position to determine what grammatical binding is. The core structures where we assume grammatical binding to apply are those described in (59), with the addition of the ECM constructions that we have just seen. Those are the structures in which a reflexive form is obligatory, and a coreferent pronoun is impossible. These structures are given in (75), with respective examples in (76).

(75) a. NP_i $[_{VP}$ V $NP_i]$

b. NP_i $[_{XP}$ ... $[_{PP}$ P* $NP_i]$ ...] (where P* = a dummy Case assigner not present at relevant level or reanalyzed P)

c. NP_i V $[_S$ NP_i ...]

(76) a. John shaved himself/*him.

b. [A man's destruction of himself/*him] is always sad to see.

b'. Ronald is proud of himself/*him.

b''. They spoke to themselves/*them.

c. John believes $[_S$ himself/*him to be intelligent]

The generalization that emerges from (75) is that grammatical binding holds between an antecedent and an anaphoric reflexive only in structures where no maximal expansion intervenes between them. For example, in (76a), we assume VP not to be a maximal expansion, as we have already stated in Chapter 1. In (76b) and (76b'), the PP node does not count as a maximal expan-

sion because the P* is not present at the relevant level: for example, one could say that the P* is inserted only at PF, so that there is no PP in the syntax. In (76b''), reanalysis has applied, so that there is no PP node at the relevant level. Finally, in (76c), $\bar{S}$ has been deleted, so that again no maximal expansion intervenes between the two NPs.

We already have a type of relation in the grammar which holds only where no maximal expansion breaks the relation. This relation, for which there is very strong independent motivation, is the notion of government presented in Chapter 1 and repeated here as (77).

(77) Government
In the structure $[_{\gamma} \ldots \underline{\beta} \ldots \underline{\alpha} \ldots \underline{\beta} \ldots]$,
$\underline{\alpha}$ governs $\underline{\beta}$ if and only if

(i) $\underline{\alpha}$ is an immediate constituent of γ

(ii) where $\underline{\psi}$ is a maximal projection, if $\underline{\psi}$ dominates $\underline{\beta}$, then $\underline{\psi}$ dominates $\underline{\alpha}$.

(where maximal projections are NP, PP, AP, $\bar{S}$ $(=V^{max})$)

In all of the structures in (75), the binder governs the bindee. This indicates that the notion of government plays a crucial role in binding and must be incorporated in the definition of binding as in (78).

(78) Binding[14]
$\underline{\alpha}$ binds $\underline{\beta}$ if and only if
$\underline{\alpha}$ governs $\underline{\beta}$ and $\underline{\alpha}$ assigns its R-index to $\underline{\beta}$.

The formulation of Binding in (78) has immediate consequen-

ces on the relation between an antecedent and an anaphor. The relation will be very local because of the government requirement; this requirement of government will also force a special structural relation to be established between the two elements. Binding will force the antecedent to be unique since it cannot apply twice to the same element: it assigns indices only to NPs that do not already have one; so if it applies once, it cannot reapply because the NP now has an index. Finally, the principle of Denotability will force the relation to be obligatory, or else the anaphor would lack a crucial property in the mapping from LF to domain D, i.e. an R-index, and thus would be uninterpretable.

2.2.3.4. Summary.

To sum up section 2.2.3., we have seen how the requirements of the principle of Denotability are met by the three different manifestations of a lexical NP. A name has an R-index and F-features intrinsically. A pronoun can freely pick any R-index at S-structure, and it must agree in F-features with the name of which it picked up the R-index. A lexical anaphor has no inherent R-index; it gets one from an antecedent that Binds it in the sense of (78), and it must agree in F-features with its antecedent. In both cases of pronoun and anaphor, agreement is determined by the principle of Agreement (16).

2.2.4 The principle of Denotability and $[_{NP}e]$.

In this section, we will see how an EC meets the require-

ments of the principle of Denotability. We have seen that a NP must have two properties to satisfy the principle of Denotability: it must have an R-index, and it must have F-features.

2.2.4.1. Assignment of R-indices.

First, consider how EC gets an R-index in Chomsky (1981a). There are three different ways in which an EC gets its R-index. In the first case, the EC gets its R-index from an antecedent to which it is related by the rule of move α: this is the case of NP trace and WH trace. The second case is that of pro in Pro Drop languages: this pro has its own index, but it must obligatorily agree with AGR. The third case is PRO in infinitives: it gets its R-index by the theory of control which either relates PRO to an antecedent, or assigns it the R-index arbitrary. Let us look at these three cases in turn. We will give a very general outline of the analysis of these cases in this section and we will return to each of them in a more detailed analysis in Chapters 3, 4, and 5.

Consider ECs that get their R-index from an antecedent to which they are related by move α. We already have an instance of this type of obligatory relation to an antecedent in order to get an R-index which is independently motivated in the grammar: it is the relation between an antecedent and an anaphor which is mediated by Binding. In line with our commitment to have general principles that apply to categories, regardless of whether they are lexical or not, we will assume that Binding as described in (78) is the relation that holds between a trace and

its antecedent, unless serious difficulties force us to adopt a weaker position.

In the case of pro, we can assume that the agreement with AGR depends on the same factors on which the agreement of other nominal elements depends, since these are independently motivated in the grammar. So we can assume that AGR is a nominal element that has an R-index, and that it is the identity of R-index between pro and AGR that triggers the agreement. The fact that this agreement is obligatory indicates that the indexing between pro and AGR is obligatory: again, this looks a lot like the relation between an antecedent and an anaphor which is mediated by Binding, so we will assume that Binding is operative in this case, too.

As for PRO, we see that the assignment of an R-index to it is not dealt with in a uniform fashion in Chomsky (1981a). We could assume that at least some cases where PRO gets its R-index from an antecedent fall along the lines of the analysis for the other ECs above: some instances of PRO are likely to be related to an antecedent by Binding. However, this cannot be the case for $PRO_{\underline{arb}}$, for example, since it does not have an antecedent. But we will see that this problem is just apparent, and that arbitrary PRO and "long distance control" PRO are strictly pronominal, not anaphoric.

This analysis of how ECs get their R-indices assumes a slightly different indexing procedure than the one presented in

Chapter 1. There, we suggested that all A-positions get an R-index at D-structure, relating this indexing to the assignment of a θ-role by inserting the R-index of a complement into the θ-grid of the head of the phrase, as proposed by Stowell (1981a). The procedure that we are now assuming is that names and quantified phrases index their A-position by being inserted in such a position at D-structure. Other A-positions, i.e. positions where pronouns or anaphors have been inserted, which do not get an R-index at D-structure by lexical insertion, either get one by binding from an antecedent when they are in the right configuration, or are freely indexed at S-structure. This free indexing of A-positions at S-structure is similar to the indexing procedure proposed in Chomsky (1981b) for all A-positions. So we have a hybrid indexing procedure: we assume Chomsky's (1981b) free indexing of A-positions at S-structure, but only for those A-positions which were not provided with an R-index previously in the derivation, either by insertion of material bearing an R-index, or by Binding. So A-positions where names and quantified phrases are inserted get an R-index at D-structure (79a); lexical anaphors get an R-index at D-structure by Binding (79b), or after application of move α by Binding if passive applied in an ECM construction for example (79b'); pro gets an R-index at S-structure by Binding by AGR, which itself is freely indexed at S-structure (79c); trace gets an R-index at the application of move α by Binding by its antecedent (79d); PRO can get an R-index at D-structure by Binding by an antece-

dent (79e), or after move α by Binding by an antecedent when it is moved in subject position (by raising or passive) (79e'); or PRO is freely assigned an R-index at S-structure, this PRO being coreferential with other NPs in the sentence (79f) or free in reference (79g); this free assignment of R-index at S-structure also covers the case of lexical pronouns (79h) and "false anaphors" (79i).

(79) a. $John_i$ said that $everyone_j$ was happy.
b. John shaved $himself_i$.
b'. John believes $himself_i$ to have been cheated t.
c. pro_i $mangia_i$ le mele.
d. $Bill_i$ was seen t_i at the movies.
e. $John_i$ tried PRO_i to leave on time.
e'. $John_i$ tried PRO_i to be allowed t_i at the meeting.
f. $John_i$ said that it would be difficult PRO_i to feed $himself_i$.
g. John said that it would be difficult PRO_i to feed $oneself_i$.
h. He_i said that $he_{i/j}$ would come.
i. $John_i$ said that a picture of him_i/$himself_i$ was hanging on the wall.

One distinctive feature of the hybrid indexing procedure adopted here is that the two manners of indexing differ in the relation that is established between the coindexed elements. The free indexing procedure is a symmetric indexing, whereas indexing by Binding is an asymmetric indexing. One consequence of this difference will be that symmetric indexing allows more than one antecedent for a given NP, whereas asymmetric indexing is restricted to only one antecedent. So this difference in

the possible number of antecedents can also be used as a clue to determine what kind of indexing is at stake. There is also a structural relation, namely that of government (cf. (78)) which is present in Binding but not in free indexing, this structural relation creating a locality condition on Binding which is not found in free indexing. Furthermore, Binding by an antecedent of an NP lacking an R-index is obligatory, whereas free indexing can assign any index to an NP, including one not previously assigned in the sentence. So the relation of Binding between an antecedent and an anaphor explains why these core antecedent-anaphor relation have the four basic properties that we saw in 2.2.3.3 when a reflexive is obligatory: 1° obligatoriness of the antecedent; 2° a one-to-one relation; 3° locality of the relation; 4° a structural condition on the relation.

What goes on in the indexing procedure is that once the narrow procedures of indexing because of lexical specifications (names and Q-phrases) and indexing by Binding have taken place, then the NPs left without an index get one by an Elsewhere assignment of index, namely free indexing at S-structure. So this provides us with a simple and conceptually attractive indexing procedure.

2.2.4.2. F-features and the EC

The second thing that we must determine is how and when the EC gets its F-features. We must assume that ECs have F-features since they are required by the principle of Denotability. In Chomsky (1981a) and current analyses in GB, it is also assumed

that the EC has F-features. Thus in GB, if the F-features are base-generated without a corresponding phonological matrix, then the EC is PRO if it is ungoverned, pro if it is governed. If the F-features are left behind by move α, then the EC is a trace; if the trace is locally $\bar{A}$-bound, it is a variable, and if not, it is an anaphor (NP trace).

So we see that some ECs get their F-features at D-structure (PRO and pro), but others (traces) get them at S-structure, since the traces were created by move α and so were not present in D-structure. The fact that all ECs have F-features gives empirical support to the claim that there is only one EC, as noted in Chomsky (1981a). We can see in (80) that the four manifestations of the EC agree with other elements.

(80) a. John seems t to be proud of himself. (NP trace)
b. Who [t saw himself in the mirror] (variable)
c. John tried [PRO to shave himself] (PRO)
d. pro se vede (pro)

We now have an answer as to where and how ECs get their F-features in GB. But notice that this approach to the F-features of ECs is dependent on what the theory of agreement is and where it applies. However, no explicit theory of agreement is given in Chomsky (1981a), where the above argument is presented. From the places in the derivation where Chomsky (1981a) assumes that ECs get their F-features, we could infer that agreement applies at S-structure, or at PF, or at LF, if, as

what seems to be assumed by Chomsky (1981a), agreement is checking that the features of coindexed elements match. In fact, agreement could be said to apply "when possible", so that it could apply at all levels, including D-structure.

In our analysis, we will assume that the how of agreement is determined by the principle of Agreement already stated in (16) (repeated here as (81)).

(81) Agreement:
$\underline{\alpha}$ assigns (redundantly) its F-features to $\underline{\beta}$ if $\underline{\alpha}$ and $\underline{\beta}$ have the same R-index.

As for when agreement applies, it will be dependent on our answer to the question "Do ECs have F-features?". If one assumes a model of grammar where what goes on in LF is "invisible" to what goes on in PF, and vice versa, then one expects the answer to the question "Do ECs have F-features or not?" not to be yes or no, but yes <u>and</u> no, that is, that ECs have F-features at LF but not at PF. The fact that an NP must have F-features at least at LF comes from the principle of Denotability. But if an NP could be assigned F-features by some operation at LF only, then this NP might not have F-features at all before LF since it could meet the requirements of the principle of Denotability anyhow. So an EC could be inserted in D-structure without any F-features and be assigned F-features at LF only. This we will call the complementary approach to ECs.

If the principle of Agreement is responsible for checking and assignment of F-features, then the complementary approach

to ECs implies that Agreement takes place on the LF side of the grammar. But we have already seen in dealing with the facts about agreement of deictic pronouns in (12-13) that this is independently motivated. So let us assume that Agreement takes place only on the LF side of the grammar.[15]

The derivation of an EC is now the following. The EC is either generated directly in the base (PRO, pro), or it is created by move α. In both cases, the EC has no F-features and no R-index. In order to satisfy the principle of Denotability on the LF side of the grammar, the EC must be provided with F-features and an R-index. If the EC has an antecedent which binds it, then it has the R-index of this antecedent by Binding, and hence it is assigned the F-features of its antecedent by Agreement. Another possibility for the EC is to be simply coreferential with some NP bearing F-features: in this case, the EC is not bound by the NP bearing F-features in the sense of (78), it is only picking its F-features from an NP with which it was randomly coindexed at S-structure. A last possibility is for the EC not to be bound to nor coreferential with any NP, in which case the EC would have no F-features at all. There is such an EC, namely what is referred to as PRO_{arb}, which is simply assigned an index at S-structure that is not coreferential with any other NP in the sentence: PRO_{arb} is essentially "free PRO". We return to the case of PRO_{arb} shortly. But first, let us look at the consequences of adopting the complementary approach to ECs on the analysis of these ECs.

2.2.4.3. Consequences of the complementary approach to ECs.

The complementary approach to ECs has a direct consequence on the distribution of ECs. First, consider how the distribution of NPs is determined in Chomsky (1981a,b) and other work in the GB framework. The distribution of lexical NPs is dependent on Case assignment: a lexical NP must bear Case (or have its Case checked) to satisfy the Case Filter at PF (or to satisfy the visibility condition at LF if the Case Filter is derived from this condition). However, Case assignment has no explicit bearing on the distribution of ECs in general in GB. ECs may have Case or not: it is assumed that variables and pro have Case, but that NP trace and PRO do not have Case.[16] The distribution of ECs in GB depends on three distinct components of the grammar, as we have already mentioned: ECP, control theory, and the Pro Drop condition. These components all have the property of allowing in some way to recover the content of the EC. So they can determine where an EC may appear, but they do not determine where the NP must be empty. This is implicitly determined by Case theory for PRO and NP trace since no Case is assigned to the positions where such ECs occur, and therefore no lexical NP could appear there. With WH-traces and pro, the explanation is different since in GB, these positions where WH-trace and pro can occur are Case marked. The reason given as for why a language will have a WH-trace rather than a resumptive pronoun is that the language will be specified in some way as having a resumptive pronoun strategy or not, or, in other words,

whether the language has a rule spelling out Case or not: so there is an indirect connection with Case here. As for pro, Chomsky (1981b) says that its distribution depends on whether a language has such an element in its lexicon or not, subject to recoverability of its content by a "rich enough" inflection on the verb: so there are no clear reasons why pro should have Case or not in this analysis.

Consider now the following idea of Jaeggli (1981). Jaeggli proposes to account for the distribtion of PRO vs lexical pronouns with a statement, given informally here, to the effect that PRO is pronounced if it has Case and is C-governed (i.e. governed structurally by an X^{O}). Suppose that we extend this idea to all the manifestations of the EC in the following way: an EC will be pronounced, i.e. lexical, if it has Case. And let us generalize this to all features that might have a morphological realization in PF: lexicalization of the NP will therefore be obligatory if the NP bears any morphological features at PF. This can be stated as in (82).

(82) Principle of Lexicalization
A noun N will be lexicalized if and only if ψ-features are present in the entry of N at PF, where ψ = person, number, gender, Case.

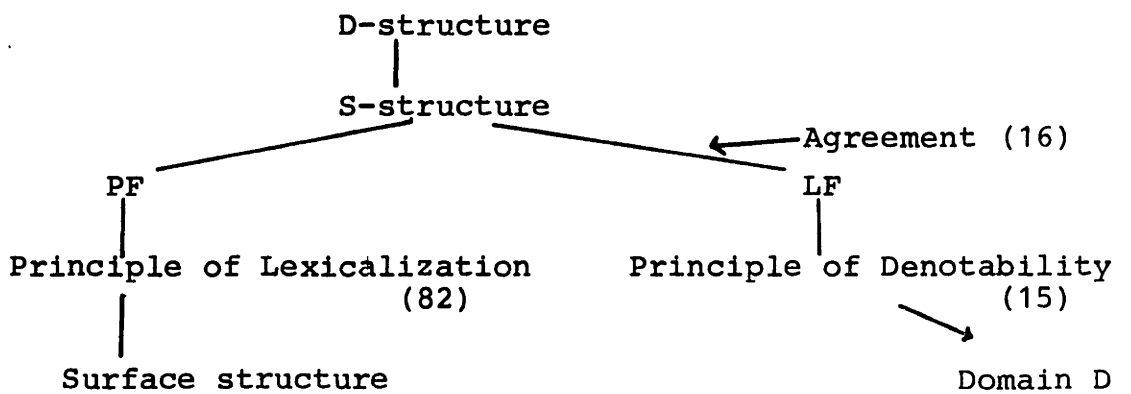
Informally, what this principle says is that a lexical noun must have all ψ-features at PF, i.e. person, number, gender, Case, whereas a non-lexical noun, i.e. an empty category, must not have any of these features at PF. The Principle of Lexical-

ization will account for the distribution of lexical and empty NPs. First, we have seen that lexical NPs all have F-features since they can directly satisfy the Principle of Denotability at LF: so by the biconditional of (82), they will be forced to have Case also. Thus we trivially derive the Case Filter. Now consider the different manifestations of the EC. PRO and NP trace do not have Case; neither do they have F-features at PF since we now assume that they are assigned these features by agreement with their antecedent on the LF side of the grammar; so PRO and NP trace are not lexicalized since they do not have any ψ-features at PF as required by (82) for an N to be an EC. As for pro, we will show in Chapter 4 that it is a property of Pro Drop languages that the subject of a tensed clause may be Caseless at PF; and pro gets its F-features by agreement with AGR at LF, so that pro has no "visible" feature at PF and is not pronounced. Note that if the inflection of the verb could not provide all of the F-features to pro on the LF side of the grammar, then pro would have to be provided with the missing feature(s) before LF in order to meet the requirements of the principle of Denotability, i.e. at S- or D-structure; hence pro would have "visible" features on the PF side of the grammar: since an N with one or more of the ψ-features in PF cannot be an empty category by (82), it would have to be lexical in order for the missing feature(s) to be provided, hence it would have all ψ-features. We return to the details of such cases in Chapter 4. Finally, WH-trace, we assume, does not have Case, but

the Case is absorbed by the WH-phrase itself; since the WH-trace also gets its F-features from its antecedent by Agreement at LF, it has no ψ-features at PF and so is not pronounced. If the Case is not absorbed by the WH-phrase, then the trace is lexicalized as a resumptive pronoun since, having at least one ψ-feature, the N cannot be an empty category, hence must be lexical and have all ψ-features by (82). (We return to the cases of resumptive pronouns and to the facts of contraction like want to/wanna in Chapter 3.)

So we see that the principle of Lexicalization, which is a very general principle applying to the category N, gives us the distribution of lexical NPs and empty NPs if we assume the complementary approach to ECs. The principles introduced so far fit into the model of the grammar as in (83).

(83)



Binding is not placed anywhere in (83) because we assume that it applies whenever it can, that is, whenever an NP is without an R-index and that NP is governed by an R-index bearing NP, then a Binding relation is established between the two NPs.[17]

Note that the principles in (83) never have to refer specifically tc EC: they are all stated on NPs or N for (82). This is clear for the principle of Lexicalization and the principle of Denotability. Agreement also applies to all Ns for which it is relevant, regardless of whether they are lexical or not: so lexical pronouns, lexical anaphors and ECs are all subject to Agreement (so are attributive names of the class like gardien/gardienne, travailleur/travailleuse, etc. in French). Similarly, Binding applies to all NPs for which it is relevant, i.e. NPs with no R-index, regardless of whether they are lexical or not: so lexical anaphors and ECs are all subject to Binding.

The general requirements to have an EC are the following: the NP must not have ψ-features at PF, but it must have F-features at LF.

The complementary approach to ECs has the direct consequence that it allows the formulation of a general principle of Lexicalization, and thus accounts for the distribution of lexical and empty NPs. The complementary approach to ECs also has for indirect consequence that, by interacting with the other components of the grammar in (83), the principle of Lexicalization allows all the statements that refer specifically to an EC to be eliminated from the grammar.

First consider the Case Filter, which applies only to lexical NPs and so should be eliminated from the grammar according to our general methodological approach. The effects of the Case Filter are essentially covered by the principle of Le-

xicalization. As for the visibility condition (1 (24)), repeated here as (84), from which the effects or the Case Filter are derived in Chomsky (1981a), it must also be eliminated from the grammar, according to our hypothesis, since it refers specifically to an EC, namely PRO, which is singled out.

(84) A chain can be θ-marked in LF if it has Case or is headed by PRO.

There are two ways to be "visible" in LF according to (84). Consider the first one which says that Case makes a chain visible. In our approach, it is the Principle of Denotability which will determine if an NP can denote in Domain D, hence bear a θ-role. If an NP has Case, by the Principle of Lexicalization it will also have F-features, and it will be headed by a lexical N. Lexical Ns either have an inherent R-index, i.e. names and quantified phrases, or they get one by Binding, i.e. anaphors, or they are freely assigned one at S-structure, i.e. pronominals. So any NP with Case will have an R-index and F-features, as required by the principle of Denotability. If such an NP is in a chain, then the chain can bear a θ-role since the position in the chain where the θ-role is assigned will either be filled by the lexical NP itself, or by an EC bound to the lexical NP, hence an EC having the appropriate features to bear a θ-role.

The second possibility of being "visible" at LF is by being PRO. Although PRO does not transmit a θ-role to a lexical NP like trace does, it can nevertheless be bound by a lexical NP or be coreferential with a lexical NP, so that it has its R-

index either by Binding or free indexing at S-structure, and gets its F-features by agreeing with its antecedent. So PRO also meets the requirements of the principle of Denotability (we return shortly to $PRO_{\underline{arb}}$). Therefore, we see that the effects of the condition of visibility are derivable from the interaction of the components in (83) (see the discussion of the condition of visibility in 3.3.2.4.).

Next, consider the three recoverability mechanisms proposed in GB. The ECP can be derived in the following way: in order for a trace to denote in domain D, and hence to bear a θ-role, it must satisfy the principle of Denotability; so a trace must have an R-index and F-features, like any NP. If we assume that a trace gets its R-index and its F-features by being bound by its antecedent and agreeing with it, then we derive the ECP since Binding involves government. We will see in Chapter 3 that, assuming that VP is not the maximal expansion of V and that Binding is as in (78), then the ECP as a recoverability condition is totally recoverable since clause (ii) of the ECP, i.e. government by an antecedent, can cover all of the cases: it reduces in fact to proper Binding. Clause (i), i.e. government by a lexical head, appears to be part of Proper Government because lexical government is necessary for a trace in some positions for reasons that are not relevant to the "proper identification" of the trace: such a trace must be lexically governed to receive a θ-role, so that lexical government is an ac-

cidental property of Proper Government. As for successive cyclicity, it will be an iteration of the Binding relation. The details of the operations involved will be the topic of Chapter 3.

The Pro Drop condition requires that pro be linked to a "rich enough" AGR. If we assume that there is a Binding relation which is established between pro and AGR, then the F-features of pro will necessarily be those of AGR, since pro and AGR will share the same R-index. If AGR lacks some feature(s) required by the principle of Denotability, then the feature(s) will have to be specified by pro itself: but then, by the principle of Lexicalization, pro will have to be pronounced. Thus we derive the effects of the Pro Drop condition, which states that a "missing subject" is possible only if AGR is rich enough. We will see in Chapter 4 that some languages can have Pro Drop independently from the richness of AGR because of other ways which they have to provide the proper F-features to pro, thus supporting our analysis of the Pro Drop phenomenon, and hence the complementary approach to ECs.

Control theory assigns an antecedent to PRO. We will see in Chapter 5 that this is not a uniform process, but rather that assignment of an antecedent can be done locally by Binding, and that "long distance" control is also possible when PRO is related to an NP by free indexing at S-structure rather than by Binding. The R-index and F-features of PRO being provided in these manners, the requirements of the principle of Denotability are

therefore met, and this part of control theory is derived from the components in (83). The case of $PRO_{\underline{arb}}$ is slightly different since it has no antecedent. But before turning to the analysis of $PRO_{\underline{arb}}$, we must discuss a redundancy that we have introduced in our grammar, since it is directly relevant to the analysis of $PRO_{\underline{arb}}$.

Consider the principle of Denotability: it states that both an R-index and F-features are necessary to denote in domain D. But we have another mechanism operating in the grammar that also deals with R-indices and F-features: Agreement, which is a well-formedness condition on coindexed nominal elements.[18] In fact, Agreement is even stronger than that since, as we have seen in (12)-(13), deictic pronouns must "agree" with elements that might never be actually mentioned in the discourse. Suppose that we say that objects in domain D have an inherent R-index in some sense: then we could assume that agreement generalizes in a straightforward fashion to these cases.[19] But then the part of the principle of Denotability that refers to F-features can be derived in the following way. We assume that objects in domain D are marked for F-features in some sense, so that names have inherent F-features. Pronouns, lexical anaphors and ECs have F-features because they have an R-index, and so they agree with the name which has the same R-index, whether the agreement is syntactic or pragmatic, as in the case of deictic pronouns. So by the very fact that the principle of Denotability states that an R-index is required for an NP to denote, it

implies, given Agreement, that F-features will be specified for such an NP. So the part of the principle of Denotability that states that F-features are required is redundant since this is already covered by Agreement. Furthermore, since to denote in domain D basically means to be an argument, and since an argument must bear a θ-role according to the θ-criterion, what the principle of Denotability amounts to is stating that some element is an argument if and only if it has an R-index. And what can be deduced from this is that θ-roles are assigned only to elements that bear an R-index. This fits nicely with Stowell's (1981a) claim that θ-roles are assigned by inserting the R-index of an NP in a θ-grid of some sort. So we can now revise the principle of Denotability as in (85).

(85) Principle of Denotability II
An NP will denote an object in domain D if and only if that NP has an R-index.

Given the assumption that Agreement takes place on the LF side of the grammar, then all of the results that we obtained above still hold.

Having properly distributed its task to each component of the grammar, we can now look into the case of $PRO_{\underline{arb}}$ with the right perspective.[20] $PRO_{\underline{arb}}$, as we have hinted at above, is simply PRO that is free, or $[_{NP}e]$that is free for that matter. This means that positions where $PRO_{\underline{arb}}$ appears are all positions which are not accessible for Binding by an antecedent. The details of the reasons for why Binding is not possible in these

cases will be given in Chapter 5. A PRO in such a position will therefore be freely indexed at S-structure. The instances when there will be a PRO with arbitrary interpretation in such a position will be when the index assigned to PRO is not shared by any other NP in the utterance. This means that PRO in such positions is always potentially ambiguous between a coreferential and a free interpretation, depending possibly on pragmatic factors and factors of agreement with other elements, as can be seen in (79f-g) above for example. These factors will be examined in more detail in Chapter 5.

When PRO is not coreferential with any element, then that means that Agreement applies vacuously, since there is no element that has F-features for PRO to agree with. Consequently, $PRO_{\underline{arb}}$ has no F-features at LF according to our analysis. At first thought, this seems a bit curious, but then we now have an explanation for why $PRO_{\underline{arb}}$ is interpreted as a variable-like element, although it is not technically a variable, i.e. it is not $\bar{A}$-bound: having no F-features, $PRO_{\underline{arb}}$ can range over all the individuals in domain D that can satisfy the predicate of which $PRO_{\underline{arb}}$ is subject, regardless of the grammatical features associated with these individuals.[21]

But then we must consider an argument put forth in Chomsky (1981a) in favor of saying that ECs have F-features: this is the fact that $PRO_{\underline{arb}}$, which cannot get its features from an antecedent since it does not have an antecedent, nevertheless must have F-features since it agrees with a reflexive, for example.

Furthermore, the features of $PRO_{\underline{arb}}$ vary from one language to another: it is singular in French and English, but plural in Italian, for example. But unanalysed data tell us nothing about what is going on. The argument that PRO has features depends crucially on what one's theory of agreement is, and where it applies. If the theory of agreement is that coindexed elements must have matching features and that this theory applies at PF, then the argument holds: PRO is going to have F-features all along. But if we assume that agreement takes place at LF, then $PRO_{\underline{arb}}$ has no F-features at LF.

The question then is what are those features that show up on the elements that agree with $PRO_{\underline{arb}}$. We propose that they are the unmarked features of the language. Note that arbitrary PRO is often used with reflexives that do not seem to have the "normal" features of the language, these features often being used only with $PRO_{\underline{arb}}$, or indefinite pronouns like <u>on</u>, <u>chacun</u> in French, for example.

(86) [PRO to praise oneself] is a bad habit.

(87) [PRO parler de <u>soi</u> continuellement] est emmerdant pour les autres.

In the case of Italian, Luigi Burzio points out to us (personal communication) that in sentences corresponding to (86) and (87), $PRO_{\underline{arb}}$ seems to be picking a single individual in domain D, suggesting that the plural marking on the reflexive is only an unmarked feature of the language.

These peculiarities in the facts about agreement give em-

pirical support to our analysis of $PRO_{\underline{arb}}$ since the interpretation given to PRO does not necessarily correspond to the one expected by the features showing up on the coindexed elements. We will also see in Chapter 5 that there are conceptual reasons to favor this analysis since it predicts the distribution of $PRO_{\underline{arb}}$ without having to add any particular mechanism to the grammar, contrary to analyses that incorporate a theory of control that specifically states where $PRO_{\underline{arb}}$ can occur.

2.2.4.4. Summary.

To recapitulate what we have seen in section 2.2.4., we assume that Binding assigns an R-index to NPs that do not have such an index and that it applies whenever it can. Names and quantified phrases have an inherent R-index when they are inserted at D-structure, but lexical pronouns and anaphors, and ECs do not have an inherent R-index. (We could assume that names and Q-phrases do not have an inherent R-index but that they must pick one freely, subject to a condition like condition C of the binding theory of Chomsky (1981a). We will discuss this topic in 2.3.) The principle of Denotability is assumed to be a well-formedness condition on R-indices for arguments, which has for consequence that θ-roles are assigned only to elements that bear an R-index. Agreement is a well-formedness condition on the F-features of coindexed elements. The interaction of these components results in a complementary approach to ECs which has for consequence that the realization of an NP as a lexical NP or an EC can be simply predicted by

the principle of Lexicalization. All of these components are stated on NPs in general, and allow to derive the components of GB where specific reference is made to ECs only or, conversely, to lexical NPs only. Thus the complementary approach to ECs is in line with the strong hypothesis that we have set as our goal, that no statement in the grammar should ever refer specifically only to ECs or to lexical NPs.

2.2.5 The partitioning of the EC.

We now turn to the different types of ECs that are manifested in the grammar. The idea that there is only one EC with different manifestations or types which are functionally determined was first presented in Chomsky (1981a). There, Chomsky observed that the types of ECs that he was proposing then virtually exhausted the partitioning of the EC. Thus, an ungoverned EC is PRO, and a governed category is trace, a variable if Case-marked, an NP trace if not. This can hardly be accidental as was observed by Chomsky, and one immediate explanation of this fact is to assume that there is only one EC, and that its different manifestations are determined functionally. Chomsky (1981a) noted, however, that this partition was not perfectly exhaustive since there remained the possibility of a governed EC which would not be <u>properly</u> governed, as required by the ECP.

Chomsky (1981b) proposes to rethink the status of the EC by considering principles A and B in the following way:

> These principles identify two categories of expressions: anaphors and pronominals. Principle (A) holds of anaphors,

> Principle (E) of pronominals. If the binding theory is correct, then in the best of all possible worlds we would expect to find four categories of expressions:
> (Chomsky 1981b, p. 44f)

(88) (i) (+anaphor, -pronominal) = overt anaphors and NP traces

(ii) (-anaphor, +pronominal) = overt pronouns and pro

(iii) (+anaphor, +pronominal) = PRO

(iv) (-anaphor, -pronominal) = names, variables

There can be no lexical element falling under (iii), says Chomsky, since it would have to be ungoverned this being derived from the fact that a pronominal anaphor is subject to Principles A and B, and hence could not get Case, thus violating the Case Filter.[22]

Chomsky (1981b) assumes that the type of the occurrence of an EC is determined by the grammatical properties of the structure in which the EC appears. We agree with this approach where there is only one EC and its different types are functionally determined. However, we disagree with Chomsky's partitioning: we believe that the premise on which it is based, namely a binding theory with his principles A and B, is wrong as we will see in 2.3. Furthermore, we do not see any convincing argument to the effect that an element like a pronominal anaphor exists: rather, PRO is either anaphoric or pronominal, and we will see in Chapter 5 that there are empirical reasons to believe this, as well as conceptual ones.

In our analysis, the prediction as to where an EC can appear is the following: an EC will be possible if it bears no

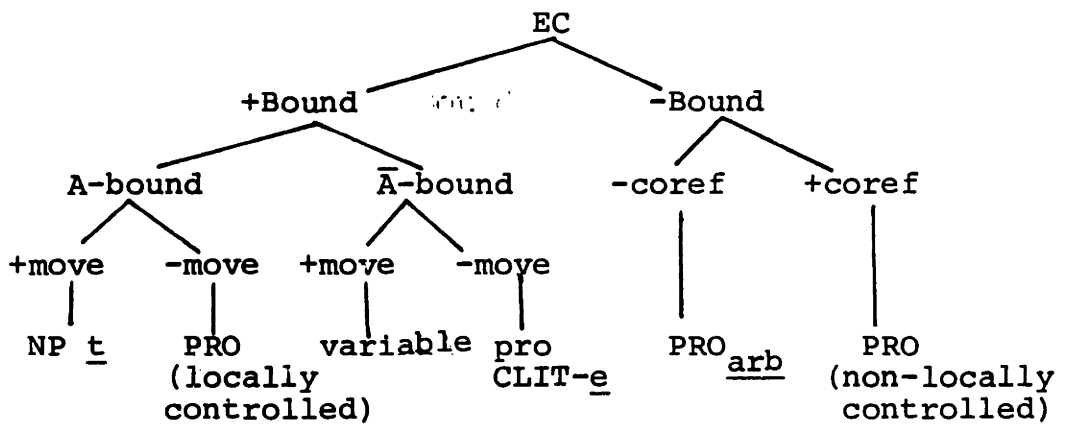
$\underline{\psi}$-features (as in (82)) when it reaches PF, and if it can be assigned F-features on the LF side of the grammar in order not to be lexicalized.

The consequences of adopting the complementary approach to ECs with the components in (83) are interesting conceptually since this allows to drop from the grammar statements that refer specifically to ECs, thus strengthening the claim that theoretical entities like ECs exist since their existence is predicted by general statements about NPs, like the principle of Denotability, the principle of Lexicalization, Binding, Agreement, which are independently needed for lexical anaphors, for example. This move will ask for some technical revisions on the functioning of some components of the grammar, and some explanations will be technically different, although not conceptually different. For example, the distribution of PRO, one realization of the EC, is dependent on whether a position is governed or not in GB. In the present analysis, the distribution of PRO will depend on a subcase of government, namely on Case, by the principle of Lexicalization. Note, incidently, that all the different realizations of EC will be dependent on Case and so will all the different types of lexical NPs, thus providing a unified account of the distribution of the EC and lexical NPs.

We assume, along the lines of Chomsky (1981a), that there is only one EC and that when we use the terms NP trace, WH-trace, PRO, QR-trace, pro, we are in fact referring to different functional uses of the EC. We propose a partitioning of the EC that

is also determined functionally since it depends on the nature of the relation that the EC has with an antecedent and how it is established in the grammar. The partition is given in (89).

(89)



The notions (±bound), (± coreferential) are not intended to be used as features in any way here, but only as descriptions of functional relations. For example, (± move) is simply to indicate how the EC was created: by non-insertion of lexical material, or by movement. Some examples of the elements in (89) are given in (90).

(90) a. John seems t to be happy. (NP trace)

a'. John was seen t at the park. (NP trace)

b. John tried PRO to go. (locally controlled PRO by Binding)

c. Who did John see t? (variable)

d. Je le$_i$ vois e$_i$. (CLIT-e)

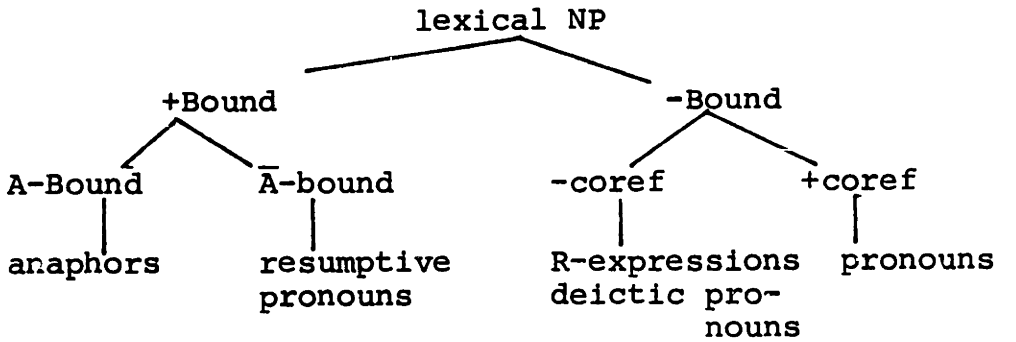
d'. e ho trovato il libro. (pro)

e. PRO to jog on Main Street is unhealthy. (PRO$_{arb}$)

f. John told Mary that PRO to leave early would be important. (PRO non-locally controlled)

The same partitioning holds for lexical NP, except for the (+move) empty elements which are created by the movement rule which vacates a position, hence cannot be lexical.

(91)



The partitioning in (89) respects the two interesting hypotheses made about the partitioning of the EC in Chomsky (1981a): there is only one EC, with different manifestations of the EC exhausting this partitioning, and the principles used to determine the different types are all independently motivated, namely Binding, coreferential index, and the rule move α.

2.3. Some consequences with respect to coreference.

In Chapter 1, we briefly reviewed the elements that participate in the processes of the binding theory in GB but without going into the details of the processes involved. We will now consider the binding theory of GB, looking at some of the reasons that are given for some choices in the analysis of coreference facts, pointing out some problems, and then presenting an analysis of coreference that is compatible with the general approach to ECs presented in the preceding sections.

Recall how facts about coreference are dealt with in GB. In Chomsky (1981a) for example, the binding conditions in (92)

govern the relations of anaphors, pronouns, names and variables to possible antecedents.

(92) Binding conditions:
A - An anaphor is A-bound in its GC.
B - A pronominal is A-free in its GC.
C - An R-expression is A-free everywhere.

The definitions of bound and governing category are as follows:

(93) a. Bound:
$\underline{\alpha}$ is X-bound by $\underline{\beta}$ iff $\underline{\alpha}$ and $\underline{\beta}$ are coindexed, $\underline{\beta}$ c-commands $\underline{\alpha}$, and $\underline{\beta}$ is in an X-position. (X = A, $\bar{A}$)

b. $\underline{\alpha}$ is locally X-bound by $\underline{\beta}$ iff $\underline{\alpha}$ is X-bound by $\underline{\beta}$, and if $\underline{\gamma}$ Y-binds $\underline{\alpha}$, then either $\underline{\gamma}$ Y-binds $\underline{\beta}$ or $\underline{\gamma}$ = $\underline{\beta}$.

(94) Governing category: $\underline{\beta}$ is a GC for $\underline{\alpha}$ iff $\underline{\beta}$ is the minimal category containing $\underline{\alpha}$, a governor of $\underline{\alpha}$, and a SUBJECT accessible to $\underline{\alpha}$. (where SUBJECT is "the most prominent nominal element" in an expansion, including AGR; accessible means that linking of $\underline{\alpha}$ and the SUBJECT must not violate well-formedness conditions like the i-within-i condition for example (cf. Chomsky 1981a)).

The notion 'accessible SUBJECT' is introduced in Chomsky (1981a) to replace a stipulation that a GC can only be NP or S of previous analyses: Chomsky says that this can be explained by the fact that NP and S are the two categories that can contain SUBJECTs, and that it is the SUBJECT that creates the opaque domain. But note that this choice of what is a GC crucially depends on the assumption that all morphological anaphors are true anaphors with respect to the binding theory.

So in this approach, all the reflexive forms in (95)-(96), for example, are true anaphors with respect to the binding theory.

(95) a. John shaved himself/*him.
b. John is always talking about himself/*him.
c. John believes himself/*him to be happy.

(96) a. They saw snakes near each other/them.
b. Victor est content de lui-même/lui.

On the other hand, it must also be assumed that some pronominals function as anaphors in this analysis, like them in (96a) and lui in (96b) for example, since there does not seem to be any structural way to distinguish between such sentences where a pronominal and an anaphoric form are equally possible.

In GB, the notion of accessible SUBJECT allows PPs to be transparent, thus explaining why reflexive forms are allowed in these contexts, but there does not appear to be any structural reason why the PPs should be opaque when a pronoun is involved (cf. the discussion of (70)-(71) above), so it is not clear how the presence of a pronoun is allowed in cases where both a pronoun and an anaphor are possible. Such an analysis also bases its typology of lexical anaphors and pronouns on strictly morphological factors, so that it does not distinguish between the two types of reflexives which we saw above: an anaphoric reflexive with obligatory, unique, local, and structurally restricted antecedent, and a pronominal reflexive with converse properties. In our analysis, some of these reflexive forms are

not anaphors, but rather pronouns: this is the case whenever the reflexive form is not governed by its antecedent, hence not Bound by it. So we have to assume that some reflexive forms are in fact false reflexives and are pronouns, not anaphors. Note that this has to be assumed both in GB and in our approadh for cases like sentence (97).

(97) That is a picture of myself.

The only difference between the two approaches in this respect is in the extension of the domain where such false anaphors can occur. Our analysis predicts where anaphoric forms are obligatory, however.

So although introducing the notion of SUBJECT in the grammar allows us to reduce the number of instances where it must be assumed that a reflexive form functions like a pronominal, and thus allowing condition A of the binding theory to cover more of the reflexive forms, though not all of them, it seems to create a problem in allowing violations of condition B by having pronouns bound in their GC in sentences where a reflexive or a reciprocal can occur in the same position where a pronoun can occur as in (70) or (96). The problem, again, is that it does not seem that a structural argument can be made to avoid this result, like postulating a PRO subject for some PPs in some cases. This lack of structural argument makes it difficult for such an analysis to predict where these violations of condition B will take place, thus weakening the binding theory proposed

in such an analysis.

In our analysis, we have one assumption to make: that is, that the notion of anaphor relevant for the binding theory is not strictly dependent on the morphological form of the elements, but rather on the manner in which they relate to their antecedent. It allows us to make the hypothesis that conditions A and B, however they are expressed in the grammar, are always respected. The only apparent exceptions like some of the reflexives in PPs for example, are not anaphors in the sense of the binding theory and do not have the four basic properties of anaphors. (We return to the definition of what is an anaphor and how conditions A and B are stated in the grammar shortly). Therefore, one should make the strongest hypothesis and adopt the analysis proposed in 2.2.3.3. over the GB analysis, unless one has strong reasons not to do so.

There are also technical problems with the notion of GC as defined in (94). For example, consider sentence (98).

(98) *[For PRO to leave] would be too bad.

In (98), PRO is governed, but it has no accessible SUBJECT since AGR is assumed to be coindexed with the subject of the sentence, which is <u>For PRO to leave</u>: thus coindexing AGR and PRO would violate the i-within-i well-formedness condition, so AGR is not an accessible SUBJECT for PRO in (98). This means that, strictly speaking, PRO has no GC in (98) if GC is defined as in (94), and therefore, PRO, which is a pronominal anaphor

in Chomsky's analysis, should be allowed here. Note that the fact that PRO must be ungoverned in GB cannot be the reason why (98) is ungrammatical since the non-government of PRO is a theorem that is derived from the fact that PRO must not have a GC. If two things are necessary for PRO to have a GC, i.e. a governor and an accessible SUBJECT, then the theorem that is derivable is that PRO must not be governed and have an accessible SUBJECT at the same time. So in order to rule out cases like (98), something additional has to be said. One possibility is to stipulate, as in Chomsky (1981a), that when a category has a governor but no accessible SUBJECT, as in (98), then the whole sentence counts as a GC: so, a root sentence is a GC for a governed element.

This makes the analysis work, but one may wonder why there is a disparity between the two elements that enter in the definition of a GC. So a higher S is considered to be the GC if there is no accessible SUBJECT, but there is a governor. However, this is not the case when there is no governor but there is an accessible SUBJECT as in (99), for example.

(99) John tried [PRO to win]

Here, PRO is not governed in Chomsky's analysis since there is no $\bar{S}$ deletion after try, but it has an accessible SUBJECT John. PRO must not have a GC in his analysis since PRO is both pronominal and an anaphor, hence is subject to both conditions A and B. So there must be a disparity between accessible SUBJECT

and governor, as in (98), but one cannot have a GC if there is only an accessible SUBJECT, as in (99).

In our analysis, the facts about PRO in (98) and (99) are accounted for in the following way. Sentence (98) is ungrammatical because the position governed by <u>for</u> is assigned Case, and therefore it must be lexical by the principle of Lexicalization. And (99) is grammatical because no Case is assigned to the position of PRO in PF, as we will see in Chapter 5 where a detailed analysis of control sentences will be given.

Another technical problem with respect to the binding theory of GB has to do with the definitions of the elements involved in the binding theory. For example, Chomsky (1981a) gives the following definitions of a variable and of a pronominal.

(100) a. $\underline{\alpha}$ is a variable if and only if it is locally $\bar{A}$-bound and in an A-position.

b. $\underline{\alpha}$ is a pronominal if and only if $\underline{\alpha}=[_{NP}\ F,(P)]$, where P is a phonological matrix and $F \subset \psi$ (person, number, gender, Case) and either (i) or (ii)

(i) $\underline{\alpha}$ is free

(ii) $\underline{\alpha}$ is locally A-bound by $\underline{\beta}$ in a θ-position.

If $\underline{\alpha} = [_{NP}\ F]$, it is PRO; otherwise, $\underline{\alpha}$ is a lexical pronoun. Chomsky (1981a) also gives the principle in (101).

(101) If $\underline{\alpha}$ is an empty category and not a variable, then it is an anaphor.

Some additional statement obviously has to be made to the effect that some lexical elements like reflexives and reciprocals

are inherently anaphors, something to the effect that anaphors do not have independent reference, suggests Chomsky. Note that (101) goes against our general approach to ECs since it is a statement that deals specifically with ECs.

Consider (102).

(102) John tried [PRO to leave]

Analyzing PRO as a pronominal anaphor in cases like (102), and deducing from this that PRO must not have a GC in order not to violate conditions A and B of the binding theory has for consequence that any anaphor having the properties in (100b) should be a pronominal anaphor, and hence have no GC. But a reflexive is such a case and should not have a GC since a reflexive can be locally A-bound by an antecedent in a θ-position or free since we have seen that there are cases where reflexive forms are used as pronouns, and hence have independent reference.

These technical problems with the notion of GC in GB are related to an attempt made in Chomsky (1981a) to replace the notion of GC by the notion of Binding category in the theory of binding.[23]

(103) Binding category:
β is a binding category for α if and only if β is the minimal category containing α and a SUBJECT accessible to α.

Chomsky shows that this simplified definition accounts for all the facts that the notion of GC accounts for, except for one

remaining problem illustrated in (104).

(104) (i) John expected [him to win]
(ii) John tried [PRO to win]
(iii) John knows [how [PRO to win]]

> In (i), him cannot be coindexed with John or (condition B) will be violated. But exactly the same argument shows that PRO cannot be coindexed with John in (ii), (iii), an incorrect result. Replacement of "binding category" by "governing category" gives the correct results, in this case. It therefore appears to be necessary to introduce a crucial reference to government in the binding theory, as in (94), though its effects are so narrow as to suggest that an error may be lurking somewhere.
> (Chomsky 1981a, p.221)

But note that if PRO is only an anaphor in (104ii) as we suggest, then the simplification of the definition to binding category is possible.[24] We return to this topic in Chapter 5.

Now consider again the theorem that states that PRO must be ungoverned. We have seen that, in fact, the theorem states that PRO cannot have a GC, and that means that PRO cannot have a governor and an accessible SUBJECT at the same time, assuming that PRO is a pronominal anaphor.

In order for the analysis to work, some stipulation has to be made to the effect that having a governor is enough to have a GC, but that having only an accessible SUBJECT is not enough to have a GC, as we have seen in the discussion of (98) and (99). Furthermore, we now see in Chomsky's discussion of (104) that the effects of the reference to government in the binding theory are extremely narrow: it is not relevant in any other case than the one of the pronominal anaphor PRO. So, first, the re-

ference to government has to be given a privileged status over that of accessible SUBJECT in the definition of GC to account for PRO, this privileged status of government holding only in the case of PRO; and furthermore, government is relevant only for PRO in the binding theory and has no effect on lexical categories. This means that the reference to government is necessary in the binding theory only to account for PRO. So in order to get the result that PRO is ungoverned, one must add to the binding theory a reference to government that has no independent motivation and is introduced in the binding theory strictly to derive the effect that PRO must be ungoverned. This means that the non-government of PRO is not a theorem of the binding theory since there are some elements of the binding theory from which such a theorem is derived which are not independently motivated: the introduction of the reference to government in the binding theory is derived from the theory-internal decision to have such an element as an ungoverned pronominal anaphor and it has no other effect in the grammar. So it seems that the non-governed status of PRO should be used in this perspective as a factor to determine the distribution of PRO, and it should not be considered as a point for or against any analysis whether or not it has an element with the properties that PRO is assumed to have in GB. In fact, since there does not seem to exist any lexical item which is a pronominal anaphor, one would have to have very strong reasons to postulate that there exists such an EC, and the burden of proof is on the side

of one who claims that such an element exists.

It is possible that these technical problems can be overcome by changing some notions or refining them. But the interaction of these problems suggests that it is the assumption about what is an anaphoric relation that is not right from the start in this type of analysis, and that an anaphoric relation should rather be approached in the way proposed in 2.2.3.3., where the anaphoric relation is much more restrictive than it is in GB.

Let us now turn to the analysis that is to follow from our general assumptions. In our analysis of coreference facts, the technical problems described above do not arise since we do not make the assumptions that create them from the start. The distribution of pronouns is straightforward in our analysis: pronouns never occur in transparent domains, they are never bound in the sense of (78).[25] As for anaphors, we have assumed that they are not to be defined only on morphological grounds, but on the grounds of their relation with their antecedent: this explains why and where some reflexive forms are obligatorily anaphoric. The remaining reflexive forms, we have assumed to function as pronouns. A class of false anaphors is necessary in any analysis, as we have seen in dealing with sentences like (97); we only extend this class of elements, and we assume that pragmatic factors, possibly of the type proposed by Zribi-Hertz, account for the distribution of these false anaphors and pronouns, with possibly additional factors that are not well under-

stood yet. The definitions of the different manifestations of NP, whether lexical or not, are the following:

(105) $\underline{\alpha}$ is an <u>anaphor</u> if and only if $\underline{\alpha}$ is Bound.

(106) $\underline{\alpha}$ is a <u>pronominal</u> if and only if $\underline{\alpha}$ is freely indexed at S-structure.

(107) $\underline{\alpha}$ is a <u>variable</u> if and only if $\underline{\alpha}$ is $\bar{A}$-Bound and in an A-position.

The notion of anaphor as defined in (105) captures the intuitive description of an anaphor given earlier: an anaphor is an element which is referentially dependent on an antecedent: it has something missing. If an anaphor is Bound as in (105), then this implies that the anaphor is missing an R-index. The notion of anaphor as defined in (105) covers the cases of NP trace and locally controlled PRO, and lexical anaphors, i.e. the A-Bound elements in the partitioning of the EC in (89) and of the lexical NP in (91). Note that it does so without having to refer to the notion of EC, as in Chomsky (1981a) in his principle given in (101), where specific reference is made to the EC. The fact that no reference is made to the EC in the definition of anaphor (105) is in line with our general conceptual approach to ECs. This notion of anaphor also covers variables, which are a subcase of anaphors; it also covers the traces in COMP when successive cyclic movement has applied. So variables are assumed to have the same relation with their antecedent that anaphors have, except that the antecedent is in an $\bar{A}$-position for variables. Bound pronouns, for which we will assume an

analysis similar to the one in Reinhart (1976) and Haik (1982), can also be interpreted as variables of LF, and the resumptive pronoun strategy will be considered as an extension of the use of bound pronouns (see 3.3.2.2 for an analysis of resumptive pronouns in Québec French).

This analysis of anaphors applying to both lexical anaphors and some ECs is in line with the original formulations of trace theory (Chomsky 1977a,b , Fiengo 1977). The basic idea was that trace theory is of very little cost since traces are anaphors and independently motivated principles are needed to account for lexical anaphors in any case. The present theory differs from those early formulations in that it claims that not all elements which have the morphological shape of a reflexive are true anaphors. For example, we claim that himself in a sentence like (108) is not an anaphor (i.e. is not Bound by its antecedent), but is rather a pronominal, a "false anaphor".

(108) Pictures of himself amused John.

In Fiengo (1977) for example, himself is considered to be an anaphor in a sentence like (108), and the condition in (109) is given to account for this.

(109) The antecedent of a reflexive pronoun must precede or asymmetrically command it (where command is the notion Kommand proposed by Lasnik: "A kommands B if the minimal cyclic node dominating A also dominates B.")

Fiengo then proposes to extend (109) to traces. He notes however that there are no known rules that yield a structure like $\ldots[\ \ldots[\underline{e}]_i\ \ldots]_{\underline{\alpha}}\ \ldots\ [X_i]$ where $\underline{\alpha}$ is a cyclic node. So the second part of (109) is relevant only for reflexives, it seems. Moreover, note that such structures with reflexives allow more than one antecedent, as in sentences like Pictures of themselves at the party amused Mary but enraged Bill.

All of this tends to show that there are not traces found in this context because this is not a structure where real anaphors are found, but rather false anaphors which behave like pronominals. Therefore, the second part of Fiengo's condition (109) should be removed from the condition on anaphors. The first part about precedence has to be refined, as has long been recognized in the literature. Our proposal is that it should be as in our definition of Binding in (78).

Pronominals, on the other hand, are elements that are freely indexed at S-structure, this being subject to a coreference condition similar to condition C of the binding theory in (Chomsky 1981a) and to which we return shortly. Pronominals can be lexical pronouns, or ECs that are not Bound, i.e. "long distance" controlled PRO and $\text{PRO}_{\underline{\text{arb}}}$, or finally reflexive forms that are not bound by an antecedent. These three elements have in common the fact that they are freely assigned an R-index as S-structure, and so fall into the class of pronominals as defined in (106).

Turning now to the binding conditions (92), we see that con-

dition A can be derived from the definition of an anaphor in (105) and the principle of Denotability, assuming that the notion of GC is changed to binding category as in (110).[26]

(110) $\underline{\beta}$ is the binding category for $\underline{\alpha}$ iff $\underline{\beta}$ contains a Binder for $\underline{\alpha}$.

Since an anaphor always gets its R-index from an antecedent which binds it by definition, then an anaphor will always be bound in its binding category: so condition A of the binding theory is derivable as a theorem.

Given that condition A is derivable as above, we expect condition B to be derivable too since it is a mirror image of condition A: pronouns can never appear in positions where true anaphors show up.

We have seen that there are three kinds of pronominals. Consider first the case of the EC pronominal, i.e. "long distance" control PRO and $PRO_{\underline{arb}}$, and also the second kind of pronominal which is the non-Bound reflexive form. The derivation of condition B is straightforward for these two kinds of pronominals since they are pronominals only when they are freely indexed at S-structure, that is, when they are not Bound by an antecedent when they reach S-structure, since in this latter case they would be functionally defined as anaphors. So the EC pronominal and the reflexive form pronominal have to be free at S-structure or else they are functionally determined to be anaphors and hence are not subject to condition B but to condition A. And we have seen that condition A is derivable from Binding

(78).

The third kind of pronominal is a lexical pronoun. Since a lexical pronoun has no R-index when it is inserted at D-structure, it could be inserted in a position where a Binding relation can be established, so that it is assigned the R-index of the antecedent, as in (111).

(111) *John$_i$ saw him$_i$.

A pronoun like him in (111) is functionally defined as an anaphor according to the assumptions made above. But we have seen that pronouns cannot function as anaphors since a coreferential pronoun is never possible in the core cases of Binding relations given in (75): thus (111) is ungrammatical. This is what is expected in condition B of GB, and this is what we want to derive. We could always state that there is a lexical requirement on pronouns to the effect that they must be freely indexed at S-structure: but this amounts to restating condition B.

What we want to express is that, although technically a derivation where him in (111) is Bound by John is possible, some additional mechanism blocks this Binding relation because of the pronominal form of the bound element. According to the analysis presented above, it is not in the technicality of the derivation that there is a hitch since the Binding relation between John and him in (111) is technically possible because him has the same bindee status that himself has: they both lack an R-index when inserted at D-structure. We could always say that pronouns

do have an R-index at D-structure, and that they differ from reflexive forms in this respect. But this would weaken the analysis of "false reflexives" which would not be on a par with true pronominal forms anymore in their indexing procedure. Furthermore, we will see in the discussion of (117) and following that the conceptual advantage of dealing with the indexing of these elements in a unified fashion is empirically motivated. So it has to be at the level of the form of the pronoun that something is blocking. Therefore, our analysis seems to push us to the formulation of a condition on the morphological form of the pronouns like (112) or (112').

(112) A Bound element (in the sense of (78)) cannot have a pronominal form.

(112') A Bound element must have the form of a reflexive.

Although it might seem a bit strange at first that condition B of the binding theory should be reduced to a morphological condition on Bound elements, there are reasons to believe that this is the case, and that the condition must be stated as in (112') rather than (112). We could see the effects of condition (112') in a language where some (or all) pronominal forms were identical to reflexive forms. In such a case, condition (112') predicts that a pronominal could appear in a Binding relation without having a reflexive interpretation: it would only be identical in form with the reflexive, not in meaning. There are some facts in French that seem to have precisely these pro-

perties.

In French, there are strong reflexive forms which are formed by adding the suffix même(s) to the strong form of the pronoun: elle-même, lui-même, nous-mêmes, etc. But there are also clitic reflexive forms. In the third person singular or plural, the clitic reflexive form is se, as opposed to the pronominal forms for objects and dative complements which are le, la, les, lui, leur.

(113) a. Jean se lave.
b. Jean se parle.

(114) a. Jean le/la/les lave.
b. Jean lui/leur parle.

As we have observed in 2.2.3.3 above, coreference between clitic and the subject in cases like (114) is impossible.

But in the first and second person, the clitic reflexive forms and the clitic pronominal forms of objects and dative complements are identical.

(115) a. Je me regarde.
b. Tu te regardes.
c. Nous nous regardons.
d. Vous vous regardez.

(116) a. Jean me regarde.
b. Jean te regarde.
c. Jean nous regarde.
d. Jean vous regarde.

The prediction that condition (112') makes is that it should be possible to have the first and second person object clitics in French Bound by an antecedent that would not receive a reflexive interpretation but rather a pronominal interpretation.

And in contrast, this situation should not be possible for third person pronominal forms.[27]

Consider the five following sets of facts about first and second person clitics in French (which are all drawn from Morin 1979). First, consider (117).

(117) Je me plains.

In (117), me can have a reflexive or a pronominal interpretation, and the sentence is ambiguous between the two readings in (118).

(118) a. I am complaining. (inherent reflexive interpretation)
b. I am taking pity on myself. (pronominal interpretation)

On the other hand, when the forms are not ambiguous as in the third person, then the reflexive form can only get a reflexive interpretation, and the pronominal form must be disjoint in reference with the subject.

(119) Jean se plaint.
= Jean complains.
≠ Jean is taking pity on himself.

(120) Jean le plaint.
≠ Jean complains about him.
= Jean takes pity on him. (him≠Jean)

In fact, there is no way to express the pronominal meaning of plaindre in the third person with this verb, and so a paraphrase like (121a) is needed, although in a few non-transparent constructions, this is possible, like in (121b).

(121) a. Jean a pitié de lui-même.
b. Jean ne plaint que lui. ('Jean takes pity only on himself.')

We see that condition (112') makes the right prediction here: a clitic pronoun can be coreferential with the subject if it has the form of a reflexive although it does not have a reflexive interpretation but it has a strictly pronominal interpretation.

A second set of facts has to do with reduced coordinate sentences as in the following:

(122) a. Je me trouvais bête, et ma soeur aussi.
b. Je me trouvais bête, et ma soeur se trouvait bête aussi.
c. Je me trouvais bête, et ma soeur me trouvait bête aussi.

(123) a. Jean se trouvait bête, et sa soeur aussi.
b. Jean se trouvait bête, et sa soeur se trouvait bête aussi.
c. ≠Jean se trouvait bête, et sa soeur le trouvait bête aussi.

Again, the non-ambiguous form se imposes a reflexive interpretation, so that (123a) cannot have the pronominal interpretation in (123c) whereas the ambiguous form me allows an interpretation where me is pronominal and Bound by the subject, the reading in (122c).

A third set of facts involves adjectives. Adjectives in French do not allow reflexive cliticization. So (124a) is ungrammatical, whereas (124b) is grammatical only if ils and leur do not corefer.

(124) a. *Ils se sont fidèles.
b. Ils leur sont fidèles.
c. ?Je me suis fidèle.

But as noted by Morin, some speakers accept quite freely sentences like (124c). This can be explained if me is not interpreted as a reflexive here, but as a pronominal Bound by the subject. Again, this is what is predicted by condition (112').

The same judgements hold for avoir constructions as in (125).

(125) a. Heureusement que je t'ai.
b. *Heureusement qu'ils s'ont.
c. Heureusement que je m'ai.

For some reason, avoir does not allow clitics with reflexive interpretation, as we see in (125b). Yet a clitic Bound by the subject is possible in (125c): so it must be that the clitic gets a pronominal interpretation here, and that it is allowed to be Bound by the subject because it satisfies condition (112').

The last set of facts has to do with some lexical items which are sensitive to the distinction between reflexive and pronominal interpretation. Morin gives the example of comprendre quelqu'un, which is ambiguous between the two meanings "to understand what someone says" and "to understand someone's nature". Morin notes that for many speakers, the second meaning can never be reflexive. If one puts comprendre in a context where only the second interpretation is possible, then the reflexive interpretation is not possible, as we see in (126a).

(126) a. *Il n'y a que lui pour se comprendre.
b. Il n'y a que moi pour me comprendre.

However, as can be seen in (126b), a clitic can be Bound by the subject in such cases if it is not a reflexive, but a pronominal with the form of a reflexive (the subject being PRO which is coreferential with moi here).

So in all these five cases, a pronominal clitic could be Bound when its morphological form was the same as the form of a reflexive clitic, although the context or restrictions on some elements in the sentence precluded a reflexive interpretation. This means that these sentences with a clitic form that is ambiguous between reflexive and pronominal all have an ambiguity as to what structure they have. They can have a D-structure where a pronominal clitic is inserted, as in (127) (assuming base-generation of the clitic here, the distinction between base-generation and derivation of the clitic by movement not being relevant here).

(127) Je me_i plaignais $[_{NP}\ e_i]$

But such sentences can also have a structure where a reflexive form is inserted. Our analysis of these cases (see fn. 8) is that the reflexive construction is a derived verb form where a reflexive form absorbs the θ-role assigned to the subject and also absorbs the Case assigned to the object (or possibly dative object). So the sentence can also have the D-structure in (128a) with the S-structure (128b) after application of move α has raised the object in subject position so that it can be assigned Case.

(128) a. $[_{NP}$ e] me plaignais je
b. Je_i me plaignais t_i

The fact that sentences containing these morphologically ambiguous items can have two different derivations explains the facts that we presented above.

1° The ambiguity between a reflexive and a non-reflexive reading of sentence (117) comes from the fact that it can have two different derivations of the type (127) or (128). On the other hand, since a sentence with se can only have a derivation as in (128), such sentences only have a reflexive interpretation, as we have seen in (119).

2° In the case of reduced coordinate sentences as in (122), we could assume that the VP of the first coordinate fills in the gap in the second coordinate. If the first coordinate can have a reflexive and a non-reflexive structure, then the sentence can get two readings as in (122). But if the first coordinate contains a se as in (123), then it can only have a reflexive structure, hence only a reflexive interpretation in the second coordinate. So the assumption that there is a certain parallelism of structure between the two coordinates at some level accounts for the facts, given the present analysis of such ambiguous forms.

3° If adjectives do not allow reflexive elements, but only pronominal clitics as we see in (124), then an ambiguous form like me is possible if it has the structure (127), not the one in (128).

4° The facts in (125) show that avoir does not enter into derived verb forms with reflexive elements. So a sentence with avoir cannot have a structure as in (128), but it can have one as in (127). Since the Binding of the pronoun is allowed by condition (112') in the case of ambiguous forms like me, sentence (125c) is grammatical.

5° The second meaning of comprendre cannot be reflexive for some speakers, says Morin. In the present analysis, this means that it does not have a reflexive derivation, presumably for reasons similar to those given with respect to the avoir facts above. So its derivation is as in (127), not (128).

It is interesting to note that there is empirical evidence that can show when we have the derivation in (127) or the derivation in (128): it can be found in the choice of auxiliaries that will be assigned. The prediction is that in a non-reflexive derivation like (127), the auxiliary will be avoir; in a reflexive derivation like (128), the auxiliary will be être. (See fn. 8 on the assignment of auxiliary.) Consider, for example, the last set of facts which is illustrated in (126). We have claimed that (126b) is grammatical because a non-reflexive derivation as in (127) is possible in this case if one assumes condition (112'). This means that the auxiliary should be avoir, according to our analysis. The prediction is borne out as can be seen in (129).

(129) a. Il n'y a que moi qui m'aie compris.
b. *Il n'y a que moi qui me sois compris.

In (129a), m' is a pronominal clitic since the auxiliary is avoir. In (129b), we have a reflexive construction since the auxiliary is être. So we see that the reason why (126a) is ungrammatical is because comprendre, in its second meaning described above, does not enter into derived verb forms with reflexive elements: in other words, it cannot have a derivation as in (128). And the reason why (126b) is grammatical is because the structure is as in (127), and me is not a reflexive but a pronominal: this is why the only auxiliary that is possible is avoir in (129), since the pronominal clitic construction is assigned the auxiliary avoir.[28]

All these facts show that some condition like condition (112') is responsible for what is going on in the binding of pronouns: the relevant condition seems to be one on the morphological form of the lexical element to be Bound, and not one about binding of pronouns in general that would be relevant to the binding theory as in GB. This means that the explanation as for why clitics in sentences like (130) are not coreferential with the subject (or rather cannot be coreferential with the subject) is that the clitic (or the EC that it binds) is Bound by the subject when it is coreferential with it in this configuration, as we observed in the discussion of (64) above, and that this is a violation of condition (112') since these clitics do not have a reflexive form.

(130) a. Jean_i lui_j parle. (i=j)
b. Jean_i en_j parle. (i=j)

These facts about binding of pronouns also show that Chomsky's (1981a) suggestion about the binding theory in languages that allow cliticization cannot be right. Recall that Chomsky (1981a) suggested that in such languages, the binding theory applies only to clitics, and not to the strong form of pronouns. But condition B, which governs the possibility of coreference for pronouns in Chomsky (1981a) is violated by pronominal clitics in all the five sets of facts discussed above.[29]

Before looking at other consequences of condition (112'), something must be said about the fact that sentences like (131) and (132) are ungrammatical, i.e. that no anaphors can appear in the subject position or a tense clause.

(131) a. *Himself came for dinner.
b. *Each other ate the apple.

(132) a. *John said that himself was rich.
b. *They said that each other would do the painting.

In (131), we could assume that the anaphors are linked to AGR, by Binding for the reflexive, and each-movement for the reciprocal. What seems to be going on is that AGR provides an R-index to the reflexive and is accessible to each-movement, but that the R-index of AGR is not "referentially strong enough" in some sense. This could come from the fact that we have been subsuming under the notion R-index two different types of indices: syntactic indices which are relevant for syntactic processes, and referential indices which are relevant for interpretive

processes. For example, we do want to say that expletive elements like it, there, il have a syntactic index, although they are not referential elements, since movement of these elements is subject to the same requirements that referential elements are subject to: namely, these expletive elements must Bind their trace in the sense of (78) as we see in the following sentences.

(133) a. It seems to be certain that Mary will come.
b. *It seems [t is certain that Mary will come]

(134) a. There seems [t to be a man in the room]
b. *There seems [t is a man in the room]

(135) a. Il semble [t y avoir trop de fumée]
b. *Il semble [t y a trop de fumée]

So expletives have a syntactic index which can be used to create a Binding relation with a trace, but they are not referential. One way to express this is to say that an R-index is in fact decomposable into two kinds of indices: a syntactic index (S-index) and a referential index (REF-index).

This way of dealing with R-indices has interesting consequences for the analysis of adjective agreement. We could assume that adjectives have an R-index that contains only an S-index since adjectives are not referential. But there is a sense in which they are coindexed with the head of NP in cases like (136) for example since they agree with the head.

(136) a. La belle fille
b. Le beau gars

Similarly, as (136) shows, determiners could also be con-

sidered as bearing an S-index, although they do not bear a REF-index. If we make this assumption about adjectives and determiners, then their agreement follows from the general principle of Agreement given in (16) above. Note that when adjectives are used as nouns in French or German, we could simply assume that the change of category from A to N involves the addition of a REF-index slot in the R-index of these forms. Note furthermore that this use of adjectives as nouns is possible only in languages where the adjective bears F-features, as is predicted from the Agreement requirement on referential elements at LF: a referential element must in some sense agree with an object in domain D (cf. the discussion of (9)-(14) above).

So the analysis of expletives and of the agreement facts of adjectives and determiners suggests that the notion of R-index in fact incorporates the two notions of S-index and REF-index. If we now return to the problem of anaphors in the subject position of a tensed clause as in (131)-(132), we have an immediate answer if we assume that AGR only has an S-index, and does not have a REF-index, which seems intuitively right. So in (131), <u>himself</u> is Bound by AGR: it therefore has an S-index but no REF-index. Since a REF-index is necessary to get a θ-role, the sentence is ungrammatical since an argument is lacking a θ-role, or conversely, a θ-role cannot be assigned. This means that we are assuming that the principle of Denotability is not stated on R-index, but on REF-index, which also seems intuitive-

ly correct if we are to make a distinction between S-index and REF-index: we do not want expletive elements, adjectives or determiners to be assigned a θ-role for example.

The case of each other in (131b) is parallel to that of himself in (131a): each is moved to the accessible nominal AGR, but AGR lacks a REF-index, hence lacks a θ-role, and therefore the interpretation of each other is impossible since it requires the linking of two θ-positions in order to be properly interpreted.

The ungrammaticality of the sentences in (132) follows in a similar fashion. In (132a), himself cannot be Bound by John because of the $\bar{S}$ boundary, so it is Bound by the AGR of the lower sentence, and this is out for the same reasons that (131a) is out. In (132b), each is moved to the nearest accessible nominal, which is the AGR of the embedded clause, and this is ungrammatical for the same reasons that (131b) is ungrammatical.

Compare now the sentences in (137) and (138).

(137) *The men expected that [each other would win]

(138) The men expected that [pictures of each other would be on sale]

The difference between the two sentences is due to the accessibility of AGR. In (137), AGR of the lower sentence is accessible for each movement, and it is therefore the nearest accessible nominal element: so the sentence is out because AGR lacks a REF-index. In (138) however, AGR of the lower sentence

is not accessible for each-movement since each other is embedded in the NP pictures of each other which is itself coindexed with AGR, being the subject of the sentence: so movement of each to AGR of the lower sentence violates the i-within-i well-formedness condition of Chomsky (1981a), and hence renders AGR inaccessible. So the first nominal element accessible to each movement is in the higher clause, and since the men is plural and has a REF-index, the sentence is grammatical.[30]

What this discussion suggests is that there are reasons to assume that R-index is in fact the combination of S-index and REF-index, and that some elements like expletives, adjectives, determiners, AGR, have only an S-index and no REF-index in their R-index, so that they enter into syntactic processes like Binding and Agreement, but they do not bear a θ-role.

Returning to condition (112'), a more important consequence of the fact that condition B as stated in GB cannot be right since it is a morphological condition as we have just shown, is that this means that condition B cannot be used as a kind of definition of the feature [+pronominal] to enter into the determination of the partitioning of the EC, as proposed by Chomsky (1981b). As we have discussed in 2.2.5 above, there are problems in using condition B in this fashion anyhow. But what we now see is that even if these problems are overcome (by modifying the formulation of condition B for example), this analysis cannot work since condition B as a binding condition cannot be right. So the claim that principles A and B identify two cate-

gories of expressions, anaphors and pronominals, and that, "if the binding theory is correct", we would expect to find four categories of expressions as in (88) above does not hold since the discussion above shows that the binding theory is not correct and that conditions A and B cannot be stated as principles of the binding theory. Therefore, what categories of expressions are to be expected in the grammar must be determined by other factors of the grammar. We have proposed that these factors be [±Bound], [±coreferential], [A,$\bar{A}$-Bound] in the discussion of the partitioning of NP in (89)-(91).

We now assume that condition A of the binding theory is derivable from Binding, and that condition B must be replaced by a morphological condition as in (112'): the effect of condition B on ECs is null since an EC is simply determined to be of the anaphor type when it is bound, hence the irrelevance of condition B in this case. We could also assume that an EC is ambiguous in form: it is either anaphoric or pronominal. So it will always be possible to Bind an EC. This is restricting our discussion at the level of the form of the element, whether lexical or not. Note, however, that at the level of discussion where the type of the NP is determined, then it is the definitions in (105)-(107) that are relevant. This means that a Bound pronominal is technically an anaphor by our definition (105). It can have a "pronominal derivation" as we have seen in the preceding discussion, and a pronominal interpretation, as in

the coordinate interpretation in (122c) for example. So for example, the D-structure of (139a) is (139b).

(139) a. Je me trouvais bête, et ma soeur aussi me trouvait bête.

b. Je me_i trouvais e_i bête, et ma soeur aussi me_i trouvait e_i bête

In the first coordinate of (139), me does not have a reflexive derivation but a pronominal derivation, because of the parallelism in structure needed for coordination, as we saw in the discussion of (122)-(123). But in the first coordinate, me is Bound by je (or the $[_{NP}\ e]$ position to which je is related), so that me is technically an anaphor by definition (105). In the second coordinate however, me is not Bound, so it is technically a pronominal by definition (106).

One could object to the conclusion that we reached above about the replacement of the binding theory condition B by a morphological condition that, although some modification of Condition B is necessary in view of the facts presented above, this could be done by keeping condition B as a genuine binding theory condition, and adding an extra stipulation to take care of the French facts presented in (117)-(129). One reason that could be invoked is the fact that condition B does not only account for the binding of pronouns by coindexing, i.e. distinct coreference, but that this condition could also be construed as a condition on disjoint reference which would then

account for the facts in (140).

(140) We saw me.

If condition B is independently needed to account for disjoint reference, then it could still be maintained as a true binding condition. Condition B can account for (140) if the indexing used is more complex than the one so far assumed. But as noted in Chomsky (1981a), a more complex indexing is required anyhow to handle sentences like (141).

(141) John_i told Bill_j that they_k should leave.

In (141), John and Bill are indexed differently, and they cannot be indexed either with John or Bill, in accordance with the general indexing convention in Chomsky (1981a). But here $\underline{k}$ can incorporate either $\underline{i}$ or $\underline{j}$, or neither, so that $\underline{k}=\underline{i}+\underline{j}$, or $\underline{i}+\underline{j}+\underline{x}$, or $\underline{i}+\underline{x}$, or $\underline{j}+\underline{x}$, or $\underline{x}$. If we suppose that an index can in fact be a sum of indices, then we can account for (141). And by the same token, we could account for (140), assuming the indexing in (142) for example.

(142) We_{i+j} saw me_i.

Here, condition B would properly exclude sentence (142) as an instance of disjoint reference. But notice however that an account of disjoint reference by condition B runs into problems. The first problem is that there seems to be a difference in the acceptibility which depends on the order of the pronouns involved. Thus, Morin (1979) noted the following contrast in dis-

joint reference in French when some expressions of opinion or perception are involved.

(143) a. Je nous trouve pas mal intelligents.
b. *Nous me trouvons pas mal intelligent.

The second problem is that this contrast is found even across sentence boundaries, i.e. governing category boundaries, as in (144), contrary to what condition B would predict.

(144) a. Tu penses que nous réussirons?
b. *Nous pensons que tu réussiras.

Here, (144b) is acceptable only if nous is interpreted as not including tu, whereas the inclusive interpretation of tu in nous is perfectly natural in (144a), where the order is reversed. Ken Hale (personal communication) informs us that this contrast in the order of the two pronouns is found quite generally in languages. The facts in (144) are particularly clear in languages that have two forms for the first person plural pronoun, depending on whether it is inclusive or exclusive. Thus, one could not appeal to some analogical interpretation with the exclusive pronominal form to account for the acceptability of (144a) since when there is a clear form of inclusive pronoun, the sentence is still good. Furthermore, such an analysis by analogy could not explain why such an analysis is not also made in (144b).

Disjoint reference is also found in third person NPs: thus there seems to be a contrast between (145) and (146), where the

underlined singular NPs are to be interpreted as being part of the reference of the underlined plural NPs.

(145) ??They think that the dean has a grudge against John/him.

(146) He/John thinks that the dean has a grudge against them.

Disjoint reference as illustrated here has been dealt with in the literature as a phenomenon closely related to non-pronominal coreference. Thus from Chomsky's (1973) observations about Langacker's conditions to Lasnik's non-coreference condition, the type of disjoint reference illustrated above is considered as a subcase of a condition which can be stated approximately as the following, which is taken from Evans (1980).

(147) A term can be referentially dependent upon an NP if and only if it does not precede and c-command that NP.

But this condition is related to condition C of the binding theory, rather than to condition B. In fact, in Chomsky (1981b), it is proposed to virtually replace condition C by Evans' condition given in (147), and Strong Cross-Over facts receive an explanation that does not rely on condition C anymore. We return to this topic shortly. Note that since the facts discussed above hold across sentence boundaries, that also suggests that condition B is not relevant, but rather condition C.

How does condition C relate to the disjoint reference facts discussed above? First, it must be assumed that the indexing convention is slightly enriched, which is independently needed

for sentences like (141), as we have seen. So for example, consider the sentences in (144), with an enriched indexing as in (148).

(148) a. Tu_i penses que $nous_{i+x}$ réussirons?
b. *$Nous_{i+x}$ pensons que tu_i réussiras.

Suppose that we interpret an inclusive nous as being referentially dependent upon tu in cases like (148), since the reference of tu is needed to determine the reference of nous in such cases. Then (148b) violates the condition in (147). On the other hand, tu is not referentially dependent upon nous, since the reference of nous is not needed to determine the reference of tu: thus (148a) is grammatical since it meets the requirements of condition (147). The same explanation holds for the other cases where there is a contrast in the order of the included referent and the including referent that we saw above.

Returning to the possibility of keeping condition B and explaining the facts that we explained by condition (112') by some other stipulation, we see that this proposal is unwarranted since there is no independent motivation for condition B: the phenomenon of disjoint reference has always been dealt with in the literature by conditions that are akin to condition C of the binding theory rather than condition B. There are also empirical reasons to believe that condition B is not responsible for disjoint reference effects since these effects are found across sentential boundaries, i.e. governing category bound-

aries in some cases. On the other hand, a condition like the one in (147) seems to provide an explanation for the contrast of ordering of NPs that show disjoint reference effects.[31]

Returning to the general discussion on binding theory, we also assume that Binding applies when it can; condition (112') applies once Binding has applied since it is a condition on elements that are Bound: it could therefore also apply when it can, or it could apply once all Binding relations have applied. This means that conditions A and B are derived from processes that do not necessarily apply at S-structure. It is interesting in this respect to note that the arguments given in the literature to support the claim that the binding theory must apply at S-structure always involve condition C, not conditions A and B (cf. Chomsky 1981a).

The common argument for applying the binding theory at S-structure has to do with the contrast between the sentences in (149).

(149) a. $[[\text{Which book}]_i$ that $\underline{\text{John}}_k$ read $\underline{t}_i]_j$ did $\underline{\text{he}}_k$ like $\underline{t}_j$?

b. *$\underline{\text{He}}_i$ likes [every book that $\underline{\text{John}}_i$ read].

If condition C did not apply at S-structure but only at LF, then sentence (149b) should be grammatical for the same reason that (149a) is grammatical: after QR of the NP every book that John likes, the pronoun he would not c-command the name John, and so they could be coreferential, as we can see in the LF structure of (149b) given in (150).

(150) $[_S$[Every book that John read]$_j$ $[_S$ he likes t_j]]

But since (149b) is ungrammatical, condition C has to apply before LF. Moreover, since (149a) is grammatical, this means that condition C must apply after WH-movement, or else the sentence would be ruled out since the structure before WH-movement violates condition C. We see this in (151), where WH-movement is blocked since the COMP is already filled.

(151) *Who said that he liked which book that John read?

The sentence is ungrammatical when he and John are taken to be coreferential. So since condition C must apply before LF, but after WH-movement, it must apply at S-structure.

It seems that, although conditions A and B are derivable from other components of the grammar, condition C is not. Let us therefore keep such a condition on names.

(152) The Binding Condition:
A name cannot be c-commanded by a coreferential element at S-structure.

This condition is essentially the condition on coreference of Evans (1980) which states that "a term can be referentially dependent upon an NP iff it does not precede and c-command that NP".

Note that the Binding condition (152) applies only for names, not for variables. This means that the facts about cross-over will have to be accounted for by other principles of the grammar. This is a good consequence since there are empirical

facts that show that strong cross-over holds in some languages where the binding condition (152) does not hold. Furthermore, cross-over cannot be a condition on variables at S-structure because it is found in languages that do not have variables at S-structure, i.e. languages that do not have overt WH-movement. So the two phenomena cannot be collapsed together. The empirical evidence is given in Mohanan (1981). First, Mohanan shows that in Malayalam, the condition on non-pronominal coreference is that pronouns may not precede their antecedents, regardless of the structure involved, i.e. regardless of whether the pronoun c-commands its antecedent or not, as we see in (153).[32]

(153) a. moohan [awante bhaaryaye] nuḷḷi
Mohan-n his wife-a pinched
(Mohan pinched his wife.) (a=accusative; n=nominative)

b. *awan [moohante bhaaryaye] nuḷḷi

(154) a. [moohante bhaaryaye] awan nuḷḷi

b. *[awante bhaaryaye] moohan nuḷḷi

And yet cross-over holds in Malayalam, regardless of the order of the variable and the pronoun, and regardless of the fact that Malayalam does not have overt WH-movement.

(155) a. aar parannu [meeri awane umma weccu ennə]
who-n said Mary-n him kiss gave that
(Who said Mary kissed him?)

b. aar parannu [awan meeriye umma weccu ennə]
who-n said he Mary-a kiss gave that
(Who said he kissed Mary?)

c. *[meeri aare umma weccu enna] awan parannu?
(Who did he say Mary kissed?)

Coindexing of the pronoun and the variable is also impossible even if neither the pronoun nor the variable c-commands the other, as when both are possessive.

(156) a. *[aarute bhaaryaye] [awante amma] nuḷḷi
whose wife-a his mother-n pinched
(Whose wife did his mother pinch?)

b. *[aarute bhaarya] [awante ammaye] nuḷḷi
whose wife-n his mother-a pinched
(Whose wife pinched his mother?)

Following Reinhart (1976), Mohanan's suggestion to account for strong cross-over is the condition in (157).

(157) Quantified antecedents must c-command pronouns.

Mohanan assumes that (157) applies at NP-structure, i.e. the level at which NP-movement has taken place but not WH-movement (this level is postulated in Riemsdijk and Williams 1980). This is to unify the account of quantified NPs and WH operators with respect to cross-over.

(158) He_i left a pamphlet near $everyone_j$ $(i \neq j)$

(159) a. Who_i did he_j say t_i likes Bill?
b. he_j said [who_i likes Bill]

In (158), the quantified NP everyone does not c-command he at NP-structure, so he cannot be bound by everyone. Similarly, in (159b), which is the NP structure (159a), who does not c-

command he, so who cannot bind he.

Following Haik (1982), we could extend the domain of the condition so that it applies to non-referential NPs (see fn. 3) as in (160).

(160) Condition on BOUND pronouns.[33]
Non-referential NP_i c-commands $pronoun_i$.

If we assume that the variable of the quantified NP is a non-referential NP, which is quite natural, then (160) allows us to account for WH-cross-over, Q-cross-over, and binding of pronouns by other non-referential NPs in a unified manner. It also allows the condition to be stated at a level other than NP structure, which takes some motivation out of the postulation of such a level. Haik (1982) assumes that (160) applies at S-structure. This means that in a sentence like (159), Haik considers that t is the BINDER of he, not who. But in the case of a WH in situ, or in languages that do not have overt WH-movement, Haik would have to assume that it is the WH-phrase itself that is the BINDER. In order to eliminate this disparity, we could assume that (160) applies at LF, and that it is the variable that is the non-referential NP BINDER in all instances of WH-constructions. Such an assumption has to be made in any case if one wants to account for cross-over in constructions where QR applies like Focus constructions.

(161) a. His mother loves JOHN.
b. $[_S$ JOHN $[_S$ his mother loves t $]]$

In (161b), the variable t does not c-command his, so it cannot BIND it. Having (160) apply at LF also accounts for cases like the following:

(162) a. *His wife likes many friends of each writer.
b. $[_S$[each writer]$_x$ $[_S$[many friends of x]$_y$ $[_S$ his wife likes y]]]

After QR has applied to derive the LF structure (162b), we see that the variable x does not c-command his, although the quantified NP each writer does. So if (160) applies at LF and if the non-referential NP BINDER is taken to be the variable, not the quantified phrase itself, then we have an explanation of the facts.[34]

There is another recent proposal to account for cross-over facts by an LF principle which is made in Koopman and Sportiche (1981) (henceforth KS). KS assume the Bijection Principle, which applies at LF.

(163) Bijection Principle (KS)
Every A-position is locally bound at most by one $\bar{A}$-position.
Every $\bar{A}$-position locally binds at most one A-position.

Their definition of local binding is the following:

(164) α locally X-binds β iff α c-commands β, and α and β are coindexed, and if γ has the same properties as α, γ c-commands α. (where X = A or $\bar{A}$)
(c-command = α c-commands β iff the first maximal projection which dominates α, also dominates β.)

So the sentences in (165) are all ruled out by the Bijection principle since two A-positions are $\bar{A}$-bound by the Q-phrase.

(165) a. *Who does his mother love t?
b. *His mother loves everyone.
b'. [$_S$ everyone [$_S$ his mother loves t]]

However, the reason why grammatical sentences like those in (166) are not ruled out by the Bijection principle is not as straightforward.

(166) a. Who [t likes his mother]?
b. Everyone likes his mother
c. [$_S$ everyone [$_S$ t likes his mother]]

In the sentences in (166), the Q-phrase $\bar{A}$-binds two A-positions at LF according to KS's definition given in (164). So the sentences should be ungrammatical with the readings where his is coindexed with the Q-phrase. But this is not the case. KS cannot appeal to the fact that his is bound by the variable and that it is the variable that locally binds his, not the Q-phrase; the reason why they cannot is because the binding which is relevant for the Bijection principle is $\bar{A}$-binding, not A-binding, so that both the variable and his are locally $\bar{A}$-bound by the same operator in (165). So some stipulation has to be made in addition to the Bijection principle to the effect that a more local A-binding "neutralizes" a less local $\bar{A}$-binding. It is not clear what the status and the effects of such a stipulation would be in the grammar.

Even assuming that such a stipulation is tenable, the Bijection principle is still violated in all cases where an operator binds two positions which are variables if the second one is not A-bound by the first one since then the two positions are locally $\bar{A}$-bound by the same operator. We know three such cases. The first one is parasitic gaps as in (167), where which articles locally $\bar{A}$-binds $\underline{t}_1$ and $\underline{t}_2$.

(167) Which articles did John file $\underline{t}_1$ without reading $\underline{t}_2$?

Chomsky (1981b) acknowledges that one must assume that the Bijection principle is bypassed in parasitic gap constructions. The second case is when there is movement from coordinated phrases as in (168).

(168) Who did John see and Bill hit:

Here again, an operator locally $\bar{A}$-binds two variables. It might be possible to circumvent this problem for the Bijection principle by using the across-the-board rule application of Williams (1978) where the two traces in (168) would be only one "factor" in the sense given there. This would ask for modifications of the Bijection principle which might not be impossible.

There is a third kind of violation of the Bijection principle which is found in sentences like (169) in French (from Damourette & Pichon 1911).

(169) La Russie dont [le bolchevisme $\underline{t}$] nuit à [la civilisation $\underline{t}$]

In (169), the relative operator dont binds both traces.

Furthermore, the Bijection principle analysis is weakened by the fact that it requires another stipulation on BOUND pronouns that we give in (170).

(170) ((23) of KS)
A pronoun may "corefer" with a variable bound by a (quasi-) quantified expression (i.e. WH-phrase, quantifiers subject to QR) only if it is in the scope of (i.e. c-commanded by) the (quasi-) quantifier at LF.

Consider the c-command condition of Reinhart (1976) in (171), which states a necessary and sufficient condition for a pronoun to be understood as coreferent with a non-definite NP.

(171) The S-structure c-command condition (Reinhart 1976)
A non-definite NP (trace, QP, focus) can be coindexed with a pronoun iff the latter is c-commanded by the former at S-structure.

In their discussion of (171), KS note that "both the Bijection principle and the CC approaches agree that (164) is a sufficient condition. Although it is formulated at LF instead of at S-structure, this is essentially the content of ..." (170) (KS p. 13). They point out that the Bijection principle approach predicts that (171) is not a necessary condition. However, they say that they were unable to find a clear configuration of data that would illustrate the difference.

Note also that the similarity of the content between (170) and (171) is not as strong as KS lead us to believe. Consider the part of the conditional that the two statements are claimed

to share according to KS.

(172) a. CC: pronoun coindexed with variable→ Q c-commands pronoun at S-structure.

b. BP: pronoun coindexed with variable→ Q c-commands pronoun at LF.

These two conditions have quite different implications in the grammar. For example, since QR moves the quantified phrase in a position that c-commands the variable, this means that the LF position of the Q-phrase will always c-command the S-structure position of the Q-phrase as in (173).

(173) S-S: $[_S$ ... QP ...]

LF: $[_S$ $QP_i[_S$ ... $\underline{t}_i$ ...]]

(If the variable is a WH-trace, it will already be c-commanded by its binder, as assumed in GB for similar reasons.)

Because of the transitivity of c-command, this means that if a Q-phrase c-commands a pronoun at S-structure, the Q-phrase will c-command both its trace and this pronoun at LF, as in (174).

(174) S-S: $[_S$ ... QP ... pronoun ...]

LF: $[_S$ QP$[_S$ ... $\underline{t}$ ... pronoun ...]]

This has some bearing on (172). It means that "Q c-commands pronoun at S-structure" implies that "Q c-commands pronoun at LF", but the converse certainly is not true. This means that there can be cases where (172b) would hold, but where (172a)

would not hold. Sentence (162), which we repeat here, is precisely such a case.

(162) a. *His wife likes many friends of each writer
b. $[_S$[each writer]$_x$ $[_S$ many friends of x]$_y$ $[_S$ his wife likes y]]]

Condition (170) of KS allows (162), whereas condition (171) of Reinhart blocks it. In order to rule out (162), KS introduce the Bijection principle. But we have seen that the Bijection principle is too strong since it rules out grammatical sentences like (166). So we conclude that a condition like Reinhart's (1976) c-command condition is to be preferred to the Bijection principle of KS. As noted in the discussion of (160), we could assume that such a condition applies at LF, and that it is the variable that is the binder in all these cases.

Returning now to the Binding condition (152) on names and cross-over effects, we note that in Chomsky (1981b), it is also recognized that strong cross-over and non-coreference facts covered by condition C cannot be collapsed. Chomsky (1981b) proposes to derive the strong cross-over effects from violations of conditions A and B at LF. Consider the sentence in (175).

(175) *Who does he think [t'[t likes Bill]]?

According to Chomsky's (1981b) analysis, if he and t are coindexed in (175), then t is an EC locally A-bound by an element in a θ-position and is therefore PRO, according to the functional definition of PRO in Chomsky (1981b). But the sentence is

ungrammatical with such a reading since, according to Chomsky's analysis, principles A and B are violated since PRO is governed by INFL. But this analysis is based on the assumption that conditions A and B are true binding conditions and not derived from other principles in the grammar. It is also crucial for this analysis that condition B applies to ECs, i.e. to PRO in this case.

But we have seen that condition B of the binding theory must be derived from the morphological condition (112'), and furthermore, that only a weak inference about the government of PRO can be made from the binding theory even if one considers PRO to be a pronominal anaphor, since the reference to government has no independent motivation in the binding theory.[35] On the other hand, the condition on BOUND pronouns (160) will account for the ungrammaticality of (175) since the variable t does not c-command he at LF.

Note that condition (160) will account in a uniform way for strong cross-over as in (175) above, and for weak cross-over as in (158).

(176) a. *Who does his mother love t?
b. *His mother loves everyone.
b'. [$_S$ everyone [$_S$ his mother loves t]]
c. *His mother loved JOHN
c'. [$_S$ JOHN [$_S$ his mother loved t]]

In all of these cases, the variable does not c-command the pronoun at LF. Chomsky (1981b) appeals to the Bijection prin-

ciple to rule out these sentences, but it also blocks some grammatical sentences.

One case which is problematic for the Bijection principle, for condition (160) as it is stated and for Reinhart's condition (171) is the case of relative clauses, since weak cross-over violations are possible in relative clauses. Thus compare the ungrammatical WH question in (176a) with the grammatical relative clause in (177).

(177) The man <u>who</u> <u>his</u> mother loved <u>t</u> best

Chomsky (1981b) proposes to account for these facts in the following way. First, he assumes that the Bijection principle applies only at LF; he then proposes that relative clauses are interpreted in terms of predication, which is done at LF'. So the LF representation of (177) is (178).

(178) The man_i who_j his_i mother loved $\underline{t}_j$ best

The rule of predication, applying to the LF representation (178), maps it into (179) at LF'.

(179) The man_i who_i his_i mother loved $\underline{t}_i$ best

Since it is assumed that the Bijection principle applies only at LF, not at LF', there is no weak cross-over effect.

Consider now how the present analysis could account for weak cross-over in relative clauses. What (177) shows is that the condition on BOUND pronouns as stated in (171) is too strong: it should not be an <u>if and only if</u> biconditional, but only an

if conditional, as in (180).

(180) Condition on BOUND pronouns
Non-coreferential NP_i can be coindexed with $pronoun_i$ if NP_i c-commands $pronoun_i$ at LF.

In (180), the c-command relation is not necessarily required for coindexing to be possible: if c-command holds, then coindexing is possible, but if the c-command relation does not hold, it is possible for NP_i and $pronoun_i$ to be coindexed if the two elements are coindexed with a third element, for example, so that an indirect relation is established between NP_i and $pronoun_i$. This is what takes place in a relative clause. For example, in (177), the head of the relative clause the man is coindexed with the variable (via the WH-operator), but it could also be coindexed with a pronoun in the sentence like his. So there is no weak cross-over effect in relative clauses because condition (180) is not violated if the head is the third element binding NP_i and $pronoun_i$.[36]

This approach seems preferable to the one that makes use of the Bijection principle for two reasons. The first reason is that there are problems with the Bijection principle, as we have seen above.

The second reason is that there is no need to create a new level of LF' in the present analysis, just like there was no need of having the level of NP-structure. There is no a priori reason why one could not postulate an extra level if it is motivated on empirical grounds. But if an analysis can be provi-

ded without postulating an extra level (but without adding extra heavy machinery to the grammar in order to avoid postulating that level, since this would only be a trade-off), then this simpler analysis is to be preferred. Of course, if that extra level is independently motivated, then it does not add to the overall cost of the grammar to make use of it. Chomsky (1981b) provides some facts about left dislocation as an argument for the postulation of the level LF'. Thus consider the left dislocation structure in (181).

(181) John_i, he_j likes him_k.

In Chomsky's analysis, the predication rule will identify John with some pronoun in the sentence. Suppose that it identifies i and k; then the open sentence in (181) is (182), which is predicated of John.

(182) he likes --

Then at LF' we have (183).

(183) John_i, he_j likes him_i

But suppose that $i = j$ at LF, say by fortuitous indexing of the pronoun. Then at LF', according to Chomsky's analysis, sentence (181) is interpreted as in (184).

(184) for x = John, x likes x

But (184) is not a possible interpretation of (181), even though (181) satisfies all conditions that hold at D- or S-struc-

ture or at LF. Chomsky's conclusion is that it must be the binding theory that rules out (183) where $\underline{i} = \underline{j}$. Therefore, the binding theory must hold at LF'.

But an account of these facts can be given without resorting to an extra level of representation, and without saying that the binding theory applies at S-structure, LF and LF', which seems unlikely if conditions A and B are derivable as we have seen above. Suppose that predication takes place at LF, and that it is not the index of the pronouns that is affected by predication, but the index of John which must match one of the indices of the pronouns. Then the fact that the two pronouns in (181) must bear distinct indices will remain constant all along since these indices remain constant, whatever the mechanism used to block coindexing of the two pronouns in this context. This blocking mechanism is already needed independently at least at S-structure anyhow. So in this analysis, the impossibility of the reading in (184) for (181) is accounted for by the usual condition that blocks coreference of the two pronouns in (181) even when predication has not taken place.

What this whole discussion of pronouns BOUND by non-referential NPs suggests is that an analysis of cross-over effects along the lines suggested in Reinhart (1976), possibly involving a condition on BOUND pronouns as in (180), seems preferable to one which makes use of the binding conditions, especially if conditions A and B are derivable from Binding and from a morphological condition. Furthermore, since condition (180)

accounts for both strong and weak cross-over in a unified fashion, it also eliminates the need of the Bijection principle. And recall that the Bijection principle analysis had to be supplemented with the condition (170), which is close enough to condition (180) to suggest that the effects of (170) and of the Bijection principle should be subsumed under (180).

2.4 Summary

In this chapter, we have seen that NPs are governed by general principles like the principle of Denotability (85), which essentially requires that arguments have an R-index to get a θ-role, and Agreement, which governs coindexed NPs. Lexical NPs fall into three main classes: R-expressions, which denote objects in domain D; pronouns, which are coreferential with R-expressions; and anaphors, which are related to an antecedent by Binding (78). Anaphors should not be defined on a strictly morphological basis, but rather on the basis of the type of relation that holds between them and their antecedent. Thus there are reflexives which are true anaphors, which are obligatorily Bound by a unique antecedent, and there are reflexives which are not anaphors and which are freely indexed at S-structure like pronouns are, and which do not have an obligatory and unique antecedent. True anaphors occur only in the contexts described in (75), indicating that government is a crucial element in the definition of Binding. As for the distribution of false anaphors and pronouns, it is governed by pragmatic factors, some

of which are described in Zribi-Hertz (1981). The present analysis predicts where true and false anaphors will appear; it also predicts that pronouns will never appear in Bound positions, unless they are allowed by the morphological condition (112'), i.e. when the pronoun has an ambiguous form which is the same as a reflexive form. So technically, an anaphor is a Bound element and a pronominal is an element freely indexed at S-structure.

Turning now to ECs, the complementary approach to ECs, which is expected in a "Y-shaped grammar" assumes that ECs have some features at LF but none at PF. This allows a simple account of the distribution of lexical and non-lexical NPs by the principle of Lexicalization which operates at PF. The complementary approach to ECs has for important consequence that it allows us to eliminate from the grammar the four statements that referred specifically to ECs; it does so by deriving them from principles and rules of the grammar independently motivated for the analysis of lexical NPs. The partitioning of the EC (89) was shown to be parallel to the partitioning of the lexical NP (91), and the different types of NPs are functionally determined by the interaction of principles of the grammar. So this analysis of ECs is in line with the hypothesis presented at the outset that no principles nor rules should refer specifically to ECs, but only to categories like N or NP.

The analysis also has consequences on the binding theory.

Condition A was shown to be derivable from Binding (78). Condition B was shown not to be an actual condition relevant to the binding theory since pronouns can be Bound in their binding category: they are then technically equivalent to anaphors as far as the relation with the antecedent is concerned. There is however the morphological condition (112') on the form of the Bound elements which governs the distribution of pronouns Bound in the sense of (78). As in Chomsky (1981b), it is assumed that pronominal non-coreference and cross-over effects cannot be collapsed under condition C. Pronominal non-coreference is governed by the Binding condition (152), which applies at S-structure, and cross-over effects are due to the condition (180) on BOUND pronouns, which applies at LF.

Now that we have presented the main components of our grammar and some ways in which they interact, we will turn to the details of the analysis with respect to the data covered by ECP in Chapter 3, the Pro Drop condition in Chapter 4, and control theory in Chapter 5.

FOOTNOTES: CHAPTER 2

1. Borer (1981) for example suggests that the index is provided by the feature [αperson] which is given by AGR.

2. A similar observation about deictic pronouns and grammatical features was also made independently in Tasmowsky-De Ryck & Verluyten (1981).

3. We have not been very specific about what the notion R-index is meant to cover. What we mean basically is to use R-indices to link NPs to one another in some sense, but it must be noted that this does not necessarily imply that the two NPs are coreferential. The reason for this is because some NPs are not referential. For example, Haik (1982) notes three types of non-referential NPs, that is, NPs that do not denote single individuals although they are marked [+sing] (ignoring the cases of inherent plural nouns like trousers, scissors). The first type is inherent quantifiers like in (i).

(i) a. *Everyone came and he had a good time.
b. Everyone said that he had a good time.

Sentence (ia) is ungrammtical because he cannot be coreferential with everyone since the variable of everyone is not referential. In (ib), the pronoun is in the scope of the quantifier (i.e. c-commanded by it) so that it can be interpreted as a bound pronoun. Such a pronoun is not coreferential with the quantifier, but rather the pronoun is assigned the value which is assigned to its antecedent at LF.

Another case of non-referential NP is an indefinite NP in the scope of a quantified NP.

(ii) Two men love a woman. She does not look nice.

(iii) Two men saw a woman when she smiled.

In (ii), a woman can only have wide scope over two men if it is coreferential with she. This is because if a woman is in the scope of two men, then it does not refer to a single individual: the NP a woman is given values which are dependent on the values of the objects denoted by two men (there could be two women in fact here). So a woman with narrow scope in (ii) is not referential, hence it cannot be coreferential with she. In (iii) on the other hand, she is in the scope of a woman even if a woman has narrow scope, so she can be interpreted as a bound pronoun: its value will be the value which is assigned to a woman at LF.

The third case of non-referential NPs noted by Haik is descriptions containing a bound pronoun.

(iv) +Everyone likes his wife, but I don't like her.

(v) Everyone thinks his wife likes her daughter.

Here, his wife cannot be coindexed with her if his wife has narrow scope, unless it c-commands her as in (v). The explanation is as in the case above: since his wife is in the scope of everyone, it does not refer to a single wife, and hence cannot be coreferential with an NP. See Haik (1982) for a detailed analysis of this type of phenomenon.

To repeat, our notion of R-index is meant to cover instances of coreference between NPs, but also instances of binding where no coreference is involved as in the cases above. So the notion of R-index is a strictly syntactic notion, which is something like the operative notion equivalent to the core notion of referential index which has to do with actual reference in domain D.

Note finally that, in order to maintain the strong version of the principle of Denotability with an <u>if and only if</u> conditional, we must say something about dummy elements like <u>it</u> and about NPs in idioms, since both of these have an R-index and F-features but do not denote an object in domain D. In the case of idioms, we could say that the restrictive rule that interprets idioms takes precedence over the general "Elsewhere Rule" which interprets non-idiom NPs, in the spirit of a Pāninian principle of rule application (cf. Kiparsky 1973). As for dummy elements like <u>it</u>, we could say that their lexical entry acts as a very restricted rule of interpretation in the sense that it specifies that dummy <u>it</u> cannot bear a θ-role for example. See 2.3 for some additional discussion.

4. It should be clear that the three F-features that we assume to be part of such lexical entries have to be considered as very broad notionswhen one considers them as part of UG. For example, <u>gender</u> is meant to cover the familiar notion of gender in Indo-European languages, but it is also meant to encompass other notions used in languages to partition the set of individuals

in domain D, like noun classes in African languages for example. Although we will continue to use notions like gender in the text since this term is sufficient for our purposes and is the one used in the languages that we will study, the reader is asked to bear in mind the above caveat.

5. Or to conditions on bound pronouns like c-command when the pronoun is not referential. See Haik (1982) and footnote 3.

6. We omit details and other conditions which are not relevant to the case at hand. Prominence in condition (i) is determined by a hierarchy of grammatical functions similar to the one proposed by Keenan & Comrie (1977). Condition (ii) is essentially equivalent to the i-within-i well-formedness condition of Chomsky (1981a); it is proposed by Koster to account for the contrast between (i) and (ii).

(i) John says that it is a picture of himself.

(ii) *John says that Mary bought a picture of himself.

7. This generalization also holds of reciprocals. So for example, Chomsky (1981a) has the following comment:

> In general, reciprocals often seem marginal in prepositional phrases, except under reanalysis, though not always. Compare the marginal examples (i) with the more acceptable (ii) (under reanalysis) and (iii).
>
> (i) a. They left with each other.
> b. They are easy for each other to talk to.
>
> (ii) a. They spoke to each other.
> b. They are sorry for each other.
>
> (iii) a. They told Mary about each other.
> b. They told Mary stories about each other.
>
> (Chomsky 1981a, p. 316)

Chomsky (1981a) also gives one example of variation across languages for each other: the Dutch equivalent of (iv) is grammatical.

(iv) *They forced me to read each other's books

So although each other is clearly anaphoric in meaning, not just in form, it seems to be allowed in some "non-core anaphor structures"; but when each other is in such a structure, the construction has the properties of marked constructions. It could be that each other is used in such PP constructions in analogy with reflexive forms used pronominally. Furthermore, the fact, to which we will return shortly, that structures where reanalysis has applied allow anaphors freely in English can also contribute to this analogical use by some speakers in non-core structures.

8. In French, cliticization of a reflexive is obligatory if the reflexive relation is established with an object or a dative complement, as we see in (i) and (ii).

(i) a. Jean se rase.
b. *Jean rase lui-même.

(ii) a. Jean s'est donné un coup de marteau sur le pouce.
b. *Jean a donné un coup de marteau à lui-même sur le pouce.

We will assume an analysis of reflexivization along the lines suggested by Grimshaw (1980) and Marantz (1981): the French reflexive construction with se is a derived verb form, and not the combination of a verb and a reflexive object clitic (whether moved or base-generated: see Kayne 1975; Burzio 1981).

We will assume with Marantz (1981) that reflexivization, being a lexical rule in French, cannot bind the logical object without affecting the predicate argument structure of the verb: since reflexivization clearly does not change the predicate structure of the verb, it must bind the predicate external argument, namely the subject. In terms of GB, this could be expressed in the following way: se absorbs the θ-role of the subject. At the same time, se absorbs a Case feature assigned by the verb to a complement. This seems to be a generalization that is recurrent in the grammar and which has consequences that have been extensively studied in Burzio (1981). Burzio states the generalization as in (iii).

(iii) T↔A where T = "assignment of a θ-role to the subject."
A = "accusative Case assignment"

So (ia) would be derived in our analysis from the D-structure (iv) by move α, with a S-structure as in (v).

(iv) $[_{NP}$ e] se rase Jean

(v) $Jean_i$ se rase t_i

This predicts that in constructions where some device can assign Case to the object, like il insertion for example, then the post-V NP is interpreted as the theme, and not as the agent.

(vi) Il se rase beaucoup de gens d'une seule main dans cet établissement.

As noted by J. Morin (1978), adverbial phrases like d'une seule main can only modify the agent of a proposition. So in (vi), d'une seule main does not relate to beaucoup de gens, but to

some unspecified agent, as predicted by our analysis of reflexivization.

This analysis of French reflexivization is also compatible with Burzio's (1981) account of the distribution of auxiliary and past participle agreement. Burzio proposes the system in (vii).

(vii) a. Essere (être) assignment:
The auxiliary will be realized as essere (être) when a binding relation exists between the subject and a nominal constituent of the predicate.

b. Past participle agreement:
A past participle will agree (in gender and number) with an element binding its direct object.

The notion nominal constituent of the predicate is required in Burzio's account to cover the reflexive si (or se) which is part of the morphology of the verb. This is not necessary in our analysis where the S-structure subject comes from a D-structure complement of the V.

Note that the fact that a dative reflexive triggers past participle agreement as in (viii) further supports the claim that these are nominal at a certain level, since PPs do not have F-features to trigger agreement in French.

(viii) Elle s'était imaginée qu'elle allait connaître ...
(M. Proust)

Grammarians condemn this agreement, but as noted by Grévisse (1964), this can be observed in "popular" use and also in the text of excellent authors, the example above being such a case.

Burzio (1981) also notes this possibility of agreement in Italian.

(ix) i ragazzi si erano parlati
(The kids to each other had talked (pl.))

Such an analysis of se where the reflexive morphology absorbs a Case feature assigned by the V and the θ-role assigned to the subject explains why se is not compatible with derived subjects, i.e. cases where verbs do not assign Case to an object or a θ-role to their subject, and also why non-dative PP complements are not subject to reflexivization.

(x) a. *Jean s'est frappé par Paul. (passive)
b. *Jean se semble être heureux. (raising)
c. *Jean s'est tombé. (ergative V)
d. *Jean se parle à Marie. (de object)

9. See the appendix to Chapter 3 on what this factor that forces reanalysis could be. Recall that there are two main trends to the analysis of stranded prepositions in the literature: some claim that prepositions stranded by WH movement and by passive are all subject to the same rule of reanalysis (cf. Hornstein & Weinberg 1981) although H&W postulate that stranding in passives is subject to an additional predication condition; others claim that there should not be a uniform account of the facts and that two separate mechanisms allow stranding in passive and in WH constructions (cf. Riemsdijk 1978; Rothstein 1981). In the latter approach, constructions in which passive can strand a preposition are a subset of constructions where

WH movement can strand a preposition. This is illustrated in (i).

(i) a. That is the knife that Pierrot cuts the cheese with.
b. *That knife is being cut with.

To account for this, Riemsdijk proposes that reanalysis takes place where passive and WH movement are possible, so that the PP does not count as a bounding node after reanalysis; in the additional cases where only WH movement is possible, he says that there is a COMP node in the PP from which the WH is extracted, thus providing an escape hatch. Rothstein accounts for the difference by assuming that passivization is a lexical rule, and that it must be fed by reanalysis, which she also considers to be a lexical process; as for the WH cases, she assumes that PP is not a bounding node, and that the marked cases are those where extraction is not possible, some independent factors accounting for these cases. What is common to these two analyses is that reanalysis has taken place for passive, and something else is to account for the extra cases of PP stranding by WH extraction. The facts about reflexive and pronominal forms support this double analysis over the uniform analysis of Hornstein & Weinberg (1980). Thus in the case where reanalysis would be postulated by Riemsdijk and by Rothstein, the reflexive form is obligatory.

(ii) a. <u>They</u> spoke to <u>themselves</u>/*<u>them</u>.
b. Mary was spoken to.

But in the case of prepositions stranded by WH movement, but not by passive, i.e. where reanalysis has not applied ac-

cording to Riemsdijk and to Rothstein, then both forms are possible.

(iii) a. Who did John hear stories about?
b. *Mary was heard stories about.
c. John heard stories about him/himself.

This suggests that constructions like (iii) are not core constructions, whereas those in (ii) are, hence that there is a structural difference between the two. Cf. Chapter 3 for discussion of reanalysis.

10. There is one thing that makes "false reflexives" look like anaphors, however: they are bound in most cases. But recall that we have seen in footnote 4 that there is an independently motivated class of bound pronouns, namely non-referential pronouns which are bound by 1° inherent quantifiers, 2° indefinite NPs in the scope of quantified NPs, 3° descriptions containing a bound pronoun. These non-referential bound pronouns have to be c-commanded by their antecedent (either directly or indirectly; see Haik 1982 for discussion). We could consider "false anaphors" as the referential equivalents of the non-referential bound pronouns. This would explain why they are bound by a c-commanding antecedent in most cases. Being referential however, they could be pragmatically bound (i.e. deictic) in some cases, as we have seen in (60).

11. There is one exception to this generalization: in causative constructions, a clitic which is the dative object of the

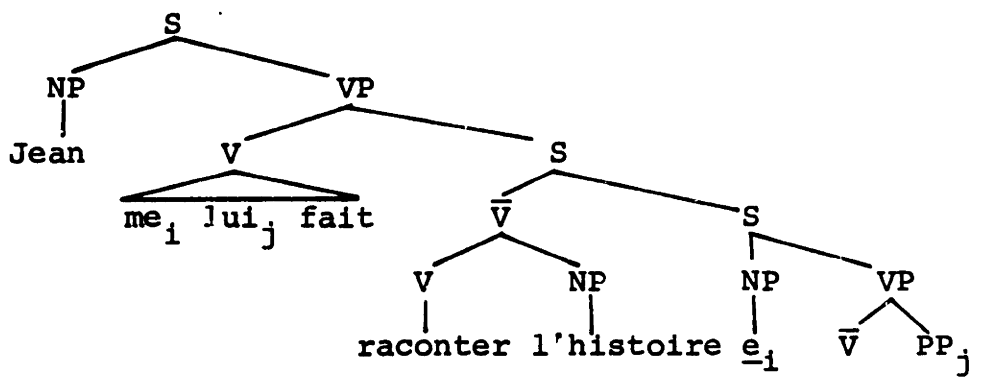
embedded V can be coreferential with the subject of faire, even if it is cliticized to faire, as noted by Kayne (1975), Morin (1979):

(i) Jean me lui fait raconter l'histoire.

(ii) Gianni mi gli la ha fatta raccontare.

L. Rizzi notes (personal communication) that this can be explained if the structure for a sentence like (i) is (iii).

(iii)



If the governing category of a chain (clitic, e) is defined on the e rather than on the clitic, then the governing category of lui in (iii) is the lower S according to the definition of governing category in Chomsky (1981a). So the chain (lui, e) can be bound by Jean which is in the higher clause and the clitic is still free in its governing category, in accord with condition B of the binding theory.

12. Traditional grammarians note that in general, reciprocals are not so good in PPs (cf. Grévisse 1964; Sandfeld 1943). It is also interesting to note that in constructions like (69a), autres is in the PP, while the structural position of l'un is unclear since it precedes the preposition. It might be that

French PPs have a COMP position (cf. Riemsdijk 1978), although preposition stranding is not possible in French. More likely, l'un is in the specifier position of the PP, since it seems to appear in the specifier position of an NP also as in (i).

(i) J'irais dans l'un ou l'autre camp.

Whatever the position of l'un in cases like (69a), it is a problem for Case theory, if indeed l'un ... l'autre must be Case marked. We leave this problem as a topic for further investigation. Note also that the form each other is a late form in English and that it used to be closer to the French un ... autre (N. Chomsky, personal communication).

13. In some dialects of English, reflexive forms can appear in subject position of a tensed clause, but these are either used as pronouns productively in the dialect, or used emphatically.

(i) Himself came.

The emphatic use of a reflexive form as a pronoun is also possible in French.

(ii) a. Jean a dit qu'il ne viendrait pas.
b. Jean ne comprends pas. Lui-même m'avait dit qu'il viendrait.

14. The notion of Binding given in (78) is not quite accurate since it allows John to Bind himself in (i), which is ungrammatical.

(i) *John believes [$_S$ Bill to like himself]

What we want to express is the fact that the anaphor must be bound in its minimal Binding category, construed as in (ii).

(ii) Binding:
In the configuration $[_{\gamma} \ldots \underline{\alpha} \ldots \underline{\beta} \ldots \underline{\alpha} \ldots]$, $\underline{\alpha}$ binds $\underline{\beta}$ if and only if
1° $\underline{\alpha}$ governs $\underline{\beta}$
2° $\underline{\alpha}$ assigns its R-index to $\underline{\beta}$
3° $\underline{\gamma}$ is the minimal category projection containing a Binder for $\underline{\beta}$

15. There are some peculiar facts of agreement in French that show that agreement sometimes does not depend only on morphological features, but is determined by reference factors hence takes place on the LF side of the grammar. We have in mind here the agreement of elements like the "polite vous" and the "royal nous". These pronouns, which are plural in form, can be used to refer to single individuals in these cases. When this is the case, an attributive adjective (or any other element) agreeing with such a pronoun is singular in form, rather than plural.

(i) Vous êtes trop libéral/*libéraux, mon ami.

(ii) "Nous sommes très libéral", dit le roi.

On the other hand, the reverse is also possible: so a noun like sa Majesté is feminine in form, but it can refer to a male, i.e. a king, and nouns and pronouns that refer to human males are usually masculine in French. But in the case of sa Majesté, it is not the "referential feature" that enters in agreement, but the morphological feature. Thus, in (iii), the person referred to can be either male or female.

(iii) Sa Majesté est bien heureuse/*heureux ce matin.

If Agreement takes place on the LF side, then reference factors like those in (i) and (ii) can be taken into account. And morphological factors like those in (iii) can also be taken into account since they are also available at LF. But if agreement takes place in PF, then it is unlikely that reference factors are available.

16. There are instances where PRO seems to have Case. For example, there are constructions in Russian where PRO seems to bear Dative Case since adjuncts that agree with the subject of an infinitival clause bear Dative Case even if no overt element bears Dative Case in the sentence (cf. Comrie 1974, Neidle 1982). There are also instances of quirky Case which remains throughout the derivation (cf. in Icelandic for example, Andrews 1976 and others). So a given V could have a Case requirement on its subject, and when the V is in an infinitive clause, the elements agreeing with its subject still bear a quirky Case, indicating that the PRO subject might have Case at this level. In Chapter 5, we will look at these cases in detail and show that PRO does not bear Case at PF in such cases.

17. Note that Binding is obligatory and applies when it can, just like other mechanisms that depend on government like Case assignment and subcategorization. The Projection Principle forces subcategorization to take place from the start in D-structure, but Case assignment (or checking) applies whenever it can as is clear in the case of raising or passive.

18. We use the term nominal element here rather than NP because we will eventually want to include in the set of elements that agree, AGR and also adjectives.

19. The fact that names have an inherent R-index could be derived from the fact that they pick the R-index of the object of domain D that they name.

20. There is another redundancy between the principle of Lexicalization and Agreement since both refer to F-features, F-features being included in ψ-features. But this is only an apparent redundancy since the two components deal with F-features in quite different a fashion. The principle of Lexicalization is a well-formedness condition on individual elements, i.e. Ns, whereas Agreement is a well-formedness condition on relations between individual elements. Furthermore, the apparent redundancy is also due to the fact that some morphological features are relevant on both the PF and the LF sides of the grammar. Thus F-features are relevant in determining the phonological shape of an N, and they are also relevant on the LF side of the grammar since they are associated with objects in domain D, as the data in (12) and (13) suggest. F-features are part of the lexical specifications of nouns, and one is not surprised that some lexical specifications are relevant in both PF and LF; since it is assumed that lexical entries provide information relevant to PF and information relevant to LF, some of this information could be relevant to both of these components.

21. There is a restriction of a selectional nature on PRO_{arb}: it seems that it must be interpreted as [+human]. We return to the reasons for why this is so in Chapter 5.

22. Chomsky notes that there could be such an element in a position where Case is not assigned under government, like the Genitive constructions. Chomsky says that "presumably, restricting an element to just such positions, if they exist, is excluded". It is difficult to see why such elements should be excluded without excluding PRO on similar grounds.

23. Given the assumptions made in Chomsky (1981a), the definition of binding category must be supplemented with a statement like (i), which is the equivalent of the one that was necessary for governing category.

(i) A root sentence is a binding category for a governed element.

This stipulation is necessary although the notion of government does not enter in the definition of binding category.

24. In (104iii), PRO is not Bound by John: it is only coreferential with John. This can be illustrated when a reflexive element can disambiguate the relation as in (i) and (ii).

(i) John knows how [PRO to behave himself] (+coref)

(ii) John knows how [PRO to behave oneself] (-coref)

25. But see the discussion of (117)-(129) below.

26. Given our assumptions, this notion of Binding category subsumes the notion of binding category of Chomsky (1981a) gi-

ven in (103).

27. We are excluding from the present discussion cases like Jean$_i$ me lui$_i$ a fait raconter l'histoire discussed in fn.11 and which are allowed for totally different reasons.

28. Recall Burzio's (1981) account of the choice of the assignment of auxiliary discussed in fn.8.

(i) Essere (être) assignment
The auxiliary will be realized as essere (être) when a binding relation exists between the subject and a nominal constituent of the predicate.

The notion 'nominal constituent of the predicate' is required in Burzio's account to cover the reflexive si, which he claims is part of the morphology of the verb. For Burzio, it is the fact that the subject binds si that forces the choice of essere. But the facts in (129) are not compatible with (i) since the subject in (129a) does not bind a nominal constituent of the predicate, namely the clitic me, and yet the auxiliary is avoir. (Note the same is true in sentences like Jean$_i$ me lui$_i$ a fait raconter l'histoire). It might be then that (i) has to be modified as in (ii).

(ii) The auxiliary will be realized as essere (être) when the subject receives its θ-role from the verb forming the predicate.

This means that the subject will originate from a VP internal position in D-structure (i.e. direct object or dative object), and have been moved in subject position by move α. It accounts

for the fact that the auxiliary is avoir in (129a) since the structure is a pronominal one as in (127) and the subject did not originate from a VP internal position. This analysis of essere (être) assignment lends support to our analysis of weak form reflexives since we claim that there is a difference in structure between a reflexive construction and a pronominal clitic construction, i.e. (128) and (127) respectively.

There is one aspect of the se constructions that we have not discussed so far: it is the fact that a sentence like (iii) is ambiguous between the readings in (iva), (ivb), and (ivc).

(iii) Les enfants se lavent souvent.

(iv) a. The children wash themselves often. (reflexive)
b. The children wash each other often. (reciprocal)
c. One washes children often. (middle construction)

In the interpretations in (iva) and (ivb), two θ-roles are assigned, whereas in (ivc), only one θ-role is assigned, that of theme, and the θ-role of agent is interpreted as unspecified and different from the surface subject. There is also a fourth type of se construction, the neuter construction as in (vb).

(v) a. Le vent a dissipé les nuages.
b. Les nuages se sont dissipés.

Neuter constructions differ from middle constructions in that they are lexically restricted, whereas middle constructions are fairly productive and also there is no unspecified agent perceived in the interpretation of neuter constructions (see

Ruwet 1972).

In all of these se-constructions, the auxiliary is always être. So if we want the principle in (ii) to account for the assignment of être, we must assume that all these sentences are derived like in (128), i.e. the subject originates from the object position. So these constructions would be derived somewhat like passives. This has been proposed for the middle construction by Gross (1968). But we must also account for the fact that in the reflexive and the reciprocal constructions, two θ-roles are assigned, but only one in the middle and neuter constructions. This could be done by assuming that se can absorb the θ-role assigned to the subject by the VP, and then it expresses this θ-role or not, depending on whether two or one θ-roles are assigned, respectively. This might force us to modify the Projection principle since se does not occupy a phrasal node position here and it is not related to the post-V position like clitics are. We could account for this difference in the number of θ-roles assigned in a way that would be more directly compatible with the Projection principle as it is stated by making the difference a structural one. Thus the reflexive and reciprocal derivation would be as in (vi), whereas the middle and neuter derivations would be as in (128), given here as (vii).

(vi) Les enfants se lavent PRO.

(vii) a. Les enfants se lavent t.
b. ___ se lavent les enfants.

Here, se is only a Case absorber, and is not related to θ-role assignment. In (vi), se absorbs the Case assigned to the object, so that an EC anaphor can appear in object position as predicted by the principle of Lexicalization. This would be exactly parallel to the English construction in (viii), with the minimal difference that the anaphors have no lexical content in the French equivalents in (vi).

(viii) a. The children wash themselves.
b. The children wash each other.

The middle construction in (vii) would parallel the English passive in (ix): in both constructions, the Case assigned to the object and the θ-role assigned to the subject are absorbed, and move α applies to save the construction.

(ix) a. ___ were washed children
b. Children were washed.

Note that both in French and in English, alternative solutions are possible to assign Case to the D-structure position, namely insertion of an expletive element as in (x).

(x) a. Il se lave beaucoup d'enfants.
b. There were washed many children.

If this analysis of reflexives and reciprocals is adopted however, principle (ii) which governs the assignment of the AUX être must be modified in a way that brings it closer to Burzio's formulation.

(xi) The AUX will be realized as être when a Binding relation exists between the subject and an argument position of the predicate.

29. Note that the condition can be kept as it is stated in (112') even if we have sentences like (i).

(i) They saw each other.

Although each other does not have a reflexive form, it is still possible in the context in (i). The reason is that each other is not Bound by they in (i), but each other is linked to they by each-movement, not by Binding. We have already seen that reciprocals are allowed in positions where reflexives are not possible because they are not related to their antecedent by the same mechanism that reflexives are (cf. (66) and (70)).

30. There remains the problem that some speakers find (i) less acceptable than (138), for which we have no explanation.

(i) ?They expected that [each other's pictures] would be on sale.

31. There is much more to be said about disjoint reference. For instance, there is a difference between nous-tu pairs and nous-je pairs as in (i).

(i) a. *Nous pensons que tu réussiras.
b. ?Nous pensons que je réussirai.

This might have to do with the fact that nous necessarily includes the speaker by lexical specification, and so that nous is not referentially dependent upon je since it is lexically specified as including the reference of je anyhow.

Our purpose was only to show that disjoint reference facts are more readily accountable in terms of conditions akin to con-

dition C than to condition B.

32. Ken Hale informs us (personal communication) that this seems to be generally the case in so-called flat-structured languages, i.e. languages with (at least) no VP, so that the object and subject c-command each other in the syntax.

33. The notion BOUND is not the same as Bound (78), but it is rather intended to cover the notion 'in the scope of' as is clear by the discussion.

34. For some discussion of this kind of facts, see Chomsky (1977a), where it was observed that cross-over restrictions hold of quantified NPs as well as of WH-phrases. See also Higginbotham (1980) where some additions are made to Chomsky (1977a) in order to account specifically for facts of the type presented in (162).

35. In the present analysis, a notion of pronominal anaphor is not possible, according to the definitions given in (105)-(107): PRO is either anaphoric or pronominal, but never both.

36. Note that the $\bar{A}$-binder of the variable cannot function as the third binder, or else there would be no explanation for the ungrammaticality of (158a). So this third binder must not be an $\bar{A}$-binder of the variable: it must be an extra binder, not one that "created" the variable, as is quite natural.

CHAPTER 3: THE EMPTY CATEGORY PRINCIPLE:

3.1 General comments.

In this chapter, we will examine in more detail the effects that are attributed to the ECP in GB, and we will see how they can be derived from independently motivated principles in the approach to empty categories presented in Chapter 2. We will take the ECP to be as defined in (1), with proper government as defined in (1b).

(1) a. Empty Category Principle
$[_{\underline{\alpha}}\ e]$ must be properly governed.

b. Proper government
$\underline{\alpha}$ properly governs $\underline{\beta}$ if and only if $\underline{\alpha}$ governs $\underline{\beta}$, and
(i) $\underline{\alpha}$ is lexical ($=X^{O}$), or
(ii) $\underline{\alpha}$ is coindexed with $\underline{\beta}$.

These are the definitions given in Chomsky (1981a). There are other variants of the ECP approach in the literature, with some modifications of the definitions in (1), but the definition in (1) can be assumed to be representative of the core notions involved. In any case, what we want to do is to show how the ECP can be derived from the general principles proposed in Chapter 2; we will therefore show how this can be done for these core notions, and the analysis should extend in a similar fashion to the differences that might arise in the coverage of slightly different definitions of the ECP. What is crucial for us is to show by what principles the effects of the ECP can be accounted for in our framework, regardless of minor details in the formulation of the ECP and its components.

As has often been noticed, the ECP is essentially a statement of local control, of recoverability on traces (Chomsky 1981 a,b, Kayne 1981a, Jaeggli 1981, Stowell 1981a, Aoun et al. 1981). The formulation of proper government contains a strange disparity as to what can be a proper governor, i.e. a "recoverer": it can either be an X^{o}, or an element coindexed with the EC. As noted by Jaeggli (1981) and Stowell (1981a), this is not a unified notion. Stowell (1981a), for example, proposes to overcome this disparity by assuming that elements that form predicates, like verbs, have a θ-grid where the indices of the verb's complements are entered. So lexical heads would therefore be coindexed with their complements in some sense by bearing these indices in their θ-grid. This allows a more unified notion of proper government, since all proper governors would now govern and be coindexed with the EC. There still remains the fact that the two coindexings are somewhat different, however.

There is also a disparity about the application of the ECP that emerges at a broader level: one could ask why it is that the ECP holds only of a subclass of ECs, namely traces, if it is essentially a statement on recoverability. Surely the content of all ECs must be recoverable. And if there is only one EC, its different manifestations being functionally determined, one might expect ideally that the content of the EC has to be recovered in a unitary fashion. We have suggested in Chapter 2 that this is the case, and that this recoverability of the content of the EC is dependent on Agreement of the EC with an

overt element bearing the same R-index at LF. The EC can get such an R-index either by being Bound by an antecedent (in the sense of 2.78), or by being freely indexed at S-structure, this allowing coreference relations like pronouns have. These ways of receiving an R-index are independently motivated for overt NPs like reflexives and pronominals, as we have argued extensively in Chapter 2. So the content to be recovered is an R-index and F-features. The fact that an R-index must be assigned to the EC (and to any argument NP) is due to the θ-criterion and the principle of Denotability II: an argument must be assigned to a θ-role, and θ-roles are assigned only to index bearing NPs. The fact that F-features are assigned to an EC (or any NP) comes from Agreement which operates on the LF side of the grammar.[1]

In the case of traces, the only mechanism of index assignment which is operative is Binding from an antecedent, for reasons that will be made clear as we look at the individual cases. Recall that in the strong form of a transformational grammar, the relation between a WH-phrase and a variable is expected to be the same as that of an NP and its trace, i.e. move α. And since one of the main motivations for introducing trace theory was to reduce the possibilities of moving phrases to those where the output was a configuration where antecedent and lexical anaphor are found, then the restrictions on the antecedent-anaphor relation, once properly determined, should be found in the antecedent-trace relation, this being the null hypothesis.

In the next sections, we will look at how the properties

of the antecedent-anaphor relation are mirrored in the NP-trace relation (3.2), the WH-phrase-trace relation (3.3), and the Q-phrase-trace relation (3.4). In all these cases, recoverability of the content is necessary whenever a gap is necessary, i.e whenever a θ-role is assigned according to the (Extended) Projection principle and no lexical element is present (including PRO and pro as we will see in the following chapters), and recoverability is dependent on the same principles in all cases.

3.2. The NP-trace relation.

In Chomsky (1981a), "local control" of NP traces is determined by clause (i) of proper government in (1b). So, for example, the underlined V is the X^0 that governs the trace in (2).

(2) a. John was <u>seen</u> <u>t</u> at the movies.
b. John <u>seems</u> $[_S$ <u>t</u> to be happy $]$.

In (2a), the trace is governed by <u>seen</u> since it is in the object position. In (2b), the trace is governed by <u>seems</u> because $\overline{S}$ deletion has taken place, and so no major boundaries block government. In contrast,sentence (3) is ungrammatical because $\overline{S}$ deletion is not possible if the embedded sentence is tensed, and so the trace is not properly governed since INFL is not a proper governor, a topic to which we will return.

(3) *Bill is said $[_{\overline{S}}[_S$ <u>t</u> is happy $]]$

In (2), <u>John</u> and <u>t</u> are related by <u>move α</u> and they form a chain: in fact, they form one discontinuous element. Each part of this discontinuous element lacks something crucial that the

other part provides: so John is an argument that has some referential content and some grammatical features, whereas the trace t is in a position that is assigned a θ-role. The discontinuous element (John, t) can therefore fulfill both requirements of the θ-criterion for the position of t and the argument John, although each individual element John and t could not fulfill these requirements on its own.

(4) θ-criterion

A. Each A-position is assigned an argument.
B. Each argument is assigned a θ-role.

But notice that in (2), the proper governors according to the definition in (1), namely seen and seems, appear to play no role in the recoverability of the content of the trace, the content being recovered from the antecedent John. However, if we assume that VP is not the maximal expansion of V, then in (2), John also governs the trace since no maximal expansion lies between John and t. So the requirement for the recoverability of the content of the trace can be simply that the trace must be Bound in the sense of (2.78): the trace would then be functionally determined to be an anaphor Bound by its antecedent. In other words, proper government could have only clause (ii), which would be a requirement to properly identify the trace in some sense by relating it to its antecedent. So, at least in the case of NP trace, it seems that we can eliminate clause (i) from proper government, and this modification essentially re-

duces proper government to our notion of Binding, i.e. government and indexing by the antecedent, this relation of Binding being independently motivated for lexical anaphors as we saw in Chapter 2.

The reason why clause (i) appears to be part of proper government in many cases is that government by a lexical head will also be necessary for the trace, but for reasons that have nothing to do strictly with the "identification" of the trace. For example, in cases like (2a), the trace must be lexically governed to receive a θ-role for the discontinuous element that it forms with John. In (2b), on the other hand, $\bar{S}$ deletion must take place for John and t to be related by Binding; since $\bar{S}$ deletion takes place only in the context [[+V]___], seems appears to govern the trace, but again for reasons that are irrelevant to the "indentification" of the EC. So all NP traces must be identified by being related to an antecedent but not all traces have to be lexically governed; for example, the subject position of a sentence is not assigned its θ-role by a lexical head, hence not lexically governed, since it is the external argument of the predicate.[2] We will come back to such cases when we look at WH-movement and the ECP in the next section.

What comes out of this discussion is that clause (i) of proper government is an accidental property of NP traces as far as their "identification" is concerned (although government by a lexical head is relevant with respect to θ-assignment for ex-

ample). The effects of ECP are derivable from the interaction of principles independently motivated for lexical NPs, namely Binding, Denotability II, the θ-criterion and Agreement. Since the relation between an NP-trace and its antecedent is essentially the same as the relation between a lexical element Bound by its antecedent, we expect the two constructions to share the basic properties that we observed in studying true anaphors. Thus we observe that the relation between NP-trace and its antecedent is the same as that of a true anaphor and its antecedent in that 1° the relation is obligatory; 2° it is a one-to-one relation, so that there is only one antecedent for a trace; 3° there is a specific structural relation with the antecedent, namely government; 4° the relation is local, because of government again. (See Koster (1978a,b) for similar observations)

If the effects of the ECP are derived from these principles, these principles should also explain why the subject of a tensed clause and the domain of a subject are the two opaque domains, i.e. the Propositional Island Constraint and the Specified Subject Condition. We have already given an answer to the first of these problems in our discussion of (3), which we repeat here.

(3) *John is said $[_{\overline{S}}[_{S}$ t is happy $]]$

In (3), John cannot Bind t because a maximal expansion blocks government, i.e. $\overline{S}$, since $\overline{S}$ deletion cannot apply when the embedded clause is finite (for reasons that presumably have to do with the relation between tense inflection on the verb

and COMP, which we will not discuss here). Thus we explain the effects of the PIC.

As for the facts that fall under the SSC, they are accounted for by the fact that the subject is always a possible local Binder, and hence prevents any NP higher in the structure from Binding the anaphor. For example, consider the sentences in (4).

(4) a. *They$_i$ expected [$_S$ me$_j$ to be seen $\underline{t}_i$ together]

b. *They$_i$ expected [$_S$ me$_j$ to be looked at $\underline{t}_i$]

In both sentences in (4), me is the Binder of the trace according to the definition of Binding (cf. fn 16, Chapter 2), so the sentences are ungrammatical under the intended readings. The EC has no index and it receives one from me, its Binder, and hence it cannot be related to they: so the effects of SSC are derived from Binding. Note that the sentences are also ruled out for a different reason in GB, namely that if they binds the trace, then it has two θ-roles since it inherits one from the trace, and it also gets one in the subject position of expected.[3] Furthermore, me will also violate the θ-criterion since it receives no θ-role under the intended readings. We see that the SSC effects are obscured in sentences like (4). It is difficult to construct examples with NP trace that illustrate the SSC effect without having other principles that also have some relevance in ruling out the sentence. For example, if we try to avoid the interaction with the θ-criterion that we find in (4) by passivizing the verb expect, thus eliminating the θ-role

assigned to the subject NP, then Case theory enters into action.

(5) *They$_i$ were expected [$_S$ me$_j$ to be seen $\underline{t}_i$ together]

We could say that (5) is ungrammatical because me is the Binder of the trace, not they. But the sentence is also ruled out by the fact that me fails to be assigned Case (and a θ-role) and hence violates the principle of Lexicalization, from which the Case Filter is derived. In fact, there does not seem to be a case where an NP trace violates the SSC effect without also violating another principle of the grammar. However, since the SSC effect holds for lexical anaphors as we can see in (6), we will assume that it also holds for NP traces as well.

(6) *They expected [$_S$ me to see themselves]

In (6), the anaphor themselves lacks an R-index, so it is assigned one by its local Binder me. But Agreement then rules out this interpretation. Note that they could govern themselves since there is $\overline{S}$ deletion here in order for me to be assigned Case by expected, but they is not the local Binder for themselves since there is the intervening me which Binds themselves in a projection which does not contain they, namely the embedded S. Therefore, there is an SSC effect, and the sentence is ungrammatical, this effect being derived from Binding and interrelating principles.

The topic of Case assignment brings us to the distribution of the NP trace. In GB, NP trace does not have Case because

the reason why the lexical NP is moved is precisely because it has been inserted in a D-structure position that is not Case marked, and hence the lexical NP must move in a Case marked position to escape the Case Filter.

(7) a. John seems [t to be happy]
b. John was seen t there.

We essentially adopt this analysis of raising and passivization. Thus in (7a), seems is not a Case assigner and so John must be raised. Similarly in (7b), passive morphology precludes Case assignment by seen, so that John must be raised to the subject position where it can be Case marked. This ties in with Burzio's (1981) observation that assignment of θ-role to the subject and assignment of Case to the object by a verb are related as T↔A (where T= assignment of θ-role to the subject, and A= assignment of accusative to the object). Thus if a verb did not assign Case to its object and yet if a θ-role was assigned to the subject, then the object could not be lexicalized since it could not have Case since if it is raised in subject position, it will get two θ-roles in violation of the θ-criterion. So languages do not seem to allow such verbs. On the other hand, if a language has a way of assigning Case in constructions like (7) by means of an expletive element like there or il for example, then the sentence is grammatical without raising taking place as in (8).

(8) a. There arrived a man.
b. Il est arrivé un homme.

Note that Burzio's observation still holds in (8): no θ-role can be assigned to the subject position since insertion of an expletive element will transmit Case to the object, and the expletive cannot bear a θ-role.

The fact that the NP trace does not have Case is consistent with our account of where an EC is possible. Recall that an EC is possible if it has no ψ-features at PF, or else that NP must be lexicalized by the principle of Lexicalization. So an EC would not be possible in a position where Case is assigned and is visible at PF. (But see the next section on WH-movement where it will be seen that an EC can appear in a position where Case is assigned in instances where the Case is borne by another element, like a WH-phrase for example.) Therefore, an NP trace can only appear in positions where Case is not assigned, like the subject of an infinitive clause (9a), the object of a passive verb (9b) or of a preposition reanalyzed with a passive verb (9c), or the object of a category that does not assign Case (9d), or an ergative verb (9e).

(9) a. John seems $[_S$ t to be happy]
b. John was seen t here.
c. John was referred to t in the paper.
d. Rome's destruction t
e. Jean est tombé t.

Notice that an NP trace is possible in the subject position of an infinitive clause only if $\overline{S}$ deletion has taken place so that the t can be Bound by the NP, and that the matrix verb

must not be a Case assigner or else the EC would be lexicalized by the principle of Lexicalization.

In the introductory comments to this Chapter, we pointed out that the only mechanism of index assignment which is operative for traces in Binding, and that free indexation is impossible. We can see why this is so for NP traces by considering the sentences in (9). In all the instances where a position is lacking Case, there is always an NP position which governs the Caseless position. So an EC in such a position is always necessarily Bound, and hence cannot be freely indexed. So for example, in a sentence like (10), it Binds the trace in the object position.

(10) It was hit t.

In (10), it Binds t. If it is referential, the sentence is grammatical. If it is expletive, then the sentence is ungrammatical since a θ-role is assigned to a discontinuous element (it, t) that has no referential content. Note that an expletive element is possible in a sentence like (11) in German since there is no EC to Bind in object position, hence no θ-role is assigned to the expletive element (see also 5.2.1.1 for more discussion).

(11) Es wird gelacht
it is laughed
(People are laughing.)

Note that we now have an explanation why sentence (10) is

not interpreted as in (12).

(12) $It_{\underline{expl}}$ was hit $PRO_{\underline{arb}}$

In GB, if it and the EC are not coindexed, then the EC is PRO by functional definition. But PRO is then governed by hit in (12), so the sentence is ungrammatical.[4] According to our analysis, the EC is Bound by it, and hence is functionally determined to be an anaphor, so that the interpretation in (12) is not possible since the EC has the R-index and the F-features of it.

In the case of the infinitive sentence as in (9a), if $\overline{S}$ deletion does not take place, then the EC in subject position is not Bindable since $\overline{S}$ blocks government by any antecedent. We will see in Chapter 5 that these are instances of long distance control PRO and of arbitrary PRO, and that these elements are indexed by free indexing at S-structure, so that they are functionally determined to be pronominals, not anaphors, and that they have properties which are radically different from those of Bound elements like NP traces. So if there is no $\overline{S}$ deletion in (9a), John gets no θ-role since it does not form a discontinuous element with t.

We must also assume that in (9c) reanalysis of the V and the P takes place so that there is no intervening PP node to block government of the t by John.

The case of (9d) is quite different from the other constructions in (9). Contrary to a sentence , an NP does not have an

obligatory subject position, so that the structures in (13) could be possible D-structures.

(13) a. $[_{NP}$ destruction Rome]

b. $[_{NP}$ destruction $[_{NP}$ e]]

In (13a), according to the principle of Lexicalization, the object Rome must have Case since it is lexical. This can be done by of-insertion as in (14a), or by movement to the subject position where Genitive is assigned as in (14b).

(14) a. destruction of Rome

b. Rome's destruction

But suppose that no Case is assigned to the object position so that it can be an EC as in (13b) and suppose that there is no subject to the NP to Bind the EC. The EC cannot be Bound by any NP outside of the NP $[_{NP}$ destruction$[_{NP}$ e]]since NP is a maximal expansion and hence blocks government by an antecedent. But then presumably the EC can be freely indexed at S-structure and hence be interpreted as a pronominal, being either coreferential to another NP in the sentence as in (15a) or free as in (15b): these interpretations would correspond to long distance control PRO and PRO_{arb}, respectively.

(15) a. *$Rome_i$ was a scene of [terrible destruction PRO_i]

b. *$[_{NP}$ the destruction PRO_{arb}] was terrible.

But these sentences are ungrammatical under these readings: destruction with no overt object is interpreted as a result ra-

ther than as an action like when it has an object. So free indexing seems to be impossible here. The reason for this according to the present analysis is that an EC is not possible in a construction like (13b). To see why, consider (13a) again. The construction is made licit by of-insertion, which roughly operates as in (16) (see Stowell 1981a).

(16) Dummy Preposition Insertion
In the configuration $[_{\alpha} \ldots \beta \ldots]$, adjoin a dummy preposition to β, where
(i) α is some projection of [+N], and
(ii) β is an immediate constituent of α, and
(iii) β = NP

Every language that has this rule will specify which of its prepositions can serve as a dummy preposition: in English, it is of; in French, de; in Italian, di; in Hebrew, šel; etc. This rule applies to (13a) and inserts the preposition of, giving the output in (14a). But then there is no reason why it should not apply to (13b) to give (17).

(17) destruction of $[_{NP}\ e]$

Recall that we assume that in general rules apply regardless of the fact that an NP is lexical or not. So (13b) satisfies the requirements of (16), and the rule applies to give (17). But (17) is ruled out by the principle of Lexicalization: the NP has Case on the PF side of the grammar, and hence must be lexical. One might argue that we seem to be taking some mo-

tivation out of the proposal to have a rule of dummy preposition insertion. Thus in GB, the rule of DP-insertion is functionally motivated by the fact that a lexical NP must escape the Case Filter, i.e. that a lexical NP must be sanctioned by Case. But there is a similar motivation in the present analysis: the rule is there to allow the presence of a lexical NP. Note that making DP-insertion obligatory explains directly why of is inserted in sentences like John was taken advantage of. The insertion of of here cannot be functionally motivated by the fact that a lexical NP must be sanctioned by Case since of does not assign Case to an NP in such a sentence: John gets nominative Case in the subject position.[5]

One problem remains with this analysis, however. If (16) applies blindly, why doesn't it apply to (14b) to yield the ungrammatical (18)?

(18) *Rome's destruction of t

In other words, why is (16) blocked from applying in (14b)? First, we see that another rule of Case assignment applied in (14b), the rule of Genitive Case assignment.

(19) Genitive Case Assignment[6]
In the configuration $[_{\alpha} \ldots \beta \ldots]$, assign Genitive Case to β, where

(i) α is some projection of [+N,-V] , and
(ii) β is an immediate constituent of α, and
(iii) β = NP

By applying or not applying (16) and (19) to (14b), there

are four possible outputs as in (20).

(20)

	(16)	(19)
a. *Rome's destruction of t	+	+
b. Rome's destruction t	-	+
c. *Rome destruction of t	+	-
d. *Rome destruction t	-	-

The only grammatical output is (20b): all the other possibilities are ruled out by the principle of Lexicalization. Thus in (20a), the t should be lexical; in (20c) the t should be lexical and Rome has no Case; and in (20d), Rome has no Case. One way to make sure that a grammatical output is possible while maintaining that the two rules are obligatory (to rule out (15) for example) could be the following, which is suggested in Stowell (1981a). We might account for the failure of DP-insertion to apply in (14b) and (20b) by invoking a principle of Kiparsky (1982)

> that a special rule R_i always takes precedence over a general rule R_j in their overlapping domain. More precisely, if R_i applies in a set of environments E_i, and R_j applies in environments E_j, and E_i is a proper subset of E_j, then R_i applies in E_i and R_j does not. Take the genitive Case assignment rule to be R_i and the of-insertion rule to be R_j; then of-insertion is blocked from applying in (contexts like (14b)) by Kiparsky's principle.
>
> (Stowell 1981a, p. 246)[7]

We see that the environments for the application of (19) are a subset of the environments of the application of (16) since α is some projection of [+N,-V] in (19) but α is in some projection of [+N] in (16). Note that in a sentence like John

was taken advantage of discussed above, of must be inserted since this is not a case where a rule of Case assignment has as the set of environments for its application a subset of the environments for the application of DP-insertion: here Case is assigned to the NP by the general process of nominative Case assignment, and the Case usually assigned by of is absorbed by the passive morphology of the V in the reanalyzed taken advantage of.

To get a sentence like (15b) under its grammatical reading as a result rather than an action, we must simply assume that there is no NP in the object position of destruction, as in (21).

(21) $[_{NP}$ the destruction] was terrible.

So as pointed out in the discussion of (13b), the reason why $[_{NP}$ destruction e] is not possible with a freely indexed EC interpretation is because an EC is not possible in such a construction since DP-insertion would apply to it. Therefore, our general observation that trace cannot get an index by free indexing at S-structure but only by Binding holds, given (16) and (19) and their mode of application.

Consider now the fact that NP traces observe Subjacency.

(22) John seems $[_{S2}$ t^2 to be certain $[_{S1}$ t^1 to win]]

(23) a. *John seems $[_{\bar{S}}$ that $[_{S}$ it will be difficult $[_{\bar{S}}$ $[_{S}$ t to feed himself]]]

b. *John seems $[_{\bar{S}}$ that $[_{S}[_{\bar{S}}[_{S}$ t to feed himself] will be difficult]]]]

In (22), the derivation is assumed to be successive cyclic, the subject John raising from one subject position to another. The reason is that it is assumed that there is a Subjacency condition which holds and which prevents an element to be moved across two bounding nodes. Since S is considered to be a bounding node in English, the only possibility to have a proper derivation in (22) is by iteration of move α.

In (23) on the other hand, the link between John and t crosses more than one bounding node, so the derivation is impossible since John cannot be assigned a θ-role.[8]

In our analysis, Subjacency facts with respect to NP traces (and also WH-traces as we will see in 3.3.3) follow from Binding and the θ-criterion. For example, in (23), since $\bar{S}$ precludes Binding, then John and t cannot form a discontinuous element (John, t), since such discontinuous elements are formed only by Binding. So John does not receive a θ-role. In (22) on the other hand, since $\bar{S}$ deletion has taken place, John can Bind t^2, which in turn can Bind t^1. Note that, although John governs t^1, it cannot Bind it directly since there is a more local Binder for t^1, namely t^2. There is evidence that there is an NP in subject position of the intermediate clause and that this NP is part of a Binding chain: this evidence shows up in languages where adjectives agree with the subject when they are attributive like in French, as is illustrated in (24).

(24) Marie semble $[_S$ t être certaine/*certain $[_S$ t de gagner]]

It is possible to test the fact that the derivation is as described, i.e. that the D-structure is (25a) rather than (25b).

(25) a. [e semble [$_S$ e être certaine [$_S$ Marie de gagner]]]

b. [e semble [$_S$ Marie être certaine [$_S$ PRO de gagner]]]

If the derivation was as in (25b), then the subject position of être certaine would be a θ-position by the Projection principle. But this is not the case as can be seen in (26) where an expletive element fills the subject position of certain.

(26) Il_{expl} est certain [que Marie va gagner]

So we conclude that the effects of Subjacency follow from Binding and the θ-criterion.

Consider now expletive elements like there and il.

(27) a. There is a man at the door.

b. $There_i$ seems [$_S$ t_i to be a man at the door]

(28) a. Il y a un homme dehors.

b. Il_i semble [$_S$ t_i y avoir un homme dehors]

These elements show that the R-indices must be conceived of as comprising two types of indices, as discussed in Chapter 2: a syntactic index S-index and a referential index REF-index. Elements like there and il would only have an S-index and no REF-index. So they can still Bind traces as in the (b) examples above, but this is not to be interpreted as saying anything about the reference of such elements.[9]

In conclusion, the properties of NP traces with respect to the ECP and Subjacency can be derived from principles which are independently motivated for lexical NPs, namely Binding, the θ-criterion, the Projection principle and the principle of Lexicalization.

3.3. The WH-phrase trace relation.

3.3.1. General comments.

We have seen above that the notion of Binding plays a crucial role in the relation of a "dependent element" to its antecedent. Thus Binding accounts for the distribution of true reflexives with their specific properties: obligatoriness of an antecedent, uniqueness of the antecedent, a specific structural relation with the antecedent, and locality. A similar account was given for the distribution of NP traces, with the additional factor that the trace being an EC, it must not have $\underline{\psi}$-features at PF according to the principle of Lexicalization.

Consider now the following data.

(29) a. $[_{\overline{S}}$ Who$_i$ $[_S$ $\underline{t}_i$ is coming for dinner$]]$

b. $[_{\overline{S}}$ Who$_i$ $[_S$ did you see $\underline{t}_i$ $]]$

c. $[_{\overline{S}}$ $[_{PP}$ in which box$]_i$ $[_S$ did you put it $\underline{t}_i]]$

(30) a. $[_{NP}$ the people $[_{\overline{S}}$ who$_i$ $[_S$ $\underline{t}_i$ are coming for dinner$]]]$

b. $[_{NP}$ the man $[_{\overline{S}}$ who$_i$ $[_S$ you saw $\underline{t}_i]]]$

c. $[_{NP}$ the box $[_{\overline{S}}$ $[_{PP}$ in which$]_i$ $[_S$ you put it $\underline{t}_i]]]$

These are the core constructions where WH-movement is said to be operative.[10] In the sentences in (29)-(30), it is assumed

that the WH-phrase is some kind of operator which is related to a coindexed EC, the trace of the WH-phrase which has been moved to COMP, assuming that there is such a position for all $\overline{S}$s as in (31) (see Bresnan 1972).

(31) $\overline{S} \rightarrow$COMP S

Note that in keeping with the assumption that there is no specific phrase structure rule and that phrase structure rules are derivable from the interaction of other components of the grammar, as we saw in Chapter 1, we can say that the COMP position in (31) is derivable from the fact that a sentential proposition must be situated in time, hence receives a tense interpretation, and that this is mediated by a position which ranges over the whole sentence, i.e. COMP (cf. Stowell 1981b for some discussion).

So assuming that there is such a COMP position, we can look at the properties of the relation between the WH-phrase and its trace.

A first property of the WH-phrase-trace relation is that it is obligatory: if the trace has no WH-phrase binding it in some sense, the sentence in ungrammatical under the relevant reading.

(32) a. *$[_{NP}$ e] saw John?

b. *Did John see $[_{NP}$ e]?

c. *Did John put it $[_{PP}$ e] ?

Conversely, the WH-phrase must be related to a position in the sentence.[11]

(33) a. *[$_{\bar{S}}$ Who [$_S$ Bill saw John]]

b. *[$_{\bar{S}}$ Who [$_S$ did John see Bill]]

c. *[$_{\bar{S}}$ Where[$_S$ did John put it in the box]] (under the relevant interpretation)

The presence of the WH-phrase is not sufficient to satisfy the obligatoriness of the relation: there is also a structural requirement that must be met as we see in (34).

(34) a. *[$\underline{t}_i$ told John[$_{\bar{S}}$ who$_i$ [$_S$ Bill would come for dinner]]]

b. * Did John see $\underline{t}_i$ and who$_i$ did Bill like $\underline{t}_i$

The relation between the WH-phrase and the EC also seems to be a one-to-one relation.

(35) a. *Who$_i$ did John give $\underline{t}_i$ to $\underline{t}_i$?

b. John gave the slave$_i$ to himself$_i$

(36) *Who$_i$ [did John tell Bill [who$_j$[$\underline{t}_{i+j}$ came for dinner]]]

Note that the reason why (35a) is ungrammatical cannot be for semantic reasons since the answer to this question in (35b) is a possible sentence. Furthermore, there are cases where an operator binds two variables at LF, so that this is semantically possible.

(37) [Which article]$_i$ did John file $\underline{e}_i^1$ without reading $\underline{e}_i^2$?

However, $\underline{e}_i^2$ is a special type of gap, a parasitic gap, and its relation with the WH-phrase is not part of the core cases

of WH-phrase-trace relations (although its distribution follows from general principles applying to other constructions in the grammar. For some discussion see Chomsky 1981b, Bouchard 1982b). For example, a parasitic gap (henceforth PG) cannot appear in the position of $\underline{e}_i^2$ in (37) without being licensed by another gap in the sentence.

(38) *Which articles$_i$ did John come back from work without filing t_i?

So we can assume that it is a property of the core relation between a WH-phrase and a trace that it is a one-to-one relation, PGs being outside of the core cases.

A fourth property of this core relation is that the WH-phrase cannot be too far from its trace in some sense.

(39) a. *Who$_i$ did John deplore [$_{NP}$ the fact that $\underline{t}_i$ came]

b. *[$_{\overline{S}}$ What$_j$ [$_S$ does John know [$_{\overline{S}}$ who$_i$ [$_S$ $\underline{t}_i$ bought $\underline{t}_j$]]]]

c. *[$_{\overline{S}}$ Quelle boîte$_i$ [$_S$ est-ce que Jean a mis le livre [$_{PP}$ dans $\underline{t}_i$]]]

What we see is that these core constructions involving WH-movement exhibit four properties that we have already attributed to the relation of true anaphors and NP traces with their antecedents: obligatoriness of the antecedent, uniqueness of the antecedent, a specific structural relation with the antecedent, and locality. It remains to be seen whether the structural relation and locality requirements involved in constructions involving WH-movement are the same as those involved in the true anaphor and NP-trace constructions. The null hypothesis would

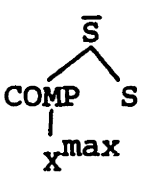
be that they are the same. The fact that these relations are basically the same was in fact one of the primary motivations for the postulation of trace theory (see Chomsky 1977b and related work). Much of the work that has been done since in the EST framework has been to discover what these properties are precisely.

We have seen that the strong form of this hypothesis would be to make use only of principles that apply indiscriminately to NPs, whether lexical or not, since this would strengthen the claim that such theoretical entities as ECs exist. We have also seen that the central notion which accounts for the four properties described above for true anaphors and NP traces is the notion of Binding as described in (2.78). In the best of worlds, we therefore expect Binding to be a central factor in accounting for the properties of constructions where WH-movement is involved.

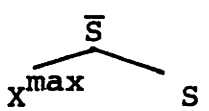
If Binding is involved in this relation between WH-phrase and trace, then the fact that the relation is obligatory and that it is a one-to-one relation follows directly. The relation is obligatory since Binding applies whenever the environment for its application is met.[12] The relation is one-to-one since the local Binder of an EC assigns it an R-index, and that Binding cannot reapply since the EC now has an R-index, Binding applying only to elements that do not have an R-index: so the first application of Binding bleeds the second application of Binding.

It can be shown that the two other properties of constructions involving WH-phrases are also the same as the properties that we assume that constructions involving true anaphors and NP traces have, namely that the structural relation and the locality involved are subsumed by government, hence that the relation involved in these constructions is also Binding. Consider again the sentences in (29)-(30). In (29a) and (30a), who is in COMP and governs the trace in subject position since there is no intervening X^{max} node. Note that we are assuming that COMP does not count as an intervening node between the X^{max} dominating the WH-phrase and $\bar{S}$ since government would not be possible in such a case according to our definition since the governor must be an immediate constituent of a node dominating the governee. What we assume is that COMP is an unlabelled position which is a sister to S, so that the configuration is not as in (40a), but rather as in (40b).

(40) a.



b.



Note that the structure in (40b) respects a strong form of $\bar{X}$-theory, namely one where nodes are labelled according to the lexical material in the head of the projection, hence must be headed. This is not so for (40a), where the structure where COMP dominates X^{max} violates this strong hypothesis since the node COMP has no lexical head. If the structure is as in (40b),

then if that is inserted under the head of X^{max}, then X^{max} = COMP, a specifier of S, if we assume that there is a lexical category 'Complementizer' to which that belongs.

Turning now to (29b) and (30b), we see that who again governs the position of the trace since there are no intervening X^{max} nodes (recall that we assume that VP is not a maximal projection). Similarly, in (29c) and (30c), the PP in which box governs the position of the trace. So in all the core constructions involving WH-movement, the relation between the moved phrase and the EC is established by Binding. This is what our theory had led us to expect since this notion is already independently needed to account for true reflexives and NP traces, and we made the hypothesis that no principle or rule would be operating only on constructions involving EC, if the existence of elements such as ECs are to be strongly motivated, which we assume must be the case if the whole approach involving ECs and the Projection principle is valid.

Consider the following sentences.

(41) a. Who_i did John speak to t_i?
b. Who_i did John look after t_i?

At first glance, these sentences seem to contradict our claim that Binding is involved here since government inside a PP should be blocked since PP is a maximal expansion. But recall that English has a process of reanalysis of the P with V which allows VP internal prepositions to become a single unit with the

V.[13] Thus true anaphors are obligatory in these cases, and passivization is possible.

(42) a. John talked to himself/*him.
b. John looked after himself/*him.

(43) a. $John_i$ was spoken to t_i
b. $John_i$ was looked after t_i

Given this rule of reanalysis, the constructions in (41) fall under the core cases of constructions involving WH-movement since the WH-phrase Binds the trace. So given the auxiliary hypothesis that there is a process of reanalysis involved in (41), the main hypothesis about WH-constructions and Binding can be maintained.

3.3.2. WH-trace and Case.

There is a crucial difference between an NP trace and a WH-trace. The NP trace is not in a Case marked position: in fact, the reason why the NP moves in most cases is to get Case in some other A-position where Case is assigned in order to meet the requirements of the principle of Lexicalization.[14] On the other hand, the trace of a WH-NP is usually in a Case marked position. So we might wonder why this position is an EC, since, given the principle of Lexicalization, an NP that is assigned Case will have a lexical head since, having a ψ-feature, it cannot be an EC. Given our general assumptions, however, we do not expect the WH-trace to differ from other ECs: it should be an EC if and only if it does not have ψ-features. But notice that there is another crucial difference between NP-

trace constructions and WH-trace constructions: whereas the NP is moved to a Case marked position, the WH-phrase is moved to an $\bar{A}$ position, and the core case for $\bar{A}$-positions is that they are not Case marked positions (we return to exceptions to this generalization shortly in 3.3.2.1). Since the WH-NP must presumably also satisfy the principle of Lexicalization, it must also be assigned Case. One way to insure that this takes place is to assume that Case is assigned to the WH-NP before it is moved to an $\bar{A}$-position, so that it carries Case along, or that WH-NP absorbs the Case of the variable that it Binds, which gives similar results.[15] If so, then we have an explanation as for why the NP position Bound by a WH-phrase is not lexicalized even if it is Case marked: it is the WH-NP that bears the Case, not the Bound NP so that this Bound NP can be an EC since it has no ψ-features at PF. Note that if the WH-NP was not moved from the position it Binds but had been base-generated in COMP, then it would not bear the Case of that position and there would be a lexical NP in that position: that is the resumptive pronoun strategy, where the variable must be lexical since it bears Case. It is interesting to note that positions where resumptive pronouns show up are usually precisely those from which extraction is impossible, i.e. positions that cannot be Bound by a WH-phrase.

There is empirical support for this analysis. The first set of data that supports it has to do with free relatives in Hebrew; the second set of data deals with a contrast between

relative clauses in colloquial French and in Standard French.

3.3.2.1. Hebrew Free Relatives.

The facts and the analysis reported in this section are essentially those reported in Borer (1981), Chapter 2. Borer shows that relative clauses in modern Hebrew can be formed by two different strategies: a movement strategy, in which all the usual constraints on movement are observed, as we see in (44)-(46), and a resumptive pronoun strategy (resumptive clitics for PPs and NPs, free standing pronouns for direct objects), where no movement is involved, as in (47)-(49).

(44) a. ha-'iš še- ('oto$_i$) pagašti $\underline{t}_i$
the-man that-him$_i$ met-I $\underline{t}_i$
(the man I met)

b. *ha-'iša še- ('ota$_i$) pagašti 'et ha-'iš$_j$ še $\underline{t}_j$
the-woman that-her$_i$ met-I acc the-man$_j$ that $\underline{t}_j$
ra'a $\underline{t}_i$
saw $\underline{t}_i$
(the woman that I met the man who saw her) (CNPC violation)

(45) a. ha-'iš še-'it-o$_i$ rakadti $\underline{t}_i$
the man that-with-him$_i$ danced-I $\underline{t}_i$
(the man with whom I danced)

b. *ha-'iša še-'it-a$_j$ ra'iti 'et ha-'iš$_i$ še-$\underline{t}_i$
the-woman that-with-her$_j$ saw-I acc the-man$_i$ that-$\underline{t}_i$
rakad $\underline{t}_j$
danced $\underline{t}_j$

(46) a. ha-'iš še-'et 'axot-o$_i$ ra'iti t_i
the-man that-acc sister-his$_i$ saw-I t_i

b. *ha-'iš še-'et 'axot-o$_i$ ra'iti 'et ha-kelev$_j$
the-man that-acc sister-his$_i$ saw-I acc the-dog$_j$
še t_j našax
that t_j bit

(47) a. ha-'iš še-ra'iti 'oto
the-man that-saw-I him

b. ha-'iš$_i$ še-pagašti 'et ha-'iša$_j$ še-t_j ra'ata 'oto$_i$
the-man$_i$ that-met-I acc the-woman$_j$ that$_j$ saw him$_i$

(48) a. ha-'iš še-rakadti 'it-o
the-man that-danced-I with-him

b. ha-'iša še-pagašti 'et ha-'iš še-rakad 'it-a
the-woman that-met-I acc the-man that-danced with-her

(49) a. ha-'iš še-ra'iti 'et 'axot-o
the-man that-saw-I acc sister-his

b. ha-'iš$_i$ še-ra'iti 'et ha-kelev še-'axot-o$_i$ 'imca
the-man$_i$ that-saw-I acc the-dog that-sister-his$_i$ adopted

In questions, however, only the movement strategy is possible, as we can see in (50)-(52).

(50) a. 'et mi ra'iti
acc who saw-I
(Who did I see?)

b. *mi ra'iti 'oto
who saw-I him

(51) a. 'im mi rakadti
with who danced-I

b. *mi rakadti 'it-o
who danced-I with-him

(52) a. 'axot mi 'imca kelev
sister who adopted dog
(whose sister adopted a dog)

b. *mi 'axot-o 'imca kelev
who sister his adopted dog

The fact that relative clauses allow resumptive pronouns whereas questions do not could be due to the fact that WH-words must have Case in Hebrew. Since there is no WH-word in the relative clauses in (47)-(49), there can be resumptive pronouns which bear Case. But in questions, if a resumptive pronoun was present, then no Case would be assigned to the WH-phrase, which would result in a violation of the principle of Lexicalization.

Consider now free relatives. Although they appear to have two options like regular relative clauses in Hebrew, Borer points out that there are two crucial differences between the two constructions. First, the resumptive pronouns in free relatives can only be clitics, and not free standing pronouns, as we can see in (53).

(53) a. ma šše-hexlatnu 'al-av
what that-decided-we on-it
(whatever we decided on)

b. mi šše-'axot-o mazkira ba-memšala
who that-sister-his secretary in-the-government

c. ma šše-raciti (*'oto)
what that-wanted-I (*it)

Furthermore, violations of the usual constraints on con-

structions involving WH-movement are impossible in free relatives, even if there is a resumptive clitic, as we can see by the ungrammaticality of the sentences in (54).

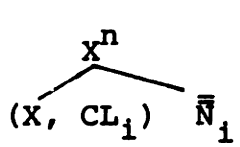
(54) a. *ma$_i$ š̌e-pagaš̌ti 'et ha-'iš̌ š̌e-hexlit 'al-av$_i$
what that-met-I acc the-man that-decided on-it
nimkar 'etmol
sold yesterday
(whatever I met the man who decided on it was sold yesterday.)

b. *keday le-hityaded 'im mi$_i$ š̌e-'e'evod be-misrad
worth to-befriend with who that-work-I in-office
š̌e-'axoto$_i$ menahelet
that-sister-his runs
(it is worth it to befriend a person whose sister runs an office in which I will work.)

What these facts lead one to assume is that free relatives are formed by movement only and that the clitics here are not "real" resumptive pronouns. To explain these facts, Borer (1981) makes the two following assumptions.

1° The constructions where a clitic shows up are clitic-doubling constructions as in (55) (for which Borer (1981) shows that there is independent motivation in Hebrew).

(55)



So the extraction site of the free relative WH-phrase is $\bar{\bar{N}}_i$ in (55).

2° Free relatives in Hebrew possess a mechanism which enables

the WH-NP to receive Case from the matrix in its landing site.

As Borer points out, these assumptions explain the difference between free relatives and questions, the latter not allowing such extractions from clitic-doubling constructions. Since Case is absorbed by the clitic in a structure like (55), the WH-word generated under the $\bar{N}$ position will not have Case. So unless it can get Case by some other means, it will be ruled out by the Case Filter, i.e. the principle of Lexicalization in the present analysis. If one assumes that there is a mechanism that can assign Case to the WH-NP after movement in free relatives but not in questions, then it follows that when Case absorbtion by a clitic takes place, only free relatives are grammatical. This explanation of the difference between free relatives and questions in Hebrew holds only if one assumes that the WH-NP itself must be Case marked. This is precisely the assumption that we make in our analysis of WH-constructions: the trace of the WH-NP is an EC because it has no ψ-features at PF since it is the WH-phrase itself that bears the Case. So this analysis of free relatives in Hebrew supports our general analysis of WH-constructions.[16]

3.3.2.2. French Relative Clauses[17]

In this section, we will look at relative clauses in colloquial French, using Québec French as an example. We will also compare relative clauses of colloquial French with those of standard French. We will show that, given the complementary approach to ECs presented in Chapter 2, a unique parameter emer-

ges which explains the difference between the two types of relative clause formation.

In standard French, relative clause formation always involves movement of a WH-phrase to COMP.

(56) L'homme $[_{\overline{S}}$ qui$_i$ $[_S$ $\underline{t}_i$ est venu ...

(57) L'homme $[_{\overline{S}}$ que$_i$ $[_S$ tu as vu $\underline{t}_i$ ...

(58) a. L'homme $[_{\overline{S}}$ à qui$_i$ $[_S$ tu penses $\underline{t}_i$ ...

b. L'homme $[_{\overline{S}}$ pour qui$_i$ $[_S$ tu as voté $\underline{t}_i$ ...

We will not give an exhaustive description of all facts related to relative clauses in Québec French: we will concentrate on the four main types of Québec relative clauses (henceforth QRC) as described in Lefebvre & Fournier (1978).[18] These four main types are illustrated in (59)-(62).

(59) Relativization of the subject:
L'homme qui est venu ...

(60) Relativization of the object:
L'homme que tu as vu ...

(61) Relativization of a "weak" preposition (<u>à</u>, <u>de</u>, ...)
a. Le gars à qui je pense ...
a'. Le gars que je pense ...
b. Le gars de qui je te parle ... [19]
b'. Le gars que je te parle ...

(62) Relativization of a "strong" preposition (<u>dessus</u>, <u>dessous</u>, <u>dedans</u>, <u>pour</u>, <u>contre</u>, <u>avec</u> ...)
a. Un gars sur qui je me fierais pas ...
a'. Un gars que je me fierais pas dessus ...
b. Le gars pour qui je vais voter ...
b'. Le gars que je vais voter pour ...

What we see from (56)-(58) and (59)-(62) is that the set of relative clause structures in standard French (henceforth SF) are a subset of relative clause structures in QRCs.

In the case of subject and object relativization, we can say, as a first descriptive attempt, that a WH-phrase is moved to COMP in both dialects, with the form qui if it is nominative, and que if it is accusative; or we could say that it is an empty WH-operator that is moved in COMP, and that the conjunctive complementizer que is obligatory in these cases, taking the form qui when it is in the environment of a WH-subject (cf. Kayne 1975, Pesetsky 1978).

In (61)-(62), we see that there are two possible strategies for relativization of a PP in QRC, whereas there is only one strategy for SF. In the first strategy, which is shared by both dialects, the whole WH-PP is moved to COMP, whereas in the second strategy, which only shows up in QRCs, the PP stays in the D-structure position. In this latter case, there are two different results depending on whether the preposition involved is "weak" or "strong". If the P is weak, then it does not show up on the surface. If the P is strong, it show up on the surface, but without an overt WH-NP complement; the strong P can vary in form with its weak form correspondent as sur/dessus, or it has an invariable form like pour.

This gives us a rough description of the facts. The difference between QRCs and SFRCs lies mainly in the fact that there are two possibilities of relativization in QRCs whereas SFRCs

have only one. Thus both dialects can form relative clauses by moving a WH-phrase in COMP, but only QRCs seem to have the additional possibility of relativization without movement of an overt WH-phrase. We must add to these data that in the cases where an overt WH-phrase is moved in QRC, there is the possibility of having a complementizer que which doubles with a WH-PP in COMP as in (63).

(63) a. Le gars à qui que je pense ...
b. Le gars de qui que je te parle ...
c. Le gars sur qui que je me fierais pas ...
d. Le gars pour qui que je vais voter ...

This optional que is not possible in SF.

Furthermore, there is another difference between the two dialects in that in QRCs, when there is no movement of an overt WH-phrase, there can be a pronoun which corresponds to the relativized position, as we can see in (64).

(64) a.??L'homme qu'il est venu ...
b.??L'homme que je l'ai vu ...
c. Le gars que je pense à lui ...
d. Un gars que je me fierais pas sur lui ...

This indicates that there is no movement in at least some of the cases where no overt WH-element is present.[20]

In the cases where WH movement does seem to take place, we will assume, in the spirit of Chomsky & Lasnik(1977), that relativization is not a specific transformational rule, but rather an instance of the general rule move α. The rule move α is op-

tional and it is factors from other components of the grammar which cause its application to be obligatory in some cases, like the need for Case marking of the lexical NP in passive constructions, for example.

Consider now the cases where relativization is done by movement of an overt WH-phrase, that is, the constructions that are common to QRCs and SFRCs. If these relative clauses are formed by the optional rule move α which would move the WH-phrase into COMP, we must assume that some factor forces the rule to be obligatory in relative clauses, since it is optional in questions, as we can see in the following sentences.

(65) a. Qui$_i$ as-tu vu $\underline{t}_i$?
b. Tu as vu qui?

(66) a. L'homme$_i$ que$_i$ tu as vu $\underline{t}_i$ ...
b. *L'homme$_i$ tu as vu que$_i$ ...

We can explain this difference if we assume that WH-elements are quantifiers of some sort. Since quantifiers are subject to a raising rule in LF (cf. May 1977, Aoun et al. 1981, Chomsky 1981a), they do not have to be moved by a syntactic rule since they will receive an adequate interpretation after this rule has applied in LF anyhow. So the LF structure of sentences like those in (65) would be identical after movement in LF (omitting non-relevant details).

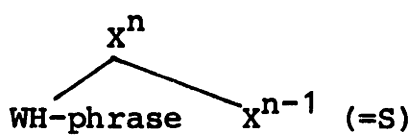
(67) [$_S$ pour quel $\underline{x}$ [$_S$ tu as vu $\underline{x}$]]

As for the relative WH-phrases, they are also some kind of

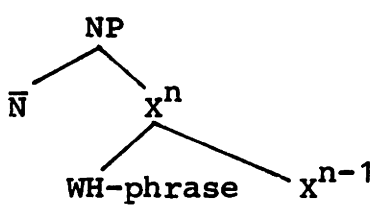
operators, but we could say that they are not subject to raising at LF', for example, since they are more closely related to pronouns or to anaphors than to operators since the variable that they bind takes the value of the element with which the relative WH-phrase is coindexed.[21] For example, in (66a), the variable $\underline{t}_i$ takes the value of l'homme since l'homme is coindexed with que, which is itself coindexed with $\underline{t}_i$. This coindexation is necessary or else the relative clause is not interpretable with respect to the rest of the sentence where it is inserted.

The obligatoriness of the syntactic movement of the relative WH-phrase could be due to the fact that the relation between the head of the relative clause like l'homme in (66a) and the WH-element que must be local in some sense, and that the COMP node is locally accessible. For example, if the structure after WH-movement is as we proposed in (40b) above, i.e. as in (68), we could assume that some features of the WH-phrase percolate to the X^n node, and hence are accessible to the head of the relative clause in (69).

(68)

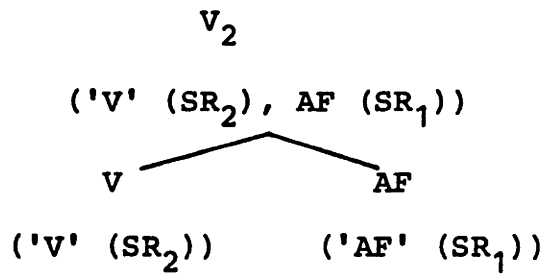


(69)



This means that in some cases, features would percolate in the syntax in the same way that it is assumed that they would percolate in morphology in Lieber (1980), for example (see also Marantz (1981)). Thus, all features of the head would normally percolate to the projection of the head, but some features from non-head nodes could also percolate when the head is not specified for these features. For example, in Marantz (1981), it is proposed that alternations in the expression of a verb's semantic dependents can be mediated by affixation of morphemes with independent argument structures. When this kind of merger takes place, the argument-taking properties of both the V and the AF may percolate to become the properties of the derived V as in (70).

(70)



(cf. Marantz 1981 for a detailed exposition of such an application of percolation theory)

There are independent reasons to believe that some feature percolation takes place between the WH-phrase and X^n in a structure like (68). For instance, Stowell (1981a) notes that a tensed $\bar{S}$ cannot occur in a position where Case is assigned. When a tensed $\bar{S}$ appears to be in such a position, i.e. when it is in object position or in the subject position of a tensed S,

Stowell provides evidence that the $\bar{S}$ is in fact in an extraposed position. He proposes to account for these facts by postulating a principle according to which a Case assigner cannot appear in a Case marked position, his Case Resistance Principle which we give in (71).

(71) Case may not be assigned to a category bearing a Case-assigning feature.

Assuming that tensed $\bar{S}$ is the projection of the category that assigns Case to the subject of the sentence, Stowell derives the distribution of tensed $\bar{S}$ from (71). He notes however that there seem to be exceptions to this generalization. Thus he observes contrasts as in (72).

(72) a. *We were talking about [$_{\bar{S}}$ [that] [we should help someone]]
b. We were talking about [$_{\bar{S}}$ who$_i$ [we should help $\underline{t}_i$]]
c. We were talking about [$_{\bar{S}}$ who$_i$ [$\underline{t}_i$ should be allowed $\underline{t}_i$ to come]]

So although normally a tensed $\bar{S}$ cannot appear as object of a PP according to (71), this is possible if the tensed $\bar{S}$ is an indirect question as in (72b,c), i.e. if there is a WH-phrase in COMP. Moreover, this is possible only if the WH-phrase is an NP, not if it is a PP as we can see in (73).

(73) a. *We were talking about [$_{\bar{S}}$ to who$_i$ [we should speak $\underline{t}_i$]]
b. We were talking about [$_{\bar{S}}$ who$_i$ [we should speak to $\underline{t}_i$]]

What seems to be going on is that some feature allowing $\bar{S}$ to bear Case seems to percolate from the WH-phrase, presumably

some [+N,-V] features, although the WH-phrase does not head the $\bar{S}$.[22] Note that the Case is not down-percolated to the WH-NP since there is no matching effect (cf. (72c)) and the WH-NP must get Case from some position inside the $\bar{S}$. So it is only a part of the features of the WH-NP that percolate to the $\bar{S}$ and seem to neutralize the Case-resisting properties of the tensed $\bar{S}$. The fact that feature percolation is only partial can also be seen by the fact, noted by Stowell, that the $\bar{S}$ in sentences like (72b,c) and (73b) does not have the categorial status of an NP since not all verbs that take them as complements allow concealed questions as we see in (74).

(74) a. I wonder what to do.
b. *I wonder the person he saw.

Summing up, we will assume that movement of the relativized element to COMP as in (69) is necessary for this element to percolate some feature(s) to $\bar{S}$ so that it becomes in some sense accessible to the head of the relative clause, and that this is what accounts for the obligatoriness of the movement in relative clauses, as opposed to questions where movement is optional (cf. the discussion of (65) above).

As for the reason why this WH-element is lexicalized, the simplest assumption is that it should depend on the same factors that account for the lexicalization of all other NPs, namely, the principle of Lexicalization: since the lexicalized relative pronoun has a WH-feature, we could assume that this feature is visible in PF, so that it forces the Lexicalization of the WH-

word. Since it is lexical, it follows from the principle of Lexicalization that this WH-phrase will have to have ψ-features, i.e. person, number, gender and Case, hence that it originated in D-structure in a position where Case was assigned since it cannot receive Case in its derived position in COMP.

Consider now the distribution of WH-phrases: in some cases, the WH-phrase is clearly overt since it has the form qui in subject position, or it is the object of a P in COMP after pied-piping. But when only que is present, it is not clear whether this que is a realization of a WH-phrase or not, since it could be the complementizer equivalent to that in English. Furthermore, there are no instances of an overt WH-phrase in QRCs that is not also part of the acceptable constructions in SF, aside from the doubled constructions in (63). So it seems that overtness of the WH-phrase is a property of SFRCs and that QRCs have an overt relative pronoun only in the constructions corresponding to the SFRCs. Let us look in detail at the constructions where there does not seem to be an overt WH-phrase in QRCs, these constructions being impossible in SF. Consider first (61a',b'), repeated here in (75).

(75) a. Le gars que je pense ...
b. Le gars que je te parle ...

In these constructions, the prepositions à and de, so-called weak prepositions, seem to have been deleted, and furthermore the WH-element is not lexicalized. We can suppose

that there is a non-lexicalized NP in the object position in (75a,b) and that these sentences have a D-structure as in (76).

(76) a. Le gars que je pense e
b. Le gars que je te parle e

In order for this EC to reach PF without being lexical, it must not bear any feature visible at that level of representation. Suppose that the EC does not have features of person, number and gender in D-structure: it will presumably get these by Agreement with the head of the relative clause in LF. We must then also conclude that the feature WH has not been inserted here since lexicalization is not forced: dialects of colloquial French would therefore have this optionality of insertion of a WH-feature in relative clauses which SF would not have. So lexicalization is forced in SF, but not necessarily in colloquial French where the relativized phrase can be an EC if the WH-feature is not inserted. In order for the EC not to be lexicalized in (76), it is also necessary that it avoids receiving a Case. We can now see the link that there is between the absence of an overt relativized phrase and the absence of Case marking in (76). Suppose that verbs like penser, parler are verbs that cannot assign Case to their object and which must do it by means of a Case marking preposition which is semantically empty. If a lexical NP is inserted in the object position of a verb of this type, then a preposition must be inserted in order for the NP to be assigned Case so that it will meet the

requirements of the principle of Lexicalization. But if an EC is inserted in such a position, then the preposition does not have to be inserted. Note that if a WH-element is inserted, it will have to be assigned Case since it will eventually be lexicalized since it has a visible feature in PF, i.e. WH: so the principle of Lexicalization forces the insertion of a preposition in such cases. Since the WH-phrase will be in a PP in such cases, the whole PP will have to be moved since there is no preposition stranding in French, i.e. no V-P reanalysis.

Consider (76) again. An EC is inserted in D-structure; it is not assigned Case, it is not in a PP: so it moved to COMP and does not have to be lexicalized since it does not have any feature visible in PF. That empty operator must be moved in order for the relative clause to be interpreted with respect to the rest of the sentence, as we have seen for sentences where the relativized element is lexical. Finally, since the dummy preposition <u>à</u>, <u>de</u> can be optionally inserted, it could be inserted even if it is not a WH-element that is in object position. In such cases, we get sentences like those in (77).

(77) a. Le gars que je pense à lui ...
b. Le gars que je t'ai parlé de lui ...

Since the NP is assigned Case here, the principle of Lexicalization forces the insertion of features of person, number, and gender, too.[23] In (77), the pronoun will have to be coindexed with the head of the relative clause in order for the re-

lative clause to receive an interpretation with respect to the whole sentence: this is the resumptive pronoun strategy, which is akin to Williams' (1980) notion of predication.[24]

Consider sentences where strong form prepositions have been inserted as in (62a',b'), repeated here in (78).

(78) a. Un gars que je me fierais pas dessus ...
b. Le gars que je vais voter pour ...

We could imagine that the analysis presented for constructions with weak prepositions could be adapted for constructions with strong prepositions. For example, we could say that the strong form of a preposition like dessus does not assign Case to its NP object, and that since the feature WH is optional in relative clauses in colloquial French, the object of dessus does not have to be lexicalized since it has no features visible at PF: so it is an EC. And then the interpretation of the relative clause forces the EC to be moved to COMP. But there is a hitch: in (76), the EC is in the object position of the verb and can be moved to COMP. But in (78), the EC is in a PP, which blocks its movement to COMP since PP is a bounding node and there is no reanalysis to allow preposition stranding in French.[25]

An alternative would be to consider the sentences in (78) as more closely related to the sentences in (77) than to those in (76), i.e. more like instances of the resumptive strategy.

This means that dessus would be a preposition that does not assign Case, and that it would have an EC object. This EC

could not be an anaphor, i.e. a trace, since it could not be Bound in our sense since PP blocks government.

The idea to consider (78) as analogous to (77) suggests that the EC would be a pronoun. There are reasons to believe that the resumptive pronoun strategy analysis is the right one since resumptive pronouns, contrary to traces, do not obey Subjacency. We can see that this is indeed the case in a sentence like (79a) where la fille is linked to the position avec+pronoun, in violation of Subjacency, whereas a violation of Subjacency when WH-movement has taken place results in ungrammaticality as in (79b).[26]

(79) a. La fille que je connais bien $[_{NP}$ le gars $[_{\overline{S}}$ qui $[_{S}$ sort avec EC]]]

b. *La fille $[_{\overline{S}}$ avec qui_i $[_{S}$ je connais bien $[_{NP}$ le gars $_{\overline{S}}$ qui sort $\underline{t}_i]]]]$

There are also independent reasons to believe that strong form prepositions are pronominal in some way since they can be used deictically as in J'ai mis le chat dedans/dessus, Je suis pour/contre.

Consider next the case of the relativization of the subject position.

(80) L'homme $[_{\overline{S}}$ que $[_{S}$ $\underline{x}$ est venu]]...

The $\underline{x}$ being in subject position here, it is assigned nominative Case. So it can be lexicalized as the clitic il as in

(81), with features of person, number and gender.

(81) L'homme $[_{\overline{S}}$ que$[_S$ il est venu$]]$...

A phonological rule will then derive the form in (82).

(82) L'homme qu'il est venu

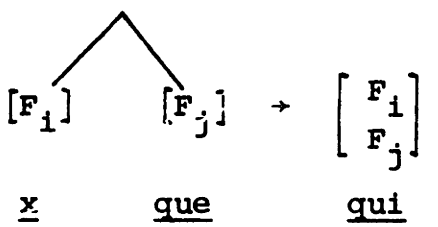
This is a very marginal construction in Québec French. Since the l of il tends to drop in general, we get the form [ki] in (82) and it is impossible to know whether the form is qui or qu'il underlyingly. However, the construction with a feminine pronoun que+elle=qu'elle allows us to see that the construction with a resumptive pronoun in subject position does exist in QRCs, but it still remains that the construction is extremely marginal. See the Appendix to this chapter for more discussion.

There is another possible derivation out of (80). The x can be inserted without a WH-feature. It gets nominative Case and is moved to COMP in order for proper interpretation of the relative clause to take place. So we have the elements $[_{[+NOM]}$X que$]$ in COMP. The rule of que-qui of Kayne (1975) can be interpreted as in (83).

(83) $[_{COMP}\ X_i$ que $I \rightarrow$ qui$_i$
$\quad\quad\quad

que combine into the element qui (cf. Pesetsky 1978). Another way to derive the effects of (83) would be to assume that there is percolation of the features of x and que to the node dominating these two elements and that this node is then realized as qui in PF as in (83').

(83')



The analysis of the relativization of the object position could be similar to the one of the subject position. So if an object position is relativized, it must be lexicalized since it is Case marked. If the NP has a WH-feature, then it is moved to COMP. It is overt because the feature forces lexicalization. If it does not have a WH-feature, then there are two possibilities. It is either lexicalized as a resumptive clitic pronoun, as in (84).

(84) La fille que Jean l'a vue.

Or it is moved to COMP and undergoes a rule similar to (83), rule (85).

(85) $[_{COMP}\ X_i\ \underline{que}] \rightarrow que_i$
$\quad\quad\quad

to the single parameter of having or not the optionality of inserting a Wh-feature in relative clauses: in SF, relative clauses must have a WH-feature, whereas such a feature is optional in QRCs.

If we assume that a WH-feature is "visible" in PF, this has for consequence that all relativized positions in SF must be linked to an element that has a visible feature in PF, i.e. WH, so that there will always be lexicalization of a relative element in this dialect. In colloquial French on the other hand, since the WH-feature is optional, it is possible to insert an EC in the relativized position, so that the relative element does not have to be lexicalized if the relativized position is not assigned a feature visible in PF by some other component of the grammar: this means that to have an EC, the relativized phrase must not be assigned Case, or else lexicalization is obligatory.

The analysis predicts that since SF must insert the feature WH, there will be only one possible form for each relativized position in this dialect, one with a WH-feature that is overt (cf. (59), (60), (61a,b), (62a,b)).[27] In colloquial French on the other hand, there are three possibilities: forms like in SF where the WH-feature has been inserted in D-structure (cf. (59), (60), (61a,b), (62a,b)); if the WH-feature is not inserted and if the position is not Case-marked, an EC is possible (cf. (61a',b'), (62a',b')); if the feature WH is not in-

serted and if the position is Case marked, then we get forms lexicalized by rules like (83) and (85) (cf. (59) and (60)), or resumptive pronouns (cf. (64)). In this last case and when the resumptive pronoun is an EC with a strong form preposition, since no element has been moved, there are no Subjacency effects. What is crucial here with respect to our general approach to WH-constructions is that it must be assumed that Case is assigned to the WH-phrase in order to account for the lexicalization or non-lexicalization of the element in the relativized position.

3.3.2.3. Problems.

We are now assuming that the trace of a WH-phrase is an EC for the same reason that any EC is not lexicalized: it has no ψ-features at PF. In order for this analysis to work, we made the natural assumption that some WH-phrases being NPs, they have to bear Case like any other NP, and that they carry Case along when they are moved, so that the trace of a WH-phrase has no Case at PF, and hence can be an EC by the principle of Lexicalization, since it bears no ψ-features at PF. The principle of Lexicalization has two kinds of effects for lexical items: on the one hand, if an element is lexical, it must bear ψ-features, and on the other hand, if an element bears a ψ-feature, it must be lexical. In this section, we are concerned with Case, and since it is a ψ-feature, we are assuming that, by the principle of Lexicalization, in order to be lexical, an

NP must bear Case on the one hand, and that if an NP bears Case, it must be lexical. Therefore, assuming that ψ-features other than Case are specified accordingly, there can be two types of empirical problems with which we can be faced with respect to Case: on the one hand, a lexical NP that seems to bear Case but yet is not well-formed at PF; and on the other hand, an NP that bears Case but is not lexicalized. The first kind of problem is found in some free relatives in Hebrew, where "Case from outside" does not seem to be sufficient for the WH-phrase to be lexical, as in (86).

(86) a. *mi_i še-'amarti le-Dan [$_{\bar{S}}$ [e_i] le-taken 'et ha-ke'ara]
who that-told-I to-Dan to-fix acc the sink

b. ma_i še-'amarti le-Dan [$_{\bar{S}}$ PRO li-knot [e]$_i$ la-tinok]
what that-told-I to-Dan to-buy to-the-baby

Recall that in 3.3.2.1. we assumed along the lines of Borer (1981) that there is a device to assign Case to the fronted WH-phrase in free relatives in Hebrew in order to explain the contrast between questions and free relatives. Yet this "Case from outside" is not sufficient for the WH-phrase since it seems that it must also get Case from inside the $\bar{S}$ in (86).

The second kind of empirical problem where an NP seems to bear Case and yet is not lexicalized is found in relative clauses formed with Ø-operators like in infinitival relative clauses as in (87), or tensed relative clauses in a few languages like English as in (88).

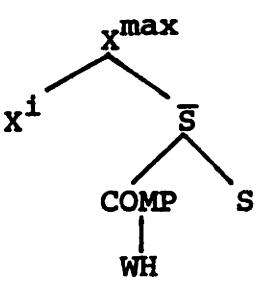
(87) John found $[_{NP}$ a book $[_{\bar{S}}$ $[\emptyset]_i [_S$ PRO to read $t_i]]]$

(88) a. The man John saw.
b. *The man you tried to win.

Let us examine these problems in turn. Consider first the problem of the Case assignment in free relatives in Hebrew. We have two problems facing us here in fact. The first one is why Case from outside is not sufficient for the WH-NP, and the second one is how Case is assigned from outside. We will see that solving this second technical problem will give us a clue to the solution to the first problem.

One way to have Case assigned to the WH-phrase in free relatives in Hebrew is to assume that the WH-phrase is in the head position of the relative clause (cf. Bresnan & Grimshaw 1978). In fact, if it is assumed that Case is assigned under government, then this is the only possibility, unless one assumes that in a structure for a free relative like (89),

(89)

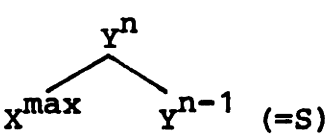


both maximal expansions X^{max} and $\bar{S}$ are for some reason transparent for government by a Case assigner in order for the WH-phrase to get Case. This is unlikely, especially if the X in (89) is N. Why wouldn't the Case be assigned to this N, rather than to

the WH-phrase? One could always say that the NP would have no head in (89) (cf. Groos & Riemsdijk 1979) so that there is nothing to bear the Case, but this is in contradiction with a strong $\bar{X}$ theory where nodes are labelled by insertion of lexical material under a head, hence must be headed. Furthermore, it is also in contradiction with the analysis of gerunds in Chomsky (1981a) where it is assumed that a gerund has the form $[_{NP*}$ NP VP], i.e. has no head for NP*, and yet it must bear Case according to Chomsky's Case Filter or Visibility condition since they are stated on NP, not on N.[28] Note also that our principle of Lexicalization predicts that the head cannot be empty in (89) and should be lexicalized since it receives Case.

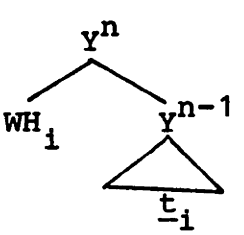
We have seen above that our analysis of core constructions involving WH-movement has led us to the assumption that COMP is an unlabelled position which is a sister to S, as in (90) (cf. 40a,b).

(90)



When WH-movement takes place, the WH-phrase is moved into the X^{max} position, so that the output of the rule is as in (91).

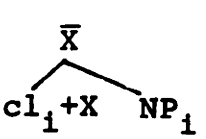
(91)



We have seen that in contructions like (91), some features from the WH-phrase can percolate to the node Y^n dominating it (cf. the discussion on percolation in the preceding section). In the cases that we have seen, only some weak form of percolation took place; for example, just enough features percolated to allow assignment of Case to the $\bar{S}$ dominating the WH-NP (cf. (72)-(74)). Suppose now that there can also be strong percolation in structures like (91) so that the WH-phrase percolates enough features to Y^n so that it can receive Case itself, i.e. Case from outside. This would have the nice property that the WH-phrase is in some sense the head of $\bar{S}$ although it has been moved there.

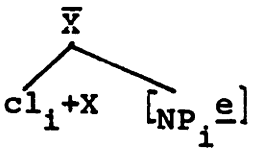
It is interesting to note when this is possible. According to Borer's data in (44)-(53) above, this is only possible in clitic-doubling constructions. So this head-like behavior of the moved WH-phrase is presumably related to some property of these constructions. What are the properties of clitic-doubling constructions? For one thing, the clitic must not absorb the θ-role, or else the full NP gets no θ-role, unless the clitic transmits its θ-role to the full NP in some way. Suppose that this is a property of clitic doubling: the full NP can Bind the clitic in (92) by assigning its R-index to it, since it governs the clitic, so that the clitic is like an object agreement element on the V. We will refer to this property as clitic Binding.

(92)



Of course, a discontinuous element (clitic, e) which bears a θ-role is still possible in these languages, so that there can be a construction as in (93) when there is no full NP Binder.

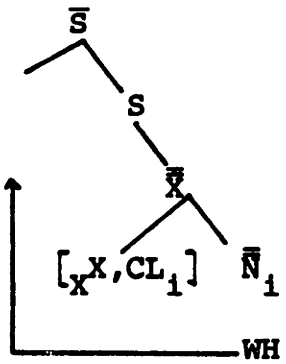
(93)



There is a second property which is necessary for the full NP to be lexicalized in clitic doubling constructions: The language must have a Case assigner that can be inserted to assign Case to the NP, since the clitic absorbs the Case that is normally assigned by the head. So Hebrew has *ʸsel*-insertion, Spanish has *a*-insertion, etc. Note that the two properties of clitic doubling constructions are independent, so that there could be constructions where only one property is realized. So for example, many languages have dummy Case assigners although they don't have clitic doubling. And we seem to get the reverse case for clitic Binding in Hebrew: in free relatives where the relativized position is inside a PP, there is clitic Binding, but no clitic doubling in the sense that there can be no lexical NP there since there is no dummy Case assigner available in this construction in Hebrew (cf. (53b)). However a WH-NP can appear in the doubled position if the WH-NP is moved to COMP where it can subsequently get Case.

The derivation that Borer (1981) assumes in the case of extraction from clitic doubling constructions is illustrated in (94).

(94)



We assume that once moved immediately under $\bar{S}$, the WH-phrase can undergo weak or strong percolation. If weak percolation takes place, no Case is assigned to the WH-NP at any point in the derivation: it can't get Case from outside because strong percolation has not taken place, and it can't get Case from inside because the clitic absorbs it. So it is ruled out by the principle of Lexicalization. If strong percolation takes place, then Case is assigned to the WH-NP which now heads the $\bar{S}$ in some sense, and we will assume that the WH-phrase also receives a θ-role as the head of the free relative. This being the case, the WH-phrase can only Bind a clitic in the sentence, and not an actual variable, since elements bearing θ-roles cannot Bind variables, as seems natural. On the other-hand, since a discontinuous element (clitic,e) can bear a θ-role independently, Binding of a clitic by the WH-phrase is not subject to this restriction.[29] This explains why it appears that Case from outside is not sufficient in sentences like (95).

(95) *mi$_i$ še-'amarti le-Dan [$_{\bar{S}}$ [e]$_i$ le-taken 'et ha-ke'ara]
who that-told-I to-Dan to-fix acc the-sink

In order to get Case from outside, strong percolation must take place, which has for consequence that the WH-NP also gets a θ-role from outside. But now if it Binds the EC in (95), the result is that a θ-bearing element Binds a variable, which we assume is impossible.

On the other hand, if weak percolation takes place in (95), the WH-NP does not get a θ-role from outside, but neither does it get Case; ans since it does not get Case inside the $\bar{S}$ either, it is ruled out by the principle of Lexicalization. Note also that Case from outside is never possible for questions since the WH-phrase is a real operator in questions, and hence cannot bear the θ-role that comes along with the Case from outside. It is also the case that Case from outside structures never involves non-clitic pronouns: the reason is that there would be no position from which the WH-phrase could originate. And the non-clitic pronoun can be bound as a resumptive pronoun only by something external to the $\bar{S}$ in Hebrew, never by a WH-phrase, either relative or question (cf. (47)-(52) above). Finally, the reason why the whole $\bar{S}$ must bear Case even if there is only weak percolation comes from the fact that an $\bar{S}$ with a WH-NP in its COMP loses its Case Resistance properties, as we have seen above in our discussion of data presented in Stowell (1981a) (cf. (72)-(74)).

So we see that an apparent problem to our claim that it is the WH-NP that bears Case, and not the trace of the WH-NP finds a

solution when the proper technical process of assignment of Case from outside is determined. We have given some possible way in which this can be done and which is consistent with our claim, and we will assume that such a solution to the assignment of Case from outside is workable. Thus the problem is no more an emipirical one, but a technical one with a suggestion of a solution. Note that if one drops the assumption that it is the WH-NP that bears the Case, then one loses the explanation for the contrast between questions and free relatives in Hebrew presented in section 3.3.2.1 above; one must also stipulate that WH-NPs are special NPs in that they do not have to bear Case in the same way that other lexical NPs have to.

Consider now the second type of problem that our analysis of Case assignment in WH-constructions faces, namely the fact that some Case marked relative operators are not lexicalized. Consider first cases like sentence (87), repeated here as (96).

(96) John found $[_{NP}$ a book$_i$ $[_{\bar{S}}$ $[\underline{e}]_i$ $[_S$ PRO to read $\underline{t}_i]]]$

In Chomsky (1981a), it is assumed that <u>a book</u> must bind $[\underline{e}]_i$ in (96) by a restriction that precludes vacuous quantification. So in (96), the variable takes the value of the element that binds the empty operator. Now consider the problem that (96) presents for our analysis, namely the fact that the operator $[\underline{e}]$ is not lexical in (96) although it comes from a Case marked position. One solution to the problem is to assume that the operator in (96) is merged with the head of the relative

clause, and that it is this combination of the head the book and the operator [e] that Binds the trace here. If so, then we could say that the complex element the book + [e] realizes the Case of the trace, somewhat like the que-qui rule creates a complex Binder that realizes nominative Case. There are reasons to believe that these constructions do involve a somewhat special process. For instance, the merging of [e] and the head occurs only in infinitival relative clauses, and infinitival clauses are known to have weaker boundaries in many instances (cf. all the facts that fall under $\bar{S}$ deletion, which is essentially a mechanism to weaken the boundary of the infinitival clause, whatever the actual mechanism is). Although a similar process takes place in English in tensed sentences like (88a), this is a marked process of English and it is rare among natural languages: the usual situation in sentences like (88a) is to have an overt element. However, the phenomenon is much more widely spread among languages when the relative clause is infinitival, so the phenomenon is presumably related to the weak status of the boundary in infinitives.

Another indication that the Binder is special in sentences like (96) is that the construction does not seem to tolerate successive cyclicity as we can see in (97). (The choice of the verb is also quite restricted.)

(97) a. John found a book that he believes Mary to like t.
a'.*John found a book to believe Mary to like t.

b. John found a book that he convinced Mary to read t.
b'.*John found a book to convince Mary to read t.
c. John found a book that he convinced Mary that she should read t.
c'.*John found a book to convince Mary that she should read t.

Judgements are not always as clear as indicated here. Thus although all speakers find (97a') impossible, some speakers find (97b') better than (97a'). The reason for this, we think, is that (97b') can receive a grammatical reading with a purposive interpretation rather than a relative clause interpretation of the infinitive clause. This reading is not possible in (97a') however. The reading as a purposive stands out more clearly in (97c). It also stands out more clearly if we change the matrix verb as in (98).

(98) a. *John bought a book to believe Mary to read.
b. *John bought a book in order to believe that Mary read it.
c.??John bought a book to convince Mary to read.
d. John bought a book in order to convince Mary to read (it).

Note that there is no trace in the object position of read under the purposive reading since no extraction has taken place: read is used as an intransitive verb here. So the weaker rejection of (97b') could be due to an analogical association with the purposive reading of the sentence. What makes us believe that this is the case is that in French, this analogical reading is not available, and that the judgements are clearer for sentences like (97b'), as we see in the French equivalent in (99b).

(99) a. Jean a trouvé un livre à lire.
b. *Jean a trouvé un livre à convaincre Marie de lire.

The reason is that the purposive and relative clause constructions differ in French: the infinitival relative clause is introduced by à as in (99), but the purposive is introduced by pour, as in (100).[30]

(100) a. Jean a acheté un livre pour lire.
b. Jean a acheté un livre pour convaincre Marie de lire.

Because of this difference, the analogical reading for sentences like (99b) is not available, hence the clearer grammaticality judgements.

So it seems that the Binder in infinitival relative clauses is not a usual Binder, and we will assume that this special status is responsible for the fact that the Case appears not to be realized. One way to capture this is to assume a merger as described above, somewhat like the effects of que-qui, or rather like the que-que rule (85) in 3.3.2.2.[31]

Borer (1981) also presents sentences like those in (101) as problems for an approach to Case in WH-constructions where the WH-phrase bears the Case since it seems that Case is required even if no lexical WH-phrase is present.

(101) a. *The man$_i$ (that) you tried [t_i to win]
b. *The man$_i$ it seems [t_i to come]

But it is false to say that Case is required in this type of construction since we can find examples similar to these which

are commonly found and are grammatical, like the well-known case discussed in Chomsky & Lasnik (1977).

(102) a man$_i$ [$\underline{t}_i$ to fix the sink]

So it seems that some other explanation must be found for the ungrammaticality of the sentences in (121). The explanation probably lies in the fact that there is an intermediate Binder in the sentences in (101), you and it, whereas a man can directly Bind the trace in (102). This is especially likely if the relation of these special complex Binders to a variable is restricted, as we have seen in (97).

Consider now another instance of what seems to be a non-lexicalization of a Case marked NP, namely sentences like (88a), repeated here as (103).

(103) The man John saw $\underline{t}$

We have already said above that these constructions of English are marked since they are rare among languages.

The reason why English allows such constructions where Case does not seem to be realized in PF might derive from processes that took place at earlier stages of the language. Consider the following facts from Old English which are now familiar (cf. Allen 1977, Bresnan & Grimshaw 1978, Grimshaw 1974, Vat 1978): 1° At that stage of the language, stranding of a preposition in a relative clause was not possible when an overt WH-phrase was present, although it was possible if the relative clause was introduced by a complementizer of the that type, or

by no overt element; 2° an optional complementizer equivalent to modern English <u>that</u> was always possible in COMP, even if a WH-phrase was also present in COMP, so that the doubly filled COMP filter of Chomsky & Lasnik (1977) was not operative then. These observations are schematized in (104)-(107).

(104) This book of which (that) I make mention <u>t</u>

(105) *This book which (that) I make mention of <u>t</u>

(106) *This book of $[_{NP}\underline{e}]$ (that) I make mention <u>t</u>

(107) This book (that) I make mention of <u>t</u>

We will not discuss the latter fact, but rather concentrate on the former about stranding possibilities. One way to explain the contrast between (105) and (107) is to assume that at that stage of the language, reanalysis implied that no Case was assigned to the NP following the reanalyzed P. So since no Case is assigned to the trace in (105) and (107), the NP cannot be lexical by the principle of Lexicalization, hence the ungrammaticality of (105) where lexical <u>which</u> receives no Case, and the grammaticality of (107) where a Ø-operator is not lexicalized. As for (104) and (106), since there is no reanalysis that takes place, Case is assigned by the P, hence it forces the lexicalization of the NP operator.

There are reasons that lead us to believe that Case was not assigned in reanalysis constructions at that stage. First, stranding was not possible in questions in Old English; if reanalysis implied that Case was not assigned, then the NP could only

be a Ø-operator by the principle of Lexicalization, and Ø-operators are only possible in relative clauses, not in questions for semantic reasons: hence the impossibility of stranding in questions. The second reason has to do with the fact, pointed out to us by Paul Kiparksy (personal communication), that reanalysis in passive constructions came about in English only after reanalysis in WH-constructions was already well established in the language. If one assumes that passive morphology is only possible when there is a Case to absorb (see the analysis of Exceptional Case Marking in 5.2.2 for more discussion), and if reanalysis blocked Case assignment at that stage, this means that passivization could not apply in a construction where reanalysis applied since this blocking of Case assignment by reanalysis has a bleeding effect on passivization. In WH-constructions on the other hand, reanalysis was possible, except that blocking of Case assignment had for effect that the NP could not be lexical by the principle of Lexicalization, hence the ungrammaticality of (105) at that stage. But when reanalysis stopped blocking the assignment of Case, as we believe happened in the passage to Modern English with consequences that we will see shortly, then passivization could apply to reanalyzed constructions since there was now a Case to absorb for the passive morphology. So the change from Case blocking to non-Case blocking in reanalysis accounts for the fact that reanalysis in passive constructions came about after reanalysis in WH-construc-

tions.

Consider now the facts of Modern English in (108)-(111).

(108) This book of which I make mention t

(109) This book which I make mention of t

(110) *This book of e I make mention t

(111) This book (that) I make mention of t

If the difference between Old English and Modern English is a passage from reanalysis with a Case assignment blocking effect to reanalysis with an optional Case assignment blocking effect, then we have an account for these facts as well as the facts about the interrelation between reanalysis and passivization just discussed. Thus, in (108), the of in COMP assigns Case, forcing lexicalization of the NP. In (109), Case can be assigned even if reanalysis has taken place at this stage of the language, so that a lexical operator is possible since it can fulfill the requirements of the principle of Lexicalization. In (110), the of assigns Case to the NP in COMP, so that the NP cannot be an EC by the principle of Lexicalization, hence the ungrammaticality of the sentence. In (111), the option to block Case assignment after reanalysis was taken, hence the possibility of having a Ø-operator. So in Modern English, it is the option to block assignment of Case or not in reanalysis constructions which accounts for the fact that either a Ø-operator or a lexical WH-word is possible in relative clauses.

As for subjects and objects, we can assume the following

analysis. The Ø-operators in object constructions would come about when there is a vacuous application of reanalysis, so that Case is optionally not assigned, and so lexicalization of the NP is optional. But this vacuous application of reanalysis is not possible in the case of the subject since no reanalysis at all is possible between the subject and the V. This means that no Ø-operator at all should be possible when the subject position is relativized: indeed this is the case as we can see in (112).

(112) a. The man who t came for dinner was tall.
b. The man that t came for dinner was tall.
c. *The man t came for dinner was tall.

As we see in (112c), a Ø-operator is not possible here, the reason being that Case is obligatorily assigned here and forces the lexicalization of the NP. In (112b), we would claim that the that is the manifestation of the Case, so that a rule similar to the French que-qui rule has applied, except that it has no morphological effect (cf. Kayne 1975, Pesetsky 1978).

So that would mean that the violation of the principle of Lexicalization in constructions where an element seems to bear Case yet not be lexicalized is only apparent since we see that, in fact, the element in question does not bear Case, so that it does obey the principle of Lexicalization. The analysis is also backed by the facts about passivization and reanalysis.

What we see, therefore, is that the apparent empirical

problems that the sentences in (86)-(88) above seem to show for the present analysis of Case in WH-constructions are reduced to technical problems, with an indication of possible solutions. Therefore, considering the advantages of dealing with the trace of a WH-phrase on a par with other ECs, we will assume that these problems can be overcome in the line of the solutions briefly presented here. However, before going to other topics, we will discuss an alternative solution that has been presented to the problem of accounting for Case in WH-constructions, namely the Visibility hypothesis that Chomsky (1981a) developed out of an idea of Aoun (1979). Our purpose in presenting this alternative solution is to show that our analysis seems to be conceptually and empirically at an advantage over the best alternative solution to this problem that we know of in the GB framework.

3.3.2.4. The Visibility Hypothesis.

The general idea of the Visibility hypothesis is that features are visible only to rules of the component to which they are relevant (cf. Aoun 1979). So phonological features are visible in PF, whereas θ-features are visible in LF. One specific proposal that came out of this hypothesis is that it is possible to derive the Case Filter by a visibility condition on θ-role assignment in LF (cf. Chomsky 1981a). The idea is to assume that in order to get a θ-role in LF, an NP has to bear visible features, one of which is Case. If the NP has no Case, then it is not assigned a θ-role, and hence does not meet the requirements

of the θ-criterion. Thus the Case Filter can be derived. However, matters are not that simple since θ-roles are sometimes assigned to NPs that do not bear Case, like PRO for example. This led to a first approximation of the visibility condition on θ-role assignment that encompassed the restriction given in (113).

(113) Assign a θ-role to an A-chain if and only if it has Case or F-features (where F-features = person, number, gender)

In fact, one could assume that Case and F-features are visible in LF and simply state the condition as in (114).

(114) Assign a θ-role to an A-chain if and only if it has relevant features.

But if there is only one EC, as it seems conceptually right in view of the partitioning of the EC as discussed in 2.2.5, then all ECs have F-features, and as noted in Chomsky (1981a), the generalization in (114) or (113) must be considered as spurious. The condition has to be stated as in (115).

(115) Assign a θ-role to an A-chain if and only if it has Case or is headed by PRO.

This is a very strange condition, however, since it claims that the features visible in LF are the feature Case, and the property of being a pronominal anaphor, which is what PRO is in Chomsky's analysis. Note that it has to be this specific property of PRO that makes it stand out, since it cannot be any

specific feature of PRO since all NPs, lexical or not, are assumed to have person, number and gender at LF.

The general notion of visibility is obscured here. If to be visible is to have features that are relevant in a certain component, then all NPs should be visible at LF, regardless of their Case marking or their status with respect to PRO. The reason is that F-features are relevant at LF: this is clear by the agreement facts of deictic pronouns discussed in (9)-(13) of Chapter 2. So (115) has to be stated over and above the general notion of visibility and it must be a strict condition on θ-role assignment, not on visibility at LF. So an NP can be visible at LF if it has F-features, for example, and yet not receive a θ-role if it does not have Case and is not PRO.

The condition in (115) suffers from conceptual problems. First, it is not a unified condition, and it does not fit well in the general idea of visibility since it is not a statement on what is visible in a certain component of the grammar: as we have just seen, it is a statement over and above visibility. A second problem is that one might wonder why PRO, one of the manifestations of the EC, has such a special status. This also goes against the strong hypothesis that no statement should refer specifically to an EC, let alone a specific manifestation of the EC.

If we now consider PF, we see that Case is relevant in PF since it shows up on the surface in lexical NPs for example. F-features are also relevant in PF since they also show up on

the surface. But it is assumed in the Visibility hypothesis that F-features are not visible in PF, since they do not block contraction for example in the want to/wanna contractions (to which we return shortly), whereas Case is visible in PF since it blocks such a contraction. So the Visibility hypothesis as given in Chomsky (1981a) has for consequence that F-features, although they are relevant in both LF and PF, are not sufficient in LF to allow θ-assignment and are not sufficient in PF to block contraction. So relevance and visibility are now separated, which weakens the Visibility hypothesis.

There are also empirical problems with the notion of visibility as presented in Chomsky (1981a). First, since the Case Filter is now accounted for by a condition on the LF side of the grammar, this entails that there can be no Case assigner that is inserted only at PF, since the Case assigned in this fashion would not be visible at LF, hence no θ-role would be assigned to the NP. But consider the following Hebrew sentences from Borer (1981).

(116) a. Itamar nitbakeš le-hachir be-'eize 'irgunim
Itamar was-requested to-declare in-which organizations

hu haya xaver $[_{PP}\underline{e}]$
he was member

b. *Itamar nitbakeš le-hachir šel 'eize 'anašim
Itamar was-requested to-declare of which people

hu haya xaver $[_{\check{s}el\text{-}phrase}\underline{e}]$
he was friend

Why can't a šel-phrase be fronted in Hebrew, like other PPs as in (116a)? Borer claims that "the impossibility of extracting šel along with the fronted WH element follows from the fact that šel simply does not exist at that level of the grammar at which extraction takes place -- namely syntax" (Borer 1981, p. 121). So šel-insertion would take place in PF, contrary to what is predicted by the GB notion of visibility given in (115). The reason why šel cannot be inserted in (116b) is because the environment of the rule is not met.

Another empirical problem for this notion of visibility is that there are many instances of sentential complements which do not have Case, but yet have a θ-role. Typical examples are given in (117).

(117) a. John is proud that he succeeded
b. John is believed t to be intelligent
c. John's attempt PRO to finish on time

In (117), all the underlined sentential complements do not have Case, yet they bear a θ-role. The converse is also possible: there can be elements that bear Case but are not assigned a θ-role, as in (118).

(118) Es wird gelacht
it is laughed

In (118), es has Case, and yet it does not bear a θ-role or transmit Case to some element bearing a θ-role, so that the Case Filter cannot be derived from an LF condition on θ-role

assignment in such cases.

As pointed out by Borer (1981), the main purpose of the notion of visibility in GB is to block variables without Case so that, for one thing, variables will be visible in LF and PF. But there are quite a few cases of variables that do not bear Case. A first case is illustrated in (119).

(119) a man$_i$ [$\underline{t}_i$ to fix the sink]

Whatever the exact Binder is in (119), there is a variable which does not bear Case here.

Second, consider the cases of quantifier lowering discussed in May (1977) of which (120) is an example.

(120) Some senator$_i$ is likely $\underline{t}_i$ to speak at every rally.

Note in particular that this sentence can have the interpretations in (121).

(121) a. It is likely that there is a senator S such that for every rally R, S speaks at R.

b. It is likely that for every rally R, there is a senator S such that S speaks at R.

In these cases where May assumes that the interpretation is derived by quantifier lowering, it is the $\underline{t}_i$ of (120) that is the variable, i.e. a non-Case marked EC. Note however, that it is only the theorem (122) that is violated here, not the visibility condition (115) since the A-chain does bear Case, Case being assigned to the position where <u>some senator</u> has been raised

to in (120).

(122) [e] is a variable if and only if it has Case.

Third, recall that in relative clauses in colloquial French, some dummy Case marking prepositions can be omitted, so that the variable does not have Case, and yet it gets a θ-role (in our analysis, the relativized element does not get Case, so it is not lexicalized).[32]

(123) a. Le gars à qui je pense ...
a'. Le gars que je pense t
b. Le gars de qui je parle
b'. Le gars que je parle t

Finally, we have seen that extraction from clitic doubling constructions is possible in free realtives in Hebrew since the WH-phrase can get Case from outside.[33]

(124) a. ma šeV-hexlantu 'al-av
what that-decided-we on-it

b. mi šeV-'axot-o mazkira ba-memšaV-la
who that-sister-his secretary in-the-government

Here the variable is not Case marked, and yet it gets a θ-role. Note that the variable and the WH-phrase do not form an A-chain, so that it should be irrelevant whether the WH-phrase gets Case or not, according to the Visibility hypothesis. In fact, all clitic doubling constructions present a problem for the Visibility hypothesis. Consider the general structure of these constructions.

(125) $[_{\overline{Y}}\ [CL_i+X]\ P^*\ NP_i]$

In (125), the clitic absorbs the Case assigned by the head X, so that a dummy Case marker P^* must be inserted in order for the lexical NP to get Case. But given the Visibility hypothesis, there is no reason as it stands for the insertion of this P^* since a θ-role should be assigned anyhow since the clitic has Case. Thus a θ-role is assigned in (126), where there is no doubling, hence no P^*.

(126) $[_{\overline{X}}\ [CL_i+X]\ [_{NP}e]]$

Chomsky (1981b) acknowledges that this is a problem and he proposes that the clitic manifests Case in (126), but not in (125). This amounts to a descriptive statement saying that sometimes Case is visible in LF, and sometimes it isn't, which weakens the Visibility hypothesis. These facts present no problem for an analysis incorporating the principle of Lexicalization: the P^* must be inserted if the NP is to be lexical or else it would be lacking a ψ-feature since the Case normally assigned to the NP is absorbed by the clitic. We have seen in previous discussions how the other problems for the Visibility hypothesis presented here can be accounted for in our analysis.

It seems therefore that the fact that the relation between Case and variable described in (122) holds in many instances is simply accidental: the variable of a quantifier phrase is in a Case marked position in most cases because it is in the S-struc-

ture position of the quantifier phrase in most cases, and since quantifier phrases are NPs in these cases, they must receive Case in these positions in order to satisfy the principle of Lexicalization at PF. Similarly, WH-NPs must bear Case, so they must get it in some position where Case is assigned. Since Case is usually assigned in A-positions, WH-NPs get Case in D-structure in the position that will ultimately be that of the variable, except in rare instances like the Hebrew free relatives where a special device is available for the WH-NP to get Case in an $\bar{A}$-position, so that the variable does not have Case here. The fact that the WH-phrase bears the Case will force its lexicalization, although this lexicalization is sometimes not readily apparent because merger with the head of the relative clause obscurs it, for example, and it will be revealed sometimes by secondary effects, like the difference in the Binding possibilities (cf. the non-successive cyclicity in (97)).

A notion of visibility that relies only on Case is also going to run into problems in explaining why some other features seem to force lexicalization in some instances. Thus it seems that Case is not the only feature relevant for visibility since we have seen that a WH-feature forces lexicalization in some cases in colloquial French relative clauses. Similarly, it is implicitly assumed in Aoun et al (1981) that a [+WH] feature is visible since they propose the following filter:

(127) *$\underset{[+WH]}{\text{COMP}}$ unless it contains a [+WH] element.

We return to this topic in section 3.4.

We will also see in Chapter 4 that in Pro Drop languages, an F-feature can also force lexicalization so that if AGR cannot provide one of the F-features to the subject at LF, the NP will have to be lexical since the missing F-feature must be provided in order to meet the Agreement requirements at LF: hence the missing feature must be inserted at D-structure, hence be visible at PF, and it forces lexicalization.

Given the conceptual weakness of the Visibility hypothesis where relevance and visibility are separated both in PF and LF, and given the empirical problems presented above, one might wonder what this notion of visibility buys us, and what the cost is.

One thing that the visibility hypothesis does is that it allows us to say that variables are visible in PF, whereas non-Case marked ECs are not. Recall however that this position is weakened by the fact that, although not visible in the sense of the Visibility hypothesis, non-Case marked ECs are considered to have F-features in PF in Chomsky's(1981a) analysis, hence features that are relevant in PF. The distinction that the Visibility hypothesis makes between variables, i.e. Case marked traces, and other ECs, is relevant to account for the contrast in (128).

(128) a. Who_i do you want $[\underline{t}_i$ [PRO to see $\underline{t}_i]]$ (wanna=OK)

b. Who_i do you want $[\underline{t}_i$ to come] (*wanna)
[+Case]

c. You_i want $[[PRO_i$ to go]] (wanna=OK)

The simplest claim to make about the contraction rule operating in (128) is that it has the form in (129), and that its application is blocked for some reason in (128b).[34]

(129) want + to → wanna

The reason why the contraction does not take place in (128b) according to the Visibility hypothesis is that want and to are not contiguous at PF in (128b): there is a Case-marked trace which is "seen" by the rules of PF, and it prevents the application of (129). But note that it must be assumed that the PF rule (129) does not see the F-features of PRO or the trace in COMP. So granting that only the feature Case makes an EC block the application of (129), then we have an explanation for the contrast in (128).

But this is not the whole story. Thus Chomsky (1981b) notes that the contraction is blocked in (130b) just like in (130a).

(130) a. Who do you want [t to see Bill] (*wanna)
b. They want, to be sure, a place under the sun. (*wanna)

It seems that some structural relation between want and to is at play in (130b) and that it is not only strict linear adjacency that is involved in contraction as appears to be the case in (130a). So we get two different answers as to why (129) seems to fail in (130). In (130a), it is due to the fact that the Case marked trace is visible in PF, hence blocks the application of (129) since there isn't linear adjacency. In (130b), Chom-

sky (1981b) says that "the answer presumably lies in a proper understanding of the domain for certain types of phonological process, a notion that may in part be a reflex of syntactic properties of government and the like" (Chomsky 1981b, p. 44a').

We can try to understand what syntactic properties are involved by looking at other structures where the contraction is blocked. For example, (129) seems to fail also in (131).

(131) They [wish and want] to go. (*wanna)

Another way to try to understand what is going on is to compare these structures with other instances where syntactic properties are involved in determining the domain for certain phonological processes. One such instance that comes to mind is the case of "liaison" in French. It is well known that liaison is either obligatory, optional or impossible, depending on the syntactic context.[35] For example, consider the following facts.

(132) Obligatory liaison:

a.	très aimable	très A
b.	vous avez	vous V
c.	sont allés	sont V
d.	grand arbre	grand N
e.	des ennemis	des N
f.	chez elle	chez N

(133) Optional liaison:

a.	donnait un cours	donnait NP
b.	écrivait à Paul	écrivait PP
c.	parait aimer	parait S

(134) Impossible liaison:

a.	donnait des lunettes à Marcel	NP PP
b.	dire qu'on partait à Marie	$\bar{S}$ PP
c.	qu'on partait était clair	$\bar{S}$ VP
d.	les lunettes étaient là	NP VP

Manzini (1981) shows that the structures involved are of the following types.

(132') 

(133') 

(134') 

To account for the respective possibilities of applications of liaison and rules of that type, i.e. phonological rules applying across words and which are constrained by considerations of juncture strength, Manzini proposes the following universal condition:

(135) in $\underline{\alpha}]\,[_{\underline{\beta}}$ or $\underline{\beta}]\,[_{\underline{\alpha}}$

if $\underline{\alpha}$ and $\underline{\beta}$ c-command each other, then rule application is OBL;
if $\underline{\alpha}$ c-commands $\underline{\beta}$, then rule application is OPT.
(where c-command = $\underline{\alpha}$ c-commands $\underline{\beta}$ if and only if neither one dominates the other and the first phrasal node which dominates $\underline{\alpha}$ also dominates $\underline{\beta}$)

We could try to account for the contraction facts in (130b) and (131) by saying that there is a notion of juncture strength which is at play here, just like in French liaison. What could this condition be? Consider the relevant structures of (130b)

and (131), given in (136).

(136) a. $[_{VP}[_{V}$ want], $[_{\overline{S}}\ [_{S}$ PRO to be sure]], $[_{NP}$a place in the sun]]

b. $[_{VP}[_{\overline{V}}$ wish and want] $[_{\alpha}$PRO to go]]

One possibility is that want must govern to in order for the rule to apply. Thus in (136a), want does not govern to since $\overline{S}$ deletion does not take place, especially since the $\overline{S}$ is a parenthetical. In (136b), want does not govern to, regardless of what α is, since according to our notion of government, the governor must be an immediate constituent of a node that dominates the governee: this is not the case here since want is an immediate constituent of $\overline{V}$, which does not dominate to.

Suppose that there is such a condition of government of to by want, and that want identifies the to by assigning an index of some sort, say the index that is usually found in a verb's grid along with its Case assigning feature. So in cases where the contraction (129) takes place, this condition has to be met, so that want should govern to in (128a) and (128c) for example. Consider these cases in turn.

(137) You want $[_{S}$ PRO to go] (wanna = OK)

In (137), according to our analysis, $\overline{S}$ deletion has taken place and the subject you Binds PRO, hence the obligatory control. But then if $\overline{S}$ deletion has taken place, want governs to and there is no visible material between the two at PF since

PRO has no features at PF in our analysis: therefore, the structural condition on contraction is met. Consider now the second case.

(138) Who_i do you want $[_{\bar{S}}\ t_i\ _S$ PRO to see $t_i]$ (wanna = OK)

In (138), although PRO and the trace in COMP have no lexical content at PF in our analysis, the structural condition is not met since $\bar{S}$ blocks the government of to by want. Yet the contraction is possible. The reason is that (138) is not the right structure for this sentence, but it is rather as in (139).

(139) Who_i do you want $[_S$ PRO to see $t_i]$

We will see in 3.3.3 that successive cyclicity should be analyzed in a way similar to Kayne's (1981a) analysis: there is no trace in COMP in sentences like (139), and the local relation between the WH-phrase and the trace is mediated through another mechanism than traces in intermediate COMPs. So if the structure is as in (139), then the structural condition on contraction is met: want governs to, and there is no intervening material which is "visible" at PF since PRO has no ψ-features at PF.

Consider now (128b). In order for the impossibility of contraction to count as an argument for the Visibility hypothesis where it is assumed that a Case marked trace is visible at PF, it must be the case that all other conditions on contrac-

tion are met in (128b), so that it is really the Case marked trace that is blocking contraction here, not some other property of the construction.

One such condition is met in (128b): there is no other lexical material between want and to, so that aside from the trace, the strict adjacency condition is met.

Note that the position of the trace must be assigned Case in (128b), so that the variable can get a θ-role in the Visibility analysis, so that the WH-NP meet the requirements of the principle of Lexicalization in our analysis. We will see in our analysis of ECM in 5.2.2 that want is not an ECM verb, which are verbs that assign Case directly to the NP subject of the embedded infinitival clause. In the case of want, Case assignment is done by the Ø-preposition in COMP so that there is a difference in structure between a believe-type construction and a want-type construction.[36]

(140) a. believe $[_S$ NP ...
b. want $[_{\bar{S}}$ Ø $[_S$ NP ...

Returning to (128b), since there is extraction of a WH-phrase from the subject position, Case must be assigned to the NP in subject position for the WH-phrase to get Case, hence the structure must be as in (141), where the Ø in COMP is a Case assigner.

(141) Who do you want $[_{\bar{S}}$ Ø $[_S$ t to come]]

But then the structural condition on contraction is not

met since $\bar{S}$ deletion does not take place: hence want does not govern to, and contraction is impossible, although there is no material visible at PF that is intervening since Ø and t have no ψ-features at PF (or else they would be lexicalized). So all the cases where contraction is blocked can get the same explanation, a structural one, and structural conditions are independently known to be involved in phonological rules applying across words, as the data on French liaison show. Thus the Visibility hypothesis, which we have seen to be a conceptually and empirically weak hypothesis, also loses its motivation with respect to contraction facts.[37]

3.3.3. Successive cyclicity and Subjacency.

At the beginning of our discussion of the WH-constructions, we have assumed that there are core constructions involving WH-movement and that the central property of these constructions is that the WH-phrase Binds its trace, this notion of Binding being independently motivated for anaphoric relations in the grammar. WH-constructions involving successive cyclicity, i.e. successive iteration of WH-movement, we claim, is a marked phenomenon. This has been recognized in the literature before. Consider the following passage from Koster (1978a).

> The whole idea of a CNPC only makes sense when we look at languages like English, and a few others. In a broader perspective, it appears to be based on the wrong assumptions. Ross's starting point was the existence of a class of unbounded transformations, for which there seemed to be ample evidence in English. So, a logical question was: when are these unbounded processes impossible? What are the exceptions to unboundedness? The CNPC was one of the answers. But if one looks at other languages, the more

> appropriate starting point appears to be just the opposite. Several languages have only bounded rules, and unbounded processes are very limited in all languages. Therefore, a better question seems to be: given the fact that rules are bounded, what are the apparent exceptions in languages like English? The answer is that extraction is only possible from complements of some verbs and adjectives (cf. Erteschik 1973). In other words, clauses are islands except complements of a subclass of categories of type [+V]. (Koster 1978a, p.43-44)

So bearing in mind that "long distance" dependencies are marked, we will expect this markedness to be reflected in the analysis of the phenomenon, or else we would expect it to be quite frequent among languages and to be quite unconstrained in languages that allow it, which seems to be contrary to fact. Granting that it is marked, it still remains that the phenomenon of "long distance" dependencies must be accounted for.

After Ross' (1967) important exposition of empirical generalizations about the phenomenon, another step was made in Chomsky's (1973) Conditions on Transformations where it was shown that movement transformations do not differ with respect to boundedness. What appeared to be unbounded transformations like WH-movement could be reinterpreted as an iteration of a local rule, constrained like all other rules. Wh-movement could be a local rule applied repeatedly in long distance dependencies because sentences have an escape hatch, i.e. COMP, through which the WH-phrase could pass from sentence to sentence. Furthermore, it was shown that different conditions like the Complex NP Constraint, the WH-Island Condition and the Subject Condition could all be reduced to a general condition of

Subjacency. Subjacency can be stated as follows.

(142) Subjacency:
A cyclic rule cannot move a phrase from position Y to position X (or conversely) in:
$\ldots X \ldots [_{\underline{\alpha}} \ldots [_{\underline{\beta}} \ldots Y \ldots] \ldots] \ldots X \ldots$
where $\underline{\alpha}$ and $\underline{\beta}$ are cyclic nodes (= S or $\overline{S}$, and NP)

If we take the option that S is a cyclic node, this gives us the four following configurations where movement is impossible.

(143) a. $\ldots X \ldots [_S \ldots [_S \ldots Y \ldots] \ldots] \ldots$
b. $\ldots X \ldots [_S \ldots [_{NP} \ldots Y \ldots] \ldots] \ldots$
c. $\ldots X \ldots [_{NP} \ldots [_S \ldots Y \ldots] \ldots] \ldots$
d. $\ldots X \ldots [_{NP} \ldots [_{NP} \ldots Y \ldots] \ldots] \ldots$

The facts from the CNPC fall under (143b) as we see in (144).

(144) a. *Who$_i$ [S did John mention [NP the fact [$\overline{S}$ [$\underline{t}_i$ that] [S Bill saw $\underline{t}_i$]]]]
b. *About whom [S did they destroy [NP a book $\underline{t}$]]

The facts from the Subject Condition also fall under (143b):

(145) a. *Who [S did [NP a picture of $\underline{t}$] please Bill]
b. *Whose [S did [NP $\underline{t}$ picture] please Bill]

The WH-island facts also follow from Subjacency if the notion of cycle is restricted to strict cyclicity, and if one assumes that there cannot be a doubly filled COMP.

(146) *Who$_i$ [S do you know [$\overline{S}$ what [S $\underline{t}_i$ saw $\underline{t}$]]]

If a language takes the option that $\bar{S}$ is a cyclic node, then this predicts that sentences like (146) ought to be grammatical in such languages, which seems to be the case.[38]

(147) Por quien$_i$ dice que no recuerda nadie que rescate$_j$ habia pagado la empresa t_j t_i?
'For who do you say that nobody remembers what ransom the company had paid.'

But as noted by Koster (1978a), this analysis where $\bar{S}$ is a bounding node makes wrong predictions with respect to extraction from NPs, since it predicts that sentences like (144b) should be grammatical in such languages, which is contrary to fact.

(148) *$[_{\bar{S}}$ [de qui]$_i$ $[_S$ ont-ils détruit $[_{NP}$ un livre t_i]]]

We return to this problem below.

One problem with Subjacency as it is stated in (142) is that it does not account for the fact that extraction out of a PP is impossible in most languages that have long distance movement. One way to correct this is to assume that PP is also a bounding node, and that something allows the PP node to be circumvented in languages like English where preposition stranding is possible. A suggestion is that reanalysis takes place in such cases, as we saw in the discussion of what we consider to be the core constructions involving WH-movement. So the PP node is not present after reanalysis, thus allowing extraction as in (149).[39]

(149) $[_{\overline{S}}$ what $[_S$ did you $[_{VP}[_V$ talk-about] $\underline{t}]]]$

The fact that not all PP nodes are permeable to WH-movement also suggests that PP is a bounding node even in languages where preposition stranding is possible, since PP blocks movement if reanalysis does not apply as we see in (150).

(150) *$[_{\overline{S}}$ what $[_S$ did you sing $[_{PP}$ during $\underline{t}]]]$

One might ask if AP isn't also a bounding node. Unfortunately, this is difficult to test since APs are usually embedded under NP nodes, so that this NP node could be responsible for the bounding effects. Adjectives can also appear in copular constructions, and extraction is possible in such cases (assuming reanalysis of the preposition with the adjective).

(151) a. John is proud of Mary.
b. Who is John proud of $\underline{t}$?

But adjectives enter quite freely into "small clauses" constructions (cf. Williams 1975, Stowell 1980a, Chomsky 1981a), so that the reason why extraction is possible in (151) could be that there is not a A^{max} node in the structure.[40]

Note that, for some reason unclear to us, it seems that pied-piping of the preposition is less than felicitous in sentences like (151b), as we see in (152).

(152) a.??Of who is John proud?
b.??About what is John worried?
c.??For what is John prepared?
d.??To who is John partial?

This could be related to the fact that the structures involved are small clauses, although it is not clear why these facts should be related. In any event, we see that the status of AP as a bounding node is unclear and we will leave it at that for the moment.

If we consider that long distance dependencies are a marked phenomenon, as we claimed at the outset of this section, then we might wonder at what is the parameter that distinguishes the languages that allow only local dependencies from those that allow long distance dependencies. It cannot be that the unmarked languages do not have an escape hatch, i.e. a COMP, since local movement to COMP is possible in those languages.[41]

One possibility could lie in the statement of Subjacency, more precisely, in the choice of bounding nodes. One could assume that the unmarked choice is to have both S and $\bar{S}$ as bounding nodes, in addition to the other bounding nodes like NP, PP and maybe AP. If both S and $\bar{S}$ are bounding nodes, then iteration of WH-movement is impossible, as we can see in (153).

(153) $[_{\bar{S}}[_{COMP}\ WH_i][_S \ldots [_{\bar{S}}\ [_{COMP}\ \underline{t}_i] \ldots \underline{t}_i \ldots$

So this choice of bounding nodes could explain why WH-movement is restricted in some languages, so that it cannot be iterated.

If we continue to examine the marked languages where long distance dependencies are possible, we expect that the difference between the marked and unmarked languages should be minimal, that

is, that the difference should only be in the choice of bounding nodes for example, the other properties of the marked WH-constructions following from other independently motivated principles of the grammar. For example, the fact that there could be no doubly filled COMP accounts for the facts in (146), repeated here as (154), if S, and not $\bar{S}$, is a bounding node.

(154) *who$_i$ [$_S$ do you know [$_{\bar{S}}$ [$\underline{t}_i$ what] [$_S$ $\underline{t}_i$ saw $\underline{t}$]]]

But this restriction against doubly filled COMP is not independently motivated and should follow from some other independently motivated principle, in an ideal hypothesis. One such candidate is the ECP, that we repeat in (155).

(155) ECP: A trace must be properly governed.
$\underline{\alpha}$ properly governs $\underline{\beta}$ iff $\underline{\alpha}$ governs $\underline{\beta}$, and
(i) $\underline{\alpha}$ is lexical (=X^0), or
(ii) $\underline{\alpha}$ is coindexed with $\underline{\beta}$

If elements inside doubly filled COMPs could not govern inside S, given a proper notion of government, then in (154), the trace in COMP cannot govern the trace in subject position, and the sentence is out by ECP, and the effects of a filter against doubly filled COMPs is derived from the ECP. A similar account can be provided for the (that-t) phenomena, which show a contrast between extraction from object and from subject position.

(156) a. Who do you believe [$_{\bar{S}}$[$\underline{t}$ that] [$_S$ Mary saw $\underline{t}$]]
b. *Who do you believe [$_{\bar{S}}$[$\underline{t}$ that] [$_S$ $\underline{t}$ saw Mary]]
c. Who do you believe [$_{\bar{S}}$ $\underline{t}$ [$_S$ $\underline{t}$ saw Mary]]

The sentence in (156b) is ungrammatical because the trace in subject position is not properly governed: there is no lexical governor available, and the trace in COMP cannot govern the trace in subject position. [42] In (156c) however, the trace in COMP can govern the trace in subject position, and the sentence is grammatical. In (156a), the trace is in object position, so it is properly governed by the lexical saw.

Note however that clause (ii) of proper government is motivated only for these cases of extraction from subject. So one cannot claim to have improved the grammar by deriving the doubly filled COMP filter if it is derived from another statement in the grammar that is not independently motivated. If, on the other hand, the ECP is interpreted as an identification principle along the lines of Stowell (1981a), so that verbs are also coindexed with their complements since they enter indices from their complements in their θ-grid, then a unified ECP can be formulated so that coindexing by a governor is an independently motivated notion in the grammar.

But note that the doubly filled COMP filter is derivable from the ECP only in the case of extraction from the subject position. In the case of extraction from object, for example, if the COMP is doubly filled as in (157), the ECP is still respected since the trace is properly governed by saw. Yet the sentence is ungrammatical.

(157) *The man $[_{\bar{S}}[_{COMP}$ who that$]\,[_{S}$ Bill saw t$]]$ ...

One way to rule out (157) could be to say that who does not bind the trace since it does not c-command it. These two ways of deriving the doubly filled COMP filter, by ECP for extraction from the subject position and by c-command requirements on binding for the extraction from object position, are very similar and suggest that some error might be lurking somewhere and that the phenomenon could presumably be dealt with in a uniform manner. Note also that the trace in COMP in (156a) is virtually without effect in the grammar since it cannot bind the trace in object position, or else there is no explanation for the ungrammaticality of (157), unless one reintroduces the actual filter against doubly filled COMP.

Returning to Stowell's θ-grid indexing, note that it also provides an explanation for the phenomenon of bridge verbs. It is well known that even in languages that do allow long distance dependencies, it is not all verbs that allow such dependencies. For example, verbs like murmur, shout, do not allow long distance dependencies.

(158) a. *Who did you murmur [t saw John]?
b. *Who did you shout [that John saw t] ?

Stowell (1980b) proposes that verbs like these do not strictly subcategorize for an $\bar{S}$ complement, despite the apparent fact that they assign a θ-role to their complement. Furthermore, Stowell assumes that the WH-phrase, or its trace, are taken as the head of $\bar{S}$ when they are in the complementized position. So

in (159) for example, the trace in COMP is properly governed by the matrix verb since it is in the head position of the $\bar{S}$, and since the matrix verb subcategorizes for an $\bar{S}$ complement, so that it is coindexed with this $\bar{S}$ complement by means of an index in its θ-grid.

(159) Who_i did you say $[t_i\ [t_i$ saw John$]]$

Thus according to Stowell's analysis, the trace in COMP is properly governed by <u>say</u> in (159) since it is governed by and coindexed with <u>say</u>, and the trace in subject position is properly governed by the trace in COMP.[43] In (158) on the other hand, the non-bridge verbs do not properly govern the trace in COMP since they do not subcategorize for an $\bar{S}$, hence are not coindexed in their θ-grid with these complements.

Suppose that some analysis of this type works. One must still account for the fact that the non-bridge verbs appear to assign a θ-role to their complements. Stowell (1980) makes the following proposal:

> It is a striking fact that most of the nonbridge verbs are manner-of-speaking verbs, such as <u>quip</u>, <u>whistle</u>, <u>shout</u>, <u>gurgle</u>, <u>laugh</u>, <u>exult</u>, <u>scream</u>, etc. Although this correlation may simply be due to some discourse-related functional principle such as Erteschik's notion of "dominance", one wonders whether some principle of core grammar might be involved (fn: for discussion, see Erteschik 1973). Notice that these verbs all intrinsically identify some aspect of the physical nature of their thematic objects. Thus "whisper" means "to utter a whisper-like sound", "shout" means "to utter a loud noise", etc. Suppose now that this property of identifying the nature of the thematic object within the lexical specification of the verb has the effect of absorbing the thematic object position, making it unavailable in principle for strict subcategori-

> zation. This would imply that the clausal complements of nonbridge verbs are not actually assigned a thematic role in the conventional sense, but rather are interpreted as adjuncts to the entire VP. Thus "John shouted to leave" would actually be interpreted as "John uttered a shout, conveying the message to leave"; "Bill whined that he was sick" would mean "Bill uttered a whine, to the effect that he was sick". (Stowell 1980b)

Stowell (1980b) also provides additional evidence to the claim that we must distinguish between thematic role assignment by bridge verbs and by nonbridge verbs. The first argument concerns the nominal counterparts to these verbs. The nominal counterparts of manner-of-speaking verbs do not refer to the action denoted by the verb, but rather to the verbs thematic object. So whereas the nominals derived from non-bridge verbs refer only to the physical utterance itself, the nominals corresponding to bridge verbs differ in that they refer to the propositional content of what was uttered. So for example, "John's claim" refers to the thing which John claimed, rather than to his act of claiming something. Stowell demonstrates this by showing that one can equate a bridge verb nominal with a propositional $\bar{S}$, as in (160), but that this is not possible for nonbridge verb nominals, as we see in (161).

(160) a. John's claim was that we should leave.
b. Bill's belief was that we would win.

(161) a. Jim's whine was very loud.
b. *Jim's whine was that we should leave.
c. *Tom's whisper was that he liked Sally.

Another argument presented by Stowell (1980b) to the effect

that the propositional $\bar{S}$ complements to nonbridge verbs are not directly linked with any thematic object position in these verbs' grids, but rather are more loosely linked with the verb, comes from differences in subcategorization for an NP complement. Almost all bridge verbs that strictly subcategorize for an $\bar{S}$ complement also allow an NP complement, as we see in (162).

(162) a. John remembered that he was sick.
John remembered his illness.
b. Bill pointed out that John had lied.
Bill pointed out John's dishonesty.
c. Jane explained how she had discovered the molecule.
Jane explained her discovery.
d. Susan knows that her boss was unfair.
Susan knows her boss' unfairness.

However, Stowell notes that nonbridge verbs do not allow NP complements as substitutes for propositional $\bar{S}$s.

(163) a. John whined that Sally had left.
*John whined Sally's departure.
b. Frank whispered that John had lied.
*Frank whispered John's dishonesty.
c. Phil screamed that his boss was unfair.
*Phil screamed his boss' unfairness.

> The correlation is not absolute, but it is significant enough to justify an account in terms of a theory of markedness. Suppose, for instance, that it is significantly less costly to associate the subcategorization features for NP with a given thematic position if the position is already associated with the subcategorization features for $\bar{S}$. This would make sense if there was a certain cost assigned to the attachment of a subcategorization frame to a thematic role, since it would be less costly to amend an existing subcategorization frame than to create a new one.
> (Stowell, 1980b)

Assuming this to be on the right track, Stowell's proposal has the nice effect that it accounts for bridge verbs with principles which seem to be independently needed in some form in the grammar. Furthermore, his notion of proper government is stated in a way that avoids the problems of the formulation in (155) where clause (ii) is not independently motivated, as we saw above.

However, there remain crucial conceptual problems with this analysis since it refers specifically to the EC and to some of its manifestations since it assumes ECP, control theory and the Pro Drop condition. This weakens a hypothesis which makes crucial use of ECs, like one which incorporates the Projection Principle, since the EC is not independently motivated in such an approach. It also misses the generalization that ECP, control and Pro Drop condition are three instances of a general process of recoverability, which is independently motivated for lexical anaphors, as we have seen in Chapter 2.[44]

Another account of long distance dependencies is presented in Kayne (1981a). Kayne proposes that instead of having one escape hatch as in analyses that stemmed out of Chomsky (1973), there are two escape hatches: successive cyclicity through COMP, as in these other analyses, and percolation projection. Kayne's notion of percolation projection comes from an intuitive approach to the ECP which he describes as follows:

> An empty category must have an antecedent; the antecedent may itself govern the empty category; if not, the empty category must, through its governor, be "closely connected" to the antecedent. (Kayne 1981a, p. 103)

Note that this intuitive approach to the ECP is very close to ours: the core notion is that the antecedent must govern the EC, i.e. Bind it in the sense of (2.78), although Kayne's government by the antecedent also allows successive cyclicity as in the analysis presented above, which is not the case for our analysis as we will see shortly.

Kayne formalizes this intuition as follows.

(164) Empty Category Principle
An Empty category $\underline{\beta}$ must have an antecedent $\underline{\alpha}$ such that (1) $\underline{\alpha}$ governs $\underline{\beta}$ or (2) $\underline{\alpha}$ c-commands $\underline{\beta}$ and there exists a lexical category X such that X governs $\underline{\beta}$ and $\underline{\alpha}$ is contained in some percolation projection of X.

(165) Percolation Projection:
A is a percolation projection of B iff A is a projection 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.

Kayne's use of percolation projection makes successive cyclicity unnecessary in sentences like (166).

(166) a. Who do you think that Mary saw $\underline{t}$?
b. Who did you vote for $\underline{t}$?

This analysis could be integrated with Stowell's (1980) analysis of bridge verbs without major modifications since both analyses make use of indexing complements by verbs. Kayne's analysis could also solve the problem for a successive cyclicity analysis shown in (167).

(167) Who does John believe $[_{\bar{S}}$ t $[_{S}$ Mary to have seen t]]?

In (167), Mary must receive exceptional Case marking from believe. But if WH-movement is successive cyclic, there must be an $\bar{S}$ here, so that government of Mary by believe is impossible, hence Case assignment is impossible, given the usual assumptions of GB. Kayne's analysis circumvents this problem since, by percolation projection, there is a possible derivation with no trace in COMP in (167), hence no $\bar{S}$, and ECM can take place as usual. Kayne's analysis also solves the problem raised by sentences like (168).

(168) What did you try $[_{\bar{S}}$ t $[_{S}$ PRO to do t]]?

In (168), the trace in COMP must be governed in order to satisfy the ECP. But if government is allowed over the maximal expansion $\bar{S}$, then PRO is also governed and the sentence should be ungrammatical, since it is a theorem of GB that PRO must be ungoverned, and Kayne assumes this analysis of PRO. In Kayne's analysis, there could simply be no trace in COMP here, given percolation projection.

Kayne's notion of ECP as a recoverability principle suffers from the same conceptual weakness that the successive cyclicity analysis suffers from: it covers only traces, and some other principles will have to account for the other manifestations of the EC, contrary to our approach where a unified analysis of recoverability of ECs is presented; and furthermore, our analysis extends to lexical anaphors, so it is independently motivated.

There is also an empirical problem with Kayne's analysis if one tries to extend it to cover WH-island phenomena. This problem stems from data recently analyzed by Torrego (1981).[45] Torrego shows that, in Spanish, when a certain class of WH-phrases (essentially, thematic arguments of the verb and the subject) are moved in the syntax, they trigger an obligatory rule of V-preposing, as we see in (169).

(169) a. Que querian esos dos?
'What did those two want?'
a'.*Que esos dos querian?
b. A quien presto Juan el diccionario?
'To whom did John lend the dictionary?'
b'.*A quien Juan presto el diccionario?

Furthermore, if there is successive cyclic movement of the WH-trigger, all the intermediate verbs must undergo this rule of V-preposing, as we can see in (170).

(170) a. Juan dice que los dos creian que Pedro habia pensado que el grupo habia aplazado la reunion.
'John says that the two of them thought that Peter believed that the group had postponed the meeting.'
b. Que dice Juan que creian los dos que habia pensado Pedro que habia aplazado el grupo?
c. *Que dice Juan que los dos creian que Pedro habia pensado que el grupo habia aplazado?

One might think that it is all verbs in the scope of such a WH trigger that must undergo V-preposing, but this is not the case: the trigger must be in or pass through the COMP of a verb's

$\bar{S}$ in order to trigger preposing, as we can see in (171).

(171) A quien (le) prometio Juan que Pedro se encargaria de que la gente sacara las entradas a tiempo?
'To whom did John promise that Peter would be in charge of people buying their tickets on time.'

Furthermore, this is also shown by the fact that it is not quite accurate to say that all intermediate verbs must prepose since, according to Torrego, the lowest verb is always free to be preposed or not. She attributes this to the fact that $\bar{S}$, and not S, is a bounding node in Spanish, so that the WH-phrase is free to pass through the lower COMP or not, as we can see in the schema in (172).

(172) $[_{\bar{S}}\ [\ ][_{S}\ldots[_{\bar{S}}\ [\ ][_{S}\ \ldots[_{\bar{S}}\ [\ ][_{S}\ \ldots$ wh-phrase $\ldots]]]]]]$

Note that the WH-phrase can only skip the first COMP, since as we can see by the dotted line in (172), a subsequent jump over a COMP results in a Subjacency violation since two $\bar{S}$ nodes are crossed. So only the first COMP can be skipped, and hence only the first verb is free to front or not.

Given these facts, an analysis like Kayne's where passage through COMP is optional since there is the additional escape hatch provided by percolation projection, makes the wrong prediction that V-preposing should always be optional, since there may or may not be a trigger in COMP. So the analysis fails on empirical grounds as it is. One possibility to save it is to assume that percolation projection is not available in Spanish,

but this would lead us far astray: all the problems solved by having percolation projection would have to be reconsidered under another light in Spanish. Instead, let us now consider what our analysis of long distance dependencies is and see how it accounts for the facts that we have discussed so far.

We have presented in some detail the analysis of Chomsky (1973) with the variants proposed in Stowell (1980b) and Kayne (1981a) because there are properties from each of these analyses that we would like to incorporate in our account of long distance dependencies. From Chomsky's (1973) proposal of successive cyclicity, we would like to retain the idea that the relation between the WH-phrase and the gap is local in the sense that it involves a succession of local relations. From Stowell's (1980b) proposal, we will retain the idea that bridge verbs incorporate the index of their $\overline{S}$ complement into their θ-grid, whereas nonbridge verbs do not. From Kayne's (1981a) analysis, we want to retain that there can be a succession of local relations without there being COMP to COMP movement.[46]

Recall that we have seen in 3.3.1 that the core constructions involving WH-movement fall under the general notion of Binding, i.e. government and coindexing by the WH-phrase. This means that Bounding theory is in fact subsumed by Binding theory. So there is no point in asking what are the bounding nodes in this approach: the bounding nodes are the "Binding nodes", i.e. they are the nodes across which Binding cannot take place, i.e.

they are the nodes which block government, i.e. they are the nodes of the form X^{max}: $\bar{S}$, NP, PP, and AP. So the reason why the WH-phrase and the trace cannot be related in structures like (173) in languages where only the core constructions involving WH-movement are found, is not that S and $\bar{S}$ are both bounding nodes: it is because the WH-phrase cannot govern into the embedded $\bar{S}$.

(173) $[_{\bar{S}}$ WH$[_S$ ...$[_{\bar{S}}$ $[_S$ ...$\underline{t}$...]]]]

And the reason why, although it is claimed that $\bar{S}$ is a bounding node in languages like French, not S, it is still impossible to extract out of an NP as in (174) is because NP blocks government of the EC by the WH-phrase.

(174) *De qui ont-ils détruit $[_{NP}$ un livre $\underline{t}$]

The same holds for PPs as in (175).

(175) *Qui Jean a-t-il voté $[_{PP}$ pour $\underline{t}$]

Binding of the EC by the WH-phrase will be possible in such constructions only if the language has a mechanism that eliminates the node that blocks government, like reanalysis in English for example.

Consider now a construction exhibiting the properties of successive movement as in (176).

(176) $[_{\bar{S}}$ Who $[_S$ did John say $[_{\bar{S}}[\underline{t}^3$ that$][_S$ Bill thought
$[_{\bar{S}}$ $[\underline{t}^2$ that$][_S$ Peter saw $\underline{t}^1$]]]]]]

We assume that the core notion of Binding and the notion of X^{max} as a barrier to government can never be tampered with in languages, i.e. that these notions are never parameterized in languages.[45] So in (176), $\underline{t}^2$ does not Bind $\underline{t}^1$, since it is not an immediate constituent of $\bar{S}$, hence does not govern $\underline{t}^1$. The trace $\underline{t}^3$ cannot Bind $\underline{t}^2$ for the same reasons. So these traces, even if they were present, could have no bearing on the relation between who and $\underline{t}^1$ if Binding is the relevant notion. Furthermore, there are reasons to believe that these traces are not present in intermediate COMPs: recall the cases involving ECM as in (177).

(177) Who does John believe $[_{\bar{S}}\ \underline{t}^2 [_S$ Mary to have seen $\underline{t}^1]]$

In (177), Mary must receive ECM from believe. But if there is a trace in COMP here, $\bar{S}$ blocks government and Mary cannot receive Case. So there must not be a trace in intermediate COMPs.

If the locality of the relation is not mediated by traces in COMP, then what could it be mediated by? The answer lies in a combination of the analyses of Kayne (1981a) and Stowell (1980b). Suppose that verbs do have a grid into which are entered the indices of their complements. If we assume that V is the head of $\bar{S}$, as we have all along, this means that the index that a matrix V enters into its grid when it has an $\bar{S}$ complement is the index of the V heading that $\bar{S}$ complement, as in (178).

(178) a. John thinks $[_{\bar{S}_i}$ that $[_{S_i}$ Mary $[_{VP_i}$ will come$_i]]]$

b. $\begin{bmatrix} \underline{\text{think}} \\ \underline{i}\text{=theme} \end{bmatrix}$

Now suppose that the embedded verb also has an $\bar{S}$ complement of its own. Then we might conceive of the indexing as in (179), where an index chain is created in the V-grid of think.

(179) a. John thinks $[_{\bar{S}_i}$ that $[_{S_i}$ Mary $[_{VP_i}$ said$_i$ $[_{\bar{S}_k}$ that $[_{S_k}$ Bill $[_{VP_k}$ came$_k]]]]]]$

b. $\begin{bmatrix} \text{think} \\ \begin{bmatrix} \underline{i}\ \text{(said)} \\ \underline{k}\text{=theme} \end{bmatrix} \text{=theme} \end{bmatrix}$

In (179b), we see that the index $\underline{k}$ is in the V-grid of said, whose index $\underline{i}$ is in the V-grid of think, so that $\underline{k}$ is in some sense in the V-grid of think. If we assume that in some languages percolation of some information is possible from one index level to another, then we capture the notion of percolation projection of Kayne (1981a). Furthermore, we see why nonbridge verbs appear to break percolation projection: as suggested in Stowell (1980b), they do not have an index for their $\bar{S}$ complement in their V-grid, so that an index chain is not established in these cases, as we see in (180).

(180) John thinks that Mary murmured$_i$ that Bill said$_k$ that Anne would come$_1$.

b. $\begin{bmatrix}\text{thinks} \\ \text{i=theme}\end{bmatrix}$ $\begin{bmatrix}\text{said} \\ \text{1=theme}\end{bmatrix}$

Since murmured does not have the index k of said in its V-grid, the index chain is broken.

So this formalism allows us to capture the central notion of percolation projection from Kayne (1981a) while incorporating Stowell's (1980b) account of Bridge verbs. Consider now a sentence like (181).

(181) John thinks that Mary saw_i Bill_k.

We assume, along with Stowell (1980b), that it is not only $\bar{S}$ complements that enter an index into a V-grid, but all complements. So the entry for think in (181) is as in (182).

(182) $\begin{bmatrix} \text{think} \\ \begin{bmatrix}\text{i (saw)} \\ \text{k=theme}\end{bmatrix} \quad \text{=theme}\end{bmatrix}$

we can now see how a local relation can be established between who and the trace in a sentence like (183).

(183) a. Who does John think $(_{\bar{S}_i}$ that$(_{S_i}$ Mary $(_{VP_i}$ saw_i t $)))$

b. $\begin{bmatrix} \text{think} \\ \begin{bmatrix}\underline{\text{i}} \text{ (saw)} \\ \text{[]=theme}\end{bmatrix} \text{=theme}\end{bmatrix}$

In (183), both who and think percolate features to $\bar{S}$ by the usual mechanisms of $\bar{X}$ theory where a head percolates some of its features to a higher node, so that they are in a close relation-

ship. Suppose that this allows who to identify the slot in the V-grid of think. Since the index of saw is in the V-grid of think, this identification is carried over to the slot in the V-grid of saw in languages that allow percolation of some information from one index level to another in the V-grid, i.e. if the language allows long distance dependencies. And since saw governs the trace in object position, we can assume that the slot in the V-grid of saw, being identified by who via an index chain, Binds the trace in object position. Thus a local relation is established between the WH-phrase and the trace. This relation holds as long as the index chain holds, i.e. as long as the chain is not broken up by a verb lacking an index in its V-grid like a nonbridge verb. So we now have a straightforward explanation for the bridge phenomenon. We also have an explanation for the possibility of ECM in sentences like (177): ECM is possible because there is no $\bar{S}$ since there is no trace in the intermediate COMP and the structure is actually as in (184).

(184) Who_i does John believe $[_S$ Mary to have seen $t_i]$

The relation between who and the trace is not mediated by an intermediate trace in COMP but by the index chain created in the V-grid of believe in the manner described above. Notice that our use of indices in V-grids is different from the use that Stowell (1980b, 1981a) makes of these indices. Stowell assumes that coindexing of the V-grid slot with the governed complement is the means to satisfy the ECP. But if we take a

very straightforward approach to the ECP and consider it as a recoverability principle, then it is not clear how the slot in the V-grid allows us to recover the content of the EC, unless one attributes something like clitic properties to these slots, assuming that they bear grammatical features for example. But this seems to be unwarranted since these slots can never license the presence of an EC on their own: there must always be some antecedent around to provide the EC with the proper grammatical features. This is what takes place in our analysis: the index in the V-grid functions as a link between the WH-phrase and the trace, but it cannot recover the content of an EC on its own. There is also a problem for Stowell's approach that shows up in sentences like (185).

(185) John seems $[_S$ t to be happy]

In Stowell's analysis, in order to satisfy the ECP, the trace in (185) must be governed by seems and coindexed with a slot in the V-grid of seems. Although there seem to be reasons to believe that the WH-phrase in COMP heads the $\bar{S}$ in some sense since it seems to percolate some of its features to $\bar{S}$ in some cases as we have seen in 3.3.2 above, it seems much less likely that the subject of the embedded sentence in (185) heads the sentence in any way, although this is the conclusion that is forced upon Stowell if the index of the trace is to be entered in the V-grid of seems.[47] In our analysis, John governs the trace directly in (185), so there is no need for seems to be co-

indexed with the trace in any sense.

We might assume that the index chains are easily identified by the fact that the V-grid slot which the WH-phrase ends up Binding is somehow marked by a WH-feature. This could be done at D-structure, i.e. at a point where the WH-phrase and the verb are still in a local relation. Thus a V-grid slot could be Bound by a WH-phrase only if the slot was identified as WH itself at D-structure. So it would therefore be impossible for a relation between a NP and a trace to be established by means of an index chain unless that NP was a WH-NP. Thus a derivation like the one in (186) is impossible.

(186) a. Mary seemed [e to be certain [e to be shot t]]
b. Marie semblait [e certaine [e d'être arrêtée t]]

As we have seen in 3.2 above, the agreement facts show that there is iteration of NP-movement in this case: the derivation must be successive cyclic, Marie being raised from subject to subject in (186). So this confirms that index chains are restricted to slots bearing a WH-feature in some sense.[48]

Consider the different positions from which a WH-phrase can be extracted. The argument positions under VP are all dealt with in the same manner that we dealt with the object position in (183) above: these arguments have a slot in the V-grid, so that the relation between a WH-phrase and these argument positions can be established via an index chain. We now understand why government by the V is related to recoverability: this struc-

tural relation is required for the index chain to be established, hence for the EC to be related to its antecedent, which reminds us of Kayne's intuitive description of the notion of ECP.

Next consider the subject position, which exhibits the familiar that-t effects as in (187).

(187) a. Who do you believe $[_{\bar{S}}\ t^2\ [_S\ t^1$ saw Mary$]]$
b. *Who do you believe $[_{\bar{S}}\ t^2$ that $[_S\ t^1$ saw Mary$]]$

In (187), since the trace in subject position is not an argument of the verb, it does not enter into the V-grid of saw. This means that the relation between who and the trace cannot be mediated by an index chain formed by entering the index of saw in the V-grid of believe since there is no slot for the subject in the V-grid of saw anyhow. If there is a trace in COMP, however, and if a phrase in COMP can percolate some features to $\bar{S}$ as we have seen above in 3.3.2, then we might assume that t^2 in (187) is accessible to the V-grid of the matrix verb believe: so who can Bind t^2 in (187a), and t^2 in turn can Bind t^1. The reason why (187b) is ungrammatical could have to do with the fact that when there is some other material in COMP like that, then the trace is not accessible to the V-grid of the matrix verb. But even if the trace was accessible in this sense, the problem with (187b) is that t^2 cannot Bind t^1, since t^2 does not govern t^1 since it is not an immediate constituent of $\bar{S}$. So we explain the that-t effect in a way similar to how Aoun et al (1981) do it for example: the trace in COMP cannot

Bind the trace in subject position if that is present. The reason why there is an asymmetry with respect to extraction between subject and object positions lies in the fact that the subject is not subcategorized by the verb, hence that its index does not enter into an index chain of which V is a part. In order to enter into such an index chain, the trace in subject position must be Bound by a trace in COMP position, which is accessible to the V-grid of the matrix verb.

Consider now the case where extraction takes place from a position which is not subcategorized for by the verb. One such instance is what is referred to as sentential adverbials, of which we see an example in (188).

(188) In which room did you say that (John saw Bill t).

We can assume that phrases like in which room in (188) are predicates which are associated to a main predicate to form a complex predicate. The structure of such a complex predicate can be schematized as in (189).

(189) $[_{\overline{X}}[_{X}\ ]\ Y]$, where X = the predicate formed by saw and Y = in which room.[49]

We could assume that it is the index of the complex predicate that enters into the V-grid of the matrix verb say in (188), so that this renders the index of the associated predicate, which is the trace t, accessible for Binding by in which room in a way that is similar to what we have proposed for complements under VP. There is a difference between associate

predicates like in which room in (188) and complements under VP however: the complements under VP are governed by the V that forms the predicate, so that their index is entered directly into the V's grid. This is probably the reason why there is a contrast between the possibilities of stranding prepositions in these two types of complements: stranding is more freely permitted for VP complements than for associate predicates, although stranding of a preposition in associate predicates is possible for some English speakers (cf. Rothstein 1981).

As for the reason why the set of possibilities of stranding a preposition in a passive construction is a subset of the possibilities of stranding a preposition in a WH-construction, it is that Binding in the WH-construction can be done via an index chain, which is not the case for the NP-trace Binder, since we assume that index chains are identified by a WH-index. So the relation between the verb and the stranded preposition does not have to be as "close" as in WH-constructions as in NP movement constructions.[50]

Another instance of a position that is not subcategorized for by the verb is an adjunct position as in (190).

(190) a. Bill looked at Paul, full of anger.
b. *What did Bill look at Paul, full of t?
c. *Of what did Bill look at Paul, full t?

As we can see in (190), WH-extraction from adjuncts is not possible. Adjuncts are different from associate predicates as

in (188) in that adjuncts do not participate in forming a complex predicate with the main predicate like associate predicates do, so that the index of an adjunct can never be part of an index chain since it is not accessible for the V-grid of any verb. This is why extraction from an adjunct is impossible.

We now have an analysis that accounts for long distance dependencies of subject position, of VP complement position and of non-subcategorized PP position, as well as an explaination as for why such dependencies are not possible for adjunct positions, from which no extraction at all is possible.

The effects of Subjacency are derived from Binding: when the relation between WH-phrase and trace is local, Binding precludes movement out of a X^{max}, which gives for example the effects of CNPC; when the dependency is not local, we assume that there is iteration of a local Binding relation by means of index-chains, but we must exclude cases that are WH-island violations. In an analysis based on Subjacency, WH-island phenomena are explained by the assumption that long distance dependencies are mediated by iteration of WH-movement from COMP to COMP, and by the assumption that a doubly filled COMP does not allow a proper relation between $\bar{A}$-binder and trace, either by assuming a filter that precludes doubly filled COMPs as in Chomsky & Lasnik (1977) for example, or by assuming that one of the $\bar{A}$-binders in COMP when there are more than one cannot Bind its trace since the proper c-command relation is not established, as in Chomsky

(1981a) for example. So in structures like (191) for example, only one element can Bind, hence the ungrammaticality of the structure.

(191) $[_{\bar{S}}\ [_{COMP}\ WH_i\ [_{COMP}\ WH_j]]\ [_S\ \underline{t}_i\ \ldots\ \underline{t}_j\ \ldots]]$

This is very similar to the <u>that-t</u> phenomena where the structure in COMP is as in (192), so that the operator cannot Bind its trace.

(192) $[_{\bar{S}}\ [_{COMP}\ \text{operator}\ [_{COMP}\ \text{that}]]\ [_S\ \ldots\ \underline{t}\ \ldots\ ]]$

So these ungrammaticalities find an explanation in a Subjacency analysis by postulating a structure in COMP and a structural relation between binder and bindee that restrict the number of binders to one for any given COMP. In our analysis, we will also account for WH-island phenomena by restricting the possibility of Binder to one. Since there is no COMP-to-COMP movement in our analysis, it is on the index-chain that the restriction must be made, not on the structure in COMP. So we will assume that only one index-chain can be passed on by a given V, so that a V can be a bridge for only one long distance dependency at a time, this restriction being the equivalent to the restriction that allows only one c-command relation to be established from COMP in a Subjacency analysis.

Assuming such an analysis of long distance dependencies, we can now return to the inversion facts of Spanish described by Torrego (1981) and see more precisely what is the trigger

for this inversion. According to Torrego, inversion is triggered by movement in the syntax of the subject and of thematic arguments of the verb. One way to capture this is to say that a verb will be preposed if it has an index in its V-grid that is Bound by a WH-phrase, i.e. a verb that is part of an index chain. The peculiar way in which a non-subcategorized WH-phrase enters its index into a V-grid by association would account for the fact that it does not trigger V-preposing.

This brings us to discussing the fact that, even with a proper trigger, the lowest verb is always free to prepose or not. As we mentioned above, Torrego attributes this to the fact that $\bar{S}$, and not S, is a bounding node in Spanish, so that the WH-phrase is free to pass through the lower COMP or not, as shown in (193).

(193) $[_{\bar{S}}[\ ][_{S} \ldots [_{\bar{S}}[\ ][_{S} \ldots [_{\bar{S}}[\ ][_{S} \ldots$ wh-phrase $\ldots]]]]]]$

According to Torrego, this is related to the fact that Spanish allows certain violations of WH-islands. As we can see in (194), a WH-phrase can always skip the lowest COMP and move directly to the second COMP since in doing so it only crosses one $\bar{S}$, in terms of Rizzi's (1978) analysis.

(194) Por quien$_j$ dice este que no recuerda nadie que rescate$_i$ habia pagado la empresa t_i t_j?
'For whom do you say that nobody remembers what ransom the company has paid?'

Notice that this kind of violation of WH-island, and hence

the optionality of V-preposing, can only take place across the most deeply embedded COMP and nowhere else in the iteration of the rule of WH-movement. In Rizzi's analysis, this is explained by the fact that any skipping of a COMP at a higher level provokes a Subjacency violation since two $\bar{S}$ bounding nodes would then have to be crossed, as we saw in the discussion of (172).

So we see that Rizzi's insight captures a good deal of what is going on. If we look at the facts in a slightly less theory oriented way, we see that extraction of a WH-phrase to the lowest COMP is always possible if that COMP is not already filled, and that in languages like Italian, Spanish, French, it is also possible to extract one more WH-phrase out of that lowest sentence. We have seen briefly how one can account for these facts in an analysis that incorporates a Bounding theory, essentially by adopting Rizzi's proposal of a parameter in Bounding nodes.

In our analysis, a WH-phrase can always move to the COMP of the lowest $\bar{S}$ dominating that phrase since the WH-phrase can always Bind positions in the S from COMP (except if there is a maximal expansion like NP or PP blocking government, of course).

(195) $[_{\bar{S}}$ WH $[_{S}$ ... $\underline{t}$ ... $]]$

Long distance dependencies are also possible when index chains are established as we saw in the discussion of (183) and the following above.

(196) $[_{\bar{S}}\ WH_i\ [_S \ldots \begin{bmatrix} V_k \\ [i] \end{bmatrix} \ldots [_{\bar{S}} \ldots V_k \ldots \underline{t}_i \ldots]]]$
[i]

So what is happening in languages like Italian, Spanish and French is that both strategies are used at the same time: one WH-phrase is moved to COMP where it locally Binds its trace, and another WH-phrase is moved out of the lowest $\bar{S}$ and related to its D-structure position by means of an index chain.

(197) a. <u>Por quien</u>$_j$ dice este que no recuerda nadie <u>que rescate</u>$_i$ habia pagado la empresa $\underline{t}_i$ $\underline{t}_j$?

b. $[_{\bar{S}}\ WH_j\ [_S \ldots \begin{bmatrix} V_k \\ \begin{bmatrix} 1 \\ [j] \end{bmatrix} \end{bmatrix} \ldots [_{\bar{S}}\ [_S \ldots \begin{bmatrix} V_k \\ [1] \\ [j] \end{bmatrix} \ldots [_{\bar{S}}\ WH_i\ [_S \ldots V_1$
[j]

$\ldots \underline{t}_i \ldots \underline{t}_j \ldots]]]]]$

This possibility of using both strategies at the same time would therefore account for the possibility of having WH-island violations in these languages, and it would also account for the quite restricted type of violations that are possible, that is, that the "extra" WH-phrase can only be in the lowest COMP so that there can only be two applications of WH-movement: one WH-phrase is moved to the lowest COMP, and one more is moved to a higher COMP. A third WH-phrase in COMP could not locally Bind its trace, and it could not be related to its trace by Binding in an index-chain since only one such index-chain can be formed at a time: so WH-extraction in syntax is restricted to

two phrases in these languages, one of which must be in the lowest COMP.

On the other hand, languages like English do not permit WH-island violations. This means that only one extraction strategy at a time can be used in English. This could be interpreted as being the parameter between the two types of languages: the possibility of having both extraction strategies or only one of them at a time.[51]

This analysis of these two types of languages is conceptually as valid as Rizzi's, if not more, since it makes do with Bounding theory by deriving it from Binding theory, so that there is only one general constraint on structural relations between elements, namely the constraint on government that no maximal expansion boundary intervene between governor and governee, hence between Binder and Bindee in the case at hand. The fact that there is only one constraint on structural relations is a factor that facilitates language acquisition.

Empirically, the present analysis explains all the facts that are explained by the Bounding theory analysis, and additionally, it accounts for facts about extraction from NP and PP in the Italian type languages that are not covered by the bounding theory analysis. Recall that extraction out of NP or PP is impossible in languages with $\bar{S}$ as a bounding node, although the WH-phrase crosses only one bounding node.[52]

(198) a. *De qui $[_S$ ont-ils détruit $[_{NP}$ un livre $\underline{t}]]$
b. *Quel film $[_S$ Jean s'est-il endormi $[_{PP}$ pendant $\underline{t}]]$

In our analysis, the sentences are ungrammatical because the NP and PP nodes block government by the WH-phrase, hence block Binding. The relation cannot be established by the index chain either since the V cannot govern the trace in both cases.

If we now return to the V-preposing facts, we note that V-preposing is obligatory in terms of the bounding theory analysis if a trigger WH-phrase or its trace is in the COMP of the $\bar{S}$ of the target verb. Translated into our analysis, this means that V-preposing is obligatory if the COMP contains a trigger WH-phrase or if the V enters in an index chain. Since both the WH-phrase in COMP and the V head the $\bar{S}$ in some sense, we can say that V-preposing takes place if $\bar{S}$ has a WH-index percolated to it, which can be either from the WH-phrase in COMP or from the head V when it enters in an index chain, so that the verb is attracted to $\bar{S}$ when it has a WH-feature matching one in the verb's own V-grid.

However, since there are two ways in which movement can take place out of the most embedded clause in the bounding theory analysis, V-preposing is optional in this lowest clause since there is a possible derivation in which there is no trigger for the rule in COMP. We can keep this analysis and duplicate its effects in our analysis of Italian type languages since we also acknowledge that there are two possible derivations for the extraction from the lowest clause: one which involves an index chain, and one which involves a more direct Binding. Con-

sider first the case involving an index chain.

$$(199)\quad WH_k \ldots \underset{\begin{bmatrix} j \\ \begin{bmatrix} i \\ [k] \end{bmatrix} \end{bmatrix}}{V^3} \ldots \underset{\begin{bmatrix} i \\ [k] \end{bmatrix}}{V^2_j} \ldots \underset{[k]}{V^1_i} \ldots \underline{t}_k$$

In (199), the fact that the WH-phrase identifies the slot $\underline{k}$ is simply carried through the index chain. As we saw above, indentification is done by percolating features from one index level to another in V-grids. Suppose that this indentification of slots from one link of the chain to another is optional. If it does not take place all the way from WH_k to $\underline{t}_k$ in (199), then the EC $\underline{t}_k$ would not be properly identified and the sentence would be ungrammatical. But if the index chain is not established between V^2 and V^1 in (199), we could assume that indentification is still possible by Binding here, the other way in which identification can be done. Thus, V^2 could Bind the slot in V^1 since V^1 is the head of the complement $\bar{S}$. On the other hand, V^3 could not Bind the slot $\underline{k}$ in V^2 since this slot is not readily accessible for Binding since it is embedded under $\underline{i}$.

Thus the fact that there are two ways of identifying the lowest V-grid slot mirrors the effect of having two ways of deriving the sentence in the bounding theory analysis. We could continue to assume that V-preposing takes place only when the V-grid is part of an index chain, so that V-preposing is obligatory for V^3 and V^2 in (199), but it is optional for V^1 since there is an alternative derivation for V^1 in which V^1 is not in

an index chain, as we can see in (200).

(200) a. A quien te imaginas que Juan ha dado el articulo?
b. A quien te imaginas que ha dado Juan el articulo?

We see that our analysis of the parameter that distinguishes between Italian type languages and English type languages with respect to bounding allows an account of the Spanish inversion facts with minimal additions to the grammar. Basically, all we have to say is that Spanish has a rule of V-preposing with a specific trigger: the rest follows from the analysis. More generally, we have seen that this analysis is conceptually interesting since it derives bounding theory from Binding theory, so that one does not have to wonder about what are the bounding nodes: they are the "Binding nodes", i.e. the nodes the block government, and this is independently motivated in the grammar. On the empirical side, the analysis provides an account for problems like the fact that extraction is impossible out of PPs and NPs even in Italian type languages (cf. (198)), and the fact that ECM can be assigned in sentences like (201) since no trace is present in COMP to block $\bar{S}$ deletion, hence block government.

(201) Who_i does John believe $[_S$ Mary to have seen $t_i]$?

It also integrates Stowell's (1980b) account of bridge verbs by means of entering the index of the $\bar{S}$ complement into the grid of the matrix verb.

3.4. ECP and LF movement.

In this section, we will look at some applications that have been made of the ECP to explain phenomena for which an analysis has been presented that makes use of movement rules taking place on the LF side of the grammar. There are two main rules of movement at LF that have been proposed in the literature: the rule QR which accounts for the scope of quantifiers with respect to one another (cf. May 1977), and the rule WH-R, which accounts for the scope of WH-operators in multiple WH-constructions (cf. Huang (1980), Aoun, Hornstein & Sportiche (1981). The study of these rules has yielded interesting results, although it is made difficult by the fact that judgements are not always sharp since there is no overt indication of the intended interpretation in most cases since LF movement has no overt effect on a sentence. Nevertheless, there are some clusterings of properties which are emerging from this study and we will present a brief overview of it and an idea of how it interacts with the framework proposed in this thesis.

Consider first the rule of QR. It applies to quantifier phrases, adjoining them to an S, and it gives their relative scope, as in (202).

(202) Three men saw four women.

a. $[_S$ [three men = $\underline{x}$] $[_S$ [four women = $\underline{y}$] $[_S$ $\underline{x}$ saw $\underline{y}$]]]

b. $[_S$ [four women = $\underline{y}$] $[_S$ [three men = $\underline{x}$] $[_S$ $\underline{x}$ saw $\underline{y}$]]]

In (202b), there are only four women involved, but in

(202a), there can be as many as twelve.

QR does not seem to obey Subjacency, as we see from the interpretations in (203).

(203) Three men thought that the idea of buying six cars was a bit crazy.

a. $[_S$[three men=$\underline{x}$] $[_S$ [six cars=y] $[_S$ $\underline{x}$ thought that the idea of buying $\underline{y}$ was a bit crazy]]]

b. $[_S$ [six cars=$\underline{y}$] $[_S$ [three men=$\underline{x}$] $[_S$ $\underline{x}$ thought that the idea of buying $\underline{y}$ was a bit crazy]]]

As for ECP, Picallo (1982) observed that it seems that LF movement of a quantifier is subject to ECP depending on the mood of the sentence. Her general observation is that ECP is observed if the sentence is in the subjunctive, but that it can be overcome if the sentence is in the indicative. Thus the underlined Q-phrase can have wide scope in (204), but not in (205).

(204) Tots els estudiants saben que <u>alguns examens</u> son dificils
IND

(All students know that some exams are difficult.)

(205) Tots els estudiants senten que <u>alguns examens</u> siguin
SUB
dificils

(All students regret that some exams are difficult.)

Similar observations hold for Focus constructions, which also involve LF movement, as we see in (206)-(207).

(206) Diuen que en JOAN arriba
IND

(They say that JOHN arrives.)

(207) *Sento que en JOAN vingui
SUB
(I regret that JOHN comes.)

Picallo also shows that there is a contrast that depends on mood in the case of negative quantifier-like NPs whose scope is determined by a pleonastic negative particle, as we see in the contrast between (208) and (209).

(208) ?En pere no creu que ningu no l'estima (IND)
(Peter does not believe that nobody loves him/her)

(209) En pere no creu que ningu no l'estimi (SUB)
*For no x, x= a person [Peter believes that [x does not love him/her]]

Picallo suggests that the contrast in LF extraction from subject position in IND and SUB clauses might be due to the fact that INFL has sufficient features when it is IND to be a proper governor, but not when it is SUB.

Consider now the rule of WH-R. Aoun et al. claim that this rule does not apply successive cyclically but that it is a one step process which moves a WH-phrase from an A-position and adjoins it to a [+WH]COMP. So WH-R does not obey Subjacency, as we see from the CNPC violation in (210) and the WH-island violation in (211).

(210) In order to foil this plot, we must find out which agent bats that are trained to kill which senator.
(from Hankamer (1974))

(211) Which men remember where to buy which books? (with the interpretation where which books is moved to the COMP containing which men.)

Aoun et al. conclude from these facts that Subjacency holds

only for syntactic movement, not for LF movement. However, they claim that ECP holds for WH-R. Thus the sentences in (212) are ungrammatical because, according to Aoun et al., WH-R being a one-step movement, there is no trace in COMP to properly govern the trace of the who in the subject position of the embedded clause.

(212) a. *Who expects that who saw John?
b. *Who expects who saw John?

Note that it should make no difference according to their analysis whether that is present or not since in any case there is no proper governor of the trace in subject position of the embedded S after WH-R has applied. But Aoun et al. mention in a footnote that there are structures like (212) which are judged better, depending on the choice of the matrix verb.

(213) No one can tell who thought who was a spy.

Furthermore, as observed by Chomsky (1981a), sentences containing multiple WH-words are always better if there are WH-phrases all across the sentence. Thus sentences like those in (214) are far better than (212).

(214) a. Who knows who put what where?
b. Who said who ate what?

The difference in judgement between (212) and (214) shows that the relative unacceptability of the sentences in (212) might be due to this fact, and that the ECP is not observed by

WH-R. Note that (215) is far worse than (214) however.

(215) a. *Who knows that who put what where?
b. *Who said that who ate what?

This could mean that the that-t phenomenon might be independent from the ECP, or at least should be viewed differently: for example, it could be related to the governing properties of INFL as suggested by Picallo. However, we would then expect INFL to be a proper governor in the indicative not just for QR and WH-R, but also for syntactic movement. But this does not seem to be the case as we see by the ungrammaticality of (216).

(216) *Who does John believe that t saw Bill?

If the effects of ECP and Subjacency are both derivable from Binding, as we suggested earlier in this chapter, then we expect them to pattern in a similar way: thus if Subjacency does not hold for LF movement, then ECP should not either. The facts are not altogether clear, but it is possible that Picallo is right in claiming that indicative sentences allow extraction from the subject at LF regardless of ECP violations. If so, then the question is why there should be a contrast between LF movement and overt movement in the syntax with respect to ECP and Subjacency. This is related to the claim of Aoun et al. that Subjacency holds only of syntactic movement, not of LF movement. The question again is why should it be so? If ECP and Subjacency are derived from Binding, and if Binding is a

way of identifying an anaphor EC, then the fact that the ECP and Subjacency does not seem to hold in these cases of movement at LF suggests that Binding does not take place in these cases, and the reason would have to be that Binding is not needed for identification of the variable at LF. One reason why this might be so could be that movement at LF is different from movement in the syntax. Movement in the syntax can be viewed as a rule that copies and then deletes the target phrase (cf. Chomsky (1955)). In LF, the rule must also copy the target phrase in order to provide the proper scope interpretation, but there is no reason to believe that the rule also deletes the target phrase: so movement in LF could be only copying without deletion. If so, then there is no need to identify the target phrase: it is not emptied of its features, so it is identifiable all along. So there would be no need for locality restrictions on LF movement since both the copied phrase and the target phrase are fully specified all along for F-features and R-index. Thus the result of movement of a WH-phrase is as in (217) if it is in the syntax, but LF movement is as in (218).

(217) $[_{\overline{S}}\ WH_i\ [_S\ \ldots\ \underline{e}\ \ldots\ ]]$

(218) $[_{\overline{S}}[_{COMP}\ WH_i\ [_{COMP}+WH]]\ [_S\ \ldots\ WH_i\ \ldots]]$

In order for the principle of Denotability to be observed in (217), the $\underline{e}$ must be provided an R-index and subsequently get F-features by Agreement. But in (218), the WH-phrase in situ meets all LF requirements since it is fully specified.

So the difference between syntactic movement and LF movement would be that the "trace" of LF movement is not an EC since when the LF structure is derived from the S-structure, there is full specification of the index and F-features of the "trace". Therefore, there is no need of a local relation between the operator and its "trace" since the content of the variable is fully specified: this is similar to the resumptive pronoun strategy which can be used in the syntax and where locality by Binding is not required. So if we see the ECP and Subjacency as instances of recoverability processes which are derived from Binding, we have an explanation for why they don't hold of LF movement: they don't have to since recoverability is possible anyhow. The same would be true of other LF movements like each-movement which we have seen must find an antecedent, hence has anaphoric properties, but does not have to have an antecedent that Binds it since the phrase from which each is extracted is fully specified at LF, as we see in (219).

(219) They saw each other.
They-each saw each other (after each movement)

This explains the contrast between reciprocals and true anaphors: reciprocals can have non-Binding antecedents (cf. the discussion in Chapter 2).[53]

So if we assume that ECP and Subjacency are derived from the recoverability process of Binding, then we have an explanation as to why they do not seem to hold in indicative sentences

at LF, although they hold for syntactic movement that is overt. Furthermore, if Binding is not involved in the relation between an operator moved at LF and its "trace", then nothing prevents multiple operators in COMP or adjoined to S since government is not involved in the structural relation between the two, whereas when Binding does hold, as in the case of syntactic movement, then if there are two WH-phrases moved to the same COMP, for example, one of them cannot Bind its trace if the structure is as in (220) since it cannot govern its trace.

(220) $[_{COMP}$ WH $[_{COMP}$ WH $]]$

As for the cases where the sentence is in the subjunctive mood, it might be that an explanation along the lines of Picallo (1982) is the solution here: proper identification of the subject position might be dependent in some way on the AGR features of the V.

Appendix: the Avoid Pronoun Principle and the Elswhere Principle.

Returning to the subject of relative clauses, it might be interesting to try to understand why resumptive pronouns are much less felicitous in subject positions and object positions in relative clauses, to the point that they are practically unacceptable as in (221) (cf. (64), (82), and (83) above).

(221) a.??La fille qu'elle est venue.
b.??La fille que Jean l'a vue.

Note first that movement out of these positions is always possible since there is no maximal expansion to block Binding by the moved phrase. So these facts could be related to the facts for which Chomsky (1981a) proposes the Avoid Pronoun Principle. Chomsky noted that in sentences like (222),

(222) John prefers [his going to the movies].

there is a very strong tendency to take his as disjoint in reference from John. His explanation is that another structure is possible with the same interpretation where PRO is in subject position of the gerund instead of his: hence the Avoid Pronoun principle, which, in Chomsky's analysis, states that lexical pronouns must be avoided when possible.

The same seems to be true in relative clauses. So constructions with pronouns like (223) are much less acceptible than forms like (224), where we could assume that a Ø-operator has moved to COMP.

(223) a.??L'homme qu'il est venu ...
b.??L'homme que je l'ai vu ...

(224) a. L'homme que est venu ...
b. L'homme que j'ai vu ...

So we could say that the Avoid Pronoun principle is operative here and that it is an empty operator that is moved in (224), although the rule of que-qui masks this in (224a).

Another fact that shows that the Avoid Pronoun principle is at work in relative clauses has to do with the "doubly filled COMP" constructions that we mentioned are possible in colloquial French (cf. (63) above). Notice that the examples that we gave all involve WH-phrases in PPs. If the relativized position is the subject or the object, however, these doubly filled COMP constructions are ungrammatical.

(225) a. *Le gars qui qui/que t est venu
b. *Le gars qui/que que j'ai vu t

This is not due to some filter which blocks doubly filled COMPs only if the WH-phrase is nominative or accusative, since the construction is possible in questions as in (226).

(226) a. Qui qui t est venu?
b. Qui que tu as vu t?

But Ø-operators are not possible in questions, whereas they are in relative clauses. So by the Avoid Pronoun principle, a Ø-operator is chosen in relative clauses in subject and object position. Note that this requires that the Avoid Pro-

noun Principle be extended to relative clause operators, which are not clearly pronouns.

We could say that the Ø-operator is in fact an anaphor, since it is Bound by the head of the relative clause, assuming that some of the features of the operator percolate to $\bar{S}$ since the operator is in COMP, and hence renders it accessible for Binding. So Binding inside $\bar{S}$, a maximal projection, is not possible, but Binding an element that percolates features to $\bar{S}$ from COMP is possible since the element then partially heads the $\bar{S}$, and heads are accessible for Binding.

Similarly in the case of gerunds like (222), the PRO is an anaphor since it is locally Bound (recall that PRO is an anaphor or a pronoun in our analysis of infinitival clauses, which we will develop in Chapter 5 and where we will see that this is also the case in gerunds). So the Avoid Pronoun principle could be interpreted not as "avoid a lexical pronoun" as in Chomsky (1981a), but rather as in (227).

(227) Don't put a pronoun in a position where an anaphor is possible, i.e. in a position where the pronoun will be interpreted as coreferential with an NP that can Bind it.

This principle can be seen as a subcase of an Elsewhere principle in the spirit of Kiparsky (1973, 1982). As we have seen in Chapter 2, anaphors have a more restricted domain than pronouns since their domain is determined by Binding; pronouns

on the other hand, are freely indexed in all other contexts, this being the general Elsewhere case.

The formulation of the principle as in (227) accounts for the facts about gerunds and relative clauses, given our assumptions. Thus a gerund can have an EC anaphor in subject position if it is sentential, since this position is then Bindable across an S node: so by the Avoid Pronoun principle in (227), the gerund must have an EC anaphor in (222). In relative clauses, if we assume that the WH-phrase is an anaphor and that the Ø-operator is also an anaphor, then these must be inserted instead of resumptive pronouns in subject and object position because these positions are always Bindable from COMP since no maximal expansion blocks Binding from COMP in these cases; and since the COMP position itself is Bindable by the head of the relative clause, this means that anaphors are always possible in these cases, hence by (227), are obligatory in these cases, which explains the unacceptability of resumptive pronouns in subject and object position.

Returning to the doubly filled COMP constructions, it is interesting to note that if the extracted element cannot be an anaphor, then doubling when extracting from subject and object position is possible. A first case of this type is when the position is questioned. The WH-phrase is not an anaphor and doubling is possible as we saw in (226) above. Doubling of this type is also possible in headless relative clauses, since

the element in COMP cannot be an anaphor, there being no head to act as a Binder of the Ø-operator. Thus sentences like (228) are grammatical in colloquial French, and so are some constructions of standard French as in (229).

(228) a. Qui qui a bu un jour va boire toute sa vie.
b. Qui que Jean a vu ce jour-là va toujours être un mystère.

(229) a. Qui que ce soit ...
b. Qui que tu aies vu ...

The formulation in (227) is also interesting in that it is in the spirit of our basic hypothesis that no specific statement should refer only to ECs, or only to lexical NPs: Avoid Pronoun as stated in (227) refers to anaphors in general, not just to ECs, and conversely, to pronouns in general, not just to lexical pronouns. So, for example, we can maintain that a single parameter differentiates SFRCs and QRCs, and that the anaphor in (224) does not have to be a Ø-operator: it can be a WH-operator too, for the SF constructions for example, since the Avoid Pronoun principle simply requires an anaphor, regardless of whether it is lexical or not.[54]

In fact, the formulation of the principle in (227) relates nicely to some facts about overt anaphors that we already have observed. First, it is related to condition (112') on Bound elements in Chapter 2, which accounted for the contrast in (230).

(230) a. *John shaved him.
b. John shaved himself.

Since the pronoun in object position is interpreted as co-referential with John in (230a), and since it is Bound by John, then the Avoid Pronoun principle as stated in (227) requires that an anaphor be inserted in object position, a lexical anaphor in this case.

A second case where (227) is also relevant is the constructions where reanalysis is possible as in (231).

(231) They spoke to themselves/*them.

We saw in Chapter 2 that, although reanalysis is usually optional, as is clear in (232),

(232) a. John was taken advantage of.
b. Advantage was taken of John.

reanalysis seems to be functionally forced to apply when there is intended coreference between the subject and the NP in the reanalyzed PP, as we see in (231). We now have an explanation for the obligatoriness of reanalysis in these constructions, namely (227). In (231), the pronoun them is interpreted as coreferential with they, and it is in a position where they can Bind it, provided reanalysis applies. So by the Avoid Pronoun Principle as stated in (227), an anaphor must be inserted instead of a pronoun, and so reanalysis must apply.

It is also most likely that the frequently observed facts about subjunctives in (233) follow from the principle in (227). (J. Guéron was the first to point out a possible link between these facts and the Avoid Pronoun principle.)

(233) a. *Je veux que j'aille voir ce film.
b. Je veux [PRO aller voir ce film].

If we assume that subjunctives and infinitives are closely related in their temporal interpretation in that both express "unrealized tenses" (cf. Bresnan 1972), then this means that the interpretation can be maintained in (233) by having an anaphor instead of a pronoun, i.e. an infinitive clause instead of a subjunctive clause since PRO is an anaphor in (233b) according to our analysis. (Cf. Chapter 5.)[55] In our analysis, the fact that PRO is an anaphor or a pronoun depends on whether the position can be Bound or not. This predicts that when PRO is a pronominal in our analysis, then a coreferent pronoun in an equivalent subjunctive construction should be possible since the distinction made by (227) is not lexical versus non-lexical, but rather pronoun versus anaphor. This prediction is borne out. We assume that in long distance control constructions like (234), the PRO is a pronoun not an anaphor since it is not Bound by its antecedent but freely indexed at S-structure (cf. Chapter 5).

(234) a. [PRO d'être menacé de mort] ne me fera pas changer d'idée.
b. Jean m'a dit qu'il serait possible [PRO d'être admis à l'académie] si nous en faisons la demande.
c. Il est préférable pour nous tous [PRO d'aller voir ce film]

As we can see in (235), coreferent pronouns in subject po-

sition are possible in the equivalent subjunctive constructions.

(235) a. [Que je sois menacé de mort] ne me fera pas changer d'idée.
b. Jean m'a dit qu'il serait possible [que nous soyons admis à l'académie] si nous en faisons la demande.
c. Il est préférable pour nous tous [que nous allions voir ce film].

What we see is that the formulation of the Avoid Pronoun principle as in (227) has the conceptual advantage that it is stated as differentiating between pronoun and anaphor, not between lexical and non-lexical pronoun; and we are assuming that this latter type of statement should not have a place in an ideal grammar. This formulation also has the empirical advantage that it brings together facts that are unrelatable under the lexical versus non-lexical formulation: the facts about gerunds (236), the condition (112') of Chapter 2 on Bound elements (237), the obligatoriness of reanalysis in some cases (238), the peculiarities of relativization when the subject and object positions are relativized with respect to resumptive pronouns (239) and que doubling (240), and finally subjunctives (241) and (242).[56]

(236) John prefers [his going to the movies] (John≠his)
(237) John shaved himself/*him.
(238) They spoke to themselves/*them.
(239) a.??La fille qu'elle est venue ...
b.??Le gars que je l'ai vu ...
(240) a. *La fille qui qui est venue ...

b. *La fille qui que j'ai vue ...
c. La fille avec qui que je suis venu ...

(241) a. *Je veux que j'aille voir Paul.
b. Je veux aller voir Paul.

(242) a. [PRO d'être menacé de mort] ne me fera pas changer d'idée.
b. [Que je sois menacé de mort] ne me fera pas changer d'idée.

Note finally that this approach to the Avoid Pronoun principle gives direct support to our analysis of Binding: the contrasts observed between anaphor and pronoun hold only if anaphors are functionally defined as elements that are governed by their antecedent, hence Bound in the sense of (2.78), pronouns being the Elsewhere case of free indexing, and if the difference between the two is not a strictly morphological one. This is especially clear in the case of like-pronouns in subjunctives and PRO of infinitives: in our analysis, PRO is either an anaphor or a pronoun, and the principle in (227) holds only in the cases where PRO is governed by its antecedent, hence where PRO is an anaphor, not a pronoun.

FOOTNOTES: CHAPTER 3.

1. Recall however the case of $PRO_{\underline{arb}}$ which has no F-features, even at LF. We return to this topic in Chapter 5.

2. For a discussion of the contrast between internal and external arguments, see Williams (1980), Marantz (1981).

3. However, we will see when we look at the analysis of locally controlled PRO in Chapter 5 that such an EC in a sentence like "John tried PRO to win" is in fact an anaphor, just like a trace, and that the "transmission" of a θ-role by a trace is simply to be interpreted as "discontinuous element with one θ-role" whereas a relation antecedent-locally controlled PRO is a "discontinuous element with two θ-roles". See also footnote 8, Chapter 4.

4. Recall however that the theorem about the non-governed status of PRO is a rather weak one: government is introduced in the Binding Theory specifically to rule out PRO in a governed position, and there is no independent motivation for government to be involved in the Binding Theory. Furthermore, the theorem is based on the assumption that PRO being a pronominal anaphor it is subject to both conditions A and B of the Binding Theory, and hence must not have a governing category, hence be ungoverned. But we have seen in Chapter 2 that the status of conditions A and B as actual Binding theory conditions is unlikely. See the discussion in 2.3.3. and the discussion of the Avoid Pronoun principle in the Appendix.

5. This observation is due to R. Kayne. In a talk given at M.I.T. in the Fall of 1981, Kayne suggested that the insertion of <u>of</u> in such cases could be due to the ECP. If one assumes that N cannot be a proper governor, then the trace would not be properly governed if <u>of</u> was not inserted. But then one is faced with the problem of explaining why <u>the city's destruction</u> is grammatical, as Kayne himself observed. Furthermore, one might presume that the unmarked case in a reanalyzed construction like these is for the V to govern the trace, but an explanation by ECP must dismiss this possibility, or else there is no need to insert <u>of</u>. Note also that we have seen that the lexical government part of ECP is just an accidental property of the grammar: a trace must be "identified" by an antecedent that governs it, and the fact that it is also lexically governed in most cases is due to θ-theory and/or Case theory. One reason why <u>of</u> might be inserted in such constructions as <u>John was taken advantage of</u> could be that passive morphology applies only if it can absorb Case, so that it could not apply here if <u>of</u> was not inserted to assign Case. For more discussion, see 3.3.2.3 and 5.2.2.

6. Since Genitive Case can be assigned to the subject of a gerund, we assume that the gerund is [+N,-V] in such cases. See the discussion on the occurrence of PRO in subject position of gerunds, and also on Accusative and Nominative on the subject of gerunds.

7. (From Stowell (1981a)) "Kiparsky's principle is actually based on a principle assumed by the Indian grammarians' theory of Sanskrit grammar in the Pāninian tradition. See Kiparsky (1982) for discussion." Note that Stowell (1981a) eventually ends up reversing his analysis and then considers that of-insertion is the special rule because of his way of stating the two rules, which differ from the ones presented here. See Stowell (1981a) for discussion.

8. However, in Chomsky (1981a), there is a visibility condition on θ-role assignment to which we return in 3.3.2.4 and it could be argued that the embedded tensed clauses in (23) do not receive Case (which is necessary for the tensed clause to get a θ-role) unless they are co-superscripted with a Case marked element as in (i), so that the sentences are ruled out on grounds independent from Subjacency.

(i) It^{i} seems $[_{\overline{S}}$ that John is happy$]^{i}$

9. We assume that expletive elements transmit Case in constructions like (27) and (28) (cf. Chomsky (1981a)). This might be technically related to the fact that these elements do not bear a REF-index. So it could be that these elements Bind referential NPs, but in a somewhat reverse fashion, the Bindee being the bearer of the REF-index here, and this could be responsible for Case transfer. It would also be related to the fact that there is a definiteness condition on the referential NP in such constructions where Case is transferred. For example, Safir (forthcoming) observes that the definiteness restriction seems to hold only if Case is transmitted. Consider the following sentences.

(i) Il est arrivé un homme/*l'homme.

(ii) Il a été tiré sur un bateau.

(iii) Il a été tiré sur le bateau.

In all of these sentences, take il with an expletive reading. In (i), un homme gets Case by being related to il and there is a definiteness effect. In (ii), the sentence is ambiguous: it can have the two readings in (iv).

(iv) a. Someone shot on a boat. (locative)
b. Someone shot at a boat. (objective)

Sentence (iii) is also ambiguous but it shows no definiteness effect, contrary to (i). The reason, according to Safir, is that le bateau does not get its Case from il in (iii), but from sur since there is no reanalysis of V-P in French, as we see in (v).

(v) *Le bateau$_i$ a été tiré sur t_i

So the definiteness effect seems to be directly related to Case transfer. Note that if (inverse) Binding of the referential NP by il is a pre-condition for the Case transfer, then this is impossible in (ii) and (iii) since the PP node blocks government and this could be what blocks the definiteness effect, not the Case transmission properties. It is possible that the fact that expletive il has no REF-index affects the possibilities of reference of the full NP in some way, hence the definiteness effect when il Binds this NP. We will not explore these possibilities any further here and leave them open to further study.

10. We will not discuss in any detail other constructions for

which WH-movement has been postulated like indirect questions, comparatives, topicalizations, clefts, tough-movement (cf. Chomsky (1973)). We will assume that the analysis presented here can be extended to these constructions without major modifications. By core constructions involving WH-movement, we do not mean to exclude these constructions like comparative, etc. These constructions are probably also part of the core constructions involving WH-movement, although some of them like tough-movement probably have some additional properties which are marked since they seem quite restricted in languages. By core constructions involving WH-movement, we rather want to distinguish constructions like those in (29) and (30) from constructions where successive iteration of the rule has applied. This property of WH-movement seems to be marked in that it does not occur in most languages and it is subject to lexical variation in the languages that allow them, like the bridge verb phenomena for example (cf. Erteschik (1973)). Although we acknowledge that these marked phenomena must be explained, we also assume that their markedness must be reflected in the analysis that the grammar provides of their properties. We return to successive cyclicity in 3.3.3.

11. This position does not have to be an EC: it can be a resumptive pronoun. We will return to this topic below.

12. There could also be independent reasons as for why this relation is obligatory, like a prohibition against free variables

or vacuous quantification. But note that, as pointed out in Chomsky (1981a), the prohibition against vacuous quantification is not a logical necessity since there could be logical systems where this would be allowed. So it would have to be stipulated that this is a property of natural languages. On the other hand, if Binding is involved here along with certain conditions on the need for an R-index and F-features for NPs at LF, then the obligatoriness of the relation would follow in most cases from the fact that the EC must get these features and an R-index from some element in the sentence. The obligatoriness of the relation could also be extended to PP traces, for example, if there are other kinds of features which are assumed to be necessary for the interpretation of a PP trace. However, it seems to be necessary to incorporate a specific statement to the effect that there must be no vacuous quantification in view of the facts about resumptive pronouns: here, the relation is obligatory, i.e. the WH-phrase must bind a pronoun (or a trace) in its domain for the sentence to be grammatical, although there is no intrinsic reason why a pronoun would have to be bound. Note, however, that this option is only possible for WH-elements, not for quantifiers like everyone, someone, etc., that is, we never find a quantifier of this type in an $\bar{A}$-position with a resumptive pronoun at S-structure. So it seems that this option is only possible if the phrase is one that can also be moved overtly in the syntax to an $\bar{A}$-position. This could mean that

the binding of a resumptive pronoun is done in analogy with the binding of a trace at S-structure, so that the restriction on vacuous quantification in natural languages is a somewhat derivative one.

13. The domain of reanalysis is broader than just V-P since we get sentences like (i), where V-NP-P is a reanalyzed sequence.

(i) a. John was taken advantage of.
b. Who did John tell Bill about?

However, we will restrict ourselves to simpler examples involving V-P for expository purposes.

Recall also that it has been noted in the literature (van Riemsdijk (1978), Hornstein & Weinberg (1981), Rothstein (1981)) that the constructions where stranding a P by passivization is possible are a proper subset of the constructions where stranding of a P by WH-movement is possible. So for instance, van Riemsdijk (1978) proposes that reanalysis takes place only in the subset where passive is possible, the other stranded Ps having an "escape hatch" that allows WH-movement. A similar distinction is made in Hornstein & Weinberg (1981), where passive is possible only with "predicates that are also semantic words", and Rothstein (1981) makes the distinction by proposing that passive is a lexical rule while WH-movement is a syntactic rule. We will also assume that such a distinction must be made: only the constructions where passivization is possible will be considered as falling under the core constructions of WH-movement

(with the help of the auxiliary hypothesis of reanalysis). The other stranded prepositions will not involve reanalysis, but a marked process which is also present in the marked constructions where successive cyclicity is possible. We return to this topic in 3.3.3.

14. We say in most cases, and not in all cases, because there are instances of NP movement which are not triggered by the need for Case. For example, consider (i).

(i) $John_i$ tried $[_S$ PRO_i to be seen t_i by Mary]

In (i), PRO does not move in order to get Case since it never does get Case. In GB, it would have to be assumed that PRO is moved in order not to remain in a governed position, since it is assumed that there is a theorem that can be derived from the binding theory to the effect that PRO must be ungoverned. (This was pointed out to us by Jane Simpson, personal communication.)

In our analysis, the EC is in the object position of seen which is a θ-position; hence, it must be assigned a proper R-index for the sentence to be grammatical. The R-index will be assigned to the EC by John via the subject position, i.e. PRO in (i), since John Binds PRO, which in turn Binds the trace.

15. Note that it is not necessary to assume an extrinsic ordering of Case assignment and WH-movement here, since the reverse ordering will result in a violation of the principle of Lexicaliza-

tion.

16. We return in 3.3.2.3 to a problem that free relatives in Hebrew seem to present for our analysis, namely the fact that "Case from outside" is not sufficient in Hebrew since the equivalent to the English sentence in (i) is also ungrammatical in Hebrew since no Case is assigned to the internal position.

(i) *Whoever I told Mary [$\underline{t}$ to fix the sink]

17. This section is a revised version of an analysis presented in Bouchard (1982a).

18. The facts presented here are found with only minor differences in colloquial French as described in Guiraud (1966), Frei (1929).

19. There is a possible variant with <u>dont</u> for this type of relativization of a <u>de NP</u>. I will disregard this form since it involves complexities which are not directly relevant to the problem at hand.

20. Many of these differences between SFRCs and QRCs are often said to be an anglicization of the Québec French. However, a study of the facts shows that this explanation by anglicization comes from a social attitude more than from a scientific study, this social attitude being brought about by factors which are obvious to whoever knows the socio-political context. However, as soon as one goes beyond this superficial level of explanation, one sees that QRCs are not the product of anglicization, but

that these constructions fit into a long evolution which saw French pass from a synthetic to an analytic form.

There are three facts that show that QRCs could not be a product of anglicization, but were brought about by much deeper changes. First, the phenomenon of "décumul", where the structure of relative clauses changes from a synthetic to an analytic form goes back to Old French, as we can see in (i).

(i) a. Et si vont les beles dames cortoises que eles ont deux amis ou trois avec leur barons. (Aucassin,6,36)
'And thus go the nice courteous ladies that they have two friends or three in addition to their barons.'

b. ...l'empereur Kyrsac de Constantinoble, que uns siens freres li avoit tolu l'empire ...
(Robert de Clari, XVII,12)
'...the emperor Kyrsac of Constantinople, that one of his brothers had tried to take from him the empire...'

c. La dame que vos avés a feu a flame soventes fois sa terre mise (Chrétien de Troyes, Guillaume d'Angleterre)
'the lady that you have to fire and flame often on her land put'

d. Il les tendroit as us et coutumes que li empereur les avoient tenuz (Ø). (Villehardouin) (Non-insertion of à)
'He would hold them to the customs that the emperor had held them (to).'

(Cf. Bouchard 1980 for more data and references. The rough translations are ours.)

A second fact is that the "décumul" can also be found in

many dialects of 20th century colloquial French, as we see in (ii).

(ii) (These examples are from Frei (1929))

a. Mon mari que je suis sans nouvelle de lui ...

b. ...un tas de farces que j'en pleurais de rire (Flaubert, Madame Bovary)

c. Mon neveu que son père est cantonnier (Hamp, Vin de Champagne)

d. Tu me diras si tu m'as envoyé le colis que tu me parlé (Ø) (non-insertion of de)

e. La jeune fille qu'il doit se marier avec.

Cf. also Guiraud (1966);also Remacle (1960) for Wallon, Haigneré (1903) for Boulonnais, Aub-Büchner (1962) for Ranrupt, Chaurand (1968) for Thiérache.

As Guiraud (1966) points out, the fact that we find these constructions in colloquial French is not surprising since it follows from a long process of historical change in the language, where the inflected relative pronoun is replaced by the complementizer que with appropriate pronouns in the sentence.

A third argument against the explanation by anglicization comes from the fact that "décumul" is found in many romance languages, which suggests that the phenomenon is very old, or at least what triggered it. Thus Diez (1876) signals in his comparative study of romance languages that the phenomenon can be found in Italian, Spanish, Provencal and Old French, among others.

21. In Aoun et al. (1981), it is assumed that raising of WH-phrases in situ by WH-R only applies to [+WH] elements, i.e. not

in relative clauses. Cf. also Huang (1982) for discussion of Chinese where only LF movement of WH-phrases takes place.

22. Stowell (1981a) claims that COMP does head the $\bar{S}$ at the same time that INFL does, which essentially comes to the same as our proposal of percolation of features. We prefer the percolation analysis since there seems to be independent evidence for this type of process in morphology (cf. Lieber 1980).

23. In the case of Dative à, the preposition cannot be omitted:

(i) a. Le gars à qui j'ai donné un cadeau
b. *Le gars que j'ai donné un cadeau

It is possible to have a Dative clitic as in (ii).

(ii) Le gars que je lui ai donné un cadeau

So it seems that there is a semantic content in the Dative marking which must be recoverable in some way: it forces the Dative to be expressed whereas the semantically null à does not have to be expressed as we see in (61a') (cf. Marantz (1981) on the assignment of semantic roles by prepositions). This means that a sentence like (iii) is not ambiguous since it cannot be derived from (v) but only from (iv) since the à complement of parler is a Dative.

(iii) Le gars que je parle ...

(iv) Le gars que je parle de PRO

(v) Le gars que je parle à PRO (≠ (iii))

24. There are two cases where a preposition is dropped: either because it is semantically empty as à, de in sentences like (75),

or because its meaning can be somewhat recovered. In the latter case, recoverability can be due to the content of the WH-word itself, or of the head of the relative clause. An example of the first case is the locatives like où, where which allow some locative prepositions to be omitted as in (i).

(i) a. La maison où j'habite. (dans)
b. L'école où il va. (à)
c. La tablette où est le dictionaire. (sur)
d. The city where he lives. (in)
e. The city where he is going tomorrow. (to)
f. The shelf where he found the dictionary. (on)

The second case where it is the head of the relative clause that allows recoverability of the missing P is illustrated in (ii).

(ii) (From Schachter (1972))
a. The time (that/at which/*which) the cat sat on the mat.
b. The way (that/in which/*which) the cat sat on the mat.
c. The reason (that/for which/*which) the cat sat on the mat.

These examples show that if the relation between the head and relativized position is mediated through a Ø-operator, then this relation allows recoverability of the P. If the relation is mediated through a WH-operator, recoverability seems to be impossible. But there does not seem to be any immediate reason why this should be impossible from a semantic point of view. However, there is a straightforward explanation in our analysis

from the perspective of Case assignment: if no P is present in (ii) and if a WH-word is present, then this WH-word does not get Case, and hence is in conflict with the principle of Lexicalization. If no WH-phrase is present however, the P can be omitted since its content is recoverable from the head of the relative clause, and it is not required for purposes of Case assignment since there is no lexical item to assign Case to.

Similar examples for colloquial French are given in (iii).

(iii) (From Bouchard (1979))

a. C'est le couteau que je coupe (le beurre) (avec)
b. C'est le fil que je coupe *(le beurre)
c. Qu'est-ce que tu coupes (*le beurre)

In (iiia), deletion of the preposition avec is possible since couteau has semantic content which identifies it as a possible instrumental for couper. In (iiib), if le beurre is not given as a direct object, then it is likely that le fil will be interpreted as the direct object of coupe rather than as an instrumental, unless the context is very clear about the intended meaning where le fil is an instrumental. In (iiic), an instrumental reading of qu'est-ce que is impossible since it has no referential content indentifying it as a possible instrumental, so the instrumental preposition cannot be omitted if the WH-phrase is to be instrumental. It can be interpreted as a direct object however.

25. See 3.3.3 for an analysis of bounding node which is consistent with the present analysis.

26. These sentences are from Vinet (1979) who also proposes to analyse relative clauses with strong form prepositions as instances of the resumptive strategy. This suggestion is also found in Kayne & Pollock (1978) who mention in their footnote 17 that the colloquial French relative clause in (i) is best analysed as akin to (ii).

(i) Le mec que tu peux compter dessus.

(ii) Le mec que je lui ai parlé.

We have already seen in 3.2 that Subjacency can be derived from independently motivated principles as far as NP movement is concerned. We will see in 3.3.3 that this is also the case for WH-movement.

27. We disregard the difference between forms like qui and lequel, assuming that they are variants of the realization of the feature WH. The difference between the two seems to be that in relative clauses, qui forms are anaphors, whereas lequel forms are pronominals. Cf. the discussion on Avoid Pronoun to which we turn in the Appendix to this chapter.

28. One way out for Groos & Riemsdijk could be to assume as in Reuland (1980) that -ing is the N head of a gerund, whereas a free relative would have no N head (see the discussion on gerunds in 5.2.1.2.1). This analysis would still not respect a strong X theory however since free relatives would not be labelled by lexical insertion under a head.

29. Note that Binding of the clitic by the WH-phrase seems to

violate government since the clitic is in a PP or an NP, and that maximal expansions usually block government. But if the clitic is on the head of the phrase, we can assume that it is accessible for government since heads are accessible for government. Cf. the discussion in 2.2.3.3, (63)-(65).

30. There are some pour relative clauses, so that (100a) can have a relative clause interpretation. As expected however, such a reading is awkward for (100b), i.e. it has the status of (99b). The crucial fact is that there are no à purposive clauses, so that there is a contrast between the sentences in (99) and those in (100).

31. This merging operation is probably also the cause of the obligatoriness of deletion of the WH-phrase when it is nominal in infinitival relative clauses, noted in Chomsky & Lasnik (1977). It could be related to the Avoid Pronoun principle discussed in the Appendix to this chapter.

32. See also some cases of Standard French and English in footnote 24.

33. Doug Pulleyblank informs us (personal communication) that similar extractions are possible in Yoruba.

34. There has been a long and lively debate about these facts. Cf. Andrews (1978), Chomsky (1980, 1981a), Chomsky & Lasnik (1977, 1978), Jaeggli (1980), Lakoff (1970), Postal & Pullum (1978) among others.

35. See Selkirk (1972, 1974) where much of the current ideas on the subject originated. Our exposition of the facts draws heavily from Manzini (1981). For another example of phonological rules which are subject to syntactic-like conditions, see the study of Ewe in Clements (1978). Manzini (1981) also proposes an account of the Ewe facts that is in line with her account of the French liaison facts.

36. This difference in structure is supported by the passivization facts as we will see in 5.2.2.

37. While revising the final version of this thesis, we came upon the latest addition to the long debate on contraction: the contribution of Postal & Pullum (1982). Postal & Pullum give more cases where string adjacency is not sufficient to allow contraction: in all these cases, there is no lexical element or Case-marked trace between want and to, and yet contraction is blocked, so that there seems to be a structural condition on contraction. The constructions involved are 1° when to is in an infinitival subject of the embedded infinitival complement of want as in (i); 2° when want is in a relative clause embedded under a subject NP which is followed by an infinitival VP as in (ii); 3° when want is followed by a purposive clause as in (iii); 4° when to is in a conjoined VP as in (iv); 5° when want is a noun subject of an infinitival clause as in (v).

(i) I don't want $[_{S}[_{\overline{S}}$ PRO to flagellate oneself in public] to become standard practice in the monastery] (*wanna)

(ii) I don't want anyone $[_{\bar{S}}$ who continues to want] to stop wanting. (*wanna)

(iii) One must want $[_{\bar{S}}$ (in order) PRO to become an effective consumer] (*wanna if under purposive reading, but OK if infinitival complement of want)

(iv) I want $[_S$ PRO$[_{VP}[_{VP}$ to dance] and $[_{VP}$ to sing]]]

(v) We cannot expect $[_S[_{NP}$ that want$][_{VP}$ to be satisfied]]

All of these cases follow from our structural condition on want+to contraction which states that want must govern to. In (i), the $\bar{S}$ of the infinitival subject cannot be deleted since it is not in a context (V ___) since there is an intervening S boundary: so government of to by want across $\bar{S}$ is impossible, hence contraction is impossible. In (ii), the embedded want clearly does not govern to. In (iii), if we assume that purposives are $\bar{S}$s attached to S (see 5.3 for more discussion), then want does not govern to and contraction is therefore blocked. In (iv), it seems that conjunction blocks government: so contraction is blocked if to is inside the conjunction as in (iv), but it is fine if to is outside the conjunction as in (vi).

(vi) I want $[_S$ PRO to [dance and sing]] (wanna=OK)

Note that the conjunction in (iv) is of VPs, not of Ss, or else PRO would be inside a conjunction and could not be governed by its antecedent I, as required by the interpretation of local control.

In (v), the N want does not govern to since want is the head of an NP and can only govern inside that NP. Even if the

infinitival clause is a complement of the N want, contraction is impossible, however, as we see in (vii).

(vii) $[_{NP}$ the want $[_{\overline{S}}$ PRO to leave$]]$ is often stronger than $[_{NP}$ the want $[_{\overline{S}}$ to stay$]]$ (*wanna)

The reason could be due to the fact that $\overline{S}$ deletion does not take place after a N, so that government by want is impossible. It could also be the case that contraction is stated only on the V want, not on the N want as suggested by Chomsky & Lasnik (1978): this is implicitly stated in Postal & Pullum's constraint on contraction since they state that want must be a V (see (xb) below).

There are two other cases briefly mentioned by Postal & Pullum where contraction is blocked which have been observed by Brame (1981) and for which Postal & Pullum have no explanation: when want is in a subjunctive clause as in (viii) and when it is in a (pseudo-)imperative clause as in (ix).

(viii) The director requires that all the actors want to give their most. (*wanna)

(ix) Want to do that and you'll be rewarded. (*wanna)

One possible explanation could be along the lines of Brame (1981): to say that want contraction is lexically restricted to want when it is indicative or infinitival. This relates to the fact that the N want also seems to be impossible to contract with to, as we just observed.

So our analysis accounts for all the facts presented by Postal & Pullum: all the structural cases that they present

where contraction is impossible are covered by the requirement that want must govern to for contraction to take place; there also seems to be a condition on the lexical content of want which excludes all instances of want which are not indicative or infinitival from contraction with to. To account for the structural conditioning on to contraction, Postal & Pullum (1982) propose the following constraint:

(x) A contraction trigger V can have a contracted form with infinitival to only if:
  a. to is the main verb of the initial direct object complement of the matrix clause whose main V is want;
  b. the final subject of the complement is identical to the final subject of the matrix.

What (a) states is that to must be in a sentence that is the complement of want: so it is a structural constraint on the relation between the two elements to be contracted. If we assume that the structural relation between a head and its complement is that of government, then (a) states that want must govern the sentence in which to is. Clause (b) is strange, however, since it says that a relation between the subject of want and the subject of the to-clause has some bearing on the contraction of want and to: so the relation between some elements A and B has some bearing on the contraction of some elements C and D. Postal & Pullum (1982) also insist that the constraint on contraction has nothing to do with adjacency since clause (b) restricts contraction contexts to local control equi

constructions and raising constructions, and this, they claim, has nothing to do with adjacency. The peculiarity of clause (b) finds an explanation if the constraint on contraction is given as a requirement that want govern to. Local control equi constructions and raising constructions happen to be the constructions where $\overline{S}$ deletion must take place in order for the antecedent to govern the anaphor EC in both cases in our analysis. If government by want is required for contraction, these are precisely the only constructions where want can govern to.

So clauses (a) and (b), which seem to be unrelated although they describe the proper contexts for contraction, receive an explanation if one assumes that the constraint on contraction is one of government of to by the verb: clause (a) follows from the fact that, if want must govern to, it must necessarily govern the sentence in which to is; clause (b) follows from the fact that $\overline{S}$ deletion must take place for the V to govern to, and if $\overline{S}$ deletion takes place, then the EC subject of the infinitival clause is governed by the NP subject of the higher clause, hence it is functionally determined to be an anaphor, i.e. locally controlled PRO.

38. This parameterization of the bounding nodes was first suggested in Rizzi (1978) for Italian. Sportiche (1979) claims that the same holds for French. Spanish would also be such a language, cf. Torrego (1981) among others.

39. As noted in footnote (13) above, the domain of reanalysis

also extends to V-NP-P sequences.

40. Recall that small clauses are not barriers to government since they allow ECM.

(i) John considers $[_A$n Bill stupid]

If Binding is the relevant notion to account for WH-movement, as we assume, and if Binding involves government, then extraction from clauses which are known for independent reasons to be permeable to government is expected. Note also that APs in copular constructions are permeable to Binding since reflexives are obligatory in cases like $\underline{\text{John}}_i \underline{\text{ is proud of himself}}_i$/ $*\underline{\text{him}}_i$.

41. We will omit from our discussion the question of when movement to COMP takes place, i.e. in the syntax or LF. For example, according to Huang (1982), Chinese would have WH-movement only at LF. So this is another parameter which is independent of the "locality parameter" that we are discussing here.

42. In order for this explanation to hold, the notion of proper government must be complemented with specific assumptions about the node INFL, which in GB is generally assumed to govern the subject position of S. If the definition of proper government is as we have seen above, then it must be assumed that INFL is not lexical, or else INFL would properly govern the subject of S by clause (i); furthermore, the indexing between INFL and the subject must be of a different kind than the usual coindexing, or else INFL would properly govern the subject by clause

(ii). In Chomsky (1981a), both of these assumptions are made: INFL is considered not to be lexical, and indexing between INFL (in fact, AGR, a subpart of INFL) and the subject is a special kind of indexing: cosuperscripting. This is a different kind of indexing for independent reasons. For example, it must not be relevant for the binding theory or else a name or a pronoun in subject position would be bound in its governing category, in violation of conditions B and C, so that sentences like (i) would be ungrammatical, contrary to fact.

(i) a. John_i INFL_i see Bill
b. He_i INFL_i see Bill

Furthermore, anaphors would be expected to be possible in subject position since they would be bound in their governing category, as required by condition A of the binding theory.

(ii) *Himself_i INFL_i come for dinner

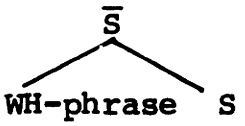
Therefore, the indexing of AGR and the subject cannot be relevant to the binding theory in this framework. In our terms, cosuperscripting would be like coindexing with an R-index that contains only an S-index and no REF-index, although the correspondance between superscript and S-index without REF-index might not always be total.

Note finally that the fact that the governor is an X^0 must be stipulated in the definition of proper government: if it is not stipulated, then one could expect that government by any node could count as satisfying the ECP, so that government of the subject position by VP would be sufficient and the contrast

in extraction from subject and object position would be lost. Recall that we saw in Chapter 1 that the fact that $\alpha=X^0$ could be derived from independent motivations in the core notion of government, so that it did not have to be mentioned. But in Chomsky's analysis, this spedification must be reintroduced in the definition of proper government.

43. Note that it must be assumed in Stowell's analysis that the node COMP does not prevent percolation of features to $\bar{S}$, so that the WH-phrase in COMP can be considered as the head of $\bar{S}$. This is similar to our analysis where the structure is as in (i) and where there is partial percolation of features from WH to $\bar{S}$.

(i)



44. For another account of bridge verbs and related facts, see Chomsky (1981a), Chapter 5. Briefly, Chomsky proposes that the bridge verb phenomenon might be amenable to the fact that the status of $\bar{S}$ and S as bounding nodes depends on the context in which they appear. Thus, bounding nodes are characterized as in (i).

(i) a. $\bar{S}$ is a bounding node iff it is in the context $\_\_[+WH]$

b. S is a bounding node iff it is in the context $[+WH]\_\_$

c. S is a bounding node when governed

We will not discuss the conceptual or empirical implications of this analysis since it requires assumptions that are quite different from the ones that we will adopt and does not relate

directly to our analysis.

45. Our discussion of V-preposing in Spanish dwells heavily on the data and analysis provided in Torrego (1981).

46. Recall from our discussion in 3.3.2.4 that it is also crucial for contraction to take place that there be no trace in COMP in a sentence like (i) so that $\bar{S}$ deletion can take place and so <u>want</u> can govern <u>to</u>, hence allowing contraction.

(i) Who do you want $[_S$ PRO to see <u>t</u>]?

47. See also Jaeggli (1981) who takes a similar position.

48. The fact that index chains are identified by a WH-index, so that index chains are only possible for WH-Binders, relates to our analysis of infinitival relative clauses (cf. the discussion of (97)-(100) above). We might wonder what property of these infinitival relative clause constructions precludes long distance dependencies. One possible answer is that a WH-feature might not be possible in these types of relative clauses where merger of the operator and the head take place, the reason being that the head cannot bear such a feature. So an index chain could not be established since the Binder is not a WH-element, and index chains can only be identified by a WH-index: thus long distance dependencies are not possible in such constructions.

49. This is similar to an analysis to which Chomsky (1981a) makes a brief allusion in his Chapter 5 and where the structure

of (183) is as in (i).

(i) In which room did you say that ($_S$ John ($_{\overline{VP}}$($_{VP}$ saw Bill) $\underline{t}$))

50. In fact, reanalysis might not have to take place for stranding by WH-movement if, assuming that Ps have an S-index, the language allows the index of the preposition to be accessible for WH in the V-grid so that the trace in the complement position of the preposition is also accessible via an index chain as in (i).

(i) a. [$_{\overline{S}}$ what [$_S$ did you [$_{VP}$ put [$_{NP_i}$ it] [$_{PP_j}$ in $\underline{t}_k$]]]]

b.
$$\begin{bmatrix} \text{put} \\ \underline{i}\text{=theme} \\ \begin{bmatrix} \underline{j}\ \text{(in)} \\ \underline{k}\text{=theme} \end{bmatrix} \quad \text{=location} \end{bmatrix}$$

51. It is possible that the parameter is more technical than this and it could be that WH-phrases in COMP in English type I languages have a stronger percolation of features to $\overline{S}$ so that it breaks an index chain. We might then expect this to have other effects in the grammar. There could be other possibilities of what exactly the parameter is, but we have no evidence for the moment that would lead us to adopt a parameter that is more specific than the one given in the text, so we will adopt it as such.

52. Some extractions out of NP are possible in French, as in (i).

(i) a. Combien est-ce qu'elle a $[_{NP}$ t d'argent]?
b. Je ne voudrais pas que tu boive $[_{NP}$ t de bière]

(Cf. Kayne (1975, 1981a), Baltin (1978), Obenauer (1976))

But in both cases, the t is the head of the NP, which might account for this extractability since heads are usually accessible for government.

Sentences like (ii) are also possible.

(ii) a. De qui as-tu vu $[_{NP}$ le portrait t]?
b. L'homme dont tu as vu $[_{NP}$ le portrait de Rembrandt t]

But only the "subject" seems to be extractable from NPs. The notion subject here is relative: as shown by Zubizarreta (1979), the characterization of the subject in NPs seems to be determined by a thematic hierarchy as in (iii).

(iii) a. possessor or source
b. agent
c. theme

The following sentences show that only subjects can be extracted. See also Cinque (1979).

(iv) a. L'homme dont le portrait d'Aristote de Rembrandt t
b. *Rembrandt, dont le portrait d'Aristote t de Pierre
c. Rembrandt, dont le portrait d'Aristote t
d. Aristote, dont le portrait t de Rembrandt
e. Aristote, dont le portrait t

So it seems that only the highest phrase in the NP structure is extractable, and it might be that it is this structural position of the phrase that makes it accessible to outside Binding, not that only one bounding node is crossed. Notice that

extractions from within NP by the WH-element dont are usually better than with a WH-phrase like de qui.

(v) a.??L'homme de qui $[_{NP}$ le portrait d'Aristote de Rembrandt t]

b. L'homme dont $[_{NP}$ le portrait d'Aristote de Rembrandt t]

Dont is also unique in that it can bind two variables even in contexts that are not Across-the-board extractions in the sense of Williams (1978).

(vi) (From Damourette et Pichon (1911))
La Russie dont $[_{NP}$ le bolchevisme t] nuit à $[_{NP}$ la civilisation t]

Finally, the contrast in (vii) shows that it is not the fact that S is not a bounding node in French that allows (viia) as claimed in Baltin (1978).

(vii) a. L'homme avec qui $[_{S}$ $[_{\alpha}$ parler t] serait difficile]

b. *L'homme avec qui $[_{S}$ $[_{\overline{S}}$ que Jean parle t] serait difficile]

Rather it seems that infinitives have a weaker boundary than tensed clauses and that α in (viia) is a weaker boundary than $\overline{S}$.

So since extractions out of NPs (or a sentence as in (viia)) seem to be involving special processes in all these cases, we will maintain our claim that extraction out of NP is generally impossible, given the core notion of Binding. On PP as a bounding node, see Baltin (1978).

53. There might also be instances of syntactic movement which

are just copying and not deletion. Thus it seems that in Cap Verde, all traces are lexicalized.

54. This formulation of the Avoid Pronoun principle would also indicate that the difference between qui forms and lequel forms in relative clauses is one of anaphor versus pronoun. As we can see in (i), lequel forms are not possible when subject or object positions have been relativized.

(i) a. *L'homme [lequel [t est venu]]
b. *L'homme [lequel [j'ai vu t]]

If lequel is a pronoun, then by the Avoid Pronoun principle, it can't appear in (i) since subject and object positions can always be Bound by the phrase in COMP, and this phrase itself being an NP, it can be Bound as an anaphor by the head. Note that lequel is possible in such positions in non-restrictive relative clauses as in (ii).

(ii) a. Cet homme, lequel [t n'a jamais fait quoi que ce soit pour son pays], nous dégoûte.
b. Cet homme, lequel [Marie n'a jamais vu t sobre] la dégoûte

The reason is that the head of a non-restrictive relative clause presumably does not govern the COMP of the relative clause since this clause is more like a parenthetical, and so the element in COMP cannot be Bound as an anaphor and so must be a pronoun: so lequel forms are allowed in such cases. Qui forms can also be pronouns, as we see in (iii).

(iii) a. Cet homme, qui [t n'a jamais fait quoi que ce soit pour son pays] nous dégoûte.

b. Cet homme, que [Marie n'a jamais vu t sobre], le dégoûte.

55. One possible explanation for the transparency of the subjunctive $\overline{S}$ is provided by Salamanca (1981). We could assume that the strong tense dependency of a subjunctive clause on the matrix clause is mediated by the percolation and indexing of the tense features of the subjunctive clause to $\overline{S}$, where they are accessible to the matrix V. Consider now (i).

(i) *Nous voulons que nous partions.

Assuming a theory of compounding and morphology similar to Lieber (1980), Williams (1981) and Selkirk (forthcoming), the external morpheme of a word is the head of the word in some sense. So suppose we have the tense morphemes of (i) as in (ii).

(ii) $(\text{-ons})_i$ $(\text{-i-ons})_j$

What we see is that the personal inflexions have to be part of the tense features that are percolated since these inflexions are external to the tense features. So this, we could assume, renders these personal inflexions in some sense transparent to the personal inflexions of the matrix verbs, so that the violation in (i) is similar to the one in (iii).

(iii) *We saw us.

See also Picallo (1982).

56. It might be argued that the facts about gerunds show only a preference in judgements, whereas the others, like the need of a lexical anaphor in (237), show a clear grammaticality con-

trast. But we believe that the contrast is as strong for gerunds and that it is an ambiquity in the structure of the gerund which creates this other weak possibility of coreference in (236). See 5.2.1.2.1. for more discussion about gerunds, some of which will be seen to be sentential, hence allow Binding across S, and others nominal, hence allow no Binding across NP. When a gerund has a lexical subject, it is nominal, hence it cannot have an anaphor in its subject position and so it does not fall under (227) in these cases. But when it is sentential, there can be a PRO anaphor in subject position and this falls under (227).

CHAPTER 4: PRO DROP LANGUAGES

4.1 General comments.

The study of the phenomenon of Pro Drop, where an overt subject is missing in a tensed clause as in (1), was slowly incorporated into generative grammar.

(1) Creo que partio.

The phenomenon is discussed in Perlmutter (1968, 1971) and taken up again in Chomsky & Lasnik (1977). But in recent years, much attention has been given to the study of the phenomenon (cf. Borer 1981, Burzio 1981, Chao 1981, Chomsky 1981a,b, Jaeggli 1981, Rizzi 1979, Taraldsen 1978 to name a few). The focus of this research has been mainly on Romance languages, especially Italian and Spanish. These languages are usually classified as configurational languages, as opposed to non-configurational languages like Japanese, Malayalam, Australian aboriginal languages like Warlpiri, and many American Indian languages.[1] However, it is also well-known that there can be missing pronouns in non-configurational languages too (cf. the work of Hale 1978, Farmer 1980, Mohanan 1981, Nash 1980, Platero 1978, and others).

Although the two phenomena of missing pronoun in configurational languages and missing pronoun in non-configurational languages seem to be quite different on the surface, one should expect that the two be related in some way since in both cases, what is allowed is to have an EC, and ECs are presumably sub-

ject to general constraints of recoverability, for instance. If we continue to assume that there is only one EC, then the missing pronoun should simply be a manifestation of the EC, and we do not expect it to be some special element that only these languages have: the missing pronoun is an EC, and its presence, type and content should be determined by the interaction of general principles just like the other ECs. Thus we will assume that the presence of the missing pronoun is determined by the extended Projection principle and the θ-criterion. There are basically two types of ECs in our analysis: an EC is an anaphor if it is Bound by an antecedent, and it is a pronoun if it is freely indexed at S-structure. We will see that missing-pronouns fit into this typology too. Finally, the content of an EC must also be determined. Recall that in our analysis, an EC is an element that has no ψ-features at PF, and yet has F-features and an R-index at LF. We will make the strong hypothesis that missing pronouns are ECs like any other EC and that they also have these properties. The rest of this chapter will be devoted to showing how this is realized in the grammar of these languages. In section 4.2, we will give an analysis of Pro Drop in configurational languages, using Italian as a case-example. We will come back to non-configurational languages in section 4.3 and show how the analysis provided by the model of grammar sketched in Chapter 2 also accounts for the missing pronoun phenomenon in these languages as well, using Japanese as a case-example.

4.2. Pro Drop in configurational languages.

There are two main sets of facts that are usually discussed about Pro Drop languages: the fact that a definite pronoun may be missing in subject position as in (2), and the fact that the subject may freely be inverted as in (3).

(2) ho trovato il libro
(I found the book.)

(3) ha mangiato Gianni
(Gianni ate.)

Other properties of Pro Drop languages are that they allow "long distance WH-movement" of the subject as in (4a), empty resumptive pronouns in embedded clauses as in (4b), and apparent violations of the that-trace filter (i.e. ECP) as in (4c).

(4) a. l'uomo [che mi domando [chi vede]]
(The man x, such that I wonder who x sees)
b. ecco la regazza [che mi domando [chi crede [che possa VP]]]
(This is the girl who I wonder who thinks that she may VP)
c. Chi crede [che partira]
(Who does he think that t will leave)

We will see below that the properties illustrated in (4) are all derivable from the principles that allow (2) and (3). Therefore, we will not say more on the properties in (4) for now and we will concentrate on the properties illustrated in (2) and (3) which seem to be the core properties to be explained. We will then come back to the sentences in (4) and show how their properties are derived from those of the sentences in (2)

and (3).

Consider first the missing subject property. In Chomsky (1981a,c), it is proposed that the parameter that accounts for this is the possibility to apply Affix Hopping (the rule R) in the syntax in Italian. Thus the subject position is ungoverned at S-structure in such a case and it can contain a PRO. In a language like English or French, R can apply only in PF, so that the subject position is always governed, and therefore no PRO can appear in this position since in GB, PRO must be ungoverned. So the Pro Drop parameter can be stated as in (5).

(5) R may apply in the syntax.

The choice of (5) as a possible option in a language is said to be dependent on the richness of the inflection on the verb:

> Following the reasoning of Taraldsen (1978), a language will have the option (5) when it can, i.e., when its inflectional system is sufficiently rich in the sense he describes, so that the pronominal subject is determined by the verb when it is not overt. (Chomsky 1981c, p.10)

In Chomsky (1981b), this analysis is rejected, one reason being that the $[_{NP}e]$ seems to be governed in WH-constructions where V-preposing has taken place in Spanish (cf. the discussion based on Torrego's work in 3.3.3).

(6) $[_{\bar{S}}\ Que_i [_S [_V\ querian] [_S [_{NP}\ e] [_{VP}\ t_v\ \ t_i]]]]$

Also, since the $[_{NP}e]$ seems to be strictly pronominal and not anaphoric, Chomsky proposes that the subject of such languages

be pro when it is not overtly realized, an EC that is [+pronominal, -anaphor].

As for the content of pro, since the fact to be established is that pro appears only in the position governed by INFL at D-structure in configurational Pro Drop languages, Chomsky (1981b) proposes that the content of pro must be locally determined, more precisely by an element of the thematic complex of which pro is part: and the only element of this complex that can determine the grammatical features of pro is AGR. As for the Pro Drop parameter, Chomsky (1981b) reverts to the original suggestion by Taraldsen (1978): the possibility of having pro as subject is related to a "rich enough" inflectional system, and it is the inflection which determines the grammatical features of the EC in subject position in a tensed clause. Note that if this condition on richness of inflection is interpreted as a recoverability condition, it weakens the claim made in Chomsky (1981a,b) that an EC has inherent grammatical features since these features now seem to be very dependent on another element in the sentence. Thus pro does not have inherent grammatical features in Chomsky's (1981b) analysis, but PRO does. This does not follow from anything and it is a curious disparity if both elements are manifestations of the EC. In our analysis, no EC has F-features at PF, so that such a problem does not arise.

There is also a problem with this approach in that it predicts that whenever a verbal form has a rich enough inflection

in a language, then the subject can be missing. For example, both forms in (7) should be grammatical in French since inflection seems rich enough to recover the content of the EC under an imperative reading.

(7) a. Nous mangeons.
b. *Mangeons. (with non-imperative reading)

One could qualify the notion "richness of inflection" and say that the whole paradigm must be rich enough: the verbal form of each person would have to be unequivocal. But this runs into problems in the account of Hebrew where the subject can be missing in past and future tenses, but only in the first and second persons (see Borer 1981).

(8) a. ('ani) 'axalti 'et ha-banana
I ate acc the-banana
b. ('atem) 'axaltem 'et ha-banana
you-pl ate acc the-banana
c. hu 'axal 'et ha-banana
he ate acc the-banana
d. *'axal 'et ha-banana

So it seems that in Chomsky's analysis, it has to be specified in the grammar of a language whether a pro is possible or not. This goes against a strong hypothesis that no statement of the grammar should refer specifically to ECs or lexical NPs exclusively.

Notice that it is also accidental in this type of analysis that the notion of government is relevant for the EC missing

subject, which must be governed by an "identifier", i.e. AGR, and that the notion of government is also relevant for traces since they must be linked to an antecedent by government either directly, if the antecedent itself governs the trace as in (9), or indirectly if it is mediated by an X^{o} governor, as in (10).

(9) Who_i $[_S$ t_i saw Bill$]$

(10) Who_i $[_S$ did Bill$|_{VP}$ see $t_i]]$

Since the two phenomena are treated independently as far as recoverability of the content is concerned, this common relevance of government is not predicted. In the framework presented in Chapter 2 on the other hand, this comes as no surprise since it is assumed that one of the two basic ways in which an EC is identified is by Binding, which involves government.[2]

Let us now consider how the missing subject property can be accounted for in the approach presented here. An EC is possible when the following properties hold: 1° the NP has no ψ-features at PF (principle of Lexicalization), and 2° the NP has an R-index and F-features at LF (principle of Denotability and Agreement). In concrete terms, this means that the NP, in order not to be lexicalized must escape Case assigning in D-structure and S-structure not to have a feature in PF, but there must be something around to provide an R-index and F-features to the NP at LF. The second property, in the case of missing subjects, is guaranteed by the richness of the inflection in most cases as we will see.

Consider first how Case is assigned. Generally, Case is assigned under government. Thus a verb assigns objective Case to the NP that it governs, and a preposition also assigns Case to the NP it governs. As for assignment of nominative Case to the subject of a tensed clause, there are two possibilities. One can assume that there is a node INFL in the expansion of S and that nominative Case is assigned by some Case assigning feature of INFL to the NP that it governs: this is essentially the position taken in the GB framework. Another possibility is to assume that INFL is not a syntactic node but that it is attached to the V in the lexicon, following an axiom of lexical phonology that all affixation is done in the lexicon (cf. Kiparsky 1982, Lieber 1980, Pesetsky 1979). Then percolation of the INFL features can take place, the features climbing to VP, so that the Case-assigning features of INFL can govern the NP in subject position and assign nominative Case to it. We will adopt the latter proposal and assume that nominative Case is assigned by the INFL features when they percolate to VP, hence govern the subject position. This analysis of nominative Case assignment preserves the parallelism that often holds between Case assignment and θ-role assignment. Thus the object of a verb is assigned Case and a θ-role by the verb that governs it. Similarly, the subject of a tensed clause will be assigned Case and a θ-role by the VP that governs it in our analysis, this VP node being a complex node where the INFL features have perco-

lated. This allows us to capture the close relation that exists between the definition of Grammatical Functions (GF), the assignment of Case and the assignment of θ-role in configurational languages. Consider first the case of the object. If we assume that a θ-role is assigned when the complement enters an index into an available thematic slot in the head (along the lines of Stowell(1981a)), then all three processes depend on government from the head, and the grid of a verb can look something like in (11).

(11) V: θ
GF
Case (with possibly additional requirements like strict adjacency, etc. for case assignment (see Stowell 1981a))

Similarly for the subject, although it does not enter its index into a V-grid, it does get coindexed in some way with the VP that assigns it its θ-role since it is coindexed with AGR, a part of INFL; and INFL, we assume, percolates its features to VP, so that the subject and the VP are in some sense coindexed. Associating the definition of GFs with Case assignment and θ-role assignment, i.e. with processes that take place under government, suggests a minor modification of the definition of GFs. Instead of saying that an OBJECT is the NP of VP, it will be the NP governed by V; and instead of saying that a subject is the NP of S, it will be the NP governed by VP. This is shown in (12), with a general schema given in (13).

(12) a. OBJ: (NP, VP) V (NP)
b. SUBJ: (NP, S) VP (NP)

(13) a. OBJ: X (NP)
b. SUBJ: $\overline{X}$ (NP)

The schema in (13) generalizes to other categories like NP, AP, PP.

The notion of GF is not directly pertinent if one has θ-roles and Case, and it is just a useful notational device to express government from an X head or an $\overline{X}$ head for example when we use OBJ and SUBJ.

If we return to the assignment of Case in Pro Drop languages like Italian, suppose that in such languages there is the option that percolation of the Case-assigning features of INFL to VP can be freely delayed to LF. Then this means that at D-structure and S-structure, the element in subject position does not get Case if the percolation is delayed since it is not governed by a node bearing Case assigning features. In other words, the relation between a Case-bearing element and a Case-assigning element is not possible before LF, since the subject position is governed only by a θ-assigning element, not a Case-assigning element. So in such languages, there is an option to assign nominative Case only at LF since percolation of the Case-assigning features of INFL can be delayed to LF. If a lexical NP is inserted at D-structure, then the option to percolate the Case-assigning features only at LF cannot be taken, and

percolation must take place in the syntax in order for the lexical NP to meet the requirements of the principle of Lexicalization at PF: that is, the lexical NP must be assigned Case in the syntax. So in such a case, INFL percolates to the VP in the syntax as in (14).

(14) NP $[_{\text{VP+INFL}}$ V+INFL $]$

So INFL and the NP are coindexed for Case assignment to take place in a way similar to the one in which a V and its object NP are coindexed. Since INFL and the subject NP are coindexed, this triggers Agreement at LF by the process that we saw in Chapter 2, repeated here in (15).

(15) Agreement:
$\underline{\alpha}$ assigns (redundantly) its F-features to $\underline{\beta}$ if $\underline{\alpha}$ and $\underline{\beta}$ have the same R-index.

Coming back to the missing subject property of languages like Italian, we now see why this is possible. Since there is the option that no Case be assigned to the subject position before LF, this meets the first condition for being an EC if no other features are inserted: there are no ψ-features at PF. As for the second condition, the fact that an R-index and F-features have to be provided at LF for the EC to be interpreted as a definite pronoun, the intuitive answer is that these will be provided by the rich INFL of the verb. So both conditions to have an EC which we postulated in Chapter 2 are met: the NP has no ψ-features at PF, but it has an R-index and F-features at LF.

This means that there are two factors that are at play in allowing missing subjects: the property of having "Case at LF" and the "richness of INFL".

(16)

Rich INFL	Case at LF	EC subject
+	+	+
+	-	-
-	+	-
-	-	-

So for example, the reason why there is no Pro Drop in French even when the INFL is rich enough is because French does not allow "Case at LF" (cf. 7b). In Italian on the other hand, both conditions are met. Consider sentence (17).

(17) $[_{\bar{S}}[_{S}[_{NP}\ \underline{e}][_{VP}$ ho+INFL trovato il libro$]]]$

If the option of percolation of the Case-assigning features only at LF is chosen, then no Case has been assigned to $[_{NP}\underline{e}]$ when it reaches PF, so that it can be an EC. It is necessary for the EC to get an R-index and F-features however, and we want to capture the intuition that these depend on the richness of the inflection on the verb. If percolation takes place only at LF, then the structure in the syntax is as in (17). Since nothing with an R-index can Bind $[_{NP}\ \underline{e}]$ in (17), $\bar{S}$ blocking government, hence Binding from outside, the NP is freely assigned an index at S-structure, just like lexical pronouns are in such A-positions. When percolation of INFL to VP takes place at LF, then the structure is like the one for a lexical NP in the syn-

tax that we saw in (14).

(18) $[_S [_{NP}\ \underline{e}] [_{VP+INFL}\ V+INFL\ \ldots]]$

So INFL and $[_{NP}\ \underline{e}]$ are coindexed at LF, just like in the case of lexical NP subjects, so that agreement takes place between the two. Thus INFL provides its F-features to pro, the EC, at LF. This explains the intuitive idea that a subject can be "missing" if the verbal inflection is rich enough to allow recoverability of some content of the missing subject.

Note that we are assuming that AGR, a part of INFL, can provide all F-features to the EC subject, although it is not always the case that all these features are expressed overtly in the morphology of the verbal inflection. Suñer (1981) gives the following examples of agreement of adjectives with a missing subject.

(19) a. $[_{NP}\ \underline{e}]$ son altos/altas.
b. $[_{NP}\ \underline{e}]$ esta abierta/abierto.

So we assume that gender in AGR in Spanish is like Case in nouns in English: it is not always overt in the morphology, but it is there at some level.

Notice that we have claimed that there are two factors that interact to give the property that a missing subject is possible in languages like Italian: Case at LF, which allows the NP to be "invisible" at PF, and richness of INFL, which allows the NP to get the required F-features at LF.

So for example, Old French had a rich verbal inflection which is still reflected in written French today.

(20) Je mange
Tu manges
Il mange(t) (cf. mange-t-il)
Nous mangeons
Vous mangez
Ils mangent

In Old French, the underlined endings in (20) were pronounced. But then a phonological rule entered the language which deleted the final consonant, so that the paradigm became opaque: the 1-2-3 singular and 3 plural are now all pronounced alike. At that stage of the language, Pro Drop stopped being possible and a clitic form became obligatory on the verb when a pronominal interpretation was wanted. So we see that the richness of the verbal inflection has some bearing on the possibility to have Pro Drop, at least in configurational languages. But if the F-features become too opaque, then an element, i.e. a Pronoun or a clitic, must be inserted to provide these features and that element will be lexical since the features are introduced at D-structure, and so are visible in PF and must be overt by the principle of Lexicalization. Once Pro Drop is made impossible by the poorness (poverty?) of the verbal inflection, then the property of Case-at-LF is lost shortly after since it is made opaque in most contexts by the need of a clitic due to the poorness of the inflection. So free inversion of the subject, which

we will see shortly depends on Case-at-LF, becomes impossible soon after the language loses the possibility to have an "empty" subject. This discussion shows that the "visibility" condition in PF must be stated on all ψ-features as in the principle of Lexicalization and not just on Case as suggested by Aoun (1979). If some F-feature is missing in INFL, and thus must be specified in the syntax when the NP is inserted since the feature is required at LF but cannot be specified by INFL at that level, then the subject must be lexicalized. So F-features are crucial in the principle of Lexicalization, hence for visibility at PF. Recall that we have also seen that the feature WH is also visible at PF since it forces lexicalization in French relative clauses (cf. 3.3.2.2).[3]

What we see is that we can account for the missing subject property of Pro Drop languages like Italian if we assume that the difference between Pro Drop languages and non-Pro Drop languages is that the former have the option to percolate the Case-assigning features from V to VP in the syntax or at LF, whereas in non-Pro Drop languages, the percolation to VP must take place in the syntax obligatorily, so that Case is always assigned to the subject position of a tensed clause in these languages, and so there can be no EC in such a position.[4]

Let us now turn to the second property of Pro Drop languages noted at the beginning of this section, namely the possibility of free inversion of the subject. There are in fact two instances

of constructions with post-verbal subjects in Italian which have quite different structures.

(21) $[_S\ \_[_{VP}[_{VP}$ ha mangiato le mele] Gianni]]

(22) $[_S\ \_[_{VP}$ arrivano molti studenti]]

In (21), we have a transitive verb which normally forms a predicate VP that assigns a θ-role to a subject in an external position: so the subject Gianni is adjoined to the VP. In (23) arrivano is an ergative verb (in the sense of Burzio 1981), so that the θ-role of molti studenti is assigned in the object position, so that molti studenti can be in object position all along the derivation.

In Chomsky (1981a,b), taking after Rizzi (1979) and Burzio (1981), it is assumed that (21) is derived from the D-structure (23) by a rule that freely adjoins the subject to VP.

(23) $[_S$ Gianni $[_{VP}$ ha mangiato le mele]]

Note that the string in (22) could also have the structure in (24c) after raising to subject and subject inversion, the intermediate structure (24b) also being a possible S-structure if subject inversion does not apply.

(24) a. $[_S\ \underline{\alpha}[_{VP}$ arrivano molti studenti]]
b. $[_S$ molti studenti $[_{VP}$ arrivano $\underline{t}]]$
c. $[_S\ \underline{\alpha}[_{VP}[_{VP}$ arrivano $\underline{t}]$ molti studenti]]

There is evidence for the need of two different structures in (21) and (22). First, ne cliticization is possible in an er-

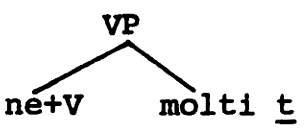
gative verb structure like (22), but not in an inversion structure like (21).

(25) $[_S\ \underline{\alpha}[_{VP}\ ne_i$ arrivano $[_{NP}$ molti $\underline{t}_i]]]$

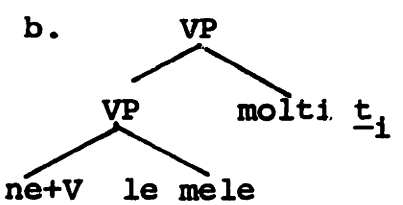
(26) $*[_S\ \underline{\alpha}[_{VP}[_{VP}\ ne_i$ mangiano le mele$][_{NP}$ molti $\underline{t}_i]]]$

In terms of our analysis of traces, the reason is that the clitic ne which is attached to the head V governs the trace in (25) hence Binds it, this trace being the head of the NP molti t_i. But in (26), ne does not govern the trace since ne+V is not an immediate constituent of a node dominating the trace since there is an intervening VP node so ne cannot Bind the trace. The trees are given in (27a) and (27b).

(27) a.



b.



The second set of facts involves the rule of "Complement-Shift" discussed in Burzio (1981). Burzio notes that there is a late rule which applies in PF and which can permute the linear order of post verbal constituents, including the inverted subject.

(28) a. Giovanni scrivera' una lunga lettera a Piero
(Giovanni will write a long letter to Piero.)

b. Scrivera' Giovanni una lunga lettera a Piero

c. Scrivera' una lunga lettera Giovanni a Piero
d. Scrivera' una lunga lettera a Piero Giovanni

However, the rule does not seem to freely permute a sentential complement and a subject adjoined to VP, as we see in (29).

(29) a. Giovanni pensa di studiare linguistica
b.??Pensa Giovanni di studiare linguistica
c. ?Pensa di studiare linguistica Giovanni

But if the verb is an ergative verb, then the order V NP $\bar{S}$ is possible since the NP can be assumed to be in its D-structure position.

(30) a. Giovanni viene a prendere il libro
(Giovanni comes to take the book.)
b. Viene Giovanni a prendere il libro
c. ?Viene a prendere il libro Giovanni

So there seem to be good reasons to postulate two different structures for the sentences in (21) and (22).

To insure that the post-verbal NP gets Case in (21) and (22), Burzio (1981) and Chomsky (1981a,b), to which we will refer here as the GB analysis, propose to deal with these constructions as if they were parallel to constructions in English and French where Case is transmitted from a pleonastic element to an NP, as in (31).

(31) a. There is a man at the door.
b. Il est arrivé un homme.

But recall that there is a definiteness effect which usually holds when Case is transmitted in such a way, so that the sentences in (32) are not acceptable because they violate this effect (cf. Safir, forthcoming, see also fn. 9, Chapter 3).

(32) a. *There is John at the door. (non-presentational)
b. *Il est arrivé Jean.

This definiteness effect does not hold in Italian however, as all the examples given above show, so that some account of this difference would have to be given if one wants to claim that the Case-assigning procedures are similar.

Note furthermore that it is not only Case but also a θ-role which is transmitted to the post-verbal position in the case where inversion has taken place as in (21). In (22), the θ-role is assigned directly to the object position.

Chomsky proposes to formalize this by assuming that α in (21) and (22) is an EC which is co-superscripted with the post-verbal NP in both cases. He then distinguishes binding, which is done by the usual coindexing, and BINDING, which holds of superscripted elements. Recall that there must be a clear distinction between binding by indexing and BINDING by superscripting: θ-roles and Case can be transmitted by BINDING, but the binding theory must apply only to binding, or else the BINDING of the post-verbal NP by α in (22) and (23) would violate Condition C or Condition B of the binding theory, depending on whether there is a name or a pronoun in subject position that has been

inverted. So it is crucial in the GB analysis that the index of the NP adjoined to VP as in (21) must not be left behind by the movement rule, or else there would be a binding-theory violation. But this is contrary to the usual assumption that both the emptied position and the moved NP keep the same index.[5]

Consider now how the present analysis can account for the inversion facts. Suppose that the VP-adjoined NP is base-generated there. What would be the consequences of this? If this position is assumed to be an $\bar{A}$-position, and if we continue to assume the indexing of A- and $\bar{A}$-positions presented in Chapter 1, this means that the index of the VP-adjoined position is only relevant at LF. Since we assume that Case assignment or checking is mediated by the index of the NP by entering this index in a V-grid or by having it coindexed with INFL for example, this means that if the index of the NP is only relevant at LF, (since the NP is base-generated in an $\bar{A}$-position), the index is not "seen" for Case assignment before LF. If we are dealing with a Pro Drop language, this poses no problem since such a language has the Case-at-LF property, as we have seen above in dealing with the missing-subject property. So the NP adjoined to VP and INFL could be matched at LF, and Case checking could be taken care of at that level. But if the language is a non-Pro Drop language, then this means that since percolation of INFL is obligatory in the syntax, Case must be launched in the syntax, as we have seen above in discussing the missing-subject

property. So the Case of the adjoined NP and the Case assigned by INFL could never be matched since the Case of INFL must be launched in the syntax, but the adjoined NP, being in an $\bar{A}$-position, is only available for Case checking at LF.

Therefore, free inversion in Pro Drop languages can be considered as free base-generation of NP in a VP-adjoined position. This is possible only in languages that have the Case-at-LF property. Since this is a necessary factor for a language to have the missing-subject property, nothing special has to be said about the possibility of having VP-adjoined NPs in Pro Drop languages: it follows from the Case-at-LF property. As for the θ-role of the adjoined NP, it is assigned as usual by the lower VP that governs it, so that the external argument receives its θ-role from the circled VP both in (33) and (34).[6]

(33) $[_S$ Giovanni $[_{\textcircled{VP}}$ ha mangiato le mele]]

(34) $[_S$ $[_{VP}$ $[_{\textcircled{VP}}$ ha mangiato le mele] Giovanni]]

We can consider that the Projection principle is respected here since the adjoined NP is in a position where θ-assignment is possible all along (see Safir (forthcoming) for a similar proposal). As for the fact that there is Agreement between the lexical NP and the verbal inflection, it is due to the coindexing of the NP and INFL at LF for Case checking, and therefore it falls into the general Agreement process between coindexed elements which takes place at LF.

This accounts for the fact that VP-adjoined NPs are poss-

ible in Pro Drop languages without any additional stipulation, given our analysis of the missing-subject property of these languages. So we now have an account for the possibility of having a post-verbal subject in structures where the NP is adjoined to VP, as in (21) above. Consider now the case where the NP is in the object position of an ergative verb as in (22), repeated here as (35).

(35) $[_S \underline{\alpha}[_{VP}[_V$ arrivano+INFL] molti studenti]]

We have seen that there are strong arguments to believe that sentences like this have a structure different from those of sentences involving NP adjunction since ne cliticization and "Complement Shift" with an $\bar{S}$ are possible in structures like (35), but not in the adjoined-NP structures. The question is: how does the NP get nominative Case in (35)? Since Pro Drop languages have the property that the Case-assigning features can remain on the verb until LF and do not have to percolate immediately in the syntax to the VP, as in non-Pro Drop languages, then this means that INFL can govern molti studenti directly in (35), and hence it can assign nominative Case to molti studenti, even if molti studenti is in object position. Molti studenti would then agree with INFL in the usual fashion for coindexed elements. This might seem to present a problem in the case of a transitive verb in a Pro Drop language. Thus, one might ask why le mele does not get nominative Case in (36).

(36) a. Giovanni [mangia+INFL le mele]
b. pro [mangia+INFL le mele]

But notice that mangia also assigns objective Case. If nominative Case is assigned to le mele in (36), then objective Case cannot be assigned. Furthermore, Giovanni would now be Caseless. Even if Giovanni is replaced by pro as in (36b), we still have the fact that objective Case is left dangling in (36) if le mele gets nominative Case. Furthermore, if we assume that θ-role assignment and Case assignment are mediated by an index which is entered in a grid, and if we assume as in (11) above that the index is entered once in a grid to get both a θ-role and Case,

(37) V: θ
GF
Case

then the assignment of objective Case and assignment of the θ-role of theme must be paired in (36), so that le mele must get objective Case in order to get the intended interpretation. If le mele was assigned nominative Case, passivization not having applied, le mele would be interpreted as the agent in (36), which is not the intended interpretation. So allowing nominative Case to be assigned directly to the object position in an ergative construction like (35) poses no problem in a Pro Drop language.

We might interpret the expletive elements there and il as being indicators that signal that the INFL features do not percolate to VP in sentences like (38).

(38) a. There is a man at the door.
b. Il y a un homme à la porte.

There and il would have an R-index with only a S-index and no REF-index, and this might be responsible for the definiteness effect: an element that has a Binder with only a S-index like there and il would have to be indefinite.

Burzio (1981) has noted that there is a contrast in the possibility of having a post-verbal subject between raising constructions and control constructions: a post-verbal subject in an embedded ergative construction is possible in raising constructions as in (39), but not in control constructions as in (40).

(39) parevano+INFL $[_S$ intervenir-ne$_i$ [molti $\underline{t}_i$]]
(40) *speravano+INFL $[_S$ di intervenir-ne$_i$ [molti $\underline{t}_i$]]

In our analysis of control, to which Chapter 5 will be devoted, local control constructions like (40) are similar to raising constructions like (39) in that in both cases, $\bar{S}$ deletion has taken place. So in (39), we can assume that the NP molti t_i gets Case because of the fact that S is not a barrier to government. Assuming as above that INFL is free to stay on the V to assign Case to an object as in (35), then we can assume that INFL can govern across to assign Case in (39). But given our assumptions about the structure of local control, the same is true in (40): nominative Case can be assigned to molti

t_i here too. The contrast therefore lies elsewhere. The main difference between (39) and (40) is that a θ-role is assigned to the subject in the matrix clause in (40), but not in (39). Now molti t_i is in a position where a θ-role is assigned by intervenir in both sentences. If we continue to assume that Case assignment and θ-assignment are both mediated by coindexing, then in (40), coindexing with INFL will result in being assigned the θ-role of the subject of speravano. In (39), on the other hand, no θ-role is associated with being coindexed with INFL. So if molti t_i is assigned Case by INFL in (40), it is assigned the θ-role of the subject of speravano and the θ-role of the object of the ergative verb intervenir: so this is ruled out by the θ-criterion. In (39) however, since parevano does not assign a θ-role to its subject, there is no θ-criterion violation since molti t_i gets only one θ-role, and the sentence is grammatical.[7,8]

So we have an analysis of the two main properties of Pro Drop languages like Italian. A missing subject is possible because requirements to have an EC can be met in subject position: the property of Case-at-LF allows the EC to be "invisible" at PF, and the required features can be assigned at LF, in most cases by the rich verbal inflection. Subjects can freely appear post-verbally because the property of Case-at-LF allows an adjoined NP to be Case marked, and it also allows direct Case marking of the object position of an ergative verb.[9]

Consider now the secondary properties of Pro Drop languages.

(41) Empty resumptive pronoun:
ecco la regazza [che mi domando [chi crede [che [$_{NP}$ e] possa VP]]]

(42) "Long WH-movement" of the subject:
l'uomo$_i$ [che mi domando [chi$_j$ t_i vede t_j]]

(43) Apparent violations of the *that-t filter:
chi crede [che t partira]

The case of the empty resumptive pronoun as illustrated in (41) is an easy one. If Italian allows resumptive pronouns, there is no reason why it would not allow an EC resumptive pronoun in the subject position of a tensed clause since a non-lexical pronoun is possible in Italian because that language has Pro Drop properties. So although the [$_{NP}$ e] in (41) cannot be the trace of WH-movement since it would violate the Subjacency effects (whether there is an actual bounding theory incorporating the notion of Subjacency or if we assume that the effects of Subjacency are derived as we have seen in 3.3.3), there can be an EC pronoun in that position since Italian is a Pro Drop language, and the pronoun could be a resumptive pronoun. So no special statement is required to account for (41): it follows from the Pro drop properties of Italian.

As for the other two problems; "long WH-movement" of the subject and apparent violations of the *that-t filter, there is a solution that has been proposed by Rizzi (1979) in which these properties are derived from the free-inversion property of Italian

(cf. also the discussion in Chomsky 1981a). Thus one could assume that "long WH-movement" of the subject and the *that-t filter violations are not possible in languages like French and English because they are ECP violations. So, for example, consider the sentence in (44).

(44) *Who do they think $[_{\bar{S}}[_{COMP}\ t_2$ that$][_S t_1$ INFL will leave soon]]

In (44), t_1 is not properly governed since INFL is said not to be a proper governor, and since the trace t_2 in COMP cannot govern t_1 either, because the that in COMP disallows the proper configuration for government (cf. our discussion of this proposal in 3.3.3). So the sentence is ungrammatical since t_1 does not respect the ECP. A similar account can be given of "long WH-movement" of the subject if the structure is as in (45).

(45) *This is the girl$_i$ that I wonder who thinks $[_{\bar{S}}[_{COMP}\ t_2$ that$]$ $[_S\ t_1$ could VP]]

Just like in (44), t_1 is also not properly governed in (45). Now suppose that Italian observes the ECP just like English. Rizzi (1979) gives the following data as support for this hypothesis (this type of contrast was first noted by Kayne for French personne).

(46) a. non voglio che tu parli con nessuno
b. *non voglio che nessuno venga (where nessuno=wide scope)
c. voglio che nessuno venga

The LF structures of these sentences would be as in (46').

(46') a. $[_S$ nessuno$[_S$ voglio $[_{\bar{S}}$ che $[_S$ tu parli con t]]]]

b. $[_S$ nessuno$[_S$ voglio $[_{\bar{S}}$ che$[_S$ $\underline{t}$ venga $]]]]$

c. $[_S$ voglio $[_{\bar{S}}$ che $[_S$ nessuno$[_S$ $\underline{t}$ venga$]]]]$

If ECP applies at LF, then the contrast in (46) could be due to ECP since the trace in (46'b) is not properly governed since INFL is not a proper governor. But Rizzi observes that the meaning of (46b) can be expressed by (47).

(47) non voglio che venga nessuno

Here, the trace of nessuno would not violate the ECP since it is governed by the verb venga, henee properly governed. What this suggests, according to Rizzi, is that the extractions illustrated in (42) and (43) are not violations of the ECP, but that in fact the extraction takes place from a post-verbal position since Italian, being a Pro Drop language, has the free inversion of the subject property. So the structures of (42) and (43) would be as in (48) and (49) respectively, where $\underline{t}_i$ is the extraction site in both cases.

(48) l'uomo$_i$ $[_{\bar{S}}$ WH$_i$ che mi domando $[_{\bar{S}}$ chi$_j$ $[_S$ $\underline{e}[_{VP}[_{VP}$ vede $\underline{t}_j]$ $\underline{t}_i]]]]$

(49) chi$_i$ crede $[_{\bar{S}}$ che $[_S$ $\underline{e}[_{VP}[_{VP}$ partira$]$ $\underline{t}_i]$

If one assumes the notion of government of Aoun and Sportiche that we discussed in Chapter 2, then the verb in (48) and (49) governs the NP adjoined to VP: so the NP is governed by an X^o, and so it is properly governed.

There are problems with this solution however. The first

one is a conceptual problem noted by Chomsky (1981a). This solution requires two nctions of c-command: under one notion of c-command, V c-commands the NP adjoined to VP since government is required for the above solution to work; but under a second notion of c-command, the V must not c-command the NP adjoined to VP since ne-cliticization is not possible from this position, as we see in the contrast between (50) and (51) (cf. (26) and (27) above), where ne-cliticization is only possible from the NP in the object position of an ergative verb, but not from an NP adjoined to VP.

(50) $[_S[_{VP}$ ne$_i$ arrivano [molti t_i]]]

(51) *$[_S$ $\alpha[_{VP}[_{VP}$ ne$_i$ mangiano le mele] [molti t_i]]]

There are also empirical problems with Rizzi's solution. This solution predicts that "long WH-movement" of the subject and the *that-t violations will only be possible in sentences where free inversion has taken place, as in (48) and (49). But there are two sets of facts that show that this is wrong.

The first counterexample comes from Brazilian Portuguese and is presented in Chao (1981). Chao shows that Portuguese allows both null subjects and *that-t filter violations freely, but that there is no free inversion of the subject in tensed embedded sentences in non-contrastive readings. This is illustrated in (52).

(52) a. missing subject: eles/Ø sairam((they) left.)

b. <u>that-t</u>: As pessoas que Joao disse [$_{\bar{S}}$ que <u>t</u> haviam saido ...]
(The people that Joao said [$_{\bar{S}}$ that <u>t</u> had left])

c. free inversion: *sairam eles (left they)

Whatever the reason why Portuguese does not allow free subject inversion[10], the data in (52) show that the explanation for the apparent violations of the ECP in this Pro Drop language cannot be that the extraction takes place from a post-verbal position where the NP would be governed by the verb, as suggested by Rizzi (1979).

The second counterexample to Rizzi's analysis comes from facts of Spanish that were brought to our attention by Esther Torrego (personal communication). It seems that, although Spanish normally allows free subject inversion just like Italian, this inversion cannot take place if the VP also contains an $\bar{S}$ complement.

(53) a. Juan dice que Maria es muy alegre.
b. *Dice Juan que Maria es muy alegre.
c. *Dice que Maria es muy alegre Juan.

An explanation for this could be given along the lines of Stowell (1981a) who claims that an $\bar{S}$ complement must be moved out of the object position and adjoined to the VP since the $\bar{S}$ must not be assigned Case by virtue of his Case-resistance principle. Assuming an analysis of this type for $\bar{S}$ complements, the structure of the sentence after movement of the $\bar{S}$ complement is as in (54).

(54) $NP[_{VP}[_{VP}\ V\ \underset{\frown}{t}]\ \bar{S}\ ]$

The reason why the subject NP cannot be inverted could be then attributed to some restriction on VP adjunction which would preclude double adjunction to the same VP. Whatever the exact explanation, the important fact is that free inversion is not possible in constructions like (54) in Spanish. The crucial fact for the question at hand is that, although free inversion is not possible in (54), extraction of the subject in contexts where the ECP is violated are still possible, as we see in (55).

(55) $Quien_i$ creia Juan $[_{\bar{S}}$ que $[_S\ t_i$ dice $[_{\bar{S}}$ que Maria es muy alegre]]]

So again, an analysis of the secondary properties of Pro Drop languages which derives these properties from the free inversion property seems to fail.

Note that, given our analysis of the ECP effects as we derived them in Chapter 3, it should make no difference whether the NP is pre- or post-verbal. Even if one resolves the problem of the need of two notions of c-command pointed out above, it still should make no difference whether the verb governs the NP adjoined to VP according to our analysis of long distance dependencies in 3.3.3. Recall that WH-movement is possible either if the WH-phrase (or its trace in COMP for subjects in certain cases) governs the trace directly in the S, or if the trace in the sentence can be related to the WH-phrase by means of an index-chain, But index-chains are possible only for elements that

enter an index in the V-grid, i.e. internal arguments, or for associate predicates which have their index incorporated in the index of a complex predicate. This means that a subject, whether it is pre- or post-verbal will never enter in an index-chain since it is an external argument. So according to our analysis, the WH-phrase in (56) could never be related to the trace with the usual properties of long distance dependencies.

(56) a. $[_{\bar{S}}$ WH $[_S \ldots [_{\bar{S}} \ldots [_{\bar{S}}$ che $[_S$ t VP $]]]]]$

b. $[_{\bar{S}}$ WH $[_S \ldots [_{\bar{S}} \ldots [_{\bar{S}}$ che $[_S$ t $[_{VP} [_{VP}$ V ...$]$ t$]]]]]]$

If free inversion of the subject is not the explanation for the apparent violations of the usual constraints on extraction from the subject position since extraction can take place even in contexts where free inversion is not possible, then what is the explanation for the facts given in (42) and (43)? The answer is that the explanation is the same as for the facts illustrated by (41): the resumptive pronoun strategy. What one can assume is that there is an empty resumptive pronoun in the subject position to which the WH-phrase is related in LF. This would account for the "long WH-movement" of the subject and the apparent violations of the *that-t filter: there would be no violation of the usual properties of long distance dependencies because the WH-phrase and the subject position are not related by means of an index-chain but by the resumptive pronoun strategy. This also gets some indirect support from the fact that the sentences analogous to the Italian (42) and (43) in non-Pro

Drop languages like French and English are acceptable with a resumptive pronoun.[11]

(57) a. This is the girl that I wonder who thinks that she will come.
b. C'est la fille que je me demande bien qui pourrait croire qu'elle ferait une chose pareille.

(58) a. Which man do you think that he will come?
b. Quelle fille est-ce qu'ils pensent qu'elle va venir?

The resumptive pronoun strategy therefore accounts in a natural way for the apparent violations of the usual constraints on long distance dependencies without running into the problem of the double notion of c-command or the empirical problems found in Portuguese and Spanish. There is one problem that remains however, and that is that this analysis cannot account for the nessuno facts presented in (46) and (47) above and repeated here as (59).

(59) a. non voglio che tu parli con nessuno
b. *non voglio che nessuno venga
c. voglio che nessuno venga
d. non voglio che venga nessuno

The resumptive pronoun strategy obviously cannot be called upon to explain these facts. But note that the embedded sentences in (59) are all subjunctives, and the same is true of the French equivalents that Kayne has discussed.[12]

(60) a. ?Je n'ai exigé qu'ils arrêtent personne.
b. *Je n'ai exigé que personne soit arrêté.
c. J'ai exigé que personne ne soit arrêté.

It seems that the importance of the subjunctive in these sentences has been overlooked and that it is crucial. According to Picallo (1982), this is a central property of these sentences and the facts are quite different with indicative sentences as we saw in 3.4. So the phenomenon illustrated in (59)-(60) seems to be of a different nature than the one that we analysed as deriving from the resumptive pronoun strategy. So, pending a better understanding of what is going on in sentences like those in (59) and (60), one should not hold these against the resumptive pronoun-strategy analysis.

Summarizing what we have ssen in this section, the properties of configurational Pro Drop languages can be derived in the present approach to ECs if we assume that these languages have the Case-at-LF property, which is obtained by delaying until LF percolation of the Case-assigning features from V to VP. So the subject position does not have Case, and if no other feature that triggers lexicalization has been inserted when we reach PF, then the NP can be an EC. If we assume that F-features are assigned to this NP at LF by the rich verbal inflection, then the EC can be interpreted as a definite pronoun: thus we derive the "missing subject" property of Pro Drop languages. The property of free inversion of the subject is derived from the fact that delaying the percolation of the Case-assigning features from V to VP allows nominative Case to be assigned to a VP-adjoined NP or to an NP in the object position of an erga-

tive verb, this being done without the need to have recourse to two notions of c-command to account for the ne-cliticization facts. Finally, the three other properties of configurational Pro Drop languages, empty resumptive pronouns, "long WH-movement" of the subject, and apparent violations of the *that-t filter were seen to fall under the resumptive pronoun strategy, so that EC pronouns are present in these structures, not traces. That these three properties could not be derived from the free inversion property as proposed by Rizzi (1979) was evidenced by the facts of Portuguese and Spanish where extraction is possible even when inversion is not possible.

4.3 Pro Drop in non-configurational languages.

We will take Japanese as a case-example of Pro Drop in non-configurational languages. In Japanese, Pro Drop is possible even though F-features are not recoverable from the inflection on the verb since this inflection is very poor. Furthermore, it seems that it is a very general property of non-configurational languages that they can have missing subjects regardless of the richness of INFL. So it is most likely that the missing subject is allowed by an intrinsic property of non-configurational languages. Let us look therefore at what we mean when we say that a language is non-configurational.

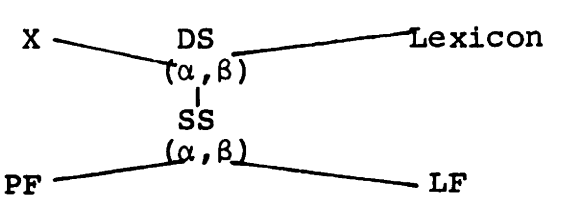
Japanese has a flat structure as in (61).

(61) $[_S\ NP_1\ NP_2$ tabe] where $S=V^{max}$

The strongest claim to make about the lexicon of Japanese

is that it is essentially the same as that of English in respects that concern us here. So for example, the verb tabe in (61) takes an NP complement: it assigns the θ-role of patient or theme to it, and it forms with it a VP that assigns the θ-role of agent to the subject. However, since Japanese has a flat structure, any of the two NPs in (61) can be forming a VP with the V: any one of these NPs can be the object of the V, and there is no apparent evidence for a VP structure in the S-structure of Japanese. To account for this, Chomsky (1981a) proposes that grammatical functions are not to be represented in D- and S-structures in Japanese in terms of the formal structures. Instead, Chomsky suggests that they are assigned randomly to D-structures and by the rule "Assume GF" to S-structures, this rule replacing move α in non-configurational languages. Thus for him, D- and S-structures in languages like Japanese are pairs (α,β), where α is a formal syntactic structure, and β is a representation of associated GFs. For languages like English, β is derived from α by abstraction from order, etc. For Japanese, α is a flat structure with no VP formed by the rule $\bar{X} \rightarrow W^* X$, and β is essentially the same as the corresponding element in English. So although Japanese does not have a VP in the β structure, it does have one in the α structure. We can schematize this proposal as in (62).

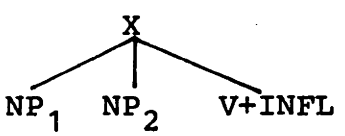
(62)



Surface structure is formed on the $\underline{\alpha}$ strueture, whereas LF structure comes from the $\underline{\beta}$ structure. There is a problem with (62) however: if we assume as in Chomsky (1981a) that Phrase Structure Rules are in fact derivable from the lexicon[13], and that all that is needed in the base besides the lexicon is a general $\overline{X}$ schema, then we see that the $\underline{\beta}$ structures can be derived from the lexicon, but the question is: where do the $\underline{\alpha}$ structures come from? What is $\underline{X}$ in (62)? This $\underline{X}$ cannot be the lexicon since in (61) for example, one of the NPs is the subject of <u>tabe</u>, hence not a complement of the V; this means that $\underline{X}$ is a set of Phrase Structure Rules.

We can solve this problem in the following way: suppose that the basic difference between configurational and non-configurational languages is not whether GFs are expressed in some structure or not, but whether the structure is ambiguous or not. For example, the structure for (61) would be (63).

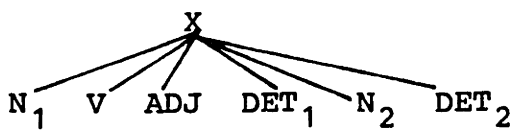
(63)



In (63), the two NPs are governed by V and also by INFL, so that either NP can be the subject or the object: both GFs are potentially expressed here, the object being governed by the V, and the subject being coindexed with the governing INFL. So the Projection principle is respected, but it is just not known what structures are to be factored out of (63). It is necessary that the structure be disambiguated at some level.

There are varying degrees in which a language can be ambiguously structured, i.e. non-configurational. For example, the difference between Japanese and a language like Warlpiri, which has even "flatter" structure, would be the following: in Warlpiri, even the elements in NPs can be "scrambled" among other elements in the sentence, and this is so because in Warlpiri, the $\underline{X}$ dominating the structure can be used to govern by any lexical head, thus making the structure wildly ambiguous as in (64) for example.[14]

(64)



So for example, any N in (64) can go and pick an ADJ or a DET under $\underline{X}$. In Japanese, however, only one lexical head can percolate its features to $\underline{X}$ and use it to govern other elements in the structure. In(63), there is ambiguity because the lexical head V+INFL is ambiguous in the sense that it contains two possible governors.

This account crucially assumes that lexical insertion determines the category of the nodes in a tree (on this topic, see Bouchard 1979, Farmer 1980, Nash 1980). Note that even in languages that allow only one head to use the $\underline{X}$ in ambiguous structure, there can be differences. For instance, we have just seen that $X=\overline{V}$ in Japanese, so that verbal complements and the subject are in free order. But as observed by Stowell (1981a), a language could have a $\overline{V}$ and a $\overline{\overline{V}}$, so that the subject would be

under $\bar{\bar{V}}$, and the VP internal complements under $\bar{V}$. If the Case system is rich enough, there could still be free order within the VP, although the subject would have a fixed position under $\bar{\bar{V}}$: German would be such a language.

The way to disambiguate structure (63) will be by Case checking. In a configurational language, which is not ambiguously structured, the relation for Case assignment is unambiguous: a governor assigns Case to a governee. But in an ambiguously structured language, this is not possible: when there is a rich Case system in a language, all the Cases of the NPs that are subcategorized for by the V cannot be assigned Case by the V if adjacency is a condition on Case assignment as proposed in Stowell (1981a), for example. Since the V cannot be responsible for all these Cases, the V will subcategorize for an NP that is intrinsically marked for a specific Case. So contrary to a non-ambiguously structured language, where Case can be assigned unambiguously by a governor to a governee, in an ambiguously structured language, the governee must <u>Assume Case</u> and check it with a governor.[15] This is crucial for the disambiguation of the structure: Case is assumed by an NP, and then the Case is matched up in a θ-grid. This again supports the idea that grids have Case specifications as we suggested in (11) above. By the above discussion, we do not want to imply that there is a difference between Case assigning and Case checking: what we really want to express is that there is a directionality in the

relation between the governor and the governee in a Case match-up: in a non-ambiguously structured language like English, the direction is from governor to governee, whereas in an ambiguously structured language like Japanese, the direction is from governee to governor.

So Case must be assumed by the NP in ambiguously structured languages, but there is no need to specify at which level the operation Assume Case takes place: the only requirement is that it must have some effect at least at LF so that the structure can be disambiguated for proper interpretation. Assume Case could not apply only at PF however, since this would not disambiguate the structure where it is relevant for interpretation.[16]

To see how Assume Case functions, consider the case of the passive derivation of a sentence with the verb atae ('award').

(65) a. $[_S$ NP NP NP atae]
b. $[_S$ NP NP atae-rare]

Let us assume that passive morphology absorbs Case (cf. Chomsky 1981a, and Marantz 1981). Since the passive morphology does not change the predicate structure of the verb, atae will still have two objects. But since one of the objects cannot have a Case relation with the V anymore, the passive morphology having absorbed one of these Cases, that object will get its Case checked by INFL. This means that no θ-role can now be assigned to the subject or that NP would get both the θ-role of the object and the one of the subject.

So in (65b), both NPs have a unique GF all along (there is no rule of Assume GF), but one of them cannot get its regular Case because of the passive morphology, so it will assume nominative Case and check it with INFL.

So ambiguously structured languages have the following properties, some of them to a varying degree (cf. Hale 1981).

(66) (1) Free word order
(2) Discontinuous expressions
(3) No movement transformations
(4) Rich Case systems
(5) No pleonastic NPs like it, there, il
(6) Free pronoun drop

(1) Free word order is possible since ambiguously structured languages allow some relations to be established where this is not possible in a non-ambiguously structured language: for example, in (63), V can govern NP_1 or NP_2. (2) If the language is more radically ambiguously structured like Warlpiri as in (64), then discontinuous expressions are possible, so that N_1 or N_2 can be related to ADJ, and/or DET_1 or DET_2 in (64). (3) No movement rules are necessary since all the proper relations and Case markings can be done without movement: it would make no sense to say that an NP is related to a trace to get a θ-role from a verb for example, since the NP can be governed by the V without moving anyhow. (4) A rich Case system is necessary to disambiguate the structure, and furthermore, assuming Stowell's (1981a) idea that there is an adjacency condition on Case assign-

ment, the NP must Assume Case since the V cannot be responsible for all the Cases assigned. The present approach to non-configurational languages also predicts that wildly ambiguous languages like Warlpiri, where structures like (64) are possible, will have on the whole a rich morphological system since discontinuous expressions will have to be identified in some way (by agreement, etc.). (5) There will be no pleonastic elements like it, there, il in such languages: if we assume that these elements are place fillers where Case is assigned, they are not present in such languages since Case is assumed by NPs, not assigned by governors. Finally, these languages have property (6): free pronoun drop. Let us now consider why this is so.

If these languages have free Pro Drop, that means that they can freely have ECs instead of lexical NPs. This means, given our assumptions, that these languages allow for Case not to be assigned to NPs. We have seen above that a crucial property of ambiguously structured languages is that they have the rule Assume Case for their NPs. Assume Case must apply to disambiguate the structure. It need not apply till LF however: assuming Case only at LF would be sufficient to disambiguate the structure for proper interpretation. We could still assume that the Projection principle is respected all along: so for example, if in (63) NP_1 must be governed by the V because it is the object of the V, it will be all along, given the structure of (63). So suppose that Assume Case applies only at LF for some NP*. This means that this NP* has no Case before LF, hence no Case at PF. If

no other feature relevant for Lexicalization is assigned to that NP*, then it can be an EC as far as the principle of Lexicalization is concerned. In order for NP* to be interpreted as a definite pronoun however, it must have an R-index and F-features on the LF side of the grammar. We can assume that the R-index is provided to NP* by free indexing at S-structure. What about the F-features? Suppose that the rule Assume Case is more general than we thought and that it generalizes to Assume Feature. We have just seen that this rule could apply at LF only: application at this "late" level would be sufficient to disambiguate the structure for proper interpretation. If Assume Feature does apply only at LF, and if it can specify all the F-features required for the interpretation of the NP* as a definite pronoun, then NP* does not have to surface in PF, regardless of the richness of INFL on the verb. Moreover, this predicts that it is not only in subject position that an NP can be "missing" in ambiguously structured languages, but that a sentence like (67) should be possible, where all three NPs are interpreted as definite pronouns.

(67) $[_{NP}\ \underline{e}]\ [_{NP}\ \underline{e}]\ [_{NP}\ \underline{e}]$ atae

This prediction is borne out: this sentence can be interpreted as in (68).

(68) He/she awarded it to her/him.

Furthermore, it has been noted that Pro Drop is possible in all ambiguously structured languages, and that it is not re-

stricted to the subject position in these languages. We give two other examples in (69) and (70).

(69) ashkii yizts'os (Navaho, from Platero 1978)
boy kissed
(The boy kissed her/him.)

(70) kotuttu (Malayalam, from Mohanan 1981)
gave
(She/he gave it to her/him.)

So our analysis of ECs allows a unified account of these facts and of Pro Drop in non-ambiguously structured languages like Italian: both types of languages, ambiguously structured or not, can have ECs as definite pronouns because in both instances, something allows Case to be assigned in such a way that it is not "visible" at PF. Ambiguously structured languages can have Pro Drop because of a basic property of these languages, namely the fact that they require a rule of Assume Case to disambiguate their structures, this rule generalizing to Assume Features. Given this basic property, it follows from independently motivated principles that ECs are allowed in all argument positions in these languages according to our analysis of ECs.

FOOTNOTES: Chapter 4.

1. Although we are sympathetic to Ken Hale's wince whenever the term 'non-configurational' is used, and we will show below that the term is inappropriate, we will continue to use it for expository purposes since it unfortunately has been around for a while.

2. The other basic way of identifying an EC is by free indexing at S-structure. We will see in Chapter 5 that the use of government for identifying an EC also extends to some instances of PRO, the other instances of the EC in subject position of an infinitive clause being identified by free indexing at S-structure.

3. Esther Torrego (personal communication) points out that we might account for the fact that emphatic pronouns must be lexical by assuming that the feature [+focus] is visible at PF, and hence forces lexicalization just like the feature [WH] for example. Thus this feature has visible effects in PF, i.e. intonational effects, in the contrast shown in (i).

(i) a. John saw HIM.
b. John saw him.

Furthermore, it is well known that clitics cannot bear such a [+focus] feature, so that a strong form pronoun must be substituted in such cases as in (ii).

(ii) a. *Jean LES a vus.
b. Jean a vu EUX.

So assuming that a feature [+focus] is added to the empha-

tic pronoun and that this feature is visible at PF since it has effects as in (i) and (ii), then in a Pro Drop language, the PF of a sentence with a [+focus] NP in subject position of a tensed clause is as in (iii).

(iii) NP VP
[+focus]

This NP cannot be an EC by the principle of Lexicalization since it bears a feature visible at PF. So it must be lexicalized as in the following Spanish example.

(iv) a. * NP vio a Jean
[+focus]

b. YO vio a Juan

4. Unless the Case is assigned to a WH-phrase which subsequently moves to COMP, carrying the Case along as we have seen in Chapter 3.

5. Rizzi (1979) proposes that the binding theory does not apply when an NP is bound from the θ-position from which that NP gets its θ-role to account for this problem. But note that the subject position is not the position from which the NP gets its θ-role in ergative verb constructions.

6. Note that if the obligatoriness of the subject in a sentence depends on the fact that tense must range over a full proposition, as we have suggested in Chapter 1, then (34) fulfills this requirement even if there is no element at all in what is the usual position of the subject of the sentence.

7. Note that in sentences like (i), John is not assigned two θ-roles as molti t_i is in (40).

(i) John tried [$_S$ PRO to go]

John Binds PRO and assigns it its R-index and F-features. PRO is an anaphor and it bears its own θ-role, just like himself does in (ii).

(ii) John shaved himself.

The difference between (i) and (40) is that in (40) an NP is actually assigned two θ-roles, whereas John and PRO get one θ-role each in (i).

8. Burzio (1981) also notes the ungrammaticality of (i).

(i) *Pareva [$_S$ Giovanni leggere molto]

He attributes this to the fact that Giovanni is not Case-governed in (i), this notion of government being different from the one of trace-government in his analysis, the latter being required for traces in the same position, as in (ii).

(ii) Giovanni pareva [t leggere molto]

Burzio relates the need of two notions of government here to the fact that Italian does not have Exceptional Case Marking which would be the case when only one notion of government is made use of in the language; so for example, "Case-government" would be identical to "trace-government" in English since English has ECM.

(iii) I expect [$_S$ John to read a lot]

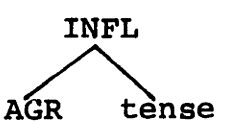
But it seems quite certain that ECM is a marked phenomenon

among languages. However, Burzio's analysis would lead us to believe that this is not the case since ECM languages make use of only one notion of government, whereas non-ECM languages require (at least) two notions of government. We will give an analysis in Chapter 5 where ECM is the marked case. Briefly, the unmarked process of objective Case assignment is to assign both Case and a θ-role at the same time, this again being consistent with the view that a V-grid is as in (iv).

(iv) V: θ
GF
Case

If a language has ECM, it is because it allows a V to assign Case without it assigning a θ-role at the same time, i.e. it allows a break in the unicity of (iv). Since Italian does not allow such a break, Case assignment and θ-role assignment by a V go together. Assigning Case to Giovanni in (i) would violate this unicity. On the other hand, in (39), we might assume that nominative Case is assigned to the head of the complement, namely intervenir, which in turn assigns it to molti t. So here the unicity requirement is respected.

9. There could be another possible way to derive these results by making use of the fact that INFL has internal structure which consists of an AGR part and a tense part, as in (i).

(i)


If we assume that tense is the nominative Case assigner,

and AGR is the element that bears the F-features, we could say that AGR percolates in the syntax, but that tense, the Case assigner, percolates only at LF as an option in Pro Drop languages. So in the syntax, AGR Binds the EC in subject position, and the F-features of AGR are then borne by the EC at LF since Agreement takes place between two coindexed elements. We see no reason to choose between this solution and the one given in the text, although it might turn out under closer scrutiny that some empirical evidence will favor one analysis over the other.

For instance, Brandi & Cordin (1981) present an analysis of two Italian dialects, Trentino (T) and Fiorentino (F), where the following facts are found.

1° Subject clitics are obligatory (except for the first sing. and second plur. in T), whether there is a lexical NP or not in the subject position.

(ii)

T		F
Vegno	I come	E' vengo
Te vegni	You come	tu vieni
El/la ven	He/she comes	E'/la viene
Vegnim	We come	Si viene
Vegni	You come	Vu' venite
I/le ven	They come	E'/le vengono

(iii)

T		F
Mi vegno	I come	Io e' vengo
Ti te vegni	You come	Te tu vieni
Mario el ven	Mario comes	Mario e' viene
Noi vegnim	We come	Noi si viene
Voi vegni	You come	Voi vu' venite
Le putele le ven	The girls come	Le regazze le vengono

2° These clitics are obligatory even in the second conjunct in cases like (iv), contrary to French where they can drop.

(iv) a. Elle danse et chante.
b. *La canta e bala (T)
c. La canta e la bala (T)
d. *La canta e balla (F)
e. La canta e la balla (F)

So as pointed out by Brandi & Cordin, the clitics in T and F seem to behave much more like an agreement form on the verb than like real clitics. This is further supported by the fact that the frequent restrictions involving specificity, definiteness and animacy found in clitic-doubling constructions (cf. Spanish, Rumanian, Hebrew) are not found in T and F.

3° In the cases where a pleonastic pro is proposed in the GB analysis, F has a "neutral" clitic.

(v) a. impersonal verbs: E' par che Mario e' sia partito.
(It seems that Mario left.)
b. extraposition: E' sara meglio anda' via
(It will be better to leave.)
c. impersonal passives: Gl' é stato trovato una borsa
(CL(neutral) has been found a bag.)
d. free inversion: Gl' é venuto una regazza
*L'é venuta una regazza
(CL has come a girl)

In T on the other hand, there is no clitic at all present in these constructions.

(vi) a. (*El) par che el Mario el sia parti
b. (*El) sara meio 'nar via
c. E sta trova 'na borsa

d. E' vegnu 'na putela
*L'ei vegnuda 'na putela

In both F & T, the verb does not agree with the inverted NP in free inversion constructions, but rather with the clitic (assuming that there is a phonetically null clitic in such cases in T).

In their analysis, Brandi & Cordin assume that INFL consists of two subparts, one which expresses the agreement with the subject, and one which is the Case assigner. This is exactly as we assume in (i) above, where AGR agrees with the subject, and tense assigns Case. The basic facts to account for are the agreement with a lexical subject as in (vii), the missing subject as in (viii), and the dummy pro as in free inversion in (ix).

(vii) a. Le putele le ven (T)
b. Le regazze le vengono (F)

(viii) a. Le ven (T)
b. Le vengano (F)

(ix) a. Ven le putele (T)
b. E' viene le regazze (F)

Suppose that we assume that AGR and tense can percolate separately as in (i) above, and furthermore suppose that AGR must percolate in the syntax. Then in (vii), both AGR and tense have percolated in the syntax, so nominative Case is assigned by tense and AGR is realized as the clitic le which agrees with the lexical subject. In (viii), AGR has percolated in the syntax, but

tense percolation is delayed until LF: so there is no Case assigned to the NP in subject position in PF, so it can be an EC and the F-features required for the interpretation of this EC as a definite pronoun are provided by AGR which Binds the EC. Finally, in (ix), AGR has percolated in the syntax, but tense percolation is again delayed until LF: so the post-verbal NP can get Case as we saw in the discussion of (34)-(35). What about AGR in such cases? We could assume that AGR, being in some sort stranded, agrees with nothing: so it is assigned some unmarked features of the language in F, and it is assigned no feature at all in T since this language presumably has no unmarked features (note the T also has some persons where no clitic is present, as we saw in (ii)).

So the difference between Italian on the one hand and F & T on the other could be that INFL always percolates as a unit in Italian, but not in F & T, this giving us the contrast just seen.

10. The reason might be, for example, that Portuguese does not allow the base generation of VP adjoined NPs.

11. The acceptability of the sentences might depend on how freely a specific dialect allows the resumptive pronoun strategy, especially in the case of (58). The fact that (58) is not as easily accepted as (57) might have to do with the Avoid Pronoun principle as in (227), Chapter 3, since both English and French have ways to have an anaphor, i.e. a variable, in (58): by non-insertion of <u>that</u>, and by the rule of <u>que-qui</u>.

12. It must be said that the judgements of acceptability for personne and nessuno sentences are very hazy. For many speakers wide scope interpretation in pre-verbal position is possible.

13. Except possibly for the expansion of S. See Stowell (1981a) who develops this idea in detail.

14. We are oversimplifying here since there are some constraints on AUX which must appear in second position in Warlpiri, and there are also constraints on infinitival complements, etc. But this is not directly relevant to the point at stake. Cf. Hale (1978), Nash (1980).

15. This is related to Kayne's (1981c) idea of unambiguous paths. In a structure like (63), if V assigns objective Case, then it can ambiguously assign it to NP_1 or NP_2 since there is an ambiguous path for government. So the NPs will assume Case instead.

16. The idea that a structure can be ambiguous is not as bizarre as it might seem at first, and there are precedents in the literature which are quite widely accepted. For example, consider the case of a lexical NP that gets its Case by Exceptional Case Marking as in (i).

(i) John believes $[_S$ Bill to have won]

In (i), Bill is governed by the matrix verb believes to get its Case, it is in some sense governed by the infinitive VP to get its θ-role, and furthermore, the whole infinitive sentence is also governed by believes and is thus assigned a θ-role. So in the ECM analysis, (i) is in some sense 3-way ambiguous structurally.

CHAPTER 5: CONTROL THEORY

5.1. General Comments.

In this chapter, we will look at another type of EC, the EC PRO which is the subject of infinitival clauses. We will see how the properties of an EC in such a construction follow from the general principles proposed in Chapter 2.

There are many analyses of these constructions in recent literature (Chomsky & Lasnik, 1977; Williams, 1980; Chomsky, 1980-1981; Koster, 1981; Manzini 1980-1982; Mohanan, 1981; Bresnan, 1982). What these analyses have in common is that they propose a specific theory of control to account for properties of the element in the subject position of infinitive clauses (whether this subject is present in the syntax as in most of these analyses, or whether it is present only at some post-syntactic level as in some cases in Bresnan (1982) for example). These theories of control specify things like what can be a controller, what can be a controllee, which NP is the controller when there are two NPs in the matrix clause, what verbs are control verbs, and what are their lexical properties. In some analyses, some of these properties are derived by other principles: for example, the fact that only the subject of infinitives is a possible controlee is derived from the fact that a controllee is a pronominal anaphor in Chomsky (1981a): this then forces it to be ungoverned for theory-internal reasons that we have seen in Chapter 2, and only the subject position of an infinitive clause or a gerund is ungoverned. These analyses also vary in that

some of them have a uniform treatment of the phenomenon (cf. Chomsky, 1981a; Chomsky & Lasnik, 1977; Manzini, 1980), whereas others assume that there are two or more subtheories of control applying differently depending on the infinitival construction involved (cf. Williams, 1980; Koster, 1981; Bresnan, 1982).[1]

What these analyses have in common is that they all propose theories of control which are meant to deal specifically with the properties of the subjects of infinitival clauses.[2] This goes against our general approach to ECs where no statement of the grammar should ever refer specifically to ECs or their properties. Instead, the properties of PRO, the EC subject of infinitives, should follow from general principles of the grammar which are also independently motivated for NPs, whether lexical or not.

The analysis that we will adopt has precisely this characteristic: assuming that there is an EC present in the syntax of infinitival constructions, we will show that its properties are parallel to some that have been observed in other NPs and that no specific statement is required to account for the properties of PRO. The analysis presents a non-uniform treatment of the EC in subject position of infinitive clauses: thus there is not a single element PRO with uniform properties as in the pronominal anaphor analysis of Chomsky (1981a) for example, but there is rather one EC with two different sets of properties which are functionally determined. This non-uniform treatment of PRO differs from

from the ones presented in Williams (1980), Koster (1981) and Bresnan (1982) however in that we do not propose two or more mechanisms to account for the properties of PRO, but rather we claim that PRO falls into two already independently motivated types of NPs. So although we agree that all ECs subject of infinitival clauses do not have the same properties, we do not make the distinction by postulating extra mechanisms in the grammar, and the division between the two types of ECs that we propose does not always fall in the same place that these analyses propose.

As we have already said in the course of preceding chapters, we claim that PRO is not a pronominal anaphor, but rather that it is either an anaphor or a pronominal, with the properties characteristic to each of these types of NPs accordingly. So the theory of control does not exist: it is derived from the principles that account for the properties of anaphors and those that account for the properties of pronominals.[3] In our analysis, locally controlled PROs are Bound anaphors, whereas long distance controlled PROs and arbitrary PROs are pronominals that freely refer. Recall Chomsky's (1981a) attempt to reduce Governing category to Binding category discussed in 2.3. Chomsky concludes that this is possible for all cases except for cases involving PRO, which he analyses as a pronominal anaphor.

(1) (i) John expected (him to win)
(ii) John tried (PRO to win)
(iii) John knows (how (PRO to win))

"In (i), him cannot be coindexed with John or (condition B) will be violated. But exactly the same argument shows that PRO cannot be coindexed with John in (ii), (iii), an incorrect result. Replacement of "binding category" by "governing category" gives the correct results, in this case. It therefore appears to be necessary to introduce a crucial reference to government in the binding theory, as in (95), though its effects are so narrow as to suggest that an error may be lurking somewhere." (Chomsky 1981a, p. 221)

But if PRO is either an anaphor or a pronominal, as we claim, then this problem does not arise. PRO is an anaphor in (ii), and so it is Bound by John. In (iii), PRO is a pronominal that corefers freely with John: this is seen in (2), where PRO is coreferential with John in (a), but not in (b).

(2) (a) John knows (how PRO to behave himself)

(b) John knows (how PRO to behave oneself)

Since there can be no $\bar{S}$ deletion in (2) and (1-iii) because COMP is filled, John cannot Bind PRO since Binding is impossible across a maximal projection, so PRO is not an anaphor: it is freely indexed like a pronominal. In (i), John and him cannot be coreferential since there is $\bar{S}$ deletion: hence John Binds him if they are coreferential, and this is proscribed by the subcase of the Elsewhere Principle given in (2.227). So if one does not consider PRO to be a pronominal anaphor, but a pronominal or an anaphor as in our analysis,

the simplification from Governing category is close to our notion of Bound anaphor, although there are some differences between the two.

We can compare how the properties of PRO are accounted for in their general outlines in Chomsky (1981a) and in the present analysis as in (3).

(3)

	Chomsky (1981a)	Present analysis
-When Pro may appear:	government and binding	Case
-Where PRO must appear:	Projection principle and Case	idem
-How PRO gets reference:	control theory	Binding or coreference

Under our analysis, a PRO can appear under the general condition which governs the appearance of any EC: it must have no $\underline{\phi}$-features at PF, hence it must be in a non-Case marked position. A PRO must appear in the subject position of an infinitive clause if the Projection principle forces the presence of an NP, and if the NP cannot be lexical since Case is not assigned. Finally, the anaphor PRO gets its reference from the antecedent that Binds it, and the pronominal PRO corefers freely, these ways of getting an R-index being independently motivated. We will see the details of how these components account for the properties of PRO by presenting the different types of PROs.

5.2. PRO as an anaphor.

If PRO is an anaphor in a certain construction, then it should exhibit the four properties that anaphors usually have: the antecedent-anaphor relation is obligatory, it is unique, it is local, and it has specific structural constraints. We have seen that these properties all follow from the fact that an anaphor is Bound by its antecedent, in the sense that the antecedent governs and assigns its R-index to the anaphor. These four properties are found in all so-called local control constructions: in these constructions, the antecedent Binds the anaphoric EC, the structure being as in (4).

(4) John tried ($_S$ PRO to leave)

Here, $\bar{S}$ deletion has taken place, assuming that $\bar{S}$ deletion takes place in the context (+V__), so that John governs PRO, hence Binds it, and PRO is functionally determined to be an anaphor. So $\bar{S}$ deletion allows government of PRO to take place, and hence Binding. We can see that $\bar{S}$ deletion is what allows the relation between controller and controllee to be established in such constructions since the relation is blocked when $\bar{S}$ deletion is impossible. For example, if the COMP is filled as in (5), the PRO has a pronominal interpretation, not an anaphoric one, as we can see by the fact that PRO freely corefers.

(5) (a) John knows ($\bar{S}$how ($_S$ PRO to behave himself))
(b) John knows ($\bar{S}$how ($_S$ PRO to behave oneself))

Similarly, one can assume that $\bar{S}$ deletion does not take place after a nonbridge verb since the $\bar{S}$ is not "identified" by the nonbridge verb since its index is not entered in the V-grid. And again PRO is not anaphoric in such cases, but pronominal, as we see in (6).

(6) (a) John shouted ($=_S$ PRO to arrest Bill)
(b) John said ($=_S$ PRO to behave)[4]

To explain the fact that PRO is not bound in sentences like (6), contrary to what her theory predicts, Manzini (1982) claims that there is a phonologically null indirect object with arbitrary reference in the matrix sentence which binds the PRO. But as she notes, this poses the problem of what this EC controller is: "It is not an NP-trace because it has no antecedent; it is not a PRO because it is in a governed position; it is not a pro because it does not have a pronominal interpretation. It can be a free variable; if however at some level of representation free variables are excluded, the mapping to this level must include a rule deleting free variables and conditions on binding like (9) must not hold at this level" (where (9) = A PRO subject of a subcategorized sentence in a sentence S is bound in S. (Manzini 1982, p. 2)

If control is possible by a subsequently deleted free variable, one might wonder why this is not also possible for verbs of local control like <u>try</u>, so that a sentence like (4) would be interpreted as in (7).

(7) (a) John tried ($PRO_{\underline{arb}}$ to leave)
(b) 'John tried that some <u>x</u> leave'

If it is a lexical property of verbs like try that they do not allow such a phonologically null indirect object, then this amounts to stating in a converse way that try is a subject control verb: so it undermines the whole attempt to reduce control theory to binding theory. And one can also wonder at what consequences this use of null elements has on the Projection principle.

This use of phonologically null indirect objects also misses the relation that exists between bridge verbs and local control verbs on the one hand, and nonbridge verbs and their property of being "pronominal control" verbs on the other hand. As noted in Chomsky (1981a), control verbs are invariably bridge verbs. We might see a corollary in the fact that nonbridge verbs are invariably non-local control verbs.[5]

In our analysis, there is nothing to say about verbs like those in (6): since they are nonbridge verbs, $\bar{S}$ deletion does not take place, and hence it follows that Binding is not possible across $\bar{S}$, hence that "local control" is impossible. On the other hand, pronominal control is predicted to be possible according to our analysis: so the PRO can freely corefer, with the possibility of having arbitrary reference as in (6), or a coreferent reading as in (8).[6]

(8) (a) Mary_i knew that John had said ($\bar{X}$ PRO_i to behave herself_i), but she_i refused to do it.

(b) Mary_i remembered that John had said/whispered ($\bar{S}$ PRO_i to behave herself_i)

If we assume that $\bar{S}$ deletion is what allows local control by allowing Binding, then the problem of our analysis lies in pairs like (9) and (10).

(9) John tried ($_S$ PRO to win)
(10) John believed ($_S$ himself to be the winner)

Conversely, the ungrammaticality of sentences like (11) and (12) raises the same problem.

(11) *John tried ($_S$ Bill to leave)
(12) *John believes ($_S$ PRO to be the winner)

Recall that in our analysis of ECs, Case is a crucial factor in their distribution: ECs cannot have any ϕ-features at PF, hence no Case. But then, if $\bar{S}$ deletion is possible after try to allow Binding of PRO, why isn't Case assigned to Bill in (11)? And conversely, if $\bar{S}$ deletion is what is allowing ECM in (10), why can't John Bind PRO in (12)? The answer to these questions lies in the proper analysis of ECM. If $\bar{S}$ deletion is the only factor that allows ECM since it allows government by a Case-assigner, then some generalizations are missed. For example, French has $\bar{S}$ deletion since it has raising as in (13).

(13) Jean semble ($_S$ t avoir fini)

Whether it be because of ECP as in the GB framework, or because of Binding of the trace by John as in the present framework, there is $\bar{S}$ deletion in (13) to allow government of the trace. But it is well known that ECM is not possible

in French in that there are no verbs like believe that assign Case to the subject NP of the embedded infinitive clause.

(14) *Jean croit ($_S$ Paul être intelligent)

But if ECM depends only on the option for a V to trigger $\bar{S}$ deletion or not, this would mean that the fact that no transitive verb in French triggers $\bar{S}$ deletion, with subsequent ECM, is purely accidental. Surely, this would be missing a generalization.

So it seems that ECM is more than just $\bar{S}$ deletion. For instance, Kayne (1981b) proposes to relate the ECM facts in (15) to the facts about Case-assignment in (16).[7]

(15) (a) John believes [$_S$ Bill to have lied].
(b) *Jean croit [$_S$ Bill avoir menti].
(16) (a) John prefers [$_S$ for [$_S$ Bill to leave].]
(a') *John prefers [$_S$ for [$_S$ PRO to leave].] (dialectal)
(b) *Jean a décidé [$_S$ de [$_S$ Pierre partir].]
(b') Jean a décidé [$_S$ de [$_S$ PRO partir].]

Kayne's suggestion is that prepositions in French do not govern structurally like prepositions in English, so that de cannot assign Case in (16b), even if there is only one S boundary; and de does not govern PRO in (16b') for the same reason, so that PRO can appear in such a position. When a preposition does seem to assign Case in French as in J'ai parlé de Jean, Kayne claims that in such cases, "P in French, rather than assigning Case, is subcategorized for (some specific) Case" (p. 364, fn. 21). Just the converse is true of for in English, so that Case is assigned in (16a), and

government precludes the presence of PRO in (16a'). Kayne then transposes this analysis to account for the facts in (15) in the following way. He proposes to account for ECM by postulating a Ø preposition in COMP which assigns Case to the subject NP, if it is itself assigned Case by some higher Case assigner. So the structures for the sentences in (15) are as in (17) according to Kayne (1981b).

(17) (a) John believes ($\bar{S}$ Ø ($_S$ Bill to have lied))
(b) *Jean croit ($\bar{S}$ Ø ($_S$ Bill avoir menti))

Since prepositions do not govern structurally in French, the facts in (17) reflect the facts in (16): the Ø preposition cannot assign Case in French, just like de cannot in (16b). In English on the other hand, the Ø preposition can assign Case although it is not subcategorized for Bill.

Furthermore, Kayne relates these facts to facts about preposition stranding in the two languages.

(18) (a) Which candidate have you voted for t?
(b) *Quel candidat as-tu voté pour t?
(19) (a) John was voted against by almost everybody.
(b) *Jean a été voté contre par presque tous.

Assuming that reanalysis of V-P is involved in preposition stranding, Kayne asks the question: why should French not have a reanalysis rule just like English? Kayne suggests that reanalysis between two lexical categories is possible only if the two govern in the same way. Since V and P do not

govern in the same way in French, reanalysis between these two categories is not possible, since V and P govern in the same way in English, reanalysis is possible. So in Kayne's analysis, the French-English contrast with respect to preposition stranding reduces to (20).

(20) In French, P and V do not govern in the same way, but in English they do. (That is, in English, P can govern structurally, as well.)

In fact, Kayne claims that reanalysis is possible in French: thus it would be involved in causative constructions, assuming an analysis of the type in Rouveret & Vergnaud (1980).

(21) Marie a fait partir Jean.

Since the reanalysis is of the V-V type, Kayne's analysis predicts that this is possible since the two elements involved govern structurally here.

Although Kayne's analysis is interesting in that it captures a relationship between ECM, preposition stranding and Case assigning by a P in COMP to the subject of an infinitive clause, it faces problems which suggest that some modifications of the analysis are warranted.

First, there is an empirical problem in that, as noted by Kayne himself, a V-N reanalysis is possible in French for idiomatic expressions like mettre fin 'put an end (to)': Pollock (1979) showed that they can be subject to reanalysis in passive sentences like (22).

(22) Je veux que soit mis fin à la guerre.

For Kayne's analysis to follow, this would require that V and N be governors of the same type. But Kayne, in work in progress, rejects the idea that N is a structural governor: this, according to him, would explain the requirement of the P of in (23) to satisfy ECP, since the N care is not a structural governor for the trace.

(23) John was taken care of t by Mary.

The need for of cannot be due to Case requirements here since John gets nominative Case.[8]

So the fact that reanalysis of the V-N type does occur in French poses a problem for Kayne's analysis.

Another problem lies at the conceptual level. If, as assumed generally in GB, θ-marking entails subcategorization, and if all of this is done under government (understood as structural government here), then this means that in Kayne's analysis, P in French would fall outside of this general pattern: it subcategorizes for some specific Case although it does not govern its complement. This would seem to require complications in the grammar that we will not try to explore here. Furthermore, it would make French the exceptional case and English the unmarked case with respect to preposition stranding and ECM: But it seems quite clear that English is the marked case in these respects (overwhelmingly so as far as statistics are concerned anyhow).

We could maintain the spirit of Kayne's analysis and keep

its nice results by restating it in a slightly different way however.[9]

We have assumed so far that θ-role assignment and Case assignment are done under government (at least for configurational languages; for nonconfigurational languages, cf. chapter 4). Furthermore, we assume that θ-role and Case assignment usually go together, so that insertion of the index of a governed complement in a V-grid, for example, results in assignment of both θ-role and Case since such a grid has the form given in (24).

(24) V: θ
|
GF
|
Case

Suppose that we strengthen this assumption and that we assume, in the spirit of Marantz (1981), that NPs in sub-categorized PPs like the underlined phrase in (25) are not assigned their θ-role by the V, but are assigned a θ-role by the P.

(25) (a) John talked about Bill.
(b) John gave it to Mary.

Consider for example the sentences in (26) and (27) from Marantz (1981).

(26) (a) Elmer gave the porcupines to Hortense.
(b) Elmer put the porcupine on the table.
(c) Elmer stole a porcupine from the zoo.

(27) (a) The train to Pittsburgh arrived at the station.
(b) The porcupine on the table slipped its leash.
(c) The porcupine from the zoo was tamer than the rest.

As pointed out by Marantz, "The underlined NPs in the sentences in (26) bear the same semantic roles as those born by the underlined NPs in the corresponding sentences in (27). If we assumed that arguments receive their semantic roles simply by virtue of filling argument slots in P-A structures (Predicate Argument structures), then it would be an accident that the items which are used to mark a verb's arguments independently assign the semantic roles that the argument they mark bear when these items are not being used to mark a verb's arguments. If the "source" argument of steal, for example, received its semantic role by occupying the second slot in P-A structure (of steal), it would be an accident that from, which assigns the source role in other constructions (see, e.g. (27c)), is used to mark steal's source argument. Since steal in this case would, in effect, be assigning the source role itself, from would be unnecessary.

If all of a verb's arguments received their semantic roles from the verb, we might expect all the arguments to be marked in the same manner, or with some arbitrary marking to specify which argument goes in which slot in a P-A structure. That arguments of a verb are marked in the same manner as NPs bearing identical semantic roles which are not arguments of a verb is the strongest evidence for viewing the assignment of semantic roles to arguments as independent of P-A structures."

(Marantz (1981), p. 52-53)

So we could generalize to all categories the proposal that assignment of θ-role and assignment of Case go together in the unmarked case. If we consider that French falls into the unmarked cases, as seems most likely, this would mean that French respects this general schema for Case and θ-role assignment: an element will assign Case to an NP only if it also assigns a θ-role to that NP. A direct consequence of this is that ECM should be impossible in French, since ECM consists precisely in separating the Case-assigning properties of an element from its θ-assigning properties. Thus, (28) is ungrammatical.

(28) *Jean croit ($_S$ Paul être intelligent)

since croit cannot assign Case to Paul since it does not assign a θ-role to Paul. Similarly, de cannot assign Case to Jean in (29) since it does not assign a θ-role to Jean.

(29) *Ils ont parlé ($_{\bar{S}}$ de ($_S$ Jean partir))

In English on the other hand, we can say that a marked option of assigning Case independently from assigning a θ-role is allowed, so that there is no one-to-one correspondence between the two anymore: government by the Case-assigner is sufficient for Case to be assigned, as we see in (30).

(30) (a) John believes ($_S$ Bill to be happy)
(b) They prefer ($_{\bar{S}}$ for ($_S$ Bill to leave))

So ECM would come about when a language allows a separation

of Case-assigning properties from θ-assigning properties.

Similarly, reanalysis is possible between V and P in English but not in French because it involves transferring the Case-assigning properties to the V and leaving the θ-assigning properties to the P. We can see this by the fact that passive morphology on the V absorbs the Case-assigning properties of the construction, as in (31).

(31) John was talked about at the conference.

So again we can assume that this is dependent on the possibility of a separation of Case-assigning properties from θ-assigning properties which is allowed in English, but not in French.

We can assume that this possibility of separating Case assignment from θ-role assignment is a marked property of lexical items in English. So for example, <u>believe</u> has this property but <u>try</u> does not.

(32) (a) John believes ($_S$ Bill to be the best)
(b) *John tried ($_S$ Bill to be the best)

Sich an analysis gives a unified account of ECM and reanalysis and it marks English as the special case with respect to ECM and preposition stranding. This analysis shows that the fact that <u>believe</u> is an ECM verb, whereas <u>try</u> is not, has nothing to do with a difference in triggering $\bar{S}$ deletion between the two Vs since both Vs trigger $\bar{S}$ deletion: it is due to a lexical property of <u>believe</u> which makes it assign Case to an NP that it governs, even if it is

not subcategorized for it. Try on the other hand does not have this property. So in a construction like (33), where $\bar{S}$ deletion has taken place since the sentential complement is infinitive and is subcategorized for by the V,

(33) V ($_S$ NP ...

the NP cannot be lexical if the V is try since, although it would be governed by the V, try does not have the property of separating its Case-assigning properties from its θ-assigning properties, and try does not assign a θ-role to this NP. A verb like try can be a Case assigner however, and it can assign Case to an NP if it also assigns a θ-role to that NP, as in (34b) and (35b).

(34) (a) *John tried ($_S$ Bill to go)
(b) John tried the boat.

(35) (a) *John prefers ($_S$ Bill to go)
(b) John prefers the red boat.

If the verb in (33) is of the believe type however, i.e. an ECM verb, then the NP cannot be an EC since it is assigned Case and hence must be lexical by the principle of Lexicalization. Thus we have the contrast in (36).

(36) (a) John believes ($_S$ Bill to be intelligent)
(b) *John believes ($_S$ PRO to be intelligent)

So we can assume that $\bar{S}$ deletion takes place for every subcategorized infinitival complement: if the verb is of the believe type, Case is assigned to the NP subject of the infinitive clause, which must therefore be lexical; if the verb is of the try type, the NP cannot be lexical since it

does not receive Case, but it is an anaphor since it can always be governed by an NP in the matrix clause: hence the obligatoriness of "control" in these constructions since Binding is always possible.

Kayne (1981b) discusses the contrast in (37).

(37) (a) *Je crois/reconnais/constate Jean être le plus intelligent de tous.
(b) Quel garcon crois/reconnais/constates-tu être le plus intelligent?

He observed that these French verbs have the property that a sequence V NP VP is ungrammatical if NP is lexical and remains in place, but grammatical if NP is a WH-phrase and moved. But this contrast is never found in "normal" verbal objects.

(38) (a) Je crois Jean intelligent.
(b) Quel garçon crois-tu intelligent?

What Kayne suggests is that in (37b), Case is assigned to quel garcon when it is in COMP, assuming a successive cyclic derivation of this sentence. So there is an intermediate stage as in (39).

(39) V ($\bar{S}$ ($_{COMP}$ quel garçon) ($_{S}$ t être...

Kayne assumes that Case can be assigned across one S-type boundary, so that V assigns Case to quel garçon in (39). Note that this stipulation that Case can be assigned across only one S-type boundary misses the generalization that Case is assigned under government (in most cases) and

that $\bar{S}$, being a maximal expansion, blocks government to an element internal to the $\bar{S}$ which is not a head of $\bar{S}$.

So assignment of Case into COMP, with the embedded subject brought "close enough" to the Case assigner, explains the contrast in (37) according to Kayne. Note that this must be specified as a property of certain verbs only, or else Case would be assigned in COMP in sentences like (40), incorrectly predicting such a sentence to be grammatical.

(40) *Who did John try ($_{\bar{S}}$ t ($_{S}$ t to go))

One solution could be as in Chomsky (1981a): to assume that try is "intransitive" when it has an infinitival complement, i.e. that it does not assign Case. In Chomsky's analysis of control verbs, $\bar{S}$ deletion is impossible after verbs like try, in order to allow the presence of PRO which must not be governed in his analysis. But then to say that such verbs are also non-Case assigners would be missing the generalization that all verbs that do not allow $\bar{S}$ deletion in English are also "intransitive" verbs. In our analysis, $\bar{S}$ deletion is a general process that takes place (at least) in the context (+V___): It is not a lexical property. Furthermore, a verb like try is a Case assigner, but it is not an ECM verb, so that it does not separate its Case-assigning properties from its θ-assigning properties. So it is because try does not assign Case to PRO that this PRO is possible.[10]

As for verbs like croire in French, they follow the

general situation in French which is that no verb is an ECM verb as we can see by the ungrammaticality of (38a). However, if the structure of (37b) is at some point as in (39), and if COMP heads $\bar{S}$ in some sense, as we have seen and assumed above, then V can assign Case in COMP in (39). So we could say that French has a class of "weak ECM verbs": these verbs can assign Case to an NP for which they are not subcategorized provided that this NP is in the COMP position of an $\bar{S}$ for which they are subcategorized, hence an NP that partly heads that $\bar{S}$ complement, so that the verb "weakly subcategorizes" for that NP. A stipulation of this type will have to be made in any analysis to account for the contrast between (37b) and (40).[11]

What we have seen so far about "local control" constructions is that the general process of $\bar{S}$ deletion, which applies in the context (+V__), allows an NP from the matrix clause to Bind the subject position of the infinitive clause. This subject position of the infinitive clause is an EC anaphor if, like all other ECs, it does not have a $\underline{\phi}$-feature when it reaches PF; otherwise, it is a lexical anaphor. Hence, the EC must not be assigned Case. We assume that the unmarked situation is for Case and Θ-assigning properties to go together, so that even if a verb governs an NP as in the local constructions, the verb does not assign Case to the NP since it does not assign a Θ-role to the NP. English has the exceptional property that some of its lexical items allow a separation of the Case and Θ-assigning features, so

that English has ECM verbs like believe, it has a preposition for in COMP which can assign Case to an NP that is governed but is not subcategorized for, and it has reanalysis between V and P, where V absorbs the Case marking properties of the construction, although P keeps its θ-assigning properties, all of these being instances of this separation of Case and θ-assigning properties.

5.2.1. The distribution of PRO: Case or government?

If, as we claim, Case is relevant for the distribution of PRO, like for all ECs, rather than government as is claimed in GB, it can be interesting to look at constructions where only one of these properties seems to be present in order to determine what is really the relevant property. Is the distribution of PRO determined by the fact that is must be ungoverned, or by the fact that it must not be Case marked? The constructions to look at are 1° when there is government but no Case is assigned, and 2° when Case is assigned but there is no government. These are the constrictions where the two analyses make different predictions.

5.2.1.1 Government but no Case.

There are (at least) four constructions where an EC can appear in a non-Case marked position which is governed, and where the EC cannot receive a PRO interpretation: raising as in (41), passive as in (42), ergative as in (43), and Adjective-NP as in (44).

(41) It seems ($_S$ e to be happy)
(42) It was hit e
(43) Il est arrivé e
(44) John is proud e

In all these constructions, the EC is interpreted as a trace, not as a PRO. So for example, if it is non-referential in (41), the sentence cannot be interpreted as (45): it is simply ungrammatical if it is non-referential.

(45) It seems (PRO_{arb} to be happy)

The same holds for (42) and (43). And in (44), the sentence cannot be interpreted as in (46).

(46) John is proud of some x.

This seems to contradict our claim that the distribution of PRO is dependent on Case, not on government. In a theory that has a theorem that PRO must be ungoverned on the other hand, PRO is not possible in (41-44) since it would be governed. But we have seen in our discussion of NP traces in 3.2 that there is a straightforward explanation for the fact that the EC is not interpreted as PRO_{arb} in (41-43). In these constructions, the EC is Bound by it, and hence it is functionally determined to be an anaphor, so that an interpretation as in (45) is not possible since the EC has the R-index and the F-features of it. One piece of evidence in favor of this analysis comes from German impersonal passives. Recall that German can have impersonal passives as in (47).

(47) Es wird gelacht
it is laughed ('People are laughing')

In (47), the verb is a non-ergative verb, and the effect of passive morphology is simply to eliminate the θ-role assigned to the subject: an expletive element es is then inserted in the subject position. The interpretation in (47) is possible because there is no EC for es to Bind in object position, so that the situation found in English in constructions like (41-43) does not arise. If the verb is an ergative verb on the other hand, then there is a θ-marked position in object position: in such a case, the expletive es Binds this θ-marked position, and this results in the same type of ungrammaticality as in the English constructions.

(48) *Es wird gefallen
it is fallen ('People fall')

So the fact that constructions like (41-43) and (48) are ungrammatical follows directly in our analysis. On the other hand, if there is a referential subject in one of these constructions as in John seems to be happy, the EC is interpreted as a trace, not as PRO, in order for John to get a θ-role.

Consider now the case of (44). We already have an explanation for the ungrammaticality of this sentence: as we saw in 3.2, of-insertion must take place in the context (+N__NP), regardless of the lexical content of NP. So (44) is not interpretable as (46) because PRO_{arb} cannot be the

object of proud since of-insertion applies and forces lexicalization of NP.

So we have explanations as for why the EC cannot receive a PRO interpretation in these four constructions where there is government but no Case marking. On the other hand, there are constructions of this type where PRO can appear. Thus, according to our analysis, all local control PROs are governed but not Case marked, hence we account for their strictly anaphoric interpretation.

(49) John tried ($_S$ PRO to go)

Similarly, in French, PRO can appear in a position governed by de when de cannot assign Case to that position, as in (50).

(50) Jean a peur ($_{\bar{S}}$ de ($_S$ PRO partir))

We have seen in our discussion of ECM above that an account can be given of this fact without having to resort to different ways of assigning Case as in Kayne's analysis, where Case can be assigned under government or by subcategorization. We saw that this creates potential problems for the Projection principle since subcategorization is no longer uniformly dependent on government in such an analysis. The problem that (50) raises for a GB analysis is that some account must be given as to why de does not govern PRO in such a construction.

Another construction which can be considered to fall together with (50) is the strong preposition construction

discussed in 3.3.2.2.

(51) (a) Un gars que je me fierais pas dessus.
(b) Le gars que je vais voter pour.
(c) Tu es assis dessus.
(d) Je suis pour.

We could assume that these Ps do not assign Case, so that the content of the PP is P-PRO (dessus PRO, pour PRO). Since reanalysis of V-P is not possible in French, as we saw above, the PRO cannot be moved out of the PP by WH-movement since it could not Bind its trace inside a PP, this being a maximal expansion which therefore blocks government. So it remains in place, which is attested by the fact that such constructions do not obey subjacency (cf. 3.79). So here, PRO is governed by the P, but it is not assigned Case.

Finally, reflexive constructions, when they are morphological as in some instances in Romance languages, could also be an instance where PRO is governed but not Case marked. Consider French se for example. We could assume that se is simply a Case absorber like all other clitics. That clitics are only Case absorbers, and not θ-role absorbers, is especially clear in languages that allow clitic doubling, where a lexical NP can appear with a clitic if a Case assigner is available, as in the general schema in (52).

(52) CL_i +V P* NP_i

If se is only a Case absorber in reflexive constructions,

then the structure can be as in (53).

(53) Jean se lave PRO souvent.

Here, se absorbs the Case normally assigned to the position where PRO is, and Jean Binds PRO, so that PRO is an anaphor. PRO is governed by lave since it gets a θ-role from lave. PRO is also governed by its antecedent Jean. So se is very similar to passive morphology: it absorbs accusative Case, but also dative Case, contrary to passive morphology. Passive morphology always eliminates the θ-role assigned to the subject: se does it optionally. Thus, in the reflexive interpretation of (53), the θ-role of the subject is not eliminated: Jean is base-generated in subject position where it gets its θ-role, and it Binds the anaphor PRO which also gets a θ-role. But (53) can also have the interpretation in (53').

(53') Jean_i se lave t_i souvent.
'One washes Jean often'

In (53') the θ-role of the subject has been eliminated: Jean was inserted in object position in D-structure, and Jean has to move to subject position to get Case since se absorbs the Case assigned to the object position. Another way for the subject to get Case in such instances is by inserting expletive il which transfers Case to the object, with the usual definiteness effect as in il se vend beaucoup de pommes à cette foire.

So we see that in the case of governed but not Case marked ECs, some of them cannot be PRO because they are either Bound by an expletive element (cf. (41-43)), or because they are not really in a non-Case marked position since of-insertion applies to assign Case. On the other hand, for some of these positions, to be interpreted as PRO allows a simple account of the strictly anaphoric properties of the EC in question, without having to revert to analyses where sub-categorization is no longer dependent on government: this is the case for locally controlled PRO, strong prepositions in French, and morphological reflexives.

5.2.1.2. Case is assigned, but there is no government.

Consider now the second type of construction where different predictions are made by analyses where government is relevant and analyses where Case is relevant for the distribution of PRO. If Case is assigned, but there is no government, a PRO is predicted to be possible in a GB analysis, in an analysis where the distribution of PRO depends on Case as ours however, the prediction is that PRO should not be possible, since no EC should be possible (except if Case is carried along like in WH-movement constructions).

In order for an NP to get Case, although it is not governed, it must be that Case is not assigned under government. We have seen two constructions where this is possible: in constructions where genitive Case is assigned (cf. 3.2), and

in nonconfigurational languages, where the rule Assume Case (generalized to Assume features) is responsible for Case marking (cf. 4.3). In the latter case, we have seen that an EC is not possible if Case is "visible" at PF, which is the level we must consider here when we say Case is assigned but there is no government, since that is the level where the Principle of Lexicalization applies. However, if the Case is only assumed at LF, hence is invisible at PF, then a pronominal EC is possible in these languages.

5.2.1.2.1. Gerunds.

Consider constructions where Genitive Case is assigned. Recall the rule of Genitive Case assignment (3.19), given here as (54).

(54) Genitive Case Assignment
In the configuration ($\underline{\alpha}$ ... $\underline{\beta}$ ...), assign Genitive Case to $\underline{\beta}$, where
(i) $\underline{\alpha}$ is some projection of (+N, -V), and
(ii) $\underline{\beta}$ is an immediate constituent of $\underline{\alpha}$,
(iii) $\underline{\beta}$ = NP

If we assume that Case assignment rules are obligatory, as seems natural, then Case is obligatorily assigned to NP* in a construction like ($_{NP}$ NP* $\bar{N}$), so that, according to our analysis, sentences like those in (55) are ungrammatical, given the principle of Lexicalization.

(55) (a) *John$_i$ read ($_{NP}$ PRO$_i$ book of poems)
(b) *The Barbarians$_i$ enjoyed ($_{NP}$ PRO$_i$ destruction of Rome)

So these sentences present no problem for our analysis: PRO is not possible here since any NP in such a position is assigned Genitive Case, hence must be lexical. They do present a problem for an analysis where government is the relevant factor for the distribution of PRO, since assignment of Case, if not under government, should have no effect on the distribution of PRO. In Chomsky (1981a), adapting from an idea of Aoun & Sportiche, Chomsky proposes to solve this problem by saying that PRO is governed in (55), although it is not governed for purposes of Case assignment. Chomsky adopts the notion of government of Aoun & Sportiche, which states that a head governs all elements in its projections (except if such an element is in another maximal expansion, of course). The structure for (55a) for example is (56).

(56) $\ldots({}_{NP}\ ({}_{NP}\ PRO)\ ({}_{\bar{N}}\ book\ ({}_{PP}\ of\ ({}_{NP}\ poems))))$

In (56), book governs PRO according to the Aoun & Sportiche notion of government. This would allow one to capture the contrast between (55) and (57).

(57) John likes $({}_{NP*}\ ({}_{NP}\ PRO)\ ({}_{VP}\ reading\ books))$

Assuming this structure for gerunds, in (57), PRO is not governed according to Aoun & Sportiche since NP* is not an expansion of reading because of the intermediate NP node.

But there are problems with this analysis. First, in constructions like $({}_{NP}\ NP^*\ ({}_{\bar{N}}\ N\ldots))$, if N governs NP*, we might expect a subcategorized element to appear in such a

position. To avoid this unwanted result, one must say that there are two notions of government: a more narrow one which is relevant for subcategorization, and a wider one such as the Aoun & Sportiche notion of government. But one might wonder at the justification of such a complication of the grammar if the second notion of government is introduced only to rule out constructions like (56), since these can be ruled out by Case theory.[12]

Note also that if *($_{NP}$ PRO($_{\bar{N}}$ ($_{N}$ book))) is not possible because book governs PRO, then one has to complicate the account for adjuncts like the one in John came home (PRO drunk). The structure of the adjunct cannot be simply ($_{AP}$ PRO($_{\bar{A}}$ ($_{A}$drunk))) in such an analysis since the PRO would be governed just like the case of the nominal above. One has to adopt some structure where the node dominating the NP PRO is not a projection of ADJ, something like ($_{ADJUNCT}$ PRO ADJECTIVE), a structure which has no independent motivation and which is a violation of strict $\bar{X}$-theory since there is no lexical category ADJUNCT.

Another problem with this notion of government is that one expects this notion of government to be paralleled by a similar notion of C-command. But then this would mean that in a construction like (58),

(58) ($_{NP}$ NP_1 ($_{\bar{N}}$ N NP_2))

NP_2 c-commands NP_1. So if a language had Genitive Case assigned to both NP_1 and NP_2, so that NP_2 is not in a PP,

since NP_2 then c-commands NP_1, it is predicted that NP_1 can be an anaphor bound by NP_2. But this prediction is not borne out, as we can see in the following example from Japanese. (Similar examples could be constructed in other languages with overt Case markings.)

(59) *(NP otagai-no karera-no hihan)
each other-Gen they-Gen criticism

Consider again the contrast between "normal" NPs and gerunds. In order for the Aoun & Sportiche analysis to hold, gerunds must have the structure given in (57), where NP* is not headed in the usual way in $\bar{X}$ theory. So this analysis works only if this violation of a very strong axiom of $\bar{X}$ theory is allowed. But this is not necessary if one adopts an analysis of gerunds similar to Reuland's (1980) analysis. What we could assume is that a gerund is headed by a (+N, -V) element, namely -ing. We can still assume that all affixation is done in the lexicon, as is assumed in lexical phonology. The structure of the head of the gerund is as in (60).

(60) $\begin{bmatrix} \text{stem} + \text{-ing} \\ (-N, +V)\ \ (+N, -V) \end{bmatrix}$

If the features of -ing percolate up in the structure, then the whole gerund phrase is labelled as a projection of a (+N, -V) element, hence it is an NP as in (61).

(61) $(_{NP}\ldots(_{\bar{N}}\ldots(_{N}\ V+\ (_{N}\underline{ing}))))$

A strong percolation of these nominal features as in (61) would give a phrase with the properties of a nominal

gerund as in Lees (1960): the gerund takes adjectives, the object is introduced by of, there can be a determiner or a Genitive subject, the gerund can be plural, and negation is indicated by the particule no. In this case, according to our analysis, PRO is never possible in subject position since Genitive Case is assigned if an NP is in subject position, forcing lexicalization. So if any of these properties of nominal gerunds is present, a locally controlled PRO subject is impossible. We give some examples in (62).

(62) (a) I_i enjoy (PRO_i singing)
(a') *I_i enjoy (PRO_i good singing)
(b) I_i enjoy (PRO_i singing operas)
(b') *I_i enjoy (PRO_i singing of operas)
(c) I_i enjoy (PRO_i reading a good book)
(c') *I_i enjoy (PRO_i readings)

But on the other hand, some gerunds do allow a PRO subject, as we saw in (57). And as observed by Lees (1960), some gerunds seem to be verbal since they have all the equivalent verbal properties of the nominal gerunds given above. Thus verbal gerunds take adverbs, the object gets accusative Case, no determiner is allowed, no number marking is possible, and negation is indicated by the particulate not. One way to account for this could be to say that -ing can also be (-N, +V): so the head of a gerund could also be of the form in (63), with percolation as in (64).

(63) $\begin{bmatrix} \text{stem} & + & \underline{\text{-ing}} \\ \text{(-N, +V)} & & \text{(-N, +V)} \end{bmatrix}$

(64) $(_{\bar{\bar{V}}} \ldots (_{\bar{V}} \ldots (V+ (_{V}\underline{\text{ing}}))))$

Since verbal gerunds do not have a COMP since they do not allow fronted WH-phrases, they are presumably Ss, not $\bar{S}$s.[13] Since they are Ss, the PRO subject can be Bound by an antecedent if the verbal gerund is in object position as in (62a-b-c).

But facts are not clear cut as this. There are ambivalent cases where a gerund seems to have properties of nominal gerunds and properties of verbal gerunds. For example, consider a sentence like (65).

(65) I hate (John's singing operas)

Here, the assignment of Genitive Case to John is a property of nominal gerunds, but the assignment of Accusative Case to operas is a property of verbal gerunds. What we can assume is that in such cases, -ing is nominal, so that its features percolate up to label the gerund as an NP, and John is assigned Genitive Case. But at the same time, it seems that some features of the verbal stem sing are also allowed to percolate to assign Accusative Case to operas. This is reminiscent of the convention on percolation presented in Lieber (1980) and Marantz (1981) which states that, in some cases, a feature of a non-head stem, as in this case for sing, can percolate up if the head is not specified for such

a feature: the feature here would be the Accusative Case assigning feature of sing, for which the head -ing is not specified. So it is the ambiguous nature of the head sing+ing as ((-N, +V) + (+N, -V)) which creates such situations. Note that if an adjective is inserted in the gerund to modify the head, then the gerund must be strictly nominal, and Accusative Case cannot be assigned by the V-stem as we see in (66).

(66) John's loud singing *(of) operas

This is due to the fact that the ADJ is attached at a low point in the structure, and that assignment of accusative Case and bearing an ADJ would clash at that level.

So the ambiguous nature of the head of the gerund explains why some nominal gerunds appear to have the verbal property of assigning Accusative Case. On the other hand, there seem to be verbal gerunds in complement position that do not exhibit Binding of their PRO subject although the analysis presented so far predicts Binding to be obligatory in such cases. Thus, as noted by Wasow & Roeper (1972) (henceforth W&R), the sentence in (67) is ambiguous between a "controlled" interpretation, and one where anyone can be doing the singing.

(67) I abhor ($_{\alpha}$ PRO singing)

If (67) is a case involving a verbal gerund, then we predict that the controlled reading is possible, where I

Binds PRO across $\underline{\alpha}$, where $\underline{\alpha}$=S. But why is there also the additional arbitrary reading? In order to understand what is going on, let us consider (68).

(68) I abhor (PRO singing operas)

As noted by W&R, (68) is not ambiguous like (67): (68) only has the controlled interpretation, where I Binds PRO. But note that (68) is verbal since Accusative Case is assigned to operas and Genitive Case is not assigned to PRO. So it seems that when the gerund is clearly verbal, Binding always takes place. What this suggests for (67) is that in the additional abritrary interpretation, the gerund is nominal. But if so, why then is PRO not lexicalized since it is assigned Genitive Case? The answer is that PRO simply is not there when the gerund is nominal in (67). Recall that NPs, contrary to Ss, do not have an obligatory subject, for whatever reason this will ultimately depend on. So there could be no subject present in the nominal gerund in (67): hence the arbitrary interpretation. This arbitrary reading is similar to other cases where the subject is not expressed, like impersonal se (il se vend beaucoup de pommes), or NPs with no specified subject (the destruction of Rome, the selling of apples), or impersonal passives (il a été découvert que S, es wird gelacht).

This can help us account for some facts noted by Thompson (1972) about control properties of gerunds. Thompson notes that some verbs taking "activity gerunds" as their complements force a controlled reading, whereas others allow a non-

controlled reading. Examples of both types are given in (69) and (70), respectively.

(69) OBL control:
  (a) Evelyn dreads singing a solo.
  (b) Max can't bear watching the tide come in.

(70) "free" control
  (a) John questioned/recommend going to the movies.
  (b) England vetoed passing such a motion.

Thompson then points out that the verbs which require a controlled interpretation denote private predicates, in the sense that "these verbs involve an individual and his private thoughts, feelings, and personal welfare: (Thompson (1972), p. 381). On the other hand, the free control verbs are "public predicates", that is, an activity is described which is generally shared. It might be that some correlation of this type is possible, but "privateness" of the predicate involved is not what forces control: the reasons are strictly syntactic. For example, consider the examples in (69) again, where the private predicate is supposed to force control. If one changes the internal structure of the gerunds, as in (71), control is no longer obligatory, contrary to Thompson's prediction that the obligatoriness of control is due to properties of the matrix verb, not of the gerund.

(71) (a) Evelyn dreads singing.
  (b) Max can't bear singing.

What seems to be going on is that some verbs take a nominal gerund as complement, others a verbal gerund, and

others allow both as complements. And the possibility of control depends on the structure of the complement: if the gerund is verbal, then Binding of PRO across S takes place; if the gerund is nominal, then PRO is impossible because of Genitive Case assignment and the principle of Lexicalization: so there is no subject, which accounts for the arbitrary reading. Thus the NP without a subject receives a reading somewhat like the NP in (72) where by <u>by</u>-phrase can be omitted.

(72) Constant denial of one's actions (by someone) is saddening.

So gerunds will have an obligatorily controlled PRO subject when they are verbal and in a position where the PRO can be Bound. If the gerund is nominal, it either has a lexical subject or no subject at all: hence it is forced to have arbitrary reading when it is clearly nominal and has no overt subject, as we can see in (73).

(73) *I_i enjoy (PRO_i singing of operas)

Note finally that there are clearly verbal gerunds in that nominal gerunds are not possible in such positions. A first case, suggested by K. Hale (personal communication), is as complement of perception verbs, as in (74).

(74) (a) I saw John coming down the street.
(b) I heard Mary singing in her room.

These gerunds do not allow Genitive Case marking, do not

take adjectives nor the preposition of, nor plural markings, nor the negative particule no. Depending on the analysis of perception verbs in English, the Case of the lexical subject can be assigned by ECM, or the structure could be one of control as in (75).

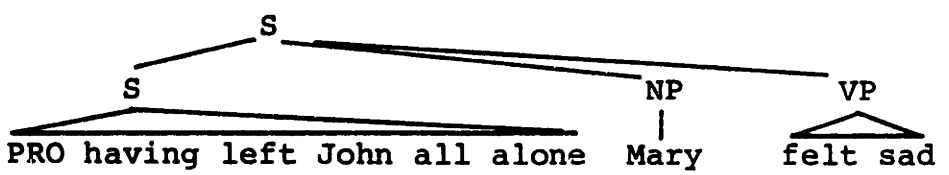
(75) V NP ($_S$ PRO V+ing ...)

A second case where gerunds are obligatorily verbal is when they are adjuncts. As we see in (76), Genitive Case assignment is impossible in such cases.

(76) *John's having left, Mary felt sad.

The other properties of nominal gerunds are also never found in adjunct gerunds. On the other hand, a PRO is always possible in these cases, and it is Bound by the subject, which is predicted in our analysis if the adjunct is attached to S: the subject then Binds the PRO.

(77)



On the other hand, other analyses of control are at a loss to explain why only subject control is possible in such cases, except by stipulating that adjuncts are "subject oriented", which is just restating the problem. Furthermore, it is false: as we will see in (79) below, when the subject of the gerund is a lexical pronoun, it can be object oriented.

The reason why adjuncts must be verbal and cannot be nominal could be due to the fact that nominal -ing must bear

Case, like all nominal elements, and that Case cannot be assigned to adjunct positions.

The subject in adjuncts can be lexical: but in such cases, it does not get Genitive Case, but rather accusative Case as in (78), and more marginally, Nominative Case as in (79).

(78) Him having left, Mary felt sad.
(79) Mary hates John, he being a bachelor.

What we can assume in such cases is that the verbal -ing has Case-assigning features and that it assigns Accusative Case like transitive verbs since it percolates and governs the subject, or it assigns Nominative Case in more marginal constructions when it is analogically associated with INFL.

A last case to consider is when a gerund is in the subject position of a sentence as in (80).

(80) (From Thompson (1972))
Trapping muskrats bothers Mary.
(a) ...She thinks it's not feminine.
(b) ...She is circulating a petition to make it illegal.

If the gerund is nominal, then no PRO is possible in (80) or else it gets Genitive Case and must be lexical: so the gerund gets an arbitrary subject reading. If the gerund is verbal, PRO is possible: it is not Bound by any NP in the sentence so that it can freely corefer. In the case at hand, it can corefer with Mary, or be arbitrary. So verbal

gerunds can be interpreted just like infinitival clauses when they are in subject position: as we will see shortly when we discuss long distance control, the interpretation of the PRO in (81) and (82) would be parallel.

(81) Mary said that (PRO trapping muskrats) is not nice.
(82) Mary said that (PRO to trap muskrats) is not nice.

So this analysis of gerunds as being potentially verbal or nominal accounts for this apparent problem where Case is assigned but there is no government. We have seen that a closer analysis shows that Case is assigned in gerunds in some instances, and that the gerund has the properties of a nominal in such cases, except for the fact that the ambiguous nature of the head of the gerund also allows Accusative Case to be assigned to the object. On the other hand, some gerunds have verbal properties, the head -ing being (-N, +V) in such cases and no Case is assigned to the subject when the gerund is verbal, with consequences for control that follow directly from our general assumptions about Binding of ECs.

5.2.1.2.2. Case agreement with PRO.

Consider now a second instance where a PRO seems to have Case, hence should be lexical according to our assumptions. There are cases in the literature where adjuncts modify a PRO agree in Case with the controller of that PRO, and cases where the adjunct has a special Case of its own, which suggest that PRO itself has a Case with which the adjunct agrees

(cf. Neidle (1982) for Russian, Andrews (1976), Levin (1981), Levin & Simpson (1982) for Icelandic, Simpson (1982) for Icelandic and Warlpiri). We will use Russian as presented in Neidle (1982) as our case-example, and show how the analysis can also account for Icelandic and Warlpiri.

In Russian, 'second predicates', a term used by Comrie (1974) to refer to modifiers which are detached from the noun phrase to which they refer, may either agree in Case with the noun they modify, or take the Instrumental Case.

(83) Ivan vernulsja ugrjumyj/ugrjumym.
Ivan (Nom) returned gloomy (Nom/Instr)

There are two words however which must obligatorily agree with their antecedent: odin 'alone' and sam 'oneself'.

(84) Ivan vernulsja odin/*odnim.
Ivan (Nom) returned alone (Nom/*Instr)

In infinitival clauses, these adjuncts odin and sam must obligatorily agree with a subject controller of the PRO that they modify as in (85).

(85) (a) Vanja xočet prijti odin/*odnomu.
Vanja (nom) wants to come alone (Nom/*Dat)
(b) Ljuda priexala pokupat' maslo sama/*samoj.
Ljuda (Nom) came to buy butter herself (Nom/*Dat)

However, when the controller of PRO is not a subject, odin and sam are invariably Dative, regardless of the Case of the controller.

(86) (a) Ja vedel emu prijti odnomu.
I (Nom) told him (Dat) to come alone (Dat)
(b) My poprosili Ivana pojti odnomu/*odnogo
We (Nom) asked Ivan (Acc) to go alone (Dat/*Acc)
(c) U Koli net sil projti samomu/*samogo
Around Kolja (Gen) (there is) not (the) strength to come alone (Dat/*Gen)
'Kolja doesn't have the strength to come alone'

Furthermore, if the infinitival clause is introduced by a complementizer, then the adjuncts odin and sam are again invariably Dative, even if the controller of the PRO is a subject in this case, as we see by the following minimal pair.

(87) (a) Ljuda priexala pokupat' maslo sama/*samoj.
Ljuda (Nom) came to buy butter herself (Nom/*Dat)
(b) Ljuda priexala, čtoby pokypat' maslo*sama/samoj.
Ljuda (Nom) came in-order-to(COMP) buy butter herself (*Nom/Dat).

Finally, if the controller of PRO is the subject of a passive construction, Dative Case is strongly favored on the adjunct.[14]

(88) On byl ungovorēn prijti *odin/?odnomu.
he (Nom) was persuaded to come alone (*Nom/?Dat)

Neidle (1982) proposes to account for these facts in the following way. Assuming a theory of control along the lines of Bresnan (1982), she suggests that in Russian, Grammatical control is only possible if the controller is a subject. Assuming that grammatical control induces full

agreement, including that of Case, this accounts for the fact that odin and sam obligatorily agree with the subject controller. On the other hand, other controllers are not determined by grammatical control, but by anaphoric control; this is also the case when a complementizer is present in the infinitive clause. So in these cases, there is not full agreement between the controller and PRO: so there is no agreement mediated between the adjunct and the controller as far as Case is concerned. The reason why the adjunct is Dative is that subjects of infinitival clauses seem to be assigned Dative invariably in Russian: thus overt NPs are found in Old Church Slavonic, and also occasionally in Modern Russian in subject position of infinitival clauses with Dative Case. Comrie (1974) gives the following example from Gorky:

(89) A nedavno, pred tem kak vzojti lune, po nebu letala
recently before COMP rise (inf) moon(Dat) about sky was flying
bol'šuščaja černaja ptica.
huge black bird
'Recently, before the moon was to rise, a huge black bird was flying about the sky'

Neidle (1982) says that Dative subjects of infinitives can also be found in expressions like (90) in Modern Russian.

(90) Kak mne skazat?
how I(Dat) say(inf)
'How could I say?'

So the simplest claim to make is that PRO has Dative Case in infinitives in Russian so that adjuncts agree with that PRO in Case. Note that in the Lexical-Functional Grammar framework assumed by Neidle, there is no PRO in the constituent structure of grammatical control construction, hence no Dative Case assigned, which would explain why odin and sam agree with the controller in this case.

As for the fact that passive subjects do not trigger agreement in Case when they are controllers, Neidle suggests that this is because the control equation needed in such a case would involve an object controller, and that object grammatical controllers are not possible in Russian.

Suppose that we were to adapt Neidle's solution to our framework. We could say that subject control involves Binding in Russian, but that some structural constraint prevents objects from being Binders in Russian: for example, we could say that $\bar{S}$ deletion only takes place if there is adjacency between the V and the $\bar{S}$ infinitive in D-structure, so that $\bar{S}$ deletion is not possible in (91), since there is an NP between the V and $\bar{S}$.

(91) My poprosili Ivana ($=_{\bar{S}}$ PRO pojti odnomu)
we (Nom) asked Ivan(Acc) to go alone (Dat)

So the non-control by an object would be parallel to non-control when COMP is filled as in (87b): there can be no $\bar{S}$ deletion, hence no Binding across a maximal expansion since government is blocked. The same explanation would

hold for passives since there is an NP object which also prevents $\bar{S}$ deletion under adjacency, hence Binding, as in (88).[15]

So we could assume that Binding triggers full agreement in Russian, including Case, whereas coreference triggers only agreement in F-features, which is the usual case. What about the Dative Case on PRO? Notice that our theory does not predict that an EC cannot have Case: it predicts that an EC cannot have Case in PF. We have seen in our analysis of Pro Drop in chapter 4 that it is possible for an EC to have Case, namely at LF, although it does not have Case at PF. Now note that Russian has structurally predictable Case, and some instances of structurally unpredictable Case, so-called quirky Case. If we assume that quirky Case is Case that is checked only, not assigned by the element which assigns the θ-role, this means that Case not being assigned, it will not be visible at PF unless some other element assigns it. For example, consider quirky Case in Icelandic.

(92) Hana vanta δi peninga
she (Acc) lacks money (Acc)

If we assume along with Marantz (1981), Burzio (1981), Levin (1981) and Levin & Simpson (1982), that quirky Case marked subjects are in fact D-structure objects, then we can assume that INFL in Icelandic simply assigns Case, any Case. The unmarked Case in Icelandic is Nominative: for example, Simpson (1982) points out that "a default assignment of

NOMINATIVE case and neuter gender will probably be needed for participles and predicate adjectives in quirky-case subject sentences and passives of quirky-case objects. In these types of sentence, passive particples and matrix predicate adjectives agree not with the quirky case subject (if there is one), but rather have NOMINATIVE case, neuter gender and singular number" (Simpson (1982), p. 9-10)

(93) (a) Steini var kastað
Stone(Dat) (masc) was thrown(Nom meut sing)
(b) Honum var kalt
(He(Dat) was cold(Nom neut sing)

So Nominative is the unmarked Case. But if the verb requires another Case to be checked to get a θ-role, then INFL will assign whatever other Case is befitting. But the V does not assign this Case: it only checks it. So languages that allow quirky Case have the property that such Cases are not assigned by the verb that assigns the θ-role, but that they are rather checked at LF.

If we return to Russian, we note that this language has this property of having the possibility of only checking a Case at LF since it has quirky Case constructions. We can assume that INFL in infinitives used to have the property that is could assign Dative Case to the subject, as is evidenced by the fact that this is found in Old Church Slavonic, archaic literary Russian and some frozen expressions of Modern Russian. If we further assume a passage from

Case assigner to Case checker for infinitival INFL in Russian, we get the result that PRO might have Dative Case, but only at LF. So the crucial fact for our theory is not that we have an EC PRO which triggers Dative Case agreement, but rather at what level that Case is present: and we see that it is possible for this Case to be present only at LF, and that it is possible for it to trigger agreement with an adjunct because Agreement takes place only at LF. But it is still possible for that NP to be an EC if it has no Case at PF.

Consider now Icelandic. In Icelandic, in constructions like those just presented for Russian, adjuncts either agree with the controller or take Nominative Case. We have seen in (93) that Nominative is the unmarked Case in Icelandic, so that it can be assigned to adjuncts by default presumably in some cases. When agreement does take place, it seems to be similar to Russian: Binding induces Case agreement, as we see in (94).

(94) (a) Eg baδ drenginn aδ vera goδur/goδan
I(Nom) asked the-boy(acc) to-be good(Nom-Acc)

(b) Eg skipa i drengnum aδ vera goδur/goδum
I(Nom) ordered the-boy(Dat) to-be good(Nom/Dat)

However, quirky Case subject verbs in infinitival clauses trigger agreement with this quirky Case, not with the Case of the controller.

(95) Eg vonast til að vanta ekki einn/einan efnii ritger ðina
I(Nom) hope till to-lack not alone(Nom/Acc) materian(Acc) for the thesis
'I hope not to lack alone material for the thesis'
(vanta is an Accusative quirky Case verb)

If we assume the analysis of quirky Case given above, then the quirky Case is checked in the object position of the verb, and the NP moves in subject position of the sentence to get a Case assigned by INFL. But this holds only in tensed sentences. Suppose that INFL cannot assign such a Case in infinitival clauses, then nothing can assign a Case to the NP: so it cannot be lexical. The NP, we can assume, can still get a θ-role from the quirky Case verb if we consider the quirky Case specification to be some kind of filtering of non-properly Case-marked NPs: since the NP has no Case at all, it is not prevented from being assigned the θ-role; only if it has a non-matching Case is it filtered out. So the NP can be an EC since it is not assigned Case at PF. And this NP cannot get Case from its Binder even by Agreement at LF since this would not pass the quirky Case filtering. So the NP has only one Case specification in its entry: that of the filtering quirky Case in the V-grid where the NP enters its index to get a θ-role, and it is this Case specification which is responsible for the Agreement of the adjunct.

A slightly different case is presented for Warlpiri in Simpson (1982). In Warlpiri, subjects can be assigned either

Absolutive or Ergative Case. The Ergative is marked and according to Hale (1978), it is semantically predictable: an agent or a perceiver is marked Ergative.

There are non-finite clauses in Warlpiri, and adjuncts can occur in such clauses modifying the PRO subject. "If the verb in a non-finite clause demands an ERGATIVE case-marked Subject, an adjunct modifying the subject gets ERGATIVE case:

(96) Warnapari-ki + Ipa-rna-rla wurrukangu kuyu
yarnunjuku-rlu dog(Dat) + PASTIMP-1sg-3Dat
Sneak-PAST meat-Abs hungry(Erg) ngarni-nja-
kurra-ku
eat-INF-COMP-Dat
'I sneaked up on the dog hungrily eating meat'

In (96), dog has Dative case in the matrix. It is the controller of the kurra clause. The verb in the kurra clause is transitive and takes an ERGATIVE Subject normally. An adjunct modifying the understood Subject of the kurra clause, 'hungry', gets ERGATIVE case." (Simpson (1982), p. 14)

We can assume that in Warlpiri, agent and perceiver are specifically marked for Ergative, somewhat like a quirky Case but attached to these semantic roles generally, instead of being lexically specified for some semantic role assigned by a verb as in the usual instances of quirky Case. Then the analysis is similar to the one for Russian and Icelandic: Warlpiri is a nonconfigurational language, so that, as we

have seen in Chapter 4, Case is not assigned in this language, but assumed at some level. So the PRO has no Case at PF, but there is this specification attached to the roles of agent and perceiver which is checked at LF: and this triggers Ergative Case Agreement of the adjunct at LF.

So we see that an objection to our deriving the distribution of PRO from Case rather than from government cannot come from unanalyzed data: one must first see how the Case is encoded in these structures, and at what level. We have seen that "PRO with Case" shows up in languages where special Case properties are present: either quirky Case, or thematically determined Case, and that these requirements are specified at a rather abstract level in the grammar, i.e. at LF. So our analysis of the distribution of PRO from Case still holds since it is Case at PF which is relevant for the Principle of Lexicalization, not Case at LF, as we already saw in the discussion of Pro Drop languages. And we have seen in the discussion of gerunds that the problems encountered by a derivation of the distribution of PRO by government with respect to different notions of government and c-command, and of strict $\bar{X}$ theory violations, in fact favor a derivation by Case theory instead.

5.2.2. Other aspects of local control.

Let us consider a few more cases involving local control, i.e. Binding of PRO. In some instances, there are two possible Binders, since both the subject and the object

govern the embedded PRO, as with verbs like promise and ask.

(97) John promised Mary ($_S$ PRO to shave himself)
(98) John asked Mary ($_S$ PRO to shave herself)

However, these sentences are not ambiguous: only John is the controller in (97), and only Mary in (98).[16] So in one case it is the subject that Binds, and in the other it is the object. However, this is not due to a structural difference between the two sentences but it is simply a matter of semantics or pragmatics, as noted by Manzini (1980): a change in the infinitival clause can reverse the choice of Binders, as we see in (99-100).

(99) John promised Mary ($_S$ PRO to be allowed to shave herself)
(100) John asked Mary ($_S$ PRO to be allowed to shave himself)

The same judgements hold when passivization of the matrix verb has taken place.

(101) (a) *Bill was promised t ($_S$ PRO to shave himself)
(b) Bill was asked t ($_S$ PRO to shave himself)
(c) Bill was promised t ($_S$ PRO to be allowed to shave himself)
(d) *Bill was asked t ($_S$ PRO to be allowed to shave himself)

This discussion brings us to another instance where passivization is involved. Nothing that we have said so far presents a derivation of (102) from the D-structure in (103).

(102) *John was tried ($_\alpha$ e to leave)
(103) __ was tried ($_\alpha$ John to leave)

But note that a sentence like (104) is also ungrammatical.

(104) *It was tried ($_{\alpha}$ e to leave)

In a GB account, the ungrammaticality of (102) could be due to ECP if $\alpha = \bar{S}$: the trace would not be properly governed, and if the e is PRO, then John does not have a θ-role. But the ungrammaticality of (104) cannot be attributed to constraints on the EC: if $\alpha = \bar{S}$ in (104), then the EC can be PRO_{arb}, and there is no violation of the θ-criterion since it can be expletive. Chomsky (1981a) proposes to account for the ungrammaticality of (104) by saying that passive morphology cannot apply to try since it does not assign Case when it has an infinivital complement. But we have seen in our discussion of try with respect to ECM in (40) above that to say that try is an "intransitive" verb when it has an infinitival complement misses the generalization that, in a GB analysis, all verbs that do now allow $\bar{S}$ deletion would also be "intransitive" verbs.

In our analysis, (104) is ungrammatical because expletive it Binds a θ-position, since $\alpha = 3$. But in (102), the ungrammaticality can only be due to Case theory: John can Bind the e as a trace, and get a θ-role. So we will also say that passive morphology cannot be applied when there is no assigned Case to be absorbed: since try is not an ECM verb, no Case can be assigned in (102), hence no Case can be absorbed, and passive morphology is not possible. Note that for a verb like believe, which is an ECM verb, a Case is assigned in the same construction, hence passive morphology

is possible, and we get a sentence like (105).

(105) John is believed ($_S$ t to be happy)

As pointed out by Chomsky (1981a), this presents a problem for the analysis of ECM of Kayne (1981b) since (105) contrasts with the ungrammatical sentences in (106) and (107).

(106) *Bill was preferred ($_{\bar{\alpha}}$ for t to leave)
(107) *Bill was wanted ($_{\bar{\alpha}}$ t to leave)

Chomsky suggests to account for (106-107) by assuming that for is not a proper governor, "perhaps a special case of the principle that the only proper governors are the lexical categories" (Chomsky (1981a), p. 298). In (107), wanted would be subcategorized for a for complementizer but with the special constraint that for is deleted after this verb. So for, although it has prepositional properties, would be a complementizer, hence not a lexical category, and hence not a proper governor. So the sentences in (106-107) are out by the ECP. In the case of believe in (105) however, Chomsky says that Kayne would be forced to claim that although for is not a proper governor, the Ø preposition in the COMP of the $\bar{S}$ complement of believe is a proper governor. So Chomsky concludes that this would virtually amount to accepting the $\bar{S}$ deletion hypothesis for ECM, although he remarks that "Kayne's approach to unifying preposition stranding and Exceptional Case Marking is, however, sufficiently attractive so that an attempt to solve the remaining difficulties surely seems in order"

(Chomsky (1981a), p. 298).

We will not enter into a discussion of the problems raised in such an analysis that claims that some elements like for are not lexical for purposes of proper government since we have seen that lexical government is an accidental property of recoverability of properties of a trace by Binding (cf. Chapter 3). On the other hand, we have presented an analysis of ECM that maintains the unified account of preposition stranding and ECM presented in Kayne (1981b), and it would be interesting if this analysis could be compatible with an account of the facts in (105-107). The account for (105) is straightforward: since $\bar{S}$ deletion takes place in the context (+V__), John can Bind the trace. Furthermore, since believe is an ECM verb, a Case can be assigned by this verb, hence passive morphology can apply to absorb it, assuming the constraint that passive morphology applies only if Case can be absorbed in English, which seems quite natural.

In (106), we can say that prefer does not assign Case to an infinitival complement, hence that it cannot take passive morphology. But even if prefer did assign Case and passive morphology was possible, the sentence would still be ungrammatical since $\alpha = \bar{S}$ since COMP is filled by for, hence $\bar{S}$ deletion cannot take place: so Bill cannot Bind the trace across $\bar{S}$, and the sentence is out by the θ-criterion since Bill has no θ-role.

This leaves us with (107). Why doesn't want allow

passive morphology? It seems to be an ECM verb since it can be followed by a lexical NP in the subject position of the embedded infinitival clause.

(108) John wants ($_{\alpha}$ Bill to leave)

So offhand, we would expect want to behave like believe. But note that want is different from believe in that it has the additional property to allow a PRO in the same position where it allows a lexical NP.

(109) John wants ($_{S}$ PRO to leave)

The solution is to adopt Kayne's use of a Ø preposition, but to restrict it to want. So the structure of (108) is in fact (110).

(110) John wants ($_{\bar{S}}$ Ø ($_{S}$Bill to leave))

Want is not an ECM verb, since it does not assign Case to an NP that it is not subcategorized for in (109) where a PRO is possible. But want can assign Case to a Ø preposition in COMP in (110), assuming that this Ø preposition, being in COMP, is the head of $\bar{S}$ in some sense. So the Case properties are not separated from the θ-properties for want, hence want is not an ECM verb, but rather something like the "weak ECM" verb croire in French (cf. the discussion of (37) above). The Ø preposition, once Case marked, then in turn assigns Case to Bill. Now since passive morphology applies only when the verb assigns Case, it will apply to want only when want assigns Case to a Ø preposition, hence

only when such a preposition will be present in the structure. This means that the structure for (107) must be as in (111).

(111) *Bill was wanted ($\bar{S}$ Ø ($_S$ t to leave))

But then Bill cannot Bind the trace in (111) since $\bar{S}$ deletion cannot take place because COMP is filled by the Ø preposition. So Bill does not get a θ-role, hence the ungrammaticality of the sentence. So the contrast in passivization between want and believe is due to the presence of this Ø preposition in the COMP of the infinitival complement of want when want assigns Case.

As for (109), the Ø preposition has not been inserted. Since want is not an ECM verb, Case is not assigned to the NP, which then can be an EC by the Principle of Lexicalization. Since $\bar{S}$ deletion takes place, COMP being empty, John Binds the PRO, and "local control" is established obligatorily, hence the ungrammaticality of the sentences in (112).

(112) (a) *John wanted ($_S$ PRO to shave oneself)
(b) *Mary knew that John wanted ($_S$ PRO to behave herself)

Thus our general principles that account for ECs can explain the contrast in passivization in (105-107) without any modification if one accepts that want is not an ECM verb but rather that it assigns Case via a Ø preposition in COMP. And this analysis is compatible with our unified account of preposition stranding and ECM. Recall that this account is also compatible with out account of contraction, discussed

in 3.3.2.4. Recall that we assumed there that contraction of want+to to wanna is not dependent on the Case-marked status of the trace since there are instances where this is not relevant and where only structural conditions seem to be at play, as in (113).

(113) (a) They want, to be sure, a place under the sun. (*wanna)
(b) They (wish and want) to go (*wanna)

In our analysis of want constructions, there is a structural difference between the complement with PRO subject and the complement with Case-marked subject, as we see in (114).

(114) (a) I want ($_S$ PRO to go)
(b) I want ($_{\bar{S}}$ Ø ($_S$ Bill to go))

When WH-extraction takes place from the subject position, Case must be assigned to the Wh-phrase by the Principle of Lexicalization: so the structure must be as in (114b). On the other hand, in order for a PRO to be possible in subject position, it must not have Case, so the Ø preposition must not be present, and then $\bar{S}$ deletion takes place since nothing blocks it, so that PRO is Bound by the subject of want; but then the structure is as in (114a). Contraction, we assume can take place only if want governs to: so it can take place in (114a), but not in (114b), where the $\bar{S}$ blocks government.

We saw in our discussion of promise and ask above that there are constructions with two potential Binders and that it is semantic or pragmatic factors that determine the choice of Binder. Another case of this type in our analysis is a

sentence like (115) (this sentence is discussed in Kayne (1981b)).

(115) Il me semble ($_S$ e avoir compris)

Since $\bar{S}$ deletion has taken place, the EC in (115) has two possible Binders: either il or me. If il is referential, then il must Bind the EC in order to get a θ-role since no θ-role is assigned to the subject of sembler. So the sentence gets the interpretation in (116).

(116) He seems to me (t to have understood)

On the other hand, il can be a nonreferential expletive element: in such a case, il cannot Bind the EC since this would result in a θ-criterion violation since expletive it cannot bear a θ-role, unless the embedded subject position is not a θ-position, as in Il me semble y avoir beaucoup de monde. But in (115), it is possible for me to Bind the EC. However, we agree with Ruwet (1976) and Kayne (1981b) that the interpretation in this case is not one where me is raised from the subject position of the infinitival clause as in (117a), a proposal made by Rouveret & Vergnaud (1980), but rather one where me "controls" the EC. Our reason to believe this is that me is related to two θ-roles, as in (117b). So the interpretation of (115) is as in (117b).

(117) (a) ≠ It seems that I have understood.
(b) It seems to me that I have understood.

So me gets a θ-role as Dative object of semble, and the

EC gets an independent θ-role from avoir compris. So it seems that me controls the EC in (115) if il is expletive. In fact, as predicted by our analysis, me must Bind the EC and (115) cannot get a reading where the EC is PRO_{arb} as in (118) since $\bar{S}$ deletion has taken place.

(118) *Il me semble ($_S$ PRO_{arb} avoir compris)

The fact that control rather than raising is involved is further confirmed by the fact that non-clitic Dative complements can control the EC.

(119) Il semble à $Jean_i$ (e_i y être allé déjà) (From Gross (1968))[17]

So in our analysis, (115) presents no particular problem. However, this presents a problem for a GB analysis. In order for the EC in (115) to be a PRO, it must not be governed in the GB analysis. But on the other hand, in order for the EC to be a trace, it must be properly governed. So sembler must trigger $\bar{S}$ deletion when there is raising, but not when there is control. So when a Dative object is present, sembler is either a raising or a control verb. On the other hand, if no Dative complement is present, sembler must be a raising verb, i.e. it must trigger $\bar{S}$ deletion, since an interpretation like (120) is not possible.

(120) *Il_{expl} semble ($_\alpha$ PRO_{arb} être parti)

In our analysis, no such problem arises as we have seen. As for (120), since $\bar{S}$ deletion takes place in this context,

expletive il would Bind the EC, and this results in a θ-violation. So an expletive il is possible in the subject position of sembler when there is an infinitival complement with a θ-marked subject, only if some other Binder is available to Bind the EC to present the il from Binding the θ-marked EC. So all the properties of sembler follow from our general principles, and there is no need to complicate the lexical entry of sembler.

5.3. PRO as a pronoun.

As NP can be an EC and a pronoun if it is in a position where no Case is assigned, and if no NP Binds it, so that it is not an anaphor but a freely indexed pronoun. Such a pronoun will not exhibit the four basic properties of anaphors which are derived from Binding by the antecedent. So a pronoun will not obligatorily have an antecedent; if it has one, it need not be local; it need not be unique; and it does not have to have a specific structural relation with the pronoun (except for the fact that the pronoun may not c-command its antecedent, this being a totally independent constraint on pronouns; cf. the Binding Condition in (2.137)).

In the case of infinitives which we are discussing here, the NP subject can always be an EC since no Case is assigned to that position, except for the special cases where the infinitival clause is the complement of a matrix verb which is an ECM verb, or when a preposition is inserted in COMP in a language that allows ECM. In all the other cases, the

NP subject can be an EC. In order for that EC to be a pronoun however, it must be the case that the EC is not Bound, or else it is an anaphor, not a pronoun, by functional definition. The EC will not be Bound if $\bar{S}$ deletion does not take place, since the $\bar{S}$ then blocks government by a Binder. There are two factors that can prevent $\bar{S}$ deletion from applying: either the $\bar{S}$ is not a context where $\bar{S}$ deletion takes place, or the COMP of this $\bar{S}$ is filled, so that $\bar{S}$ deletion is not possible. Consider the latter. The obvious candidate is when COMP is filled by a WH-phrase.

(121) (a) They don't know ($_{\bar{S}}$ what ($_{S}$ PRO to do $\underline{t}$))
(b) They don't know ($_{\bar{S}}$ how ($_{S}$ PRO to behave themselves/ oneself)

Here PRO is pronominal since there is no Binder for it (the WH-phrase Binding another EC). Since it gets no R-index by a Binder, PRO is freely indexed at S-structure. PRO can pick the index of an NP in the sentence, so that it can be coreferential with <u>they</u> in these sentences. But PRO can also pick some other index, and have no coreferent in the sentence as in the reading with <u>oneself</u> in (121b). There might be various prigmatic factors that will determine the choice of coreferent.

The COMP might be filled, and hence block $\bar{S}$ deletion, because a relative clause is formed, as in (122).

(122) John gave (Mary) ($_{NP}$ a book$_i$ ($_{\bar{S}}$ $\underline{e}_i$ ($_{S}$ PRO$_k$ to read $\underline{t}_i$)))

Here PRO$_k$ is pronominal since it has no accessible Binder.

For pragmatic reasons which have to do with the meaning of give, if an object like Mary is present, this object will be the preferred antecedent for PRO_k. But Mary does not not Bind PRO_k: it is simply coreferent with it. This is clearer in a sentence like (123).

(123) John found another book ($_{\bar{S}}$ e_i ($_S$ PRO_k to read t_i))

The preferred reading is where PRO_k is interpreted as coreferent with John, but it need not be so if a context is given where John has the habit of suggesting books to read to his friends or students.

Consider now the second case where PRO is pronominal, i.e. where $\bar{S}$ deletion does not take place because the $\bar{S}$ is not in the proper context. We will look at four different constructions where this happens.

1° Nonbridge verbs.

(124) John shouted/whispered ($_{\bar{S}}$ ($_S$ PRO to leave))

We have already discussed this case at the beginning of 5.2. We assume that $\bar{S}$ deletion does not take place after a nonbridge verb since the $\bar{S}$ is not "identified" by the nonbridge verb since its index is not entered in the V-grid. So the PRO is pronominal since it cannot be Bound, and it freely refers like a pronoun. In (124), PRO cannot corefer with John under normal circumstances, but it can corefer with some other NP in the sentence if one is present, as in (125).

(125) (a) John whispered to Bill ($_{\bar{S}}$ PRO to leave)
(b) They said that John shouted ($_{\bar{S}}$ to leave)

In (125a), PRO can be Bill or Bill and some x. In (125b), PRO can be they, or they and some x, or some x.

2° Sentential subjects.

(126) ($_{\bar{S}}$ ($_{S}$ PRO to finish the work on time)) is important for me.

Here, $\bar{S}$ deletion does not take place since the $\bar{S}$ is not in a context (+V__). Even if it was in fact, there would still be no Binder for PRO, so PRO is pronominal. It can corefer with me, but it need not, as we see from the contrast in (127).

(127) (a) To finish my work on time is important for me.
(b) To finish one's work on time is important for me.

A similar account for the referential possibilities of PRO holds if the sentential subject is in an embedded sentence, as in (128).

(128) (a) John thinks that ($_{\bar{S}}$ ($_{S}$ PRO to feed himself/oneself)) will be difficult.
(b) John thinks that ($_{\bar{S}}$ ($_{S}$ PRO to leave early)) will be difficult.

In (128b), PRO can be John, John and some x, or some x. Note that all cases of "long distance" control given in the literature involve pronominal PRO, where $\bar{S}$ deletion is not possible.

3° Purposives.

(129) John bought a book ($_{\bar{S}}$ ($_{S}$ PRO to please his teacher))

In (129), John seems to Bind PRO, but it is not the case that a subject Binds the PRO of a purposive, since sentences like (130) are also possible.

(130) The book was sold ($_{\bar{S}}$ ($_{S}$ PRO to help the refugees))

As noted by Manzini (1980), in sentences like (130), the PRO is arbitrary, but it is still perceived as being related to the agent of sold, although this agent is not expressed. So this is a matter of pragmatics, not of Binding. In fact, it seems that the PRO can be related to any NP which can be a "potential agent", not only to one that bears that thematic role in the sentence. This would explain the ambiguity of (131), where either John (the agent of hired), or Mary (the patient of hired), or even John and Mary can be doing the firing.[18]

(131) John hired Mary ($_{\bar{S}}$ PRO to fire Bill)

The reason why a PRO in a purposive is not Bound is because $\bar{S}$ deletion does not take place since, just like in the case of nonbridge verbs, the $\bar{S}$ is not "identified" by any verb since it is not governed by any verb and its index is not entered in any V-grid. So the PRO is pronominal in such constructions.

4° Extraposed infinitival clauses.

Consider the contrast between (132) and (133).

(132) (a) It is important ($_{\alpha}$ PRO to leave early)
(b) It is impossible ($_{\alpha}$ PRO to leave early)

(133) (a) *It is certain ($_{\alpha}$ PRO to leave early)
(b) *It is sure ($_{\alpha}$ PRO to leave early)

Our theory tells us that the reason why the sentences in (133) are ungrammatical is due to the fact that expletive it Binds a θ-position here, so that $\bar{S}$ deletion has taken place and α = S. If this is right, then we expect raising to be possible in these constructions, since the subject of the matrix clause can Bind the subject of the embedded clause across S.

(134) (a) John is certain ($_{S}$ t to leave early)
(b) John is sure ($_{S}$ t to leave early)

On the other hand, since PRO gets an arbitrary interpretation in (132), it is a pronominal, and so it is not Bound by it: so $\bar{S}$ deletion is not possible here, and α = $\bar{S}$. We then expect raising not to be possible in such constructions, since the raised NP cannot Bind its trace across $\bar{S}$, and hence has no θ-role. This is indeed the case.

(135) (a) *John is important ($_{\bar{S}}$ t to leave early)
(b) *John is impossible ($_{\bar{S}}$ t to leave early)

So the contrast between (132) and (133) is the consequence of a difference in the possibility for $\bar{S}$ deletion. The reason why there might be such a difference could be due to the fact that these two types of constructions have different D-structures. The simplest assumption to make about (133) where $\bar{S}$ deletion

takes place is that the D-structure is (136).

(136) NP is $(_{AP}\ \bar{A}\ \bar{S})$

So the $\bar{S}$ is in a context (+V__) and $\bar{S}$ deletion takes place. But (132) on the other hand would involve a derivation by extraposition from a D-structure like (137).

(137) $\bar{S}$ is AP

So this reduces to the case of sentential subjects: $\bar{S}$ deletion cannot take place, even after extraposition since the $\bar{S}$ is not "identified" by the ADJ, so PRO is pronominal.[19] If extraposition does not take place, then we get the S-structure in (138).

(138) (a) $(_{\bar{S}}$ PRO to leave early) is important.
(b) $(_{\bar{S}}$ PRO to leave early) is impossible.

These are the four major contexts in which $\bar{S}$ deletion is not possible. In these contexts, PRO is a pronominal, so that it exhibits the converse of the four basic properties of anaphors derived from Binding. Thus the PRO does not have to have an antecedent, as in (139).

(139) (a) It is important $(_{\bar{S}}$ PRO to leave early)
(b) John knows $(_{\bar{S}}$ how $(_{S}$ PRO to behave oneself))

There can be more than one antecedent as in (140).

(140) John told Mary that $(_{\bar{S}}$ PRO to leave early)) is important.
(PRO = John, Mary, John and Mary, John and x, Mary and x, John and Mary and x, x)

We also see in (140) that the antecedent need not be

local. In fact, the strictly pronominal nature of PRO in these constructions can be seen from the fact that PRO can have an antecedent even across sentences in a text, just like lexical pronouns, as noted in Mohanan (1981, Bresnan (1982).

(141) Tom felt embarrassed.
A - (PRO pinching elephants) was a mistake.
B - It was shameful, (PRO exhibiting himself in public) like that.

Finally, there is no structural constraint of the type found for anaphors, as we see in (142).

(142) ($\bar{S}$ PRO to finish his work on time) is important for a child's development.

Thus, while on the one hand only an NP that can govern the PRO can be a "local controller", therefore only a subject or an object, on the other hand any NP can be a "long distance controller" since only coreference is involved in this case, but Binding, and no special structural relation is taking place.

A point which we have delayed to address is why pronominal PRO has the properties of what is known as arbitrary PRO when PRO is a free pronominal, i.e. non-coreferential. What are the properties of this free pronominal EC? Since it has an R-index of its own, that means that at LF, Agreement applies vacuously for free pronominal PRO since there is no element that has F-features for PRO to agree with. So since this PRO has no F-features inserted in D-structure and S-structure,

or else it would be lexical by the Principle of Lexicalization, and since this PRO does not get any F-features by Agreement at LF either, this has for consequence that free PRO has no F-features at LF. If, as we have seen in 2.2.2., an element must have the F-features of some specific object in domain D in order to denote that object, then free PRO cannot denote such a specific object: but then we have an explanation for why $PRO_{\underline{arb}}$ is interpreted as a variable-like element, although it is not technically a variable, i.e. it is not $\bar{A}$-Bound: having no F-features, free PRO can range over all the individuals in domain D that can satisfy the predicate of which free PRO is subject, regardless of the grammatical features associated with these individuals.

There is one constraint on free PRO however: it seems to get a preferred reading with the selectional feature (+animate). Thus (146) is very awkward if one has in mind rocks or trees, but fine if one has in mind humans of animals.[20]

(146) (a) ($\bar{S}$ PRO rolling down the hill) would be dangerous.
(b) ($\bar{S}$ PRO rolling up that hill) is impossible.

A possible explanation for this restriction can be found in Marantz (1981). Marantz makes a distinction between "nominative-accusative languages" like English, French, and "ergative languages" like Dyirbal and Central Artic Eskimo. "In addition to appearing in some P-A structures, semantic-role classes seem necessary to express generalizations about the organization of P-A structures (Predicate Argument structures) within a language. In English and many other

languages, it is generally true that if one of the inherent roles associated with a verb is an "agent" role, i.e., is the role of an active, animate being who intentionally causes something, then this role will be assigned to the logical subject of the predicate that the verb produces. It is also generally true in these languages that "theme" inherent roles -- roles of objects that the verb specifies to undergo a change of state -- and "patient" inherent roles -- roles of objects that bear the brunt of the action described by the verb -- are assigned by verbs, i.e., are borne by logical objects. These generalizations must be stated within the grammar of a language. A straightforward statement is given in (147)." (Marantz (1981), p. 54-55)

(147) (a) agent roles - logical subject
(b) theme/patient roles - logical object

Marantz tested this hypothesis with children 3-4 years old (cf. Marantz (1980)). He presented the children with made up verbs like moak, meaning "to pound with the elbow", the predicate argument structure violating the generalization in (147) as in (148).

(148) moak: 'pound with the elbow'" object = agent
subject = patient

So sentences like (149) containing moak violate the generalization for English-speaking children that agents come preverbally, patients post-verbally.

(149) The book is moaking Larry.

Marantz's experiments show that English-speaking children have more difficulty learning to use verbs like moak which do not conform to the generalization in (147). For example, when shown Cindy pounding a ball with her elbow, the children tend to say "Cindy is moaking the ball".

So it could the fact that the unmarked case for a subject in English is to be an agent, hence an active, animate being who intentionally causes something, that induces the preferred reading for non-specified subjects like free PRO to be (+Animate).

5.4. On differences between trace and PRO.

In Chomsky (1981a), Chomsky assumes the following differences between trace and PRO.

(150) (i) trace is governed, PRO is ungoverned.
(ii) the antecedent of trace is in a $\bar{\theta}$-position, the antecedent of PRO is in a θ-position.
(iii) trace obeys Subjacency, PRO does not.

The point in (i) is a question of choice of analysis: in our analysis, PRO is governed when it is proximate, and its distribution depends on Case, not government: arguments to this effect have been given in the preceding sections of this chapter. As for (ii), we agree with Chomsky. Consider now (iii). Since proximate PRO is Bound by its antecedent, just like a trace, Subjacency holds for anaphoric PRO in our analysis. Thus in (151).

(151) *Ronald seems ($\bar{S}_1$ that it is certain ($\bar{S}_2$ $\underline{t}$ to be back))

$\bar{S}_2$ is deleted, and $\underline{it}$ Binds the trace, so the sentence is out by the θ-criterion because expletive $\underline{it}$ has a θ-role, and referential $\underline{Ronald}$ does not have a θ-role. Similarly, in (152),

(152) *John tried ($\bar{S}_1$ that it is sure ($\bar{S}_2$ $\underline{e}$ to leave early))

$\bar{S}_2$ is deleted, and $\underline{it}$ Binds the EC, which is a θ-criterion violation. On the other hand, all the examples given in the literature where PRO violates Subjacency involve pronominal PRO, not anaphoric PRO.

(153) (a) John thinks (that ($\bar{S}$ PRO to feed himself) will be difficult)

(b) *John seems (that ($\bar{S}$ $\underline{t}$ to feed himself) will be difficult)

In (153a), the relation between PRO and $\underline{John}$ violates Subjacency, as we can see from the ungrammaticality of (153b). Indeed it does, but because this is a case of pronominal PRO, and pronominal PRO can corefer with any NP. We can see that PRO is pronominal in (153a) since it can be free PRO, as in (154).[21]

(154) John thinks that to feed $\underline{oneself}$ will be difficult.

Another difference between trace and PRO is illustrated in the following sentences (from Burzio (1981), discussed also in Chomsky (1981a)).

(155) One interpreter each seem (S $\underline{t}$ to have been assigned to the visiting diplomats)

(156) *One interpreter each tried ($_S$ PRO to have been assigned to the visiting diplomats)

The difference according to Chomsky is that "the D-structure position of one interpreter each must be "close enough" to the phrase the visiting diplomats for each to be interpreted appropriately as a quantifier related to the latter phrase." (Chomsky (1981a), p. 62)

We agree with this explanation, but the question is: what is "close enough"? Suppose that it is the following: the NP each must be part of the same Predicate Structure as the phrase to which each is related. For example, in (155), one interpreter each is an argument of the predicate formed by assign, and the visiting diplomats also is an argument of the predicate formed by assign. But in (156), one interpreter each is an argument of the predicate formed by tried, whereas the visiting diplomats is an argument fo the predicate formed by assign. So we have "closeness" in (155), but not in (156). The difference between trace and PRO that comes out of the contrast between (155) and (156) is that the antecedent of PRO is in a θ-position, whereas the antecedent of trace is not in a θ-position. And this is precisely the difference that we claim that there is between trace and PRO.

5.5. Concluding remarks.

Any analysis which derives control from the binding theory and where PRO always has anaphoric properties, as when it is

a pronominal anaphor, will always be faced with the problem of multiple antecedents for PRO and antecedents across sentences in a text, these properties never being found for "true anaphors". And any analysis that has an actual theory of control will be weakened by the fact that this theory of control will always restate the properties of coreference for pronouns when multiple antecedents are present in a construction, i.e. when PRO is pronominal in our analysis. As for the fact that the behavior of PRO is sometimes identical to that of an anaphor, and sometimes identical to that of a pronoun, this being determined by the precise context in which PRO appears, such theories will either fail to capture this fact, or they will require principles that essentially reproduce effects that are independently accounted for by principles accounting for properties of anaphors and pronouns in general.

In our analysis, there is no theory of control: PRO is either an anaphor or a pronoun, this being determined functionally as we have just seen, and the properties of PRO as an anaphor or as a pronoun are determined by the same general principles which are independently motivated for all instances of anaphors or pronouns. And this analysis is consistent with our general methodological principle which states that no statement in the grammar should refer specifically only to ECs or only to lexical NPs.

Recall that, besides these conceptual reasons to adopt an analysis with anaphoric or pronominal PRO, there

are also empirical reasons to distinguish between the two types of PRO which have to do with the subcase of the Elsewhere Principle in (3.227) discussed in 3.3.2.2 and repeated here as (157).

(157) Don't put a pronoun in a position where an anaphor is possible, i.e. in a position where the pronoun will be interpreted as coreferential with an NP that can Bind it.

We saw that "like-pronouns" are not possible in subjunctives when a corresponding infinitival clause is possible with an anaphoric PRO, as in (158).

(158) (a) *Je veux que j'aille voir ce film.
(b) Je veux ($_S$ PRO aller voir ce film)

But if PRO is a pronominal, then a "like-pronoun" in an equivalent subjunctive construction is possible since the distinction made by (157) is not lexical versus non-lexical, but rather pronoun versus anaphor. Thus, PRO is a pronominal in the sentences in (159) according to our analysis, and we see that the corresponding subjunctives with "like-pronouns" are possible as in (160).

(159) (a) ($_{\bar{S}}$ PRO d'être menacé de mort) ne me fera pas changer d'idée.
(b) Jean m'a dit qu'il serait possible ($_{\bar{S}}$ PRO d'être admis à l'académie) si nous en faisons la demande.
(c) Il est préférable pour nous tous ($_{\bar{S}}$ PRO d'aller voir ce flim)

(160) (a) (Que je sois menacé de mort) ne me fera pas changer d'idée.

(b) Jean m'a dit qu'il serait possible (que nous soyons admis à l'académie) si nous en faisons la demande.

(c) Il est préférable pour nous tous (que nous allions voir ce film).

One might argue that there are still minor specifications that must be made about which NP can be a controller. For example, Grinder's (1970) Intervention Constraint might be brought back to surface to account for the possibility of controller in the following sentence.

(161) Eric said that Roxanne knew that it would be difficult (PRO to criticize herself/*himself)

If PRO is a pronoun here, something additional must be said about its possibilities to corefer. We see that PRO is indeed a pronoun here since it can be free, i.e. arbitrary, as in (162).

(162) Eric said that Roxanne knew that it would be difficult (PRO to criticize oneself)

As for the constraint on the interpretation of (161), it has nothing to do with distance from the controller or any factor of this type: it is simply due to semantic properties of the verb know, and inverting the verbs in (161) as in (163) changes the possibilities of coreference for the pronominal PRO.

(163) Eric knew that Roxanne said that it would be difficult (PRO to criticize herself/himself)

We do recognize that some semantic and pragmatic factors are at play ir possibilities of coreference, but these are not part of a "theory of control".

5.6. Appendix: control in French.

We will not give an exhaustive analysis of infinitival clauses in French, this is a topic for another whole dissertation, but we will look at some leads in the analysis of French infinitival clauses.[22] We saw in footnote (7) that Kayne (1981b) provides some arguments to the effect that de (and presumably à and par in some cases) is in COMP in French infinitival clauses, just like English for. Thus, 1° de does not co-occur with a WH-phrase in COMP; 2° de appears with infinitival complements, but not NP complements after certain verbs; 3° de is not possible after raising verbs; 4° de precedes the negative particule ne.

These observations are not quite accurate however. Thus, as noted by Kayne (1981b), some verbs which are traditionally considered to be raising verbs take a preposition in their infinitival complement's COMP, like commencer (à), finir (de, par), menacer (de), risquer (de), continuer (à), cesser (de), arrêter (de)... But if these verbs are raising verbs, then in Kayne's analysis, the trace is not properly governed by the preposition in COMP, since he assumes that P is not a proper governor in French. His proposal is to say that these

are in fact control verbs. A similar conclusion, for different reasons, is presented in Rouveret & Vergnaud (1980) for menacer. Rouveret & Vergnaud give the following paradigm to show that menacer is in fact a control verb, since it does not behave like sembler in some instances.

(164) (a) Il semble/menace d'y avoir beaucoup de monde à la fête.
(b) Il semble/menace de pleuvoir
(c) Il semble/*menace de falloir partir
(d) Il semble/*menace de s'avérer que Jean est idiot.

Rouveret & Vergnaud claim that subjects of atmospheric verbs and of expressions like y avoir are controlable, thus explaining the acceptability of (164a-b). On the other hand, they say that expletive subjects cannot be controlled as is evidenced by the contrast between (165) and (166), where pleuvoir would have a controllable subject, but not the "ergative" verb tomber.

(165) Ici, il tombe rarement beaucoup de neige sans pleuvoir.

(166) *Ici, il pleut rarement sans tomber beaucoup de neige.

But if the expletive subject of tomber is not controllable, then Rouveret & Vergnaud have no explanation for the grammaticality of (167) if menacer is not a raising verb.

(167) Il menace de tomber beaucoup de neige.

The contrast between (165) and (166) is probably due to the fact that il can transmit Case to the complement of tomber in (165), but that a PRO controlled by il as in (166)

cannot. On the other hand, (167) shows that a raised il can transmit such a Case, as we must say anyhow for sentences like (168).

(168) Il semble tomber beaucoup de neige chez vous.

The reason why menacer is not very good in (164c-d) might be due to its double entry as both a control verb and a raising verb. Thus a nearly synonymous verb like risquer, which is only a raising verb, is much better in these contexts.

(169) (a) Il risque de falloir partir.
(b) Il risque de s'avérer que Jean est idiot.

Furthermore, all the verbs in the list above are compatible with ergative verbs (in the sense of Burzio (1981)) in the sentential complements, and such verbs are not supposed to have controllable subjects according to Rouveret & Vergnaud.

(170) Il a commencé à/fini par/menacé de/ risqué de/ continué à/ cessé de/ arrêté de tomber beaucoup de neige/ arriver beaucoup de monde.

All these verbs also allow the EN-AVANT rule of Ruwet (1972) (see also Couquaux (1981)), so that in a control analysis, an alternative explanation will have to be found for a sentence like (171).

(171) (a) La solution du probleme a semblé/risqué d'/ cessé d'/... être révisée
(b) ($_{NP}$ La solution t_i) a semblé/ risqué d'/cessé d'/... en_i être révisée.

En is like Italian ne: it can be related to a trace in a

lexical NP that is in an object position only. So it is possible in passive constructions, raising constructions, ergative verb constructions, but en cannot be related to a trace in an NP that is in a subject position. This generalization is lost for all these verbs in a control analysis, and it is disturbing that the violation of the generalization is found only with these verbs.

Another argument for the raising status of these verbs comes from idioms: idiomatic expressions can be passivized or raised, but they cannot control since they would receive a θ-role in the controller position which they cannot bear.

(172) (a) On a donné tort à la police.
(b) Tort a été donné à la police.
(c) Tort semble avoir été donné à la police.
(d) *Tort a essayé d'être donné à la police.

The verbs in our list above do allow idioms as subjects, and this confirms that they are raising verbs.

(173) Tort risque d'/finira par/ continue d'/...
être donné à la police.

What all this seems to indicate is that these verbs are better dealt with if one assumes that they are raising verbs. But the fact that they have a preposition in COMP is troubling, both for a GB analysis since it entails an ECP violation, and for the present analysis since this means that $\bar{S}$ is not deleted, and hence that Binding by the subject of the matrix clause should be impossible. A solution is to say that the

preposition is not in COMP in these cases, but rather that it is like English to. The matrix verb can still subcategorize for a specific preposition in the lower clause since this preposition is attached to the lower verb, so that it is accessible for subcategorization since the lower verb is the head of $\bar{S}$. As for the reason why there is no WH-phrase possible in these cases, it could come from subcategorization and semantic reasons. Thus although sembler does not have a preposition in the COMP of its infinitival complement, it still does not allow a WH-phrase in the embedded clause. The same is true of control verbs like vouloir. So it might be that the preposition is now always in COMP and that it is sometimes lower in the embedded clause.

There are also instances where the P is not in COMP but where it seems that the P is higher than COMP rather than lower. Consider verbs like penser and parler, which can be control verbs as in (174).

(174) (a) J'ai pensé à partir tôt.
(b) J'ai parlé de partir tôt.

Note that these constructions involve pronominal coreference rather than Binding since in both cases, PRO can be je, je and x, or even just x (at least for parler). the P seems to be higher than COMP since WH-extraction is possible to the COMP in sentences like (175).

(175) (a) As-tu pensé à(qui on pourrait inviter t)?
(b) Avez-vous parlez de(qui on pourrait inviter t)?

So if the infinitival clause is under PP, or if the P is in COMP blocking $\bar{S}$ deletion here, a pronominal PRO rather than an anaphoric PRO is expected according to our analysis.

Some verbs like dire enter in constructions where a P is present in the infinitival complement, and also in constructions where no P is present.

(176) J'ai dit de PRO partir tôt.

(177) J'ai dit (PRO vouloir partir tôt/préférer la verte)

In (176), PRO is arbitrary. In (177), PRO is necessarily je. If we assume that de is in COMP here, then $\bar{S}$ is not deleted in (176), hence the pronominal reading. On the other hand, $\bar{S}$ deletion takes place in (177), and je Binds PRO, giving a controlled reading. Note that (176) cannot have an interpretation where je is coreferential with the pronominal PRO: this could be due to the Elsewhere Principle (cf. (3.227)) given the availability of a construction with an anaphoric PRO like (177). The de seems to be in COMP in (176), given the following contrast noted by Kayne (1981b).

(178) (a) Je lui ai dit (où aller)
(b) *Je lui ai dit (où d'aller)

Dire can also take a Dative object.

(179) Je lui ai dit (PRO préférer la verte)

(180) Je lui ai dit (de PRO partir tôt)

In (180), PRO can be coreferential with lui, not with je: the reason might again be due to the Elsewhere Principle. In (179), only je is the Binder: this is probably due to semantic reasons related to the meaning of dire which has drifted depending on whether it has a de+infinitive complement or an infinitive complement: in the first case, dire has an imperative meaning, and in the second case, it has a meaning close to admit. Thus interchanging of the complements in (179-180) as in (181-182) gives semantically odd results.

(181) ??Je lui ai dit (PRO partir à 5 heures)

(182) ??He lui ai dit (de PRO préférer la verte)

There are subject control verbs that take a de+infinitive complement.

(183) J'ai essayé/oublié/décidé (de PRO partir)

The de seems to be in COMP, given the facts in (184).

(184) (a) J'ai oublié (où PRO aller)
(b) *J'ai oublié (où d'aller)

The reason why there is no WH-phrase with de cannot be attributed to the subcategorization or semantics of oublier since (184a) is grammatical: so it must be due to the presence of de in COMP. But then de is not obligatory in the infinitival complement of verbs like oublier since a sentence like (184a) is grammatical. On the other hand, if no WH-phrase is present, the presence of de seems to be required, as we

see by the ungrammaticality of (185).

(185) *J'ai oublié/essayé/décidé (PRO partir)

So it seems that the requirement for a de in COMP is a somewhat "surfacy" requirement" if the COMP is not filled, something must fill it at some level. A similar fact is noted in Rochette (1980) with respect to sentences like (186-187).

(186) (a) Je force Marie à manger.
(b) Je persuade Marie de manger.

(187) Je permets à Marie de partir.

Rochette points out that in sentences like (186-187), "the presence of a preposition is syntactically predictable since French cannot have a sequence of more than one verbal complement without introducing the second (and any other following the second one) by a preposition." (Rochette (1980), p. 15).

A similar kind of phenomenon is also found with verbs like avouer, certifier, admettre. These verbs are subject control verbs and take a bare infinitival complement.

(188) Jean a avoué/certifié/admis (PRO avoir tout vu)

But if the infinitival clause becomes a sentential subject, or if impersonal passivization applies, then de is required.

(189) (From Rochette (1980))
(a) De pouvoir participer à la conférence nous avait été certifié à maintes reprises par les authorités.
(b) Il nous avait été certifié (de pouvoir participer au meeting)

Now suppose that insertion of de in an empty COMP in such cases is done at a late level, say only at PF, when certain syntactic or lexical conditions are met. Then all the facts above follow: de and a WH-phrase do not co-occur in COMP because de is inserted only if COMP is empty. As for the subject control cases in (183), if de is inserted only at PF (due to lexical conditions in this case), then the S-structure is as in (190).

(190) NP V ($\bar{S}$ ...($_S$ PRO ...))

If this is a case of subject control, then according to our analysis, NP ought to Bind PRO in (190). But it seems that it should not be able to do it since an $\bar{S}$ is present, blocking government. However, if we interpret $\bar{S}$ deletion in a way similar to Lasnik & Kupin (1977) so that ($\bar{S}$ ($_S$ counts only as ($_S$ when COMP is empty, then at S-structure, government of PRO by NP is possible in (190), since COMP is null at that level, the de being inserted only at PF.[23] So this explains why these verbs are subject control verbs.

So it seems that the analysis to derive control from general principles applying on pronouns and anaphors extends to French rather naturally, although a much more extensive study of French will be needed to see if no modification is needed in the end.

FOOTNOTES; Chapter 5

1. We will not give a systematic presentation of each of these analyses since they are fairly well known, but we will point out some properties of specific analyses when they differ in crucial respects from the others with respect to some point at hand.

2. There is one exception in the recent literature in Manzini (1982), to which we will return.

3. The idea that there is no independent theory of control and that its effects are derivable from principles that account for properties of anaphors and pronominals was presented in (Bouchard (1981)). Manzini (1982) also proposes to wholly derive the theory of control: she claims that PRO being a pronominal anaphor, its properties with respect to control can be derived from the binding theory conditions A and B, given an extended notion of binding theory as in (i) (her (55)).

(i) (A) If $\underline{\alpha}$ is an anaphor and $\underline{\gamma}$ is a governing category for $\underline{\alpha}$, $\underline{\alpha}$ is bound in $\underline{\gamma}$.

(A') If $\underline{\alpha}$ is an anaphor, there is no governing category for $\underline{\alpha}$ and $\underline{\gamma}$ is a domain-governing category for $\underline{\alpha}$
$\underline{\alpha}$ is bound in $\underline{\gamma}$

(B) If $\underline{\alpha}$ is a Pronominal and $\underline{\gamma}$ is a governing category for $\underline{\alpha}$ $\underline{\alpha}$ is free in $\underline{\gamma}$

She proposes the following definitions:

(ii) Governing category:
$\underline{\gamma}$ is a governing category for $\underline{\alpha}$ iff
$\underline{\gamma}$ is the minimal category domain of a subject containing $\underline{\alpha}$ and a governor for $\underline{\alpha}$
$\underline{\gamma}$ contains a nominal accessible to $\underline{\alpha}$

(iii) Domain-governing category:
$\underline{\gamma}$ is a domain-governing category for $\underline{\alpha}$ iff
$\underline{\gamma}$ is the minimal category domain of a subject containing (the domain of) $\underline{\alpha}$ and a governor for the domain of $\underline{\alpha}$
$\underline{\gamma}$ contains a nominal accessible to $\underline{\alpha}$

Note that given (i-A'), there is less reason to believe that PRO is a pronominal: it could be just an anaphor in some cases. Manzini's way of deriving control theory from binding theory is very close to ours: however, her analysis depends on a binding theory where conditions A and B (with some extensions like (A')) are stated, and on the theorem that PRO must be ungoverned. We have seen that conditions A and B do not seem to be independent components of the grammar: condition A derives from the functional definition of an anaphor, and condition B follows from a restatement of the Avoid Pronoun Principle as the Elsewhere Principle (see the appendix of Chapter 3). We will also see that there is no reason to add a statement like (i-A'). Although Manzini seems to adopt a uniform treatment of PRO as a pronominal anaphor, what her analysis reproduces is in fact the properties of anaphors for some PROs, and the properties of pronominals for others: thus some PROs are locally bound by a unique antecedent, whereas others freely corefer.

4. We assume that there are two verbs say: one which is a bridge verb as in John said that Bill was coming, and one which is a nonbridge verb as in (6b). There is a difference in meaning between the two which correlates with the difference in bridge status: the first say is a statement verb, the second one is an "imperative" verb.

5. It is not clear whether Manzini (1982) intends to extend her use of this phonologically null indirect object to the cases of "implicit control" noted by Postal (1970), Wasow & Roeper (1972) and Thompson (1972) as in (i).

(i) (a) Hunting elephants can be dangerous.
(b) Singing loudly is good therapy.
(c) Eating vegetables is healthful.
(d) Playing volleyball is fun.

For example, Thompson notes that (i.a) cannot be interpreted as in (ii), but only as in (iii).

(ii) (a) His hunting elephants can be dangerous for us.
(b) Their hunting elephants can be dangerous for you.

(iii) (X_i hunting elephants) can be dangerous (for X_i)

Although the (for X_i) can be interpreted as a free variable in generic sentences like (i), Thompson also gives examples of "implicit control" in nongeneric sentences as in (iv) where (for X_i) can no longer be interpreted as a free variable.

(iv) (a) Hitting Jack over the head was unwise. (X = me, you)

(b) Going there was fun. (X = me, us)

(c) Tearing up my new paper dolls was mean. (X = you)

(d) Filling out our income tax form was easy this year. (X = us)

If in (iv-d), PRO gets its features from the free variable controller, and if our agrees with PRO, this means that PRO has specific features, hence the free variable too. This poses the question of recoverability of the deletion of the free variable with such features at LF', if a free variable with such features is possible at all.

6. In Manzini's (1982) analysis, these readings, contrary to fact, are given as ungrammatical since there is a phonologically null indirect object to said which obligatorily binds the PRO in (8): the only way in which such readings would be possible in Manzini's analysis is if the phonologically null controller is coreferent with Mary: but then it is not a free variable anymore. Furthermore, it would presumably force the interpretation of the sentence to be something like (8')

(8') Mary knew that John had said to Mary that Mary should behave herself.

But this is not the only possible reading and Mary could have found out from someone else what John had said, for example.

One solution for Manzini could be to say that the indirect object with a free variable is optional here. But then one might expect it to be optional in sentences like (6) also, with subject control, but this is not possible. So this optionality of indirect object would have to be lexically specified, which again amounts to restating that some Vs are subject control Vs and others not.

7. Kayne gives arguments to the effect that de is in COMP just like English for.

1° de does not co-occur with a WH-phrase in COMP.

(i) (a) Je lui ai dit où aller.
(b) Je lui ai dit d'aller là
(c) *Je lui ai dit où d'aller

2° de appears only with infinitival complements after some verbs, not with NP complements.

(ii) (a) Je lui ai dit de partir.
(b) *Je lui ai dit de quelque chose.

3° de is not possible after raising verbs.

(iii) *Jean semble/paraît d'être parti.

4° de precedes the negative particule ne and so parallels English for rather than to.

(iv) (a) Je lui ai dit de ne voir personne.
(b) *Je lui ai dit ne de voir personne.

See the appendix to this chapter for further discussion.

8. Recall that in our analysis of of-insertion, of is inserted obligatorily unless a more specific rule like Genitive Case Assignment takes precedence over it. Since there is

no such rule here, of is inserted, and the Case that it would normally assign is absorbed by passive morphology in the reanalysed taken-care-of.

9. The following analysis was arrived at independently from Kayne's work, which we subsequently discovered to be closely related to our analysis in many interesting aspects.

10. Stowell (1981a) provides some evidence to the effect that matrix verbs seem not to assign Case to infinitival complements in general. Thus he contrasts the sentences in (i) and in (ii).

(i) (a) John knew that the water would be clean.
(b) It was known that the water would be clean.
(c) That the water would be clean was known.

(ii) (a) John knew how to fix the sink.
(b) *It was known how to fix the sink.
(c) *How to fix the sink was known.

The sentences in (ii-b-c) cannot be ruled out because of the need for a controller since PRO can be arbitrary here. Stowell suggests that the reason for this ungrammaticality could be that the matrix verb generally does not assign Case when it has an infinitival complement as in (ii), so that it cannot take passive morphology, which absorbs Case assigning features.

11. Note that in our analysis, considérer in French cannot have a complement structure as in (i) since no V is an ECM V in French, i.e. no V allows a separation of Case and θ-assignment by a verb.

(i) Jean considère ($_S$ Marie heureuse)

Rather, the structure must be something like (ii).

(ii) Jean considère ($_{NP}$ Marie) ($_{AP}$ PRO heureuse)

Burzio (1981) comes to the same conclusion for Italian. He claims that ECM does not exist in Italian, so that, for example, alternations comparable to the English ones in (iii) are altogether lacking (or at least extremely rare) in Italian. The same is true for French.

(iii) (a) I want him captured.
(b) He kicked the door shut.
(c) He pulled the curtains open.
(d) He ordered a monument erected.

If these sentences are cases of ECM, and if ECM does not exist in Italian and French, then we have a straightforward explanation for the contrast between English on the one hand, and Italian and French on the other hand. But then this means that the structure of the complement of verbs like considérer must be as in (ii).

There is also a contrast between French and English as in (iv) which also suggests that the languages differ in this respect.

(iv) (a) I consider it impossible that John did that.
(b) Je considère impossible que Jean aie fait cela.

12. The Aoun & Sportiche notion of government is also used in Chomsky (1981a) to account for the fact that VP adjoined NPs are governed, since extraction from such a position seems

to be possible in view of apparent that-t violations in Italian. See the discussion of this proposal by Rizzi in 4.2 above. But we have seen there that this does not seem to be the right analysis since extraction is possible in configurational Pro Drop languages even when inversion of the subject is not possible. A better analysis seems to be along the lines of the resumptive pronoun strategy. Recall also that, because of the ne cliticization facts, which is possible only from real object positions, not from adjoined positions, this lead to double notions of c-command for the analysis to work.

13. The reason why verbal gerunds do not have a COMP could have to do with their tense interpretation.

14. There are nevertheless sentences like (i), so that this last instance is only a preference for Dative Case.

(i) Ja byl prinuždën borot's ja odin
I(Nom) was forced to fight alone(Nom)

15. The fact that cases of Binding seem to be possible sometimes in passive constructions as in footnote (14) suggests that it is linear adjacency that is relevant for some speakers, not structural adjacency.

16. Some speakers do accept sentences with ask where the subject is the controller even when an object is present.

17. If Dative complements are NPs at some level as we saw in the discussion of (2.52-53), then they can Bind, hence

control, the EC.

18. The contrast between (i) and (ii) presented in Manzini (1982)

(i) The prices were decreased (PRO to help the refugees)
(ii) *The prices decreased (PRO to help the refugees)

can be due to the fact that some agent must be overtly or covertly present in the sentence toward which the purposive can be oriented. So the predicate formed by a passive verb would still have a covert agent role that it could assign, which can be overtly expressed for example by a by-phrase. But in (ii), the predicate has been tampered with, and this agent role is no longer avilable for the purposive to be directed to. Thus a by-phrase is not possible with constructions like (ii).

(iii) (a) The prices were decreased by the board.
(b) *The prices decreased by the board.

19. Note that WH-extraction is possible out of an extraposed infinitival clause as in (i).

(i) Who_i is it important/impossible ($\bar{S}$ ($_S$ PRO to invite t_i))?

So although A-Binding is not possible into such a construction because $\bar{S}$ blocks government, $\bar{A}$-Binding seems to be possible. The explanation is that there is movement to COMP in the embedded $\bar{S}$, so that the structure is in fact as in (ii).

(ii) Who_i is it important/impossible ($\bar{S}$ e_i ($_S$ PRO to invite t_i))?

The e in COMP can be Bound by who, assuming again that elements in COMP head the $\bar{S}$ in some sense: so who Binds e, which in turn Binds t.

20. Thompson (1972) gives the following example from Lord (1972), in which an inanimate reading seems to be acceptable.

(i) PRO colliding with another weather satellite would be next to impossible, according to Dr. Schwitzel.

But note that this example is at the very limit between coreferential pronominal PRO and free PRO, since the use of another clearly biases the interpretation of PRO. If another is replaced by a, then the preferred reading for PRO is (+animate) again.

21. A similar observation is found in Wehrli (1981).

22. On control in French and related topics, see Kayne (1975-1981b), Long (1976a-b), Pollock (1978), Rochette (1980), Ruwet (1972-1976), Wehrli(1980).

23. This means that the fact that $(_{\bar{S}}$ $(_{S}$ cannot be crossed for government even if COMP is empty after a nonbridge verb would have to be slightly revised and some emphasis would have to be put on the fact that some "identification" by the matrix V is necessary, that is, that government across $(_{\bar{S}}$ $(_{S}$ is possible only if the $\bar{S}$ is a complement of the V. Further research will be necessary to give a more substantive description of this "identification".

BIBLIOGRAPHY

Allen, C. (1977), Topics in English Diachronic Syntax, doctoral dissertation, University of Massachusetts at Amherst, Mass.

Andrews, A. (1978), "Remarks on To Adjunction", Linguistic Inquiry, 9.2.

Aoun, Y. (1979), "On government, Case-marking, and clitic placement," mimeo, MIT.

Aoun, Y., N. Hornstein and D. Sportiche (1981), "Some aspects of wide scope quantification", mimeographed, Journal of Linguistic Research.

Aub-Büchner (1962), Le parler rural de Ranrupt, Klincksieck.

Baltin, M. (1978), Toward a Theory of Movement Rules, doctoral dissertation, Massachusetts Institute of Technology.

Borer, H. (1981), Parametric Variations in Clitic Constructions, doctoral dissertation, Massachusetts Institute of Technology.

Bouchard, D.(1979), Conjectures sur une grammaire indépendante du contexte pour les langues naturelles, M.A. dissertation, Université de Montréal.

Bouchard, D. (1980), "L'origine et l'évolution des phrases relatives à tête nominale du francais," Recherches linguistiques à Montréal.

Bouchard, D. (1981), "On some redundancies", mimeo, MIT.

Bouchard, D. (1982a), "Relatives en francais de Montréal et en francais standard: étude d'un paramètre", in C. Lefebvre, ed. (1981).

Bouchard, D. (1982b), "A note on parasitic gaps", mimeo, MIT.

Brame, M. (1981), "Trace Theory with Filters vs. Lexically Based Syntax Without", Linguistic Inquiry, 10.

Bresnan, J. (1972), Theory of Complementation in English Syntax, doctoral dissertation, Massachusetts Institute of Technology.

Bresnan, J. (1982), The Mental Representation of Grammatical Relations, MIT Press.

Bresnan, J. and G. Grimshaw (1978), "The Syntax of Free Relatives in English", Linguistic Inquiry, 9.3.

Burzio, L. (1981), Intransitive Verbs and Italian Auxiliares, doctoral dissertation, Massachusetts Institute of Technology.

Chao, W. (1981), "Pro Drop Languages and Nonobligatory Control", mimeo, University of Massachusetts, Amherst.

Chaurand, M.(1968), Les parlers de la Thiérache et du Laonnois, Klincksieck.

Chomsky, N. (1955), The Logical Structure of Linguistic Theory, Plenum , 1975.

---------- (1965), Aspects of the Theory of Syntax, MIT.

---------- (1970), "Remarks on nominalization", in R. Jacobs and P. Rosenbaum Readings in English Transformational Grammar, Ginn.

---------- (1973), "Conditions on Transformations" , in S. Anderson and P. Kiparsky, A Festschrift for Morris Halle, Holt, Rinehart and Winston.

---------- (1975), Reflections on Language, Pantheon.

---------- (1976), "Conditions on Rules of Grammar", Linguistic Analysis 2

---------- (1977a), Essays on Form and Interpretation, North-Holland.

---------- (1977b), "On WH-movement", in Culicover, Wasow and Akmajian, eds.,

---------- (1980), "On Binding", Linguistic Inquiry, 11.1.

---------- (1981a), Lectures on Government and Binding, Foris.

---------- (1981b), "Some concepts and consequences of the theory of Government and Binding", mimeo, MIT.

---------- (1981c), "A note on non-control PRO", Journal of Linguistic Research.

Chomsky, N. and H. Lasnik (1977), "Filters and Control", Linguistic Inquiry, 8.3.

---------- (1978), "A Remark on Contraction", Linguistic Inquiry, 9.2.

Cinque, G. (1979), "On extraction from NP in Italian", mimeo, MIT.

Couquaux, D. (1981), "French Predication and Linguistic Theory", mimeo, Marseille-Luminy.

Culicover, P., T. Wasow and A. Akmajian, eds. (1977), Formal Syntax, Academic Press.

Damourette & Pichon (1911), Des mots à la pensée, Bibliothèque du "Français moderne".

Diez, A.(1876), Grammaire des langues romanes, Vierweg.

Erteschik, N. (1973), On the Nature of Island Constraints, doctoral dissertation, Massachusetts Institute of Technology.

Evans, C. (1980), "Pronouns", Linguistic Inquiry, 11.2.

Farmer, A. (1980), On the Interaction of Morphology and Syntax, doctoral dissertation, Massachusetts Institute of Technology.

Fiengo, R. (1977), "On trace theory", Linguistic Inquiry, 8.1.

Frai, H. (1929), La grammaire des fautes, Paul Geuthner.

Freidin, R. (1978), "Cyclicity and the theory of grammar", Linguistic Inquiry, 9.4.

Greenberg, J.(1963), Universals of Language, MIT Press.

Grevisse, M. (1964), Le bon usage , Duculot.

Grimshaw, J. (1974), "Relativization by deletion in Chaucerian Middle English", in E. Kaisse and V. Hankamer, eds., Papers from the Fifth Annual Meeting of the North Eastern Linguistic Society, Harvard University.

---------- (1980), "On the lexical representation of Romance reflexive clitics", Center for Cognitive Science, MIT, Occasional Paper #5.

Groos, A. and H. vanRiemsdijk (1979), "Matching effects in free relatives", mimeo, University of Amsterdam.

Gross, M. (1968), Grammaire Transformationnelle du Français, Syntaxe du Verbe, Larousse.

Gruber, J.S. (1965), Studies in Lexical Relations, doctoral dissertation, Massachusetts Institute of Technology.

Guiraud, (1966), "Le système du relatif en français populaire", Languages 3.

Haigneré, (1903), Le patois boulonnais, Slatkine Reprints, 1969.

Haïk, I. (1982), "Indirect Binding", mimeo, MIT.

Hale, K. (1978), "On the position of Walpiri in a typology of the base", mimeo, MIT.

-------- (1981), "Remarks on so-called non-configurational languages," presented at NELS, MIT.

Hankamer, J. (1974), "On WH-indexing", NELS 5.

Helke, M. (1970), The Grammar of English Reflexives, doctoral dissertation, Massachusetts Institute of Technology.

Higginbotham, J. (1980), "Pronouns and bound variables", Linguistic Inquiry, 11.4.

Hornstein, N. and A. Weinberg (1981), "Case theory and preposition stranding", Linguistic Inquiry, 12.1.

Huang, J. (1981), "Move WH in a language without WH-movement", mimeo, MIT.

--------- (1982), Between Syntax and Logical Form: A Case Study in Chinese, doctoral dissertation, Massachusetts Institute of Technology.

Jackendoff, R. (1972), Semantic Interpretation in Generative Grammar, doctoral dissertation, Massachusetts Institute of Technology.

---------- (1977), $\bar{X}$-Syntax, A Study of Phrase Structure, Linguistic Inquiry Monograph 2, MIT.

Jaeggli, O. (1980), "Remarks on To contraction", Linguistic Inquiry, 11.2.

---------- (1981), On Some Phonologically-null Elements in Syntax, doctoral dissertation, Massachusetts Institute of Technology.

Kayne, R. (1975), French Syntax: The Transformational Cycle, MIT.

--------- (1981a), "ECP extensions", Linguistic Inquiry, 12.1

--------- (1981b), "On certain differences between French and English", Linguistic Inquiry, 12.3.

Kayne, R. (1981c), "Comments on N. Chomsky", a Rationalist Approach to language and cognition", mimeo, Université de Paris VIII.

Kayne, R. and Pollock (1978), "Stylistic inversion, successive cyclicity, and move NP in French", Linguistic Inquiry, 9.4.

Kiparsky, P. (1982), Lexical "Morphology and Phonology", mimeo MIT.

Koopman, H. and D. Sportiche (1981), "Variables and the Bijection Principle", mimeo, U.Q.A.M.

Koster, J. (1978a), Locality Principles in Syntax, Foris.

---------- (1978b), "Conditions, empty nodes, and markedness", Linguistic Inquiry, 9.4.

Lakoff, C. (1970), "Global Rules", Language 48.

Lees, R. (1960), The Grammar of English Nominals, Mouton.

Lefebvre, C. (1982), Le français de Montréal, Office de la langue française du Québec.

Lefebvre & Fournier (1978), "Les relatives en français de Montréal", in Recherches linguistiques à Montréal.

Levin, L. and J. Simpson (1981), "Quirky case and lexical representations of Icelandic verbs", presented at C.L.S.

Lieber, R. (1980), On the Organization of the Lexicon, doctoral dissertation, Massachusetts Institute of Technology.

Long, M. (1976a), "French infinitival complementizers and their place in a generative grammar", in Lujan & Hensey.

-------- (1976b), Semantic Verb Classes and Their Role in French Predicate Complementation, doctoral dissertation, Indiana University.

Manzini, M.R. (1980), "On control", mimeo, MIT.

---------- (1981), "Syntactic-like conditions on phonological rules", mimeo, MIT.

---------- (1982), "On control and control theory", mimeo, MIT.

Marantz, A. (1981), On The Nature of Grammatical Relations, doctoral dissertation, Massachusetts Institute of Technology.

May, R. (1977), The Grammar of Quantification, doctoral dissertation, Massachusetts Institute of Technology.

Mohanan, K.P. (1981), "On pronouns and their antecedents", mimeo, MIT.

Morin, J.Y. (1978), "Une théorie interpretative des causatives en français", Linguisticae Investigations 2.

---------- (1979), "Interpretation des pronouns et des réfléchis en français", mimeo, Université de Montréal.

Nash, D. (1980), Topics in Warlpiri Grammar, doctoral dissertation, Massachusetts Institute of Technology.

Neidle, C. (1982), "Case agreement in Russian", in J. Bresnan, ed.,

Obenauer, H. (1976), Etudes de syntaxe interrogative du français, Niemeyer.

Perlmutter, D. (1968), Deep and Surface Contraints in Syntax, doctoral dissertation, Massachusetts Institute of Technology.

---------- (1971), Deep and Surface Constraints in Syntax, Holt, Rinehart and Winston.

Pesetsky, D. (1982), "Complementizer-trace phenomena and the nominative island condition", mimeo, MIT.

---------- (1979), "Russian morphology and lexical theory", mimeo, MIT.

Picallo, C. (1982), "The INFL-node and the Pro-drop parameter", mimeo, CUNY.

Platero, P. (1978), Missing Noun Phrases in Navajo, doctoral dissertation, Massachusetts Institute of Technology.

Pollock, J.-Y. (1974), "Réanalyse et constructions impersonnelles", Recherches linguistiques 8, Universite de Paris VIII.

Postal, P. and G. Pullum (1978), "Traces and the description of English complements and contraction", Linguistic Inquiry, 9.1.

Postal, P. and G. Pullum (1982), "The contraction debate ", Linguistic Inquiry, 13.1.

Reinhart, T. (1976), The Syntactic Domain of Anaphora, doctoral dissertation, Massachusetts Institute of Technology.

Remocke, N. (1960), Syntaxe du parler Wallon, Les Belles lettres.

Riemsdijk, H. van (1978), A Case Study in Syntactic Markedness, Foris Publications.

Riemsdijk, H. van and E. Williams (1981), "NP-structure", The Linguistic Review, 1.3.

Rizzi, L. (1978), "Violations of the WH-Island constraint in Italian and the subjacency condition", mimeo, MIT.

--------- (1979), "WH-movement, negation and the pro-drop parameter", mimeo, Scuola Normale Superiore.

Ross, P. (1967), Constraints on Variables in Syntax, doctoral dissertation, Massachusetts Institute of Technology.

Rothstein, S. (1981), "Preposition stranding and the status of PP as a bounding node in English", mimeo, MIT.

Rouveret, A. and J.R. Vergnaud (1980), "Specifying reference to the subject", Linguistic Inquiry, 11.1.

Ruwet, N. (1972), Théorie syntaxique et syntaxe du français, Seuil.

-------- (1976), "Subject raising and extraposition", M. Lujan, F. Hensey, eds., Current Studies in Romance Linguistics, Georgetown University Press.

Safir, K. (1982), Syntactic Chains and the Definiteness Effect, doctoral dissertation, Massachusetts Institute of Technology.

Salamanca, D. (1981), "Subjunctive syntax", mimeo, MIT

Sanfeld, K. (1943), Syntaxe du français contemporain, Droz.

Schachter, P. (1973), "Focus and Relativization", Language 49.

Selkirk, E. (1972), The Phrase Phonology of English and French, doctoral dissertation, Massachusetts Institute of Technology.

---------- (forthcoming), Phonology and Syntax: The Relation Between Sound and Structure, MIT Press.

Sportiche, D. (1979), "On bounding nodes in French", mimeo, MIT.

Stowell, T. (1980a), "Subjects across categories", mimeo, MIT.

---------- (1980b), "Complementizers and the empty catetory principle", Proceedings of NELS XI, University of Ottawa.

---------- (1981a), Origins of Phrase Structure, doctoral dissertation, Massachusetts Institute of Technology.

---------- (1981b), "The tense of infinitives", mimeo, UCLA.

Suñer, M. (1981), "On null subjects", mimeo, Cornell.

Taraldsen, T. (1978), "On NIC, vacuous application and the that-trace filter", mimeo, MIT.

Thompson, S. (1972), "On subjectless gerunds in English", Linguistic Inquiry, 3.

Torrego, E. (1981), "On inversion in Spanish and some of its effects", mimeo, University of Massachusetts, Boston.

Vat, J. (1978), "On footnote 2: evidence for the pronominal status of poer in Old English relatives", Linguistic Inquiry, 9.4.

Vergnaud, J.R. (1974), French Relative Clauses, doctoral dissertation, Massachusetts Institute of Technology.

Vinet, M.T. (1979), "Core grammar and intransitive prepositions in one variety of French", mimeo, McGill University.

Wasow, T. and R. Roeper (1972), "On the subject of gerunds", Foundations of Language 8:1.

Wehrli, E. (1980), "Construction infinitive: compléments VP et leurs implications theoriques" , Montreal Working Papers 14.

---------- (1981), "On thematic roles and exceptional case marking", mimeo, MIT.

Williams, E. (1975), "Small clauses in English", in J. Kimball, ed., Syntax and Semantics, vol. 4, Academic Press.

---------- (1978), "Across-the-board rule application", Linguistic Inquiry, 9.1.

Williams, E. (1980), "Predication", Linguistic Inquiry, 11.1.

---------- (1981), "'Lexically related' and 'Head of a word'", Linguistic Inquiry, 12.2.

Zribi-Hertz, A. (1980), "Coréférence et pronoms réflechis: Note sur le contraste lui/lui-même en français", Linguisticae Investigationes IV,1.

Zubizarreta, M.L. (1979), "Remarks on Portuguese infinitives", mimeo, MIT.