STRUCTURAL INVARIANCE AND SYMMETRY

IN

SYNTAX

by

DOMINIQUE SPORTICHE

Maîtrise de Mathématiques
Université de Paris 7

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENT
FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Octobre 1983

ⓒ Dominique Sportiche

Signature of author_____
Department of Linguistics and Philosophy
October 5, 1983

Certified by_____
Noam Chomsky
Thesis Supervisor

Accepted by_____
Samuel Jay Keyser
Chairman, Departmental Graduate Committee

STRUCTURAL INVARIANCE AND SYMMETRY

IN SYNTAX

by

DOMINIQUE SPORTICHE

Submitted on October 5, 1983
to the Department of Linguistics and Philosophy
in partial fulfillment of the requirements for the
Degree of Doctor in Philosophy.

ABSTRACT

This essay investigates the incidence of the Isomorphy Principle, a principle of thematic invariance across levels of syntactic representations, on the nature of the relations between these levels, within the model of Universal Grammar proposed by the Transformational Generative Theory (the Government and Binding framework). This leads us to undertake a reanalysis of various syntactic dependencies -move NP, move-wh...- and to develop a theory of $\dagger$-categories and correlatively a theory of Binding relations.

Move NP is exclusively studied from the point of view of syntactic chains, from which its properties will be shown to be entirely derivative: this result entails primarily that D-structure is not an independent level of representation.

Move-wh and more generally the theory of the set of A'/A relations is investigated. We show that this set is symmetric with respect to the value of any binary classificatory features used. In particular, we conclude that invariance across levels is one such feature so that A'/A relation types partition equally depending on whether they remain invariant across levels or not: we also deduce that clitic constructions do not involve an A'/A relation.

The set of $\dagger$-categories is also shown to be closed under symmetry. From this, we conclude that there is no type distinction between expletive PRO and NP-trace, and between pronouns, resumptive pronouns, wh-traces and pro. This last result is the conceptual cornerstone of our treatment of Weak Crossover, Strong Crossover and Parasitic Gap structures. We conclude as well that PRO is a "pure" anaphor and that the theory of its referential properties -Control Theory- partly reduces to Binding Theory, partly to the theory of the range of non-overt operators.

Thesis Supervisor: Noam Chomsky
          Title: Institute Professor

PREFACE

TABLE OF CONTENTS

INTRODUCTION

1. Setting

This study presupposes a rather rich background of assumptions about the nature of human grammatical knowledge - or linguistic competence - and about how to proceed to investigate it, describe it and explain its ontogenetic development in the mind.

This inquiry is basically assigned the task of constructing explicit models of the linguistic knowledge of individual speakers, models usually called grammars, and beyond that of constructing an explicit model of human linguistic competence, understood to be a theory of formal and substantive universals by biological necessity, a theory often referred to as Linguistic Theory or (the theory of) Universal Grammar (U.G.).

I will not attempt to provide a detailed description of this background of assumptions here, except for some of the rather specific technical apparatus which has a direct bearing on the conduct of our investigation.

The epistemological and methodological assumptions underlying this enterprise are presented in many recent publications. They expound and justify fundamental hypotheses concerning necessary idealizations, the appropriate level of abstraction at which such investigations must be conducted (cf. Chomsky (1975; 1980a)), and also discuss boundary conditions that a successful model of UG must meet - e.g. questions of

empirical, descriptive and explanatory adequacy (cf. Chomsky, op. cit; Horstein & Lightfoot, 1981, Introduction).

The basic model of UG has remained unchanged in some fundamental features since its earliest formulation in Chomsky (1955). Linguistic theory still attempts to reduce the immense complexity of linguistic or grammatical knowledge to manageable proportions by modeling it in constructing a system of <u>levels</u> <u>of</u> <u>representations</u> and still regards itself as the abstract study of these levels and of their interrelations. However, the internal structure of these levels of representation and the theory of their relations has considerably evolved (cf. Chomsky (1965; 1982); Bresnan (1982a)).

In this study, we adopt and presuppose most the of model of grammar presented in <u>Lectures</u> <u>on</u> <u>Government</u> <u>and</u> <u>Binding</u> (LGB, Chomsky (1981) which develops a global and integrated view of the particular research program initiated in Chomsky (1973). This model is sometimes referred to as the Government/Binding theory (GB theory) within the general framework of <u>Transformational</u> <u>Generative</u> <u>Grammar</u>.

This model conceives UG as organized into a set of levels of representation whose nature and interactions are regulated by a number of parametrized principles. Acquisition of knowledge of grammar consists partly in setting the values of these parameters on the basis of presented experience. This process is sometimes said to yield a "core grammar". An actual grammar, representing full grammatical knowledge of a language, consists of a core grammar extended to a periphery incorporating more idiosyncratic elements.

2. Assumptions

We now give a rough outline of the organization of the relevant aspects of the model of the U.G. and we list some of our starting assumptions, espectially when they differ from those of Chomsky (1981) to which and to whose bibiliography we refer the reader for detailed discussion.

As we have said, in the GB theory, U.G. is conceived as consisting of levels of representations and relations between these levels. Levels cluster naturally depending on the nature of the primitive vocabulary and relations each uses. We can thus distinguish several components, each consisting of one or more levels:

(1)   i.   The syntactic component: D-structure
                                            S-structure
       ii. The Logical Form component: Logical Form (LF)
       iii. The Phonetic Form component: . PF

The syntactic component, properly speaking, comprises two levels of representations, D-structure and S-structure, as indicated in (1i), and the LF component One level: the level of LF. We will designate the reunion of these two components as _syntax_, and the levels they include as _levels_ _of_ _syntactic_ _representations_.

The relation between levels is implemented by mapping operations which, as far as syntax is concerned, conform to the general schema:

(2)   Move x

Depending on the two levels it relates, a _rule_ or mapping of the type Move x has specific properties. We assume, following Chomsky & Lasnik (1977) that the various levels are organized as in (3):

```
                        D-structure
                            |
(3)                     S-structure
              PF-levels                      LF-level
```
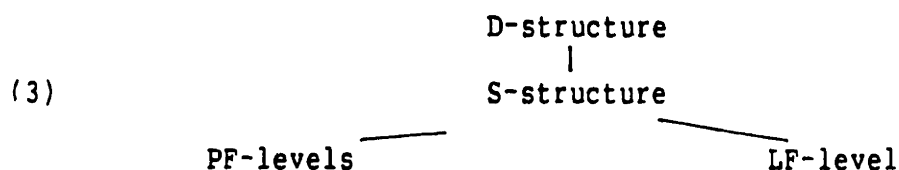
(3) means that any mapping between levels belonging to different components can be factorized into two mappings, each one involving S-structure.

We sometimes understand component C as referring to both the levels in C and to the various mappings relating these levels between them and to the S-structure level. In particular, we use the symbol LF to refer ambiguously to the level of Logical Form, to the mappings between S-structure and LF or to both.

Given a "sentence" S, U.G. assigns it a set of structural descriptions, each at some level of representation. In syntax, these structural descriptions take partly the form of a set of strings, representable as labelled bracketing which assigns a "Phrase marker" interpretation to S. The properties of these structural descriptions, i.e. the internal structure of these levels is determined by the joint consequences of a number of subtheories, each applying at one (or more) level:

(4)   i.   The Base
        ii.   Government Theory
        iii. θ-theory
        iv.   Case theory
        v.   Binding theory
        vi.   ECP
        vii. Bounding theory
        viii. Control theory

Let us now roughly outline each of them (except Control theory, to which we return to in Chapter IV).

2.1. The Base

The base contains the lexicon and its theory, and the categorial component. The lexicon specifies the abstract morpho-phonological structure of each lexical item, its syntactic features, including categorial features and contextual features. Amongst these it is specified whether a lexical item is of the "argument" type or of the "predicate" type. If it is of this latter type, it is specified how many places this predicate has. In such a case, the lexical entry of the predicate further states amongst its contextual features the categorial nature of the arguments (that it governs) which it can take, features sometimes called subcategorization features.

The theory of the lexicon contains generalizations about the nature and properties of contextual features and their cross-categorial invariance, e.g. perhaps X'-theory - although X'-theory might possibly be generalized to non lexical categories, e.g. INFL, S - (cf. Chomsky (1970); Jackendoff (1977)).

Finally, we assume that the categorial component reduces essentially to the single rule S ---> NP INFL VP and that the properties of constituent ordering follow from other properties of grammar (cf. Stowell (1981); Koopman (1983)). We assume INFL to be the head of S, i.e. to contain the properties (mood, tense, AGR...) determining those of S. More generally, given an X^0 in the X'-system, the maximal phrase containing X^0, whose properties are determined or projected from those of X^0 is called a maximal projection of X^0 and noted X^{max}. X^0 is called the head of X^{max}.

2.2. Government Theory

The central notion of Government Theory is the relation between the head of a construction and the categories dependent on it. This relation, already isolated in Chomsky (1965), has come to play a prominent role recently, starting with Rouveret & Vergnaud (1980) and Chomsky (1980). Following Aoun & Sportiche (1981), we will define it as:

(5)   x governs y iff   ↑, ↑ a maximal projection
                              ↑ dominates x iff ↑ dominates y

Properly speaking, the government relation holds only of a pair (x,y) where x is an X° and Y is an X^{max}.

However, we will also use the term government to denote the structural relation in (5) regardless of the categorial nature of X and Y.

2.3. θ-Theory

We will take θ-theory to be concerned with the relation between argument-places of a predicate and arguments. Note first that the notions argument and argument-of must be carefully distinguished. The notion argument is a semantic notion referring to a property of phrases "denoting" objects, concrete or abstract, state of affairs... Argument-of is a formal relational notion holding of a pair (x,y) where x is a predicate and y an argument linked to an argument place of x.

The semantic function assigned to a particular argument place of a predicate $\underline{P}$ is determined in part by the lexical properties of $\underline{P}$ and is often called thematic role or θ-role.

The syntactic representation of an argument-place is called a

θ-position. A syntactic position which is not an argument place is called a θ'-position. We will use interchangeably "receive a θ-role from" and "is linked to an argument place of". A syntactic position in a Phrase marker that can receive a θ-role, given appropriate choice of lexical material this phrase marker contains is called an A-position (≠ from argument position, which means argument-place). A syntactic position which is not an A-position is called an A'-position.

The fundamental principle of θ-theory is the θ-criterion - a generalization of the principles of functional uniqueness and relatedness of Freidin (1978) - which we assume is stated as follows:

(6)   θ-criterion
          Each argument is linked to one and only one argument-place
          Each argument-place is linked to one and only one argument

It is usually assumed that all argument-places of a predicate must be syntactically represented (except, perhaps, for some morphologically derived predicates, e.g. passives). Notice incidentally that some syntactic position S can be an argument-place of some predicate P only if P governs S: θ-role assignment requires government.

The formalization of the notion linked used in (6) requires a statement of (6) in terms of well-formedness conditions on objects called syntactic chains, to which we return in detail in Chapter II (For the definition of chain, and BIND, cf. Chomsky (1981, p. 333)).

The Projection Principle states that the θ-criterion must hold at every level of syntactic representation, i.e. D-structure, S-structure and LF. In particular, it is assumed that it holds in a particular way at D-structure in that each argument occupies an argument-place and each argument-place contains an argument at that level. The Projection

Principle implies the existence of syntactic positions representing the argument places of predicates at all syntactic levels, and only of such positions. It is noteworthy that the Projection Principle by itself implies the existence of subject positions of clauses only when they are θ-positions. Since the rule S --> NP INFL VP of 2.1. above also implies the obligatoriness of a syntactic position - the subject position of clauses - at all syntactic levels, the Projection Principle and the effects of this rule are sometimes conjoined under the name of Extended Projection Principle. Note that it follows essentially that only subject positions of clausal structures can be θ'-positions, since they are the only positions whose existence is not a consequence of the Projection Principle.

2.4. Case Theory

Case theory deals with the assignment of Abstract Case and its morphological realization (cf. Vergnaud (1982)). Case assignment (or, rather Case checking, as in Jaeggli, 1978) to x by y requires government of x by y.

We name the particular Cases assigned in accordance with traditional usage, nominative, objective, genitive...These distinctions will be considered to be a matter of morphological realization and will play no role in this scudy.

The fundamental principle of Case Theory is the Case Filter:

(7)   The Case Filter

           *NP where NP is Caseless and phonologically non-null

In LGB Chapter 6, Chomsky proposes to reduce the Case Filter to the
θ-criterion by essentially requiring that arguments be Case-marked in
order to be visible for θ-role assignment. This is the Visibility
Hypothesis. (More precisely, this reduction is achieved by requiring of
syntactic chains containing an argument to contain a Case-position to
validate the θ-position it has to contain).

2.5. Binding Theory

The central relation of Binding Theory is the binary relation binding
derived from the notion of c-command introduced in Reinhart (1976). Here,
we will adopt the version of this notion advocated in Aoun & Sportiche
(1981):

(8)   x c-commands y   iff   ∀γ, γ a maximal projection
                                    γ dominates x only if γ
                                    dominates y

(9)   z X-binds y   iff   z c-commands y and z is coindexed
                                    with y and z and y are of the same
                                    categorial nature and z is in an
                                    X-position (X=A or A')

If z (X)-binds y, y is said to be (X)-bound by z. If y is not (X)-bound
by anything (in some syntactic domain D), it is said to be (X)-free (in
D).

As M. Brody has first remarked, the useful notion of Binding theory
is actually local binding, which we will define as in (10):

(10)   z locally X-binds y iff   z X-binds y and ∀t/t binds y,
                                       the first constituent
                                       containing t and y contains z.

Intuitively speaking, z locally X-binds y if z is the "closest binder" of

y and z is in an X-position.

The Binding Theory is concerned with relations of pronominals, anaphors and names, and corresponding non overt categories, in A-position to possible antecedents in A-position (for definitions of the terms involved: Ɣ-category, empty category, variable, trace, pronominal, cf. Chomsky (1981, p. 330)). It is stated as follows:

(11)   Principle A: An anaphor must be (locally) A-bound in its
                        Governing Category

          Principle B: A pronominal must be (locally) A-free in its
                        Governing Category

          Principle C: R-expressions (i.e. names and variables)
                        must be locally A-free

Where we define Governing Category as in (12):

(12)   The Governing Category of y is the first NP or S
          containing y and some X° governing y.

We assume that the Binding Principles in (11) constrain one or both of S-structure and LF levels. We leave it open whether it does some other level.

2.6. The Empty Category Principle (ECP)

We start by distinguishing empty categories from the null category. The null category is the non argument non overt category appearing at D-structure. An empty category, whose nature, we assume for the time being, is "functionally" determined (as in Chomsky, 1981, p.330), is a non overt category which is not the null category. (We will drop that distinction later favor of the expletive/non expletive distinction). The

distribution of empty categories is constrained by the ECP at the level of LF (and, perhaps S-structure):

(13)   ECP: an empty category must be properly governed

If suffices here to note that <u>proper</u> <u>government</u> <u>is</u> <u>a</u> <u>weaker</u> <u>notion</u> <u>than</u> <u>government</u> <u>by</u> <u>the</u> <u>category</u> <u>V</u>: if V governs some category K, it properly governs it.

2.7. Bounding Theory

Bounding Theory imposes locality conditons on a subset of the inter level mapping operations falling under the schema Move x, or on the resulting configurations at S-structure or at LF. Its central principle is the <u>Subjacency</u> <u>Condition</u>:

(14)   The Subjacency Condition

  * $z_i...y_i$ where z and y are related by Move x
            z locally binds y
            and...contains more than one
            bounding node.

The theory of <u>bounding</u> <u>nodes</u> assumes that at least S' and NP are bounding nodes, and perhaps S in English...

3. Summary of Contents

Finally, let us present a general outline of the content of this study, which might help to keep track of the thread of the argumentation, beyond sometimes fairly technical discussions. Before we proceed, let us

make two remarks about the general structure of the argumentation. Reflecting the modular organization of the model itself, it does not develop linearly. Rather, it can be pictured as a loop-shaped tree: starting from some point on the loop, we encounter branches (sometimes branching themselves) whose nature is relevant to the structure of the loop. So that we explore them, before pursuing along the loop itself. Secondly, and mostly for expository reasons - an axiomatic type of presentation would be cumbersome - the assumptions underlying our argumentation at some point do not remain constant throughout the study. When we begin altering certain assumptions of the LGB model, we keep certain others constant, which we will ultimately modify. We try to point out, as we proceed, why or why not these later modifications do not influence the validity of the earlier arguments. Hopefully, as the loop closes, we end up with a consistent theory.

As can be expected in a model of Linguistic Theory which is becoming richly structured, unsettled interrelated questions arise about every single specific hypothesis, both of empirical and explanatory adequacy, but of internal coherence and consistency as well.

We begin in Chapter I precisely with a question of internal consistency. We try to demonstrate that a general requirement of compatibility between levels of representations, which we call the Isomorphy Principle, is not respected by the LGB model and that this is due to the interaction of two factors:

  i. The formalization of move NP (e.g. Passive...) as a
      mapping from D-structures to S-structure
  ii. The adoption of contextual definition for non overt
      categories

In Chapter II, we give (i) up, postponing until Chapter IV reasons

for not giving up (ii). We explore the possibility of formalizing Move NP by deriving and/or embedding its properties (into) a theory of S-structure chain formation. Correlatively, we argue that properties that have been specifically ascribed to D-structure representations can be derived from S-structure configurations, i.e. that S-structure is the "basic level of syntactic representations" and that D-structure is not an independent level of representation, but rather a projection of certain S-structure properties. This in turn leads to a provisional reorganization of the classification of f-categories, as derivational characterizations are no longer available. In the course of this discussion, we ascertain the truth of some basic propositions (e.g. the Case Filter is independent from the θ-criterion), in order to be able to proceed.

Obviously, if the properties of Move NP are derived from the theory of chain formation, the question arises of how the properties of Move wh should be represented. This topic is investigated in Chapter III. More generally, we establish a basic classification of A'/A relations and correlate it with different criteria: behaviour in Weak Crossover constructions, Parasitic Gap constructions and with respect to the Subjacency Condition. We show that the properties of A'/A relations depend on the nature of A' and on whether the relation is established at S-structure, or at LF, i.e. whether the relation is invariant across levels or not.

Importantly, we argue that there is no type distinction between "wh-traces" and resumptive pronouns. In this light, we explore in more detail properties ascribed to A'/A relations: Strong Cross over, Weak Cross over and Parasitic Gap constructions.

Finally, we discuss how the Subjacency Condition can be formulated

and argue that it is a property of "movement to COMP".

In the Appendix to Chapter III, we come back to pronominal and reflexive Clitic constructions. In Chapter III, we have argued that they cannot be analyzed as involving A'-binding. Here we argue that do not involve A-binding either, by pursuing the consequences of our general assumptions. We conclude that their properties are in fact characteristic of Agreement relations.

In Chapter IV, we return to the problem of ʏ-category classification and of developing a Binding Theory congruent with our earlier conclusions. We first compare two alternative classifications of ʏ-categories. We reach the conclusion that a context dependent classification based on Case is to be preferred, thus justifying a posteriori the choice of (ii) over (i) above in Chapter II. From this, we conclude that PRO may be governed. We also conclude that the properties of ʏ-categories can be predicted, regardless of whether they are overt or not, and that there is no type distinction between pronouns, resumptive pronouns, "wh-traces" and pro. This permits us to embed our approach to strong and weak crossover of Chapter III in a conceptually more coherent general theory of ʏ-categories. We then examine how this classification determines the formulation of the Binding Principles, and turn to the consequences of these conclusions. About Control Theory, we argue that it partly reduces to Binding Theory, partly to the theory of empty operators. We then discuss the nature of NP-traces and argue that they are subcases of expletive PRO. Consequences for the formulation of the ECP are drawn. Finally, we show that assuming that ʏ-categories constitute a symmetric system entails the existence of Caseless variables, from where we derive solutions to several outstanding problems.

CHAPTER I: INVARIANCE OF THEMATIC STRUCTURE: THE ISOMORPHY PRINCIPLE

1. Congruence between levels of representations

As we have seen, one of the important features of the model of grammar presented in Chomsky (1981) and in the Introduction is its modular character: properties of what appears to be a complex system are factored out into several autonomous subcomponents, whose interaction insures that the relevant properties are met.

One case that is particularly interesting to us is the problem of correspondance between levels.

Given a sentence S, the grammar attributes to S structural descriptions $L_i(S)$, each at some linguistic level L^i The grammar must be so structured as to insure that the various structural descriptions $L_i(S)$ of S are compatible with each other and reflect true properties of the mental representation of S. One natural way to insure that the various $L_i(S)$ are compatible is to organize the various linguistic levels as related through well defined mapping operations, which "commute" as the following diagram indicates:

$$
\begin{array}{ccc}
m_1 & \longrightarrow & L_1 \\
 & & \downarrow m_{12} \\
S \longleftarrow m_2 & \longrightarrow & L_2 \\
 & & \downarrow m_{23} \\
m_3 & \longrightarrow & L_3
\end{array}
$$

where, by definition

$m_2 = m_{12} \cdot m_1$

$m_3 = m_{23} \cdot m_2 \; \ldots$

and the properties of the m_{ij} are fixed.
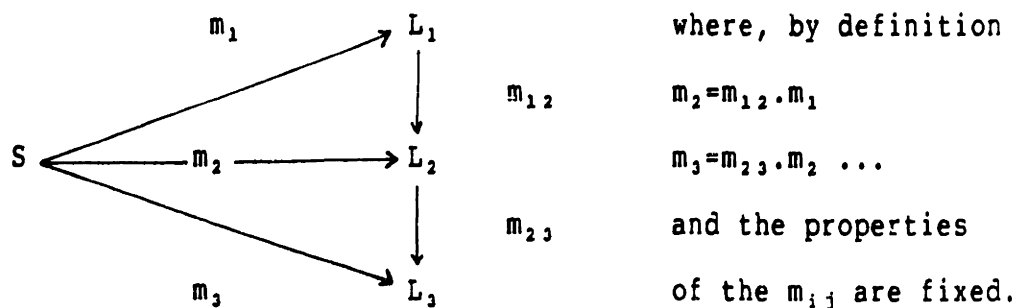
As illustration, consider D-structure and S-structure. D-structure and

S-structure representations are related by the mapping Move x. It is assumed that the respective well-formedness conditions constraining the representations at each level are sufficient to make it possible to require very few intrinsic properties of Move x. Amongst the intrinsic properties of transformational rules that "Move x" meets are perhaps various conditions on possible structural descriptions (e.g. Boolean conditions on analyzability...) and structural changes (e.g. substitution, adjunction...) as well as general conditions such as <u>Recoverability</u> <u>of</u> <u>Deletion</u>.

As illustrative example of the latter, consider the following derivation from D-structure (1i) to S-structure (1ii) via substitution of $[_{NP} e]_i$ = PRO to $[_{NP}$ John$]_i$:

(1)   i.   It is difficult $[_{NP}$ John$]$ to find how $[_{NP}* e]_i$ to solve his problem

        ii.   It is difficult $[_{NP*} e]_i$ to find how $[_{NP*} e]_i$ to solve his problem

Clearly, such a substitution is permitted by Movement Theory per se. Moreover, each representation is well-formed: both empty categories in (1i) are interpreted as PRO's, since the first one is free, and the local binder of the second one is in a θ-position. Since both positions are ungoverned, this is permitted by the Binding theory, and since (1i) is a well-formed S-structure (that could correspond to an identical D-structure) it is not ru<s>l</s>led out either by Case-theory or θ theory.

Some independent means has to be provided to rule out this derivation, namely, the condition of <u>Recoverability</u> <u>of</u> <u>deletion</u>, since the disappearance of <u>John</u> is not recoverable from the S-structure (1i). Considering this example more closely, one may wonder why such a

derivation should be considered ungrammatical.

In a model of grammar where "semantic interpretation" is performed off D-structure (as used to be the case when the Recoverability Condition was introduced), the answer is straightforward, since the resulting pairing (sound, meaning) to put it informally, would be ill-formed.

In a model where interpretation is read off (or off-off) S-structure, the reason for excluding such a derivation, although conceptually similar, must come from somewhat more abstract considerations.

Intuitively speaking, what seems wrong with this derivation is the fact that the argument bearing the θ-role "subject of find " is not the same at the two levels of representations, i.e., it is the fact that the θ-criterion, although it is satisfied at each level, is not satisfied in the same fashion. Note that the Projection Principle is not violated here: it only requires that the θ-criterion be met at all syntactic levels, without paying attention to the actual pairing (argument, θ-role) at each level. Actually, a stronger principle is involved here.

If we call thematic structure at a given level L_i, the set of pairs $(X_i, [Y]_j)$, where X_i is a particular argument, and $[Y_j]$ the θ-position with which it is linked at L_i, we might formulate this principle as in (2):

(2)   The Isomorphy Principle: Thematic structure is
                                  syntax invariant

Thematic structure is syntax invariant means: if some argument A_i bears the particular θ-role $θ_j$ at the syntactic level of representation L_k, then this is also true at any other level of representation L_1.

It is worth pointing out that this principle recalls that part of the Katz-Postal Hypothesis (cf. Katz & Postal, 1964) that has resisted

criticisms resulting in the Extended Standard Theory Model. In essentials, the Katz-Postal hypothesis stated that all sentence grammar semantic information was determined by D-structure configurations, with the corollary assumption, sometimes misleadingly termed "meaning preservingness of transformations", that those structural aspects of a D-structure phrase marker determining thematic structure, quantifier scope..., and other relevant "semantic" information could not be altered in the course of a derivation. Although this hypothesis has been shown to be inappropriate (cf. Chomsky, 1972, Jackendoff, 1972), it is remarkable that no argument to that effect dealt with thematic structure (in the sense adopted here, cf. Jackendoff, 1972, for a different view), but rather with notions such as scope of negative elements, quantifiers, etc...

There are two ways to think about the Isomorphy Principle. We can think of it as an axiom of the theory of grammar, much the same way the Projection Principle is, i.e. as a stipulated property. Or we can think of it as a derived property, either trivially (say, in a model with only one level of representation) or as a real theorem of the theory of grammar. If it is a theorem, it must be, as we pointed out earlier, that the various mappings between syntactic levels are sufficiently constrained.

We will try to show in the next section of this chapter that they are not sufficiently constrained to achieve this result.

Note incidentaly that the Isomorphy principle is not obviously correct, and, if correct could be interpreted in different ways.

The first, and strongest interpretation is the one we have introduced, which takes thematic structure to be a set of pairs (X_i, Y_j),

where X_i is a particular occurrence of a phrasal category with its content.

The second interpretation would only require of X_i to be a particular occurrence of a string which is the content of a phrasal category, without specifying which category-type. For example, recent work by Pesetsky (1982) and Higginbotham (1982) suggests that there might be (different types of) derivations violating the Isomorphy Principle under the first interpretation, but not under the second interpretation, for example by assuming that some instances of movement of a category X may leave a trace of different categorial status Y.

Since nothing in what follows will bear on this question, we will assume here the strongest version.

2. Non isomorphic cases of movement

2.1. Let us now turn to some cases suggesting that the Isomorphy Principle has to be stipulated, if correct.

We have to exhibit well-formed derivations relating two well-formed representations, each at some level, which are such that their respective thematic structures are distinct.

In terms of mapping, there are only two possible candidates: mappings from D-structure onto S-structre, and mappings from S-structure onto LF. Consider first derivations from S-structure to LF and the kind of operations mapping one onto the other: Quantifier raising, wh-raising, perhaps Focus interpretation... All these operations share the property that they relate some A-position to an A'-position (a relation usually

expressing the scope of the item in the A-position). Since thematic structure is a property of the distribution of arguments relative to A-positions, none of these operations will affect thematic structure. Consider next derivations from D-structure to S-structure. Given the remark above, the range of relevant cases i.e. non isomorphic derivations, can be a priori narrowed down to instances of well-formed derivations which involve a movement relation between at least two A-positions. Furthermore, we may start by restricting our attention to cases involving at most two A-positions, i.e. one step derivations in which a single item has moved only once. Moreover, we may assume that this item is an argument (for non-arguments do not enter into thematic structure). Let S be some sentence with distinct thematic structure at D-structure and S-structure. This situation may arise if some moved item, say X, acquires a thematic function that it did not possess prior to movement by virtue of the position it occupies after movement, which may arise in two ways: either movement of X has been to a θ-position, or movement of X has been to a θ'-position which is not "properly related" to the trace of X; that is, it is not recoverable from the S-structure configuration that the trace of X, which indicates the thematic function of X at D-structure, $\underline{is}$, the trace of X.

In fact, cases of the latter sort can be easily constructed. Consider for example, the following pair (3i, 3ii), where (3i) is a D-structure and (3ii) an S-structure:

(3)   i. $[_{S_i}$ that $[_S$ e was told e* $[_{S_i}$ that S $]]]$ indicated
                      $[_{S_i}$ how $[_S$ Bill to VP$]]$

       ii. $[_{S_i}$ that $[_S$ $[Bill]_j$ was told e* $[S_i$ that S$]]]$ indicated
                      $[_{S_i}$ how $[_S$ e_j to VP$]]$

Starting with (3i), consider _e_*. It is in θ-position, hence must be an argument. Consequently, it cannot be the null category, since it is not an argument. So it must be an empty category in the technical sense given in the Introduction and since it is, by necessity at this level, free, it is interpreted as PRO. By exactly the same line of argument, _e_ can only be the null category, or we would have a violation of the θ-criterion. With _e_ and _e_* so defined, it is easy to see that (3i) is a well-formed D-structure. Consider now (3ii). Provided that _e_* is accidentally coindexed with _Bill_ by the procedure of free indexing, (3ii) is a well-formed S-structure. (Otherwise, we would end up with a violation of the Binding Theory since _e_*, interpreted as PRO, would be in a governed position; alternatively, _e_* could be bound outside its clause and thus could be interpreted as a trace, but would fail to be bound in the minimal relevant domain, here its clause, as required by principle A of the Binding Theory).

At this level, it is interpreted as a trace, namely that of _Bill_; and e_i, while it is the trace of _Bill_, it is interpreted as PRO, since it is free. Now, is the movement of _Bill_ from its base position to its surface position permitted? Since we deal here with a case of substitution and we have shown that there is at least one derivation from (3i) to (3ii) that will substitute _Bill_ into a null category, the principle of _Recoverability_ _of_ _Deletion_ is not violated. However, this movement violates a principle of Bounding theory, namely the Subjacency Condition, if we take it to be a constraint on rule application (otherwise, it is irrelevant) for at least two S-nodes and two S'-nodes, all bounding nodes in English, intervene between the two positions.

If we try to find derivations displaying similar properties to those

of (3i, 3ii), and in particular in which there is no c-command relation between the moved phrase and its trace, we will always end up with a violation of the Subjacency Condition: indeed, the moved NP must move to a non-θ A-position, i.e. a subject position, and out of its clause (since the subject of a clause c-commands every position in it). It cannot move to the subject position of the next clause up either for it too would c-command the original trace). So it must move into subject position of yet another clause: we see that, in a way, the Subjacency condition violation is a "consequence" of the non-c-command requirement that we have imposed on the movement process.

We could claim then that this derivation is ruled out by the Bounding theory. Note however, that the nature of the violation might suggest that it is more of an accidental consequence of the formulation of the Bounding theory: the structure in (3ii) is totally uninterpretable with the reading that would be associated with (3i). This is not usually the case with Subjacency violations of equal strength, e.g. violations of wh-islands, or PP extraposition outside its clause.

Furthermore, there are examples showing that the Subjacency Condition would not eliminate all the possible non isomorphic derivations. Before turning to these, let us consider two more relevant possibilities. First, it might be argued that (the LF representation for) (3ii) is ruled out on grounds of Control Theory, because the interpretation for (3ii) where the subject of the infinitive corefers with Bill does not seem to be readily available, a problem that we can easily circumvent by replacing indicated with show him in (3i) and (3ii), with him coindexed with e_i.

A second objection could be based precisely on the fact that movement has taken place to a non-c-commanding position. So it could be argued

that non isomorphic derivations could be avoided if we require that movement only take place to c-commanding positions. Besides it being redundant with the Binding theory in most cases, this option would be insufficient: we now turn to examples of non isomorphic derivations which respect both the c-command requirement and escape a Subjacency violation. Consider the pair below:

(4)   i.   e was told e* [$_{s'}$ how [$_s$ Bill to leave]]

       ii. Bill$_i$ was told e*$_i$ [$_{s'}$ how [$_s$ e$_i$ to leave]]

Exactly the same reasoning holds for this pair, as the one we presented above '3i, 3ii) which shows that both (4i) and (4ii) are well-formed at their respective levels of representation. In this case, however, the movement of the NP Bill neither violates Subjacency (Subjacency is irrelevant here because of similar examples in Italian or French where S is not a bounding node, cf. Rizzi, 1982; Sportiche, 1981) nor any c-command requirement, yet is not isomorphic since the thematic structure changes between D-structure and S-structure: the NP Bill assumes the θ-role"object of tell " at S-structure, but not at D-structure.

But, since examples of the sort (3) are ruled out by the Subjacency condition, we could argue at this point that the Isomorphy Principle is wrong and may be violated in derivations such as those in (4). Although the thematic structure as we have defined it does change, it could be argued that it must be defined in terms of referent rather than argument occurrences. Consider (4) for example. Assuming that indices are present in D-structure and that coindexation indicates identity of reference, we see that the set of pairs (referent, θ-role) does remain unchanged (e* must be already coindexed with Bill at D-structure).

Consider, however, the following examples:

(5)   i.   It is unclear how Bill to tell Mary that e seems e* sick

         ii. It is unclear how e_i to tell Mary that $Bill_i$ seems e* sick

At D-structure, e is the null category and e* is PRO. At S-structure, after movement of <u>Bill</u> and coindexing with e*, e* is interpreted as a trace and e_i as PRO.

But here, the interpretation associated with (5ii) is not availabe.: if the derivation indicated was possible, we would expect a reading of (5ii) in which e_i is referentially dependent upon $Bill_i$. Of course, this reading is not permitted. e_i must be arbitrary in reference in (5ii). Furthermore, we could not invoke a violation of principle C of the Binding Theory – e_i is referentially dependent upon a name it c-commands – given that the same observation holds if we replace <u>Bill</u> by a pronoun, as in (6):

(6)   It is unclear how e_i to tell Mary that he_i seems sick.

Although one might argue that it is ruled out by some version of Control theory, it would seem unnatural, given that Control theory deals with overt element not inheriting indices by virtue of some syntactic process (e.g. coindexing under movement).

It is, on the other hand, excluded straightforwardly by the Isomorphy Principle.

2.2 We have so far restricted our attention to non-isomorphic derivations comprising only one step. It is quite clear that nothing in the examples we have discussed places particular restrictions on the number of

applications of Move x to a given category. There are examples of non

isomorphic derivations which are particularly interesting since they also

are cases of improper movement. So consider the following derivations:

(7)   i.   e were bought e* for the dog [[PRO to play with bones]]

        ii. Bones$_i$ were bought e* for the dog [e$_i$ [PRO to play
             with e$_i$]]

Again, taking $\underline{e}$ to be the null category and $\underline{e}$* to get the index $\underline{i}$ at

S-structure, we get a lawful derivation relating two well-formed

representations, yet one that is not isomorphic. Note that Case conflict

could not be taken to rule (7ii) out, under the hypothesis that NP's are

base-generated with Case and Case is checked in place (as suggested in

Jaeggli, 1978 and Chomsky, 1981), for we can manufacture an example in

which the S-structure position of the moved phrase and its original trace

(on which Case is left, assuming Jaeggl and Chomsky's system) require the

same Case. Consider:

(8)   I believe bones$_i$ to have been bought e* for the dog
             [e$_i$ [to eat e$_i$]]

where both bones and (the) e$_i$ (in A-position) are objective. So the

theory as it stands allows for certain types of improper movement.

2.3 We have mentioned that there were two types of derivations that would

yield violations of the Isomorphy Principle. The first type has been

illustrated in the preceding section: they were cases of derivations in

which some argument acquires some new θ-role by virtue of being moved to a

position where it is linked back to some empty category which is not its

trace. The second type are cases of movement into θ-position.

As things stand in our model of grammar, movement to a θ-position is possible. Let us briefly see why.

Consider some θ-position P. Since D-structure has, among other things, the property of being a pure representation of GF-θ, P contains an argument at D-structure, say A.

Suppose A does not move. Then, if some other argument B moves into P by substitution, this will erase A, which is not possible, assuming that deletion of arguments is not recoverable.

Suppose now that A has been moved. Its trace, now occupying P, is not an argument, so B can move into P. Now A is an argument, and must therefore be assigned a θ-role. But this is clearly possible as we have shown in the preceding section, if A moves to some θ'-position where it acquires a θ-role by being coindexed with some empty category in a θ-position (alternatively, A moves to some θ-position P', in which case let A be B and let the argument formerly occupying P' be A, and we can repeat the argumentation...). An example is given below:

(9)   i.   e was told e* that Bill wondered how Mary to leave.
        ii. Bill$_i$ was told e* that Mary$_j$ wondered how e$_j$ to leave

A similar example can be constructed where the movement to a θ-position is furthermore improper:[1]

(10)   i.   e was told e* that Bill brought John to the baby
             [to play with Henry]
        ii. John was told e* that Bill brought Henry$_i$ to the baby
             [e$_i$ [PRO to play with e$_i$]]

Let us pause at this point and examine the first part of this argument.

Movement to a θ-position P occupied by argument A is excluded by the Principle of Recoverability of Deletion if A does not move. Consider the

scope of that principle.

Deletion may occur either because of some deletion process, or as a result of substitution into a non-null category (which partly reduces to deletion as well).

It is not clear whether there is any need for deletion rules within U.G. any longer: deletion of designated elements such as self in the case of Equi constructions have been reanalysed as involving control of PRO. Deletion of wh-elements in COMP has been reanalysed as movement of phonologically zero wh-phrases and deletion of complementizers has been reanalysed as failure of base generation of the target items or optional non phonetic realization of their feature matrices. Finally, deletion under identity, e.g. gapping or VP deletion, can be viewed as interpretive rules instead.

So we are in fact left with deletion occuring as part of substitution. We have two cases to consider: if we substitute into a position containing a lexical argument (say a proper name), we can say that the Principle of Recoverability of Deletion is violated because of the content of the erased argument. Consider, however, the case of substitution into an argument that is an empty category (recall empty ≠ null), i.e. PRO. Then we cannot say that Recoverability of Deletion is violated because of the erased content of the argument, since a PRO cannot be intrinsically distinguished from an NP-trace, as we have assumed, following Chomsky (1981), and NP-traces are erasable as far as Recoverability of Deletion is concerned.

This casts some doubts on the formulation of the principle involved in blocking these derivations. But it is clear that any derivation involving substitution into a PRO will be non isomorphic (as would be any

derivation involving substitution into a lexical argument). So we can entirely eliminate the Principle of Recoverability of Deletion, which is now subsumed under the Isomorphy Principle.

3. Extended non-isomorphic derivations

We see that, if we accept the idea that the relation between the various syntactic levels of representations should meet the Isomorphy Principle, some modification has to be introduced in the model of grammar we have so far assumed. Before exploring the various alternatives open to us, let us turn to some derivations involving movement to an A'-position, which, although isomorphic, display the same abstract structure as non-isomorphic derivations.

Consider the syntactic rule of wh-movement. Let us suppose that [+wh] is an optional member of the matrix of any NP (or perhaps argument NP). What does it mean for some NP in position P to be marked [+wh] at D-structure?

It is reasonable to suppose that D-structure indicates the semantic function, the thematic role of the variable ultimately bound by a wh-phrase in COMP (at S-structure if syntactic wh-movement has applied, at LF in any case).

As an illustration, consider the following derivations (D-structure/S-structure):

(11)   i.   He wondered [[you say who]]
        ii. He wondered [who$_i$ [you saw e$_i$]]

(12)   i.   She is asking [[e has been bought what]]
        ii. She is asking [what$_i$ [e$_i$ has been bought e$_i$]]

In both cases, the wh-phrase assumes the θ-role "object of the embedded verb". In (11ii), this is directly represented by the S-structure position of the variable e_i. In (12ii), it is indirectly represented by the coindexing of the variable e_j in subject position with the empty category in object position of buy. Consider the following derivations:

(13)   i.   e wonders [how [who to leave]]
        ii. who$_i$ e$_i$ wonders [how [e$_i$ to leave]]

(14)   i.   he tried [who to leave]
        ii. who$_i$ he$_i$ tried [e$_i$ to leave]

Here, in both cases, the wh-phrase has been moved directly into the matrix COMP position and we have accidentally coindexed it with the matrix subject. The S-structure (13ii) is well-formed, yet the thematic information given, respectively, by (13i) and (13ii) is different. (14ii) illustrates the same point, although it is not grammatical, due to the accidental property of English of not allowing resumptive pronouns in subject position, but the same derivations would be allowed in Italian, where subject pronouns may be left phonologically unrealised, and therefore are not observationally distinguishable from wh-traces (Note incidentally that Bounding Theory for Italian permits this type of wh-movement illustrated in (13) and (14), cf. Rizzi, 1982).

Let us refer to these derivations as extended non-isomorphic derivations.[2]

4. Why are there non isomorphic derivations.

The existence of extended non isomorphic derivations shows that the

mere stipulation that derivations must meet the Isomorphy Principle would not be sufficient to ensure a proper correspondance between levels, since the thematic structure as we have defined it is left invariant by extended non-isomorphic derivations. There are several possible options that would remedy the problem. The first, and, in a sense, most radical would be to assume that there is no independent level of D-structure: The correspondance problem would not arise because the relevant level simply does not exist. This is not the only option, however.

Consider why (extended) non isomorphic derivations arise. First, because we adopt a version of the indexing mechanism that permits accidental coindexing. Secondly, because we allow the identification of (empty or "pronominal") categories to change across a derivation: as can be checked, the essential reason why (extended) non isomorphic derivations exist is the duplication of possible origins for notions operative at S-structure, such as PRO, NP-trace...

Consequently, we might first try to adopt a different version of the indexing mechanism. A full discussion of the possible alternatives would take us too far afield. We may however, sketch what form they would take. Essentially, they would amount to replacing conditions (such as disjoint reference, i.e. principle B of the Binding Theory) on impossible coindexing configurations by rules of coindexing which would be defined to apply only under lawful circumstances. Consider, in particular, the coindexing relation of some pronoun to some other NP. In effect, these alternatives would be equivalent to reinstating pronominalisation rules of the type that were assumed in the Standard Theory model: the environments in which the coindexing rules of the alternatives would be permitted to apply would be exactly those meeting the structural descriptions of the

pronominalisation rules mentioned. So, that one important criticism against this latter approach, namely the observation that it is easier to characterize impossible NP/pronoun coindexing situations than possible configurations carries over to these alternatives (cf. Reinhart, 1976 and the references cited therein).

A more interesting option would assume that A-chains, i.e. the objects to which θ-roles are assigned at S-structure, must be the projection of the actual movement history, as suggested in Chomsky (1981). In fact, such a move is not sufficient, as the existence of extended non isomorphic derivations shows. In (13i) and (13ii), the A-chains are identical. Assuming a stronger hypothesis, namely that both A-chains and A'-chains, to use Chomsky (1981) terminology, must be the projection of the actual movement history, would yield the desired result. However, this hypothesis essentially comes down to giving up any contextual definition for empty categories, which we have so far assumed, and replacing it by a system of intrinsic identification.

One way of getting the same result, i.e. that A-chains are the actual projection of the derivation, is to intrinsically identify the empty categories, e.g. PRO is marked with a feature F_1, NP-trace by a feature F_2... It is clear that transformational mappings from D-structure to S-structure would be, in that case, unambiguously projected onto S-structure A-chains.

Alternatively, we could keep to some kind of contextual identification for empty categories and assume that the proper relation between the levels of D-structure and S-structure is not defined through transformational mappings from one onto the other. Rather, we could assume that D-structure is not an independent level of representation but

rather reflects particular properties encoded at the S-structure level.

There are arguments going in each direction. In favor of the first alternative, we find arguments supporting the existence of an independent level of D-structure. One important argument, presented in Chomsky (1981) is the following observation: if we require that the θ-criterion holds at D-structure, we can adequately constrain the process of θ-inheritance, e.g. θ-roles are always transmitted upward[3]...

In favor of the second alternative, we find a powerful argument derived from Chomsky's functional characterization for empty categories; Empty categories partition differently with respect to different sets of properties: e.g. functional properties, binding properties, referential properties. Clearly a better theory of Empty categories is one which allows one to derive, in a simple way, all these properties on the basis of minimal stipulations.

We will try to show in Chapter IV that context dependent definitions give a superior theory of empty categories than context-free definitions. Of course, if the second alternative is adopted, it must be able to explain or derive the properties of θ-inheritance somehow. We will try to show that this goal is attainable without the help of no otherwise unnecessary principles. This in turn can be seen as an additional argument in favor of the second alternative.

FOOTNOTES CHAPTER I

1. Note that these derivations could be excluded by the Strict Cycle Condition on transformational rules. But cf. Freidin (1978), for discussion of why that condition is not available.

2. Clearly, we might as well construct extended non isomorphic derivations between S-structure and LF, which will have some theoretical bearing cf. Chapter III.7.

3. Note however that this argument in favor of the existence of D-structure is considerably weakened by the fact that it does <u>not</u> follow from the requirement that the θ-criterion holds at D-structure that D-structure is a pure representation of GF-θ: it must further be <u>stipulated</u> that A-chains are singletons.

CHAPTER II: CHAIN THEORY

0. S-structure as the basic level of representation

We now try to develop a different approach to the relation between the levels of D-structure and S-structure. What we basically aim for is to have <u>all properties of D-structure projectable from S-structure</u>.

Pursuing with this assumption, we must define all properties of S-structure directly from S-structure configurations and not derivatively from other levels. Keeping everything else in the model presented in the Introduction (henceforth M) unchanged, let us, for concreteness, assume base generation of S-structure representations. By this, we mean that the set of phrase markers is generated by a set of context free rewriting rules of the familiar kind, probably reducible to a bare minimum given the Projection Principle, X'-theory and perhaps Case Theory along the lines suggested in Chomsky (1981) and Stowell (1981).

A priori, note that no particular loss of generalization arises by taking this step, in particular with respect to the surface distribution of phrasal categories: substitution transformations play no role in this issue, by definition and the theory and effects of adjunction rules (e.g., Baltin, 1978, 1982, landing site theory) could a priori just as well be incorporated in a phrase-structure grammar. Finally, note that the fact that D-structure phrase-markers, in a way, represent X'-theory canonically, may just as well be expressed at S-structure (or LF) by stating that X'-theory is a property of the A-positions network. We furthermore assume that the indexing algorithm is simply:

(1) Index

i.e. is a free procedure, subject to various filtering mechanisms (e.g. the Binding Theory) that we will discuss later.

Let us now turn to the analogue of Move x in this alternative model (call it MA). It is worth stressing that the point here is not to exhibit a model of grammar doing away with movement rules; rather we ask the question of what the proper formalization of movement rules must be. Movement rules play a variety of different roles. Some movement rules to A'-positions, e.g. wh-movement, have as a result the overt expression of the scope properties of the moved phrase, while others do not such as, for example, Heavy-NP shift. Movement to an A-position, on the other hand always expresses θ-role inheritance: the link between a moved category and its furthest trace expresses that the θ-role assumed by the moved category is that assigned to the position occupied by its furthest trace. Before considering the problem of how chains and related notions should be defined, note that there are four types of instances of movement:

(2) i.   Movement from an A'-position to an A'-position
          (e.g. COMP to COMP)
     ii. Movement from an A'-position to an A-position
          (e.g. improper movement)
     iii. Movement from an A-position to an A'-position
          (e.g. wh-movement)
     iv. Movement from an A-position to an A-position
          (e.g. NP-movement)

In the following chapters, we will deal with each of them in turn, starting, in the present chapter, with (2iv), which we take to be the core case of movement. We will therefore develop all the relevant notions for this case and modify, if and when necessary, as we go along to the other types of movement relations. Considering (2iv) further, we may

distinguish among A/A movements between upgrading, i.e. movement to a c-commanding position, e.g. Passive, Raising... and downgrading movements, i.e. movement to a non-c-commanding position illustrated by free inversion in Italian and French Stylistic inversion... Here, we will concentrate on movements of the first type only. As for the second type, we will assume without arguments that introduction of appropriate notations and modifications much along the lines of Chomsky (1981, Chapter 6) can be performed, so that the theory of upgrading movements can be appropriately seen as a subtheory of A/A relations in general.

1. On the theory of chains

1.1. Introduction

Given that the S-structure object playing the relevant role in terms of θ-inheritance is the chain, we may straightforwardly replace the rule schema Move x, as far as NP-movement is concerned, by the S-structure algorithm:

(3)   Build chains

i.e., considering a fully indexed S-structure phrase markers, freely construct a set of objects, each one meeting the relevant properties that chains must possess. This means that structural descriptions at S-structure are made of pairs (x, y), where x is a phrase marker and y is a set of chains associated with x. We will return in the course of the section to the important question of whether y is uniquely determined by

x.

1.2. Some remarks on Chomsky's theory of chains

Let us first consider the existing notions of chains, namely those proposed in Chomsky (1981, Chapter VI.). Chomsky (1981) defines a chain as follows:[1]

(5) $C=(x_1, x_2..., x_n)$ is a chain iff
   i. x_i is an NP
   ii. x_i locally A-binds x_{i+1}
   iii. i>1 (a) x_i is a non pronominal empty category or
            (b) x_i is A-free
   iv. C is maximal, i.e. is not a proper subsequence of a chain
        meeting (i-iii)

Restricting ourselves as we did to upgrading A/A relations comes down to ignore BINDING relations and consider only binding relations. With respect to chains, it means replacing BINDS in (5ii) by binds and dropping (5iii. b), since if x_i is A-free, it is necessarily BOUND but not bound, because of (5ii); hence it is cosuperscripted with a c-commanding element, that is, falls outside our present domain of concern. So the relevant notion of chain reduces to (6):

(6) $C=(x_1,..., x_n)$ is a chain iff
   i. x_i is an NP
   ii. x_i locally A-binds x_{i+1} 1≤i<n
   iii. i>1, x_i is a non-pronominal empty category
   iv. C is maximal

This characterization of chains displays a number of undesirable properties that we now review.

First, recall that a chain is an object to which θ-roles are assigned. The θ-criterion basically requires that any chain contain at

most one argument. If it contains one argument, either it is a PRO, or it is not and must be Case-marked, and in either case must contain one and only one θ-position. If the chain contains no argument, then it may not contain a θ-position.

Note first, as a simple remark, that, since x_i locally A-binds x_{i+1} for $1 \leq i < n$, x_i for $1 \leq i < n$ is by definition in an A-position. This, however, does not say anything about x_n. In fact, x_n could perfectly well be in an A'-position (in the extreme case, we could have a chain with a single member in an A'-position). Recall that a chain is meant, intuitively speaking, to be a S-structure projection of possible NP-movement derivations. Let us simply exclude the possibility mentioned by stipulation, as seems natural. We therefore replace (6i) by x_i is an NP in an A-position and, as corollary, (6ii) by x_i locally binds x_{i+1}, $1 \leq i < n$.

A more substantial remark can be made about clause (6iii). Putting aside the case n=1, where it plays no role, assume n>1. Then, x_i i>1 in C cannot be a variable, because by (6ii), it is locally A-bound (and we admit without discussion that it cannot be both locally A-bound and A'-bound).[2] Therefore, by (6iii), it must be an NP-trace: it is an empty category which is not a variable, i.e. either a pronominal, a possibility ruled out by (6iii), or an NP-trace.

Now, since it is an NP-trace, it is by definition locally A-bound by some phrase with no independent θ-role. Indeed, recall that:

(7)   If x is an empty category in an A-position
   i.   x is a variable iff it is locally A'-bound, otherwise it is an anaphor

   ii. x is a pronominal iff it is free or locally A-bound by y with an independent θ-role

Consequently, if an empty category is neither a variable, nor a

pronominal, it must be locally A-bound by y, y θ-dependent on x, i.e. it is an NP-trace. Let us pause here a moment and examine what <u>independent</u> <u>θ-role</u> in (7ii) means. Although intuitively clear, this notion is not easy to characterize precisely within this framework of assumptions. It seems reasonable to construe it as follows:

(8)   y, locally A-binding x, has an independent θ-role iff it
       is either in θ-position, or does not inherit its θ-role
       through the position x occupies.

Ignoring the conceptual problem first noted by M. Brody (cf. Brody, 1983) underlying (7) and (8) to which we return, and coming back to our main line of discussion, we see that no x_i, $1 \le i < n$, can be in a θ-position without violating (6iii). Otherwise, x_{i+1} would be locally A-bound by an θ-independent item, i.e. it would be a PRO. Assume now that C contains an argument. By the θ-criterion it must contain a θ-position that may only be x_n. So the definition (6) implies that the only θ-position a chain may contain is its most deeply embedded position x_n. However, if a chain C is the projection of the movement history of the phrase marker, movement has taken place from the D-structure position x_n, which must be a θ-position, if C contains an argument, because of the θ-criterion applying at D-structure. In other words, we see that (6iii) seems to be redundant with the property of D-structure to meet the θ-criterion, i.e. to be a pure representation of GF-θ, a consequence which we will elaborate upon shortly.

Is (6iv) also a redundant specification? Assume C' to be a strict subpart of C meeting (6i, ii, iii). Because of (6ii), it must be of the form $C'=(x_r,\ldots, x_t)$, that is, a continuous subpart of C, (admitting that any phrase has at most one local A-binder). Suppose first that C is an

A-chain, i.e. <u>a</u> <u>chain</u> <u>containing</u> <u>an</u> <u>argument</u> (or a Θ-chain in Safir, 1982, terminology). Since, as we have just seen, for $i \neq 1$, x_i, in C is an NP-trace, the argument must be x_i, and we have seen that the (only) Θ-position C (may and) must contain to obey the Θ-criterion is x_n. Therefore if C' contains x_i, it must also contain x_n, so is therefore equal to C.

If there were a principled way to exclude NP-traces from chain initial position, it would follow that x_r of C' must be x_i, i.e. by the above argument, that C'=C.

Suppose now that C is an A'-chain (or Θ'-chain), i.e. not an A-chain (Θ-chain). The principle blocking NP-traces in chain initial position would still imply that $x_r = x_i$. But the rest, i.e. $x_t = x_n$, would not follow.

Pending the introduction of a principled reason why NP-traces cannot head chains, we may derive the maximality requirement for A-chains, but not for A'-chains.

2. Formal properties of chains

2.1. Chains and NP-types

Putting all the above remarks together, and sticking to minimal assumptions, let us assume the following definition for chains:

(9) $C = (x_1, x_2 ..., x_n)$ is a chain iff
   i. x_i is an NP in an A-position
   ii. x_i locally binds x_{i+1}
   iii. x_i is a $\dagger$-category

Recall that ϒ-features are grammatical features of pronouns. We make here a distinction between ϒ categories and empty categories. An empty category is a phonologically zero category A ϒ-category is a category which has no features other than ϒ-features, with the exception of, perhaps, phonological features, and is specified for every ϒ-feature.

This definition would include amongst ϒ-categories wh-traces, NP-traces, PRO, pro, pronouns (resumptive or not), to use the usual terminology. We further specify that a category x binds a category y iff x and y are coindexed, x c-commands y and x and y are identical in terms of ϒ-features (cf. also III.7, for a more detailed presentation). Of course, we also have to modify the partitioning of ϒ-categories. Recall Brody's observation about pronominals. Chomsky (1981) defines a pronominal in terms of the notion of Θ-independence (cf. (7)). Yet, at S-structure, Θ-independence between two categories only means membership to two distinct chains. And chains were defined (cf. (6)) with the use of the notion of pronominal category.

In order to avoid circularity, we have avoided reference to pronominal category in our definition of chains (9). We can now adopt the most natural way to characterize Θ-independency, namely: x and y are said to be Θ-independent from each other iff they belong to different chains.

In particular, we could now define pronominals as ϒ-categories (perhaps of a certain type) in chain initial position (a position that we may also call the head of a chain), and as corollary, we may call non-heads NP-traces. Before exploring the full range of consequences of this move for the classification of NP-categories, let us continue to explore the properties of chains.

Given a chain $C = (x_1, \ldots, x_n)$ it might contain a Case position,

θ-position, an argument. Are there any generalizations about where they might occur, given our definition of chain, beyond the obvious consequence of the θ-criterion that A-chains contain at most one θ-position and at most one argument, and, if it does, at least one θ-position?

2.2. The position GF-θ in a chain

We noted earlier that requiring the θ-criterion to hold at D-structure was, as far as A/A upgrading movement was concerned, implied by the requirment that non-heads of chains be non-pronominal empty categories, in the terminology of Chomsky (1981), which yields the conclusion that only the most deeply embedded position of a chain may be a θ-position. (cf I.fn 3). Is such a property desirable and if yes, do we need to stipulate it (i.e. that the θ-criterion holds at D-structure) or can we derive it from other properties of grammar.

Let $(x_1,..., x_n)$ be an A-chain. To formulate the question above in other words, can we derive that x_n in C is the θ-position C must contain?

Suppose x_p, $p < $ n is the θ-position in C, and consider x_{p+1}. x_{p+1}. cannot be a θ-position because a chain cannot contain two such positions. Therefore it is a θ'-position.

We have noted in the Introduction that only subject positions of a clause or of a gerund are non theta positions.[3]

Furthermore, since it cannot be null because of agreement under local binding, it is a $\dagger$-category. As such, it is what we call an NP-trace, i.e. an anaphor subject to condition A of the Binding Theory and the ECP (or equivalent).

Suppose first it is ungoverned. Then it is excluded by the ECP (or,

alternatively, by the Binding Conditions). Suppose next that it is governed, but only from the inside of the category (or the maximal projection of the category) it is the subject of. Then it is ruled out again by the Binding Condition A: an anaphor must be bound in its governing category (this is essentially the case of an NP-trace in subject position of a tensed clause). Suppose next that it is governed from the outside of the (maximal projection of the) category it is a subject of. This situation only occurs after S'-deletion, and perhaps "small clauses", predicates,[4] i.e. only in the following type of structures (x_{p+1} =NP*):

(10) ...[$_S$ NP [$_{VP}$ V [$_K$ NP*...   K=S or small clause

Nothing blocks such a possibility so far. Let us explore it further.

NP* is a subject θ'-position. Let us examine the kind of predicate NP* could be the subject of. Given that we take weather "it" to be an argument (which is a constant), i.e. that the subject position of weather verbs is considered a θ-position we are left with impersonal subjects of ergative verbs (cf. Burzio, 1981) viz. French (11i), subjects of predicates followed by clausal arguments viz (11ii, 11iii) and, most importantly, subjects of impersonal passives as (11iv) in Dutch (and German):

(11) i.   il est arrivé 3 hommes
      ii.   il semble [$_S$' que Jean soit parti]
      iii. il est clair [$_S$' que tu as faim]
      iv.   er werd gedanst (= it has been danced)

If we somehow could argue that these impersonal subjects are not available as members of the chain C (e.g. because they are members of another chain containing an argument or a θ-position), we would have excluded the possibility under discussion entirely (x_F = a θ-position). A plausible

line of argument would put forth that the impersonal subject must in fact belong to a chain containing the postverbal argument, so we would get the chains (il, 3 hommes) for (11i) and (il, S') for (11ii) and (11iii), either for reason of Case transmission - (11i) and perhaps (11ii) - or perhaps for reasons of θ-role transmission if the argument is in a θ'-position - (11iii). Of course, we would also need some reason why the impersonal subject il cannot belong to two distinct chains, which, in these last two cases, could perhaps follow from the requirement that elements in a chain must be non distinct with respect to ϒ-features.

However, supposing that impersonal subjects, as in (11ii) and (11iii) are necessarily included in a chain with the post-verbal clausal argument poses some problems, as Chomsky (1981,Chapter VI) notes, in view of examples such as:

(12)   i.   It is believed/seems that S   (*...to seem that S)
        ii.   Il faut partir   (*...falloir partir)
        iii. It seems/is believed [John to VP]
        iv.   John$_i$ seems/is believed [e$_i$ to VP]
        v.   My belief [that S]

For it would mean that, if an impersonal subject is available, as in (12i), (12ii) and (12iii), then it must be incorporated in a chain with the post-verbal clausal argument; because furthermore, a chain headed by an NP≠PRO must have Case, this would exclude the examples in parentheses: there is an impersonal subject but it has no Case.

Now, if no impersonal subject is available, as in (12iv) and (12v) (especially (12iv): Stowell, 1981 argues that (12v) does contain an NP heading a chain containing the post nominal clausal arguments), then the clausal argument is not in an NP-headed chain and requires no Case: in other words, it would mean that a clausal argument needs Case only by

virtue of being in an NP-headed chain, a suspicious conclusion. (cf. Stowell, 1981, and Safir, 1982 for opposing views on this matter).

Furthermore, even if such an approach is correct in some form, a possibility we do not dismiss, it would still leave out the cases like (11iv), in which the impersonal subject is not linked to any syntactically represented category. (We return to some important properties of such constructions in 2.6 below).

We must therefore pursue along different lines. Note first that no argument in terms of the Binding Theory could be invoked to rule out chains containing NP* in (10), at least within the assumptions that we have made so far (but cf. 2.3.2 below and IV.2): as non head of chain, NP* would count as an anaphor. It thus suffices that it be bound in its Governing Category, namely S of (10), to avoid a Bounding Theory violation. However, if S in (10) is in the domain in which NP* must be bound, it may only be bound by NP of (10), so that NP locally-binds NP*, i.e. (NP,NP*) is a sub-part of the chain under discussion, and, in particular, NP is in a θ-position by assumption.

One might wonder at this point whether we should try to exclude such A-chains $(x_1, ..., x_n)$ where x_n is not a GF-θ. Some considerations suggest that we should. Consider the following S-structure:

(13) *...how John$_i$ to believe [$_s$ e$_i$ to seem [$_s$· that S]]

where (John$_i$, e$_i$) forms a chain C. If such a chain is permitted, this structure should be well-formed,[5] since C is an A-chain, contains a θ-position (subject of <u>believe</u>) and a Case-marked position (subject of <u>seem</u>) and is therefore visible. The ungrammaticality of (13) suggests therefore that such chains should not be permitted.

Let us, in passing, point out that no argument in favor of a "D-structure as a pure representation of GF-θ" theory could be constructed on the basis of the ungrammaticality of structures like (13). It could be argued that, taking chains to be projections of the movement history and taking D-structure to obey the θ-criterion, such chains as C for (13) could not be constructed. But this seems to me to be inexact, if some further assumptions are not brought into play; there does not seem to be any natural move that would prevent the following derivation: D-structure (14i), intermediate structure (14ii), S-structure (14iii):

(14)   i.   ...how John$_i$ to believe [e$_i$ to seem that S]
        ii. ...how e$_i$ to believe [John$_i$ to seem that S]
        iii. ...how John$_i$ to believe [e$_i$ to seem that S]

which has the same effect as making up the chain C of (13) off S-structure. It is important at this point to recall that the functional characterizations for empty categories proposed in Chomsky (1981) are circular: it could not be argued that e$_i$ of (14iii), having a θ-local antecendent, would be interpreted as a PRO and thus excluded because it is governed.[6]

Returning to our main line of discussion, we see that the type of chains in (10) we want to exclude contains (NP,NP*) as a sub-part, where NP is the subject of V and in a θ-position and NP* is a subject, governed by V. Now, it turns out that NP* must be a Case-position. Indeed, Chomsky (1981, p. 125) points out the following observational generalization (cf. also Burzio, 1981):

(15)   If some NP governed by V is not assigned Case from V,
        then the VP headed by V assigns no θ-role.[7]

Since NP in (10) receives a θ-role, (15) implies that NP* must be a Case

position.

Suppose now we   make the following assumption, that we will justify on independent grounds in the next sections:

(16)   If a chain contains a Case position, it is its head position

What (16) claims is that the presence of a Case position breaks a chain and in particular, that there cannot be Case-inheritance upward.

Given (16), we can actually <u>derive</u> that only x_n may be a θ-position in a chain $C=(x_1,\ldots, x_n)$. It is worth pointing out that this result does not rest on the truth of (15) but rather on that of a weaker generalization, namely that in: (i) NP_1 $[_{VP}$ V $[_S NP_2 \ldots,$ if VP θ-marks NP_1, it must Case mark NP_2 if NP_2 can be a θ'-position. Indeed, if NP_2 is always a θ-position it cannot form a chain with NP_1 We will in fact see that Reflexive Clitic constructions are of the type (i), but with NP_2 always a θ-position.[9]

Tempering this result a little, note that it is achieved at the cost of making the hypothesis (16), and can be considered a real result and not a simple trade-off between stipulations as to the position of Case and position of GF-θ in an A-chain only if (16) is justified by independent considerations (which, we will see it is).

Note finally that (16) cannot be derived from an assumption that is sometimes made, that Case-marked non overt categories must be variables: first because of the existence of <u>pro</u>; secondly, because it would not have to be non overt (since it is Case-marked: cf. next section). Note also that (16) is a strong version of the notion of chain-internal Case conflict, since it implies that a chain can contain at most one Case position, (given that a chain has at most one head and if every NP belongs

to one and only one chain).

It is worth noticing in particular that the notion of chain internal
Case conflict (perhaps implied by Binding if Case is a Υ-feature, given
non distinctness of elements belonging to the same chain) is not
sufficient to exclude the chains illustrated in (10), as the following
structures show:

(17)   i. * It is unclear how John to believe [NP* to seem that S]
        ii. I expect John to believe [NP* to seem that S]

Indeed in (17i), the chain (John, NP*) receives only one Case, and in
(17ii), it receives the same Case twice (i.e. the same type of Case: it
could be argued that two token Case count for ruling out (17ii)). Note
also that examples similar to (17) can be constructed with an argumentless
predicate in the embedded clause, e.g. impersonal passives, avoiding the
possible necessity to form a chain (NP*, that S).

2.3. The Case position in a chain

2.3.1 Let us now turn to a more systematic examination of the Case
positions inside a chain, and start with the question: can we determine
where is/are the Case positions in a Case-marked chain C. Let (NP_1, NP_2)
be two consecutive positions of C, i.e. what Chomsky (1981) calls a link
and let us ask the question of whether NP_2 can be a Case-position. First,
we may assume that NP_1 is necessarily subject of VP[10] for

        a.   either NP_2 is a θ-position, in which case NP_1 must be
             a θ-position by the θ-criterion, i.e. an [NP, S]

    b.   or NP_2 is not in a θ-position in which case
      i. either NP_1 is not in a θ-position, hence is an [NP, S]
      ii. or NP_1 is in a θ-position, in which case
        we are in the type of configuration by (10) and
        NP_1 is an [NP, S].

Secondly, we may run through all the possible structural configurations according to whether NP_1 and NP_2 belong to all the same clausal structures (we assume gerunds to have a clausal internal structure - say to be dominated by S, for concreteness -) or not, i.e. whether NP_1 and NP_2 are dominated by all and only the same S's or not. Suppose first that they do not. By assumption NP_2 is an NP-trace, thus an anaphor which must be bound in its governing category. Before pursuing, let us note that the definition of governing category that we adopt for NP-traces, which we will justify in Chapter IV is essentially as follows:

(18)   A governing category for x, x an NP trace, is the
          minimal NP or S containing x and a governor of x.

Given this definition, it is clear that NP_2 may only be the subject of a non-finite clause. In particular, it cannot be the subject of a tensed clause, which excludes NP-movement cases such as:

(19) $John_i$ seems that NP_i* left

Where NP* is the trace of John. Clearly, if NP* is an empty trace, this is no surprise because, for example, of the ECP. However, we hold this is true regardless of the realization of NP* and in particular, it is excluded even if NP* is realized as the analogue for NP-traces of resumptive pronouns for wh-movement.

2.3.2. Let us briefly digress on this question. There has been some

discussion in the literature (cf. e.g. Koopman, 1980, on Vata, Ingria, 1982, on Greek...) on whether such types of chains should be excluded. Without going into details, we see that they are excluded by the Binding Theory on the assumption that NP-traces are always anaphors, or more precisely, that non-chain-initial elements are always subject to Principle A of Binding Theory. Suppose however, that there are not, contrary to what we have assumed; suppose furthermore, that we allow Case-marked NP-traces and that, precisely when they are Case-marked, they are not subject to Principle A of the Binding Theory but rather to Principle B (i.e. they must be free in their Governing Category). Then, such NP-traces would act both as "θ-transmitters", being inside a chain, and as "pronouns" in that they would have to be locally free. Under such assumptions, structures like (19) should be allowed, and more generally chains (NP_1, NP_2) where NP_2 is Case-marked and NP_1 not in its Governing Category. This would allow an equivalent of unbounded NP-movement [in fact, in cases of Vata similar to (19), this is exactly what we find cf Sportiche. (forthc b)]. Of course, if such a possibility was correct we would have to modify our discussion of section 2.2. accordingly.

Although very appealing, these assumptions raise some questions. First, why would such structures as (19) be ungrammatical in English? In languages like English lacking overt resumptive elements, we could argue that NP* has to be empty and thus violates the ECP (same if NP* is not in subject position). However, in languages like Italian, NP* could simply be pro, permitted by the ECP to occur in subject position of tensed clauses. However, structures like (19) in Italian are ungrammatical. Secondly, as D. Pesetsky points out, such constructions as (19) when they seem to occur in some language (e.g. Vata) systematically fail to permit

"Raising" of idiom chunks and other non-arguments, i.e. the analogue of (20):

(20)   *[Close tabs]$_{NP}$ seem that NP* have been kept on John

Notice that (20) could not be excluded by some principle ruling out co-reference of pronouns and idiom chunks in the cases where NP* is physically realized as a pronoun, because NP*, being a member of the chain (NP,NP*) not in head position would <u>not</u> be a pronoun.

We will therefore assume for the time being that structures like (19) are indeed excluded with (John,NP*) forming a chain, although we recognize it is on rather narrow grounds that this conclusion is based. (cf. Sportiche, fortcoming b, for additional discussion).

2.3.3 Returning to our main line of discussion, we see that NP$_2$ cannot be either inside the subject of a tensed clause, as is permitted for lexical anaphors, viz (21):

(21) a. They think that pictures of each other are on sale
     b. *They seem that NP* left

where <u>they</u>= NP$_1$ and NP*=NP$_2$, again because of a Binding Theory violation if NP* is non zero, or both for reasons of Binding and for the fact that NP* would not be properly governed, hence would violate the ECP, if NP* is phonetically null. For the same reasons NP$_2$ cannot be the subject of a gerund nor that of a regular NP.

So, as far as cases in which NP$_1$ and NP$_2$ do not belong to all the same clausal structures, only cases of type (22) are not ruled out independently:

(22) NP_1 $[_{VP}$ V $[_S$ NP_2 [-tense]...]]

Consider next the cases in which NP_1 and NP_2 do belong to all and only the same clausal structures. Then NP_1 is the subject of a VP containing NP_2, i.e. we are dealing with the following type of structural configuration:

(23) NP_1 $[_{VP}$ V...NP_2...]   where...contains no S boundary

Suppose first that NP_2 is Case marked by V. It could be argued that such a case cannot arise by using the following line of reasoning. First we might invoke the converse of observation (15), stated in Chomsky (1981, p. 113) as in (24) (cf. also Burzio, 1981):

(24)   A verb that assigns Case to the NP it governs assigns
          a θ-role to the subject of its VP

If (24) is correct, (which would not be if the analysis of verbs like impress in Chomsky, 1981, is (but cf. fn. 9), NP would be θ-position, just as NP_2 , which is not a subject by assumption, so that the link (NP_1, NP_2) could be part of a well-formed chain.

So, besides case like (22), we are left with cases like (23) in which NP_2 is not governed by V (i.e. is not an object of V or of a reanalysed P).

We now turn to such cases. As we shall see, one plausible principle that we will invoke for these cases carries over to the full range of structures that we have examined in this section.

Let us first consider abstractly what the relevant structures must be like. They must meet the description (25):

(25) NP_1 INFL $[_{VP}$ V (NP)...$[_{PP}$ P NP_2]...

Where NP_1 is a θ'-position (since NP_2 is a θ-position). What prevents the making up of a chain (NP_1, NP_2) ?

Suppose first that NP_2 is empty. Then we may for example appeal to the Empty Category Principle, as Kayne (1981) suggests, i.e. assume that P's are not proper governors.

Suppose however, that NP_2 is not empty. Then, this configuration is not excluded by the ECP, for it is one of the essential properties of the ECP that it only constrains the distribution of empty categories (cf. Koopman, 1980, 1982, Aoun, 1980, for important arguments to that effect).

This configuration could not be ruled out by Case-conflict either, since NP_1 is not necessarily a Case position. Nor could it be a condition to the effect that a token Case can only be realized onto one NP at most, for NP_1 is not necessarily lexical, and as such, does not necessarily need Case. In fact, if we make NP_1 =PRO in an ungoverned context, i.e. as in (26):

(26)   ...[$_{S'}$ why [$_S$ PRO [$_{VP}$ V [$_{PP}$ P NP]...

Nothing excludes the chain (PRO, NP_2) if NP_2 is a resumptive "NP-trace". Note in particular that it could not be Bounding theory, for NP_2 is subjacent to PRO. Let us now exhibit such cases. Consider (27):

(27)   Il a été tiré sur Jean
            It has been shot at Jean

Note first that there is no reanalysis of the sequence _tirer_ _sur_, as the pied-piping of the preposition under wh-movement indicates (cf. Pollock, 1979; 1981 for extensive discussion of these constructions[11]). We can now sum up our argument by claiming that the theory of grammar, as it stands, does not exclude all of the ungrammatical structures in 28:

(28)   i   ...*Pourquoi [$_{NP}$ Jean]$_i$ a été tiré sur NP$_i$*
       ii ...*Pourquoi [$_{NP}$ Jean]$_i$ etre tiré sur NP$_i$*
       iii ...*Pourquoi [$_{NP}$ PRO]$_i$ etre tiré sur NP$_i$*

where NP* is a lexicalized trace of the NP in subject position, i.e. where (NP$_i$, NP*$_i$) forms a chain. We therefore need to introduce some principle to exclude these chains. One possibility would be to adopt a (modified version of a) generalization offered in Hornstein & Weinberg (1981) namely the Filter:[12]

(29)   *[trace]
             +oblique

(where [+oblique] simply means that the P assigning Case has not been reanalyzed). This is even more plausible if the criticisms of the analysis of P-stranding in terms of the ECP as proposed by Kayne (1981) offered recently (cf. Huang, 1982a; Aoun, 1982) stand. (29) would apply to traces of NP whether lexicalized or not, and similarly to traces of movement to an A'-position. However, this approach seems to make the wrong predictions in some cases of lexicalized wh-traces. For example in Haïtian (cf. Koopman, 1982a) wh-movement obeys the usual constraint on movement, yet the extraction of the object of a P (P's are never reanalyzed) is only possible if the wh-trace is lexicalized, i.e. is a resumptive pronoun. This type of construction would be incorrectly excluded by (29) (cf. III.5 for additional discussion).

So, let us propose instead that the following clause be included in the definition of a chain:

(30)   if C=(x_1,..., x_n) is a chain x_i, i>1 is not a Case position.

Of course, the underlying assumption here is that there is a direct

relation between the ability for NP to be lexicalized and to be Case-marked: this is essentially the content of the Case Filter.

(30) is what we have listed as (16) in the preceding section. As we have said, if (16) was not to be a simple trade off between stipulations as to where GF-θ or the Case position is in a chain, some independent motivation had to be provided for it. As is now clear, we see that (16) = (30) is indeed independently motivated by its ability to rule out structures like (25) (or (28)), which cannot be ruled out by stipulating that GF-θ is the most embedded position of a chain.

As (28iii) makes clear, the type of configuration that we want to rule out does not depend on Case conflict, or on Case-inheritance upward. However, the issue of Case inheritance downward does arise: for (30) can have the desired effects of ruling out overt NP-traces only if they cannot inherit Case from the head position of the chain.

Note also that (30), now incorporated in the definition of chain, covers all the cases discussed in this section (2.3) so that we need not rely on the truth of (24) nor on the particular arguments in terms of the Binding Theory that we have presented.

Concluding this section, let us remark that the necessity of (30) does not depend on our basic assumption that properties of D-structure are projected from S-structure. The standard theory must also prevent generation of examples like (28). The net effect, however, is that from (30), we can derive the essential property of D-structure: θ-inheritance is upward.

We have also adopted that NP-traces cannot inherit Case from the top down, again, clearly an independently necessary assumption.

2.4. Position of the argument in a chain

2.4.1. So far, we have tried to answer the questions of where the Case-positions and θ-positions should be located in a well-formed chain. We now turn to the third question we started with: if a chain contains an argument, i.e. is an A-chain, is there any generalization as to which position this argument occupies.

Suppose first that the chain contains only γ empty categories. Then its head is, by definition, PRO which, because of the θ-criterion can only be an argument PRO, and the other categories if any, are NP-traces, i.e. by definition non arguments.

Suppose next the chain contains a lexical argument. By the very definition of chain that we have given, and more particularly clause (9iii), this argument can only occupy the head position x_1, of a chain $C=(x_1,\ldots, x_n)$. This conclusion only holds by definition. We might wonder whether it may follow from somewhere.

Recall that we have argued that only x_1 may be a Case position, and only x_n) may be a θ-position. If NP* is a lexical argument, the Case Filter will require of NP* to be Case-marked (alternatively the θ-criterion, if the Case Filter can be derived from the θ-criterion). We have furthermore proposed in the preceding section that NP-traces could not inherit Case from some other position inside a chain. Suppose we generalize this to prevent any sort of Case inheritance inside a chain. If so, it follows that NP* must be in a Case position in order to conform to the Case Filter, i.e. must be in chain initial position.

In other words, if we can maintain this strong version of non-Case Inheritance, we may derive the distribution of arguments with respect to

chains.

2.5. Maximality and Uniqueness of Chains

Since we have modified the way chains are defined since we last made the argument, we might wonder whether A/A upgrading chains must be maximal, and, as a corollary question, whether a given NP may belong to two distinct chains.

An examination of the argument given in section 1.2 above reveals that it remains intact despite the modifications that we have introduced: it still follows that A-chains are maximal, although not necessarily A'-chains, since our argument was based on x_n being a θ-position in a chain $C=(x_1,..., x_n)$ and on the fact that x_1 could not be an NP-trace. We can now make this last statement clearer. Let $C=(x_1,..., x_n)$ be a well formed chain and $C'=(x_r,..., x_p)$ be a strict subchain of C. We must show that $x_r = x_1$ (hence $x_p = x_n$ for A-chains, since x_1 hosts an argument and x_n is the only θ-position in C).

Assume that $r \neq 1$, then x_r is empty (since it is not in a Case position) and is an NP-trace in C. But, as head of chain C', it is a PRO in C'.[13] If NP-traces and PRO's have disjoint distribution, we would get a contradiction. Indeed, if the observation made in Chomsky (1981, Chapter 6) is correct, NP-traces must be governed (in fact, properly), while PRO must be ungoverned. So r=1.

In conclusion, A-chains are maximal: there cannot be a well-formed strict subchain of a well-formed A-chain. We might wonder now whether chains are allowed to overlap, or equivalently whether a given NP may belong to more than one chain.

Clearly, every NP belongs to at least one chain (i.e. the chain with itself as a single member, if no other alternative is possible) in a well-formed structure. Can a given NP belong to two (or more) distinct chains.

Let $C=(x_1,\ldots, x_n)$ and $C'=(y_1,\ldots, y_m)$ and assume $x_r = y_q$, $r<q$. We first show that $x_1 = y_1$. Suppose $r\neq 1$:

If $q=1$, $y_q = x_r$ is empty, and is both a PRO in C' and an NP-trace in C, which is impossible. So $q\neq 1$. Now both x_{r-1} and y_{q-1} locally bind both x_r and y_q and are in θ'-positions. However θ'-positions are subject positions and two subject positions can locally bind a third same position only if they are identical. So $x_{r-1} = y_{q-1}$. By iteration, we get $x_{r-2} = y_{q-2}$ and $x_1 = y_{q-r+1}$. If $q-r+1\neq 1$, we get a contradiction, for x_1 must be empty, just as y_{q-r+1}, and be both a PRO (head of C), since $x_1 = y_{q-r+1}$ is empty, and an NP-trace (non head of C'). So $q-r+1=1$, i.e. $q=r$. In other words, $x_1 = y\underline{1}$. So far, we were able to show that two chains which overlap at some point must be identical above this point. Let us now consider the problem of whether it is also true below it. Consider now the links (x_r, x_{r+1}), (y_r, y_{r+1}), where $y_r = x_r$ and we may assume that r is maximal (i.e., is the greatest such that $i<r \rightarrow x_i = y_i$). Assume first that neither x_{r+1}, nor y_{r+1} is a θ-position. Then x_{r+1} and y_{r+1} are both subject positions, both locally A-bound by $x_r = y_r$. Given that both x_{r+1} and y_{r+1}, as NP-traces, must be bound in their governing category (cf. (18)), they both must be governed from the outside of the category they are subject of (otherwise their governing category would include x_r). They may therefore be subjects of S'-deletion complements but not of small clauses, since a subject of a small clause is always a θ-position. If such a situation may arise, we cannot go further. Consider the archetypal

situation illustrated below:

(31) i. $[_{NP}\ x_r\ ]$ INFL $[_{VP}\ V...[\ [_{NP}\ x_{r+1}]\ ...]_S\ [\ [_{NP}\ y_{r+1}]..]..]$

Certainly, there is no predicate selecting two S'-deletion clauses. We have no explanation for this observation. It is certainly related to the observation that no predicate selects two clausal arguments, a fact that might follow from a substantive theory of thematic structure.[14] So let us assume that structures like (31i) are principally excluded. With respect to A'-chains, which may contain no θ-position, the above argument allows us to conclude that overlapping implies inclusion. Indeed, in order to escape a contradiction, we must have either $x_r = x_n$ or $y_r = y_m$, i.e. one of C and C' is included in the other With respect to A-chains, we may conclude that either x_{r+1} or y_{r+1} is a θ-position, i.e. is respectively either x_n or y_m. Say $x_{r+1} = x\underline{n}$ (i.e. n<m). Before exploring further, note that we may, without loss of generality assume n=m=2 (if n>2 or m>2, it simply adds more intermediate traces). In other words, we may assume that $C=(x_1,\ x_2)$, $C'=(y_1,\ y_2)$, $x_1 = y_1$, $x_2 \neq y_2$. Let us now examine what kind of structure may meet this description. $x_1 = y_1$, being θ'-position is a subject, locally binding both x_2 and y_2. So either x_2 and y_2 belong to the same clause as x, or they do not. Furthermore, there must be no c-command relation between y_2 and x_2. So x_2 and y_2 cannot belong to the same clause as x_1: the only possibility would be for them to be inside NP's or PP's (in order that none c-command the other), but they would then be in Case-marking positions, which is excluded for NP-traces by (30) (cf IV.3). So x_2 and y_2 do not both belong to the same clause as x_1. But, for reasons of Binding, x_1's clause must be their governing category. Therefore, either x_2 belongs to the same clause as x_1 and y_2 is the

subject (anything else is excluded by Binding Theory) of an S'-deletion complement clause or a small clause, or both x_2 and y_2 satisfy this last condition. We illustrate the only possibility below:

(31)   ii. $[_S [_N P\ x_1]\ INFL\ [_{VP}\ V\ [_{NP}\ x_2]\ ]\ [[_{NP}\ y_2]\ ...]]$
        iii. $[_S [_N P\ x_1]\ INFL\ [_{VP}\ V\ ]\ [_P [_{NP}\ x_2]...]\ [_K [_{NP}\ y_2]\ ...]]$

where $\underline{K}$ and $\underline{P}$ are small clauses or S'-deletion clauses. Note that in (31iii), x_2 must be a Caseless position, and K must hang from S, otherwise y_2 is c-commanded by x_2. Similarly, in (31iii), both y2 and x_2 must be Caseless positions, and furthermore, since no predicate selects for two clausal complements, at least one of K and P hangs from S, say K. It is reasonable to assume that K cannot be an S'-deletion clause, for this property of clauses is selected by lexical heads, and K, hanging from S, is not lexically selected. So K is a small clause. If such a configuration is possible, there does not seem to me to be any non ad-hoc way of ruling out the chains under discussion (It could not be ECP, for, e.g. in (31ii), y_2 is properly governed and coindexed with x_2 - the configuration x_1 /y_2 is similar to that of who/t in who t left). But if such chains are allowed, it means that there are parasitic gap structures with respect to NP-movement in certain (restricted) environments: both x_2 and y_2 are interpreted as traces of x_1. Note that if the θ-criterion were defined on NP's rather than on chains, this configuration would be excluded since x_1 would assume two θ-roles: one transmitted by x_2 and one transmitted by y_2. Such an argument is not possible if the θ-criterion holds of chains, for the two chains C and C' are distinct objects, each assigned one and only one θ-role.

There might be, however, a principled reason why such structures as (31ii) and (31iii) do not arise. One plausible reason could be as follows

(cf. also footnote 14). Assume that there are in fact two types of small clauses: opaque small clauses and transparent small clauses (the analogue of S' clauses and S'-deletion clauses) and suppose furthermore that small clauses may become transparent only in the same (type of) contexts S'-clauses may: it would follow that transparent small clauses, having the same distribution as S'-deletion clauses, would only appear in lexically governed contexts. So that K in (31) will always be an opaque small clause, i.e. a governing category for y_2 if there is one, which yields a Binding Theory violation. If on the other hand, y_2 is not governed, hence without governing category, we get an ECP violation.

We may therefore conclude that the situation described is not possible, hence, there cannot be well-formed chains of the type C and C'. Consequently, there cannot be two overlapping A-chains which are not identical. In other words: Every NP in an A-chain belongs to one and only one A-chain.

However, as we have seen above, we cannot extend this conclusion to all NPs. We could only show that, if two A'-chains overlap, one is included in the other and they share heads, but we cannot derive that they are identical. Thus, given an A'-chain like C in (32i), chains such as those of (32ii) may be postulated without violation:

(32)   i. $C = (x_1, x_2 ..., x_n)$
         ii. $C = (x_1, ..., x_{n-1})$
              $C = (x_1, x_2) ...$

This negative result seems to suggest that we are perhaps forced to modify our assumption that chain-building is free. In fact, if we examine both the argument of A-chain maximality and the argument that A-chains are non overlapping, we see that we make use of an assumption that is dubious,

namely that PRO and NP-traces are in complementary distribution. More precisely, although it might be true that argument PRO and NP-trace have disjoint distributions at S-structure, it is not clear that we can use this assumption for our argumentation unless we further assume that argumenthood is contextually determined. Indeed, consider a chain $C=(x_1,\ldots, x_n)$; let x_p, $p\neq1$, $p\neq n$, a member of C. x_p in C is an NP-trace hence a non-argument. Considering another chain $C'=(x_p,\ldots)$, x_p in C' is PRO iff it is an argument: argumenthood must be contextually defined. It is unclear however, why x_p could not be a non-argument in chain initial position especially if C' contains no θ-position (i.e. does not contain x_n). This shows that the argumentation would be correct only if NP-traces and these expletive Caseless elements have a disjoint distribution, an unsupported assumption. Consequently, I will assume that the simplest possible algorithm of chain formation that we have adopted is inadequate and should be replaced by:

> Principle of chain formation:
> Partition the set of NP's into chains

It is clear that this principle will have all the right consequences. It not only implies maximality for A-chains, but also guarantees that every NP is in one chain and one chain only.[15]

2.6. Empty Categories and Category Classification.

We have tentatively put forth in 2.1 above a characterization of empty categories which established a basic dichotomy between heads of chains and non heads of chains (NP-traces). We now proceed to a preliminary examination of category classification, reserving a fuller

discussion for Chapter IV. We wish simply to establish the nature of some of the principles governing the distribution of such categories such as The Case Filter, etc.

2.6.1. Heads of Chains, Case Filter and Visibility

Let us first explore the various types of elements that may appear in chain initial position. There are a number of dimensions along which such elements may be classified, e.g. phonetic/non-phonetic, argument/non-argument that we might assume are randomly assigned to certain categories, part of their lexical entry for others. Recall the distinction introduced in Chomsky (1981) between γ-features and the complementary set, call it γ'-features. We define γ-categories as categories which are specified for every γ-feature and only γ-features (except phonological and Case features. perhaps). As a first approximation, we may take γ-features to include [x person], [y number], [z gender], and perhaps [± animate], γ'-categories are categories which are not γ-categories. Intuitively speaking, γ'-categories posses some idiosyncatic properties expressed by some feature F, F not in γ. The γ'-γ content of these categories must somehow be recoverable and we can a priori divide γ'-categories into two different classes depending on the way this content is recovered from the string.

We first find γ'-categories whose content is recovered from themselves. Consider for example, an NP like <u>John</u>. It is recognized as being an argument referring to some particular individual through the physical realization of its phonetic matrix [jan], to which all these properties are (perhaps not uniquely because of homonyms) associated.

We find next, †'-categories whose content is not recoverable from their physical shape: what we might call †'-empty categories. It might be that we find empty categories physically unrealized which contain some †'-features. In such a case, some signalling item (grammatical morpheme or otherwise) must be present in the string that makes these properties recoverable. I argue in Sportiche (1982) that there might be such a case in Bete. Clearly, we must find some principled criterion limiting the set of possible †'-features recoverable "externally", once we have a better idea of the inventory of such items, if they exist. Because it is also conceivable that, as a matter of principle, phonetically zero categories cannot be †'-categories.

We can bring out the distinction from a different point of view. We could term the first type of †'-categories lexical, in that the information their lexical entry contain suffices to determine their semantic and grammatical import. And we could term the second type grammatical, in that their content can be recovered in a construction, through the presence of some grammatical morpheme. As a concrete example, consider the following French form:

(33)   Les enfants se rasent [$_{NP*}$ e]

and assume that NP* is an argument, for the sake of this illustration. How does the reflexive or the reciprocal meaning arise? Is it a property of NP$_i$, signalled by the presence of se (on top of se other properties)? There is some arbitrariness to any decision. We may assume the possibility just outlined. Alternatively, we may exclude empty †'-categories and introduce a specific rule of interpretation in LF that makes NP* into a reciprocal or reflexive phrase at LF, when it is linked

with se

Consider next Υ-categories. We propose provisionally that the relevant dimensions along which they should be classified are [±argument], [±phonetic], [±Case] and [±locally A'-bound].[14] It is worth pausing on any particular choice of classifying features and wondering whether there is some naturalness to the system adopted. There are, of course, a posterio justifications: simplicity of the classification, exhaustiveness... We shall also see that the system of locally A'-bound categories is not symmetrical to that of non-locally A'-bound categories, which requires the introduction of a distinguishing feature [±locally A'-bound]. But one would like to offer reasons why it is these particular dimensions that are the relevant dimensions.

There is also the question of the level at which this particular classification applies. In fact, it could either be S-structure or LF. Apart from A'-binding that may arise as a result of some LF rule (e.g. Quantifier Raising), the other features do not change between the two levels.

2.6.1.1 Let us first briefly consider Υ'-categories.

We may organize the a priori possibilities in the following table:

(34)   +loc A'-bound   (i)

        -loc A'-bound [+Case] +Ph +θ (ii) [-Case] +Ph +θ (v)

                                    -θ (iii)   -θ (vi)

                            -Ph +θ (iv)   -Ph +θ (viii)

                                    -θ (v)   -θ (xi)

As we stated above, whether Υ'-empty categories exist or not depends on

what appears to be an arbitrary decision. Since in any case, we have little to say on this topic, let us assume they do not. So we may ignore (34iv, v, viii, ix).

Similarly, we have little to say about (34i). There are clearly semantic constraints on the possible binding of an NP by an A'-binder: for example, if the A'-binder is an operator defining a domain over which the NP it binds must range the bound NP internal semantics must allow ranging, so a constant, for example is not allowed. On the other hand, this is not clearly true for A'-binders which are not operators, such as, perhaps, clitics, since they seem to allow binding of names, or cons.ants (e.g. Clitic doubling).

Let us turn to the other cases (34ii) is the usual case of a lexical argument, e.g. John, table or idiom chunks. They must have Case because of the Case Filter, or, equivalently because of the LF-visibility requirement on θ-assignment, which excludes their Caseless counterpart, namely (34vi).

We are left with (34ii) and (34vii). These are phonetic non arguments. English there and similar items might be a priori plausible candidates, but, again, we might attribute their properties -here the existential reading- to the nature of the syntactic constructions in which they appear, rather than to themselves. That is what we shall assume here, noting further that it makes all ɤ'-categories arguments, a very natural conclusion that we might formulate as:

(35)   Only ɤ-categories may be non-arguments.

Indeed, we have first eliminated ɤ'-empty categories, then ɤ'-categories which are no., arguments.

2.6.1.2. We now turn to ¥-categories. Again, listing all the possible combinations but restricting ourselves to arguments, we get the following array:

(36)

	+locally A'-bound		-locally A'-bound		
	+Case	-Case	+Case	-Case	
+Ph	(i)	(iii)	(v)	(vii)	+Ph
-Ph	(ii)	(iv)	(vi)	(viii)	-Ph

Of these eight possible elements, three do not seem to exist. Let us first run through the various existing elements.

(36i) are resumptive pronouns. Note that they may come in two distinct varieties: they may be A'-bound only at LF, or both at LF and S-structure (in the first case, they would fall under (36v) at S-structure).[17] (36ii) is the non-overt counterpart to resumptive pronouns. They may also fall in the two distinct categories we just mentioned. They are exemplified by e.g. "wh-traces".

(36v) and (36vi) are their non locally A'-bound counterparts, namely lexicalized pronouns (i.e. regular pronouns) and empty pronouns (i.e. pro of Chomsky, 1982, or empty pronouns of Rizzi, 1982,IV) In (36viii), we get PRO. Note that its Case-marked parallel is pro. and not some kind of Case-marked PRO that would have the distributional properties of PRO.[18]

Let us now turn to the gaps: (36iii), (36iv), (36vii). Let us first recall Chomsky (1981) proposal, that can be summarized as follows:

(37)   i.   A chain is assigned a θ-role only if it is visible

        ii. A chain is visible iff its head
            is visible[19]

        iii. An NP is visible iff it is PRO or
                     in a Case marking position[20]

iv. Locally A'-bound elements are arguments

The idea of this proposal is to reduce the Case Filter to the θ-criterion by requiring of arguments (≠PRO) that they be Case-marked (or more precisely in Case-marked chains) in order to be θ-marked.

The gaps (36iii, iv) follow: As locally A'-bound elements, they are arguments (by (37iv). Hence, by the θ-criterion, they must receive a θ-role (more precisely, must be in a θ-marked chain). Therefore they must be visible, hence must be in a Case position (by 37ii, iii): As Caseless arguments ≠PRO, (36iii) and (36iv) are excluded.

The non-existence of (36vii) follows also from the same argument. Indeed, if such an element existed, it would be a phonetic equivalent to PRO with the same distribution and, we may assume, the same range of interpretation, since it is true of all pairs +Phonetic/-Phonetic, e.g. pronoun/pro, resumptive pronouns/variables. Call it PRO*. PRO*, as argument distinct from PRO must get Case. So it is ruled out. Note however, that this result is in fact stipulative. Indeed, in order to distinguish PRO and PRO*, which would by assumption, satisfy the same definition, a clause excluding the [+Phonetic] version from consideration must be added. There does not appear to be any principled reason why (37iii) could not be generalized to make both PRO and PRO* visible. We certainly would like a more principled account of this state of affairs. The problem arises from the fact that the systems of arguments in head of chain position is not symmetrical, as the following summary table shows:

(38)

	+locally A'-bound			-locally A'-bound	
	+Case	-Case		+Case	-Case
	resumptive pronouns	*	+Ph	pronouns	*
	variables	*	-Ph	pro	PRO

i.e. there is no locally A'-bound Caseless empty element, which would be the A'-bound counterpart of PRO.[21]

So far, we have only looked at argument ϒ-categories in chain-initial position. In order to get a complete picture of the various possibilities instantiated, let us turn to non arguments. Again, we would get a table similar to (36).

Corresponding to (36i) through (36iv), we get no non arguments. This follows directly from (37iv). As we shall see momentarily, the system is entirely asymmetrical in that we do find non-locally A'-bound elements in chain initial position which are non-arguments so that this discrepancy obviously requires some stipulation like (37iv). Note that (37iv) should in fact be slightly modified or further specified. Indeed, we have distinguished between severals types of arguments:

> -Regular arguments that meet the selectional restrictions of θ-roles allowing for a range of θ-bearer e.g. the predicate "be nice" may assign its subject θ-role to anything that can be nice: call these arguments.

> -What we may call constant arguments or quasi-arguments, i.e. arguments meeting the selectional restriction of predicates which may assign their θ-role only to a particular individual argument. For example, the "it" of "it rains", certain idiom chunks. These differ both from regular arguments and from non-arguments in their syntactic behaviour.[22]

> -Finally non arguments.

If we consider A'-binding, we might subdivide it into operator binding and

non operator binding. It is clear that, as far as operator binding is concerned, (37iv) is trivially true. By definition, operators either select an argument in a given non trivial domain, or require their bindee to range over some domain, so that an operator bound NP must receive a θ-role allowing an infinite range of θ-bearers, i.e. must be a regular argument. Otherwise the resulting sentence is semantically ill-formed. (We will slightly modify this conclusion in IV.6). In the case of non operator A'-binding, the matter is more complex, for it depends on the particular semantic analysis of the binding involved. We may assume here that they parallel operators in the relevant respect, so that (37iv) appears to be a very natural principle.

Corresponding to (36v) through (36viii), we get a more interesting situation. If non arguments corresponding to the first two cases can be easily illustrated:

(39)　　i.　he left　　　　　　(36v, "he" argument)
　　　　ii.　it seems that S　　(36v, "it" non argument)
　　　　iii.　[$_{NP}$ e] ha mangiato　(36vi, argument)
　　　　iv.　[$_{NP}$ e] sembra che S　(36vi, non argument)

The matter is not as easily settled concerning the last two cases. Consider the various types of non arguments that one may find in chain initial position:

(40)　　i.　il est arrivé trois hommes
　　　　ii.　there have arrived may people
　　　　iii.　it seems that S

How can we tell whether these elements, e.g. il, there, it require Case or not, and whether they have non-overt counterparts or not. A priori, it would seem sufficient to examine structures in which they occur in Caseless positions. But in fact, it is not.

Let us start with there. As a non argument, there may appear in θ'-positions only, i.e. subject position of clauses. Furthermore, due to locality requirements that must hold of the relation between there and the NP it quantifies (call it NP*), the possible configurations in which there may appear reduce to (41):

(41)   ...[there V NP*...]
          where no NP ≠ "there" c-commands NP* in S and
          V governs NP*

(In fact, there are also configurations with a "trace" of there in the position of there in (41), which have the same properties).

Now, as we mentioned, it is no obvious that there has to have Case. If NP* requires Case, it could be argued that there does not need Case, but must "transmit" Case to NP* somehow.

Suppose first that NP* does not need Case. Because it is an argument (there only quantifies over arguments) it must be PRO given (38). However, this is exluded since NP* is in a governed position.[23] Suppose next NP* needs Case. If it is assumed not to be in a Case position, we get the usual situation with there in a Case position transmitting its Case to NP*.[24]

Suppose finally that NP* is in a Case position. This is the crucial situation for, if there needs no Case, it should be able to appear in a Caseless position. Otherwise it should not. Given (35), the test situation is one in which some verb V assigns Case to the NP it governs call it NP*, but assigns no θ-role to its subject so as to allow the occurrence of there (or of its trace). But this is in contradiction with generalization (24). If (24) is correct, we must look elsewhere for an argument. But suppose (24) is incorrect, an assumption we may put forward

since we do not rely on (24)'s truth. Then, ergative verbs - in the terminology of Burzio (1981) - are good candidates to consider since they are transitive, with no subject θ-role. If they are, the following structures show that <u>there</u> needs Case.[25]

(42)   i.   there have arrived 3 men
        ii. *there to have arrived 3 men is unbelievable

We leave the question unsettled here. The answer depends on whether (24) holds. Note that this is an important issue to decide, if <u>there</u> needs Case "for itself", it means that the Case Filter holds of non arguments and therefore, cannot follow in toto from the θ-criterion.

Let us now turn to impersonal "it".[26] Similarly to <u>there</u>, the basic observation concerning expletive <u>it</u> is that it always co-occurs with a clausal argument, as in (40iii).

The same reasoning we have given for <u>there</u> holds for expletive <u>it</u>. It could a priori be argued that it may appear in Caseless positions. The situation here is different in that there is no equivalent to (24) with respect to clausal complements. However, Chomsky (1981) discusses a possibility that will have similar effects.

We have noted that an expletive subject position (at S-structure) always co-occurs with a clausal complement. Chomsky argues that whenever a postverbal clausal complement co-occurs with an expletive subject, they must be linked to form a chain (for reasons of chain maximization) whose head is the expletive; it follows that all expletives will be in such chains. From the visibility convention (37), this chain must be headed by a Case-marked position (expletive ≠ PRO); no expletive can appear in a Caseless initial position of a chain. Such an account predicts that there are no corresponding non arguments for (36vii) and (36viii), a consequence

corroborated by the ungrammaticality of the forms below:

(43)　　i.　*Seeming that S, John decided to leave
　　　　ii.　*It turned out without ever appearing, that S

Coupled with our assumption that weather it is a (quasi-) argument, the observation that expletive it always co-occurs with a postverbal clausal argument means that there are no predicates taking no syntactic arguments at all, at least not in French and English, a curious fact which is certainly not logically necessary (and false for German or Dutch, cf. below).

However, as Safir (1982) remarks, this is not true of all languages. He points out that contrary to English and French, Dutch and German allow certain morphological processes to "strip down some predicates of all their arguments". This situation is particularly interesting to us since it will permit us to check whether expletive elements may appear in Caseless contexts or not when they head a chain. So consider (44):

(44)　　i.　　Jan heeft gedanst　　　　(John has danced)
　　　　ii.　　er werd gedanst　　　　(there was danced)

　　　　iii.　*er werd gedanst zonder [e] gegeten te hebben
　　　　　　　(there was danced without e being eaten)

　　　　iv.　*er werd gedanst zonder er gegeten te hebben (er≠loc)

As (44iii) illustrates, despite the absence of any syntactically represented argument to the predicate eat, the subject position is still not allowed to be Caseless. Notice incidentally that this is a powerful argument for having empty subjects (particularly in (41iii)), and, more generally, an obligatory subject position for otherwise, the ungrammaticality of (41iii) would be quite difficult to account for.

　　　　Similarly, (44iv) illustrates the impossibility of a non zero

expletive element appearing in a Caseless position. Note also that we have to proceed in quite a roundabout way to establish this pattern of facts. If the assumption that (it, S') or (there, NP*) form a chain is dropped as Safir (1982) does, the conclusion we will now draw is straightforward, and indeed Safir (1982) draws it (cf. in particular his Chapter 3, section 2).

We are now in a position to answer negatively the question we started with on the existence of expletive elements corresponding to (36vii) and (36viii). Structures like (44iii,ii) show that their non-occurrence cannot be reduced to the non-occurrence of Caseless arguments (≠PRO) but rather must be ascribed to some other principle, yet to be formulated, independent of the argument status of NP's.[27] Call this principle P. Then P carries over to all the cases involving non arguments that we have reviewed so far i.e. (42ii), (43i), (43ii) and (44iii, iv). In particular, with respect to the distribution of expletive elements, the existence of P makes an appeal to the existence of chains (there, NP*) and (it, S') redundant.

We can now summarize the preceding discussion in table (45), which represents the occurring ϒ-categories appearing in chain initial position.

(45)   +locally A'-bound

+arg		-arg	
+Case	-Case	+Case	-Case
resumptive pronouns	* +Ph	*	*
variables	* -Ph	*	*

- locally A'-bound

+arg			-arg	
+Case	-Case		+Case	-Case
pronoun	*	+Ph	expletive	*
pro	PRO	-Ph	expletive pro	*

The organization of this table suggest several remarks. First of all, we notice that it is doubly asymmetrical. Contrary to the A'-bound system, the non A'-bound system allows for the existence of a Caseless element, namely PRO. Secondly, this element is treated as an argument. Were it not, the two systems in question would be parallel at lease as far as arguments are concerned. Even this much is incorrect, if (45) exhaustively represents Ɣ-categories in chain initial position. This asymmetry is reflected in the presence of the disjunctive statement (37iii) singling out PRO as an intrinsically visible category. Secondly, the systematic non existence of Caseless overt categories, and in particular of overt Caseless non arguments suggests that the Case Filter cannot be entirely reduced to the θ-criterion.

In fact, we return directly (cf. 2.6.2) to an independent argument suggesting that the Case Filter requires to be stated apait. If we restrict out attention to the table (45) (and given (35)), we can account for the range of existing and non existing elements by adopting the following hypothesis:

(46)   A category in chain initial position is either
         Care-marked or PRO

which does the work, for lack of a better alternative. The fact that we

need a disjunctive statement suggests that we may be missing something or that the invenory listed in (45) is inadequate.

2.6.2. Non-heads of chains

Turning now to non heads, which are Ÿ-categories by definition, (47) illustrates the a priori possibilities, given that non heads are not arguments and cannot be locally A'-bound:[24]

(47)	+Case		-Case
* (i)	+Ph	(iii) *	
* (ii)	-Ph	(iv) NP-trace	

Recall that we have defended the hypothesis that non heads (NP-traces) could not occupy Case-marked positions (cf. (30)), that we repeat here (in a slightly different form) as (48):

(48)   A non-head of a chain cannot be a Case position.

Given (48), both (47i) and (47ii) are excluded. (47iv) is the usual situation for an NP-trace. Finally, (47iii) does not occur. However, it is not ruled out by any principle presented so far. The visibility hypothesis (37) does not apply to it since it is a non argument, nor does (46), which we adopt in replacement of (37), since it only deals with heads of chains. Nor can (47iii) be excluded by the Binding Theory, since it should be able to appear in the same positions as a regular NP-trace without binding violation. Obviously, it is the notion of PF-visibility that is involved here as expressed by the Case Filter in Rouveret & Vergnaud (1980) and Chomsky (1980) formulation. In order to be

phonologically interpretable or visible, a category must be Case-marked. Consequently, we adopt the Case Filter coupled with (46) as a replacement for (37). Before analyzing the system of principles we end up with to deal with the whole range of existing ϯ-categories, it is worth noting that the existence of NP-trace partially fills a gap in the possible types of ϯ-categories. This can be better exemplified if we list the whole range of existing ϯ-categories, regardless of whether they appear in chain initial position or not. Putting (45) and (47) together, we get:

(49)　　　　　　　+locally A'-bound

	+arg		-arg		
	+Case	-Case	+Case	-Case	
	1. res. pron.	3. *	5. *	7. *	+Ph
	2. vbls	4. *	6. *	8. *	-Ph

　　　　　　　　　-locally A'-bound

	+arg		-arg	
	+Case	-Case	+Case	-Case
	9. pron.	11. *	13. exp. it	15. *
	10. pro.	12. PRO	14. exp. pro	16. NP-trace

We have made the following three hypotheses in order to explain the structure of this table:

(50)　　i.　locally A'-bound ϯ-categories are arguments
　　　　ii.　chain-initial categories are either Case-marked
　　　　　　or PRO
　　　　iii. Case Filter: Only Case-marked NP's are PF-visible.

Note first that we have placed, as is natural, NP-traces amongst non

arguments. If the way the Binding Theory has been formulated for the past few years is correct, which establishes a significant overlap in the properties of the relations antecedent/lexical anaphor (between one category with referential properties - the anaphor - and one which may have some or not depending on whether it is an NP-trace or not viz: the men$_i$ seem t$_i$ to like each other$_i$) , antecedent/trace (between one category without referential properties - the trace - and one with or without, depending on whether it is a trace itself or not), it follows that Binding does not mean referential dependence, and that no well-founded principled argument can be levelled against the assumption that expletive elements may enter into chain-formation, as we have so far assumed. An important conceptual distinction may still be made with respect to A-binding between the purely formal intra-chain A-binding, with no relation to (co)-reference properties, and extra-chain A-binding whose relation to questions of reference is obvious.[29],[30]. With respect to the latter, it makes little sense to talk about coindexing or binding between, say, two manifestations of expletive it.

Note secondly that there is a certain amount of redundancy between the three propositions listed in (50). (51) sums them all up:

(51) i. 49.3: ii,iii   49.5: i 49.7: i,ii,iii 49.11: ii,iii
          49.4: ii   49.6: i 49.9: i,ii

          chain initial   non chain initial

          49.15: ii,iii   iii
          49.16: ii

     ii. (50i):   49.5, 49.6, 49.8
          (50ii):   49.3, 49.4, 49.7, 49.8, 49.11, 49.15
                    (in chain initial position), 49.16
          (50iii): 49.3, 49.7, 49.11, 49.15

In (51i), each ruled out case is followed by those principles which rule

it out. 49.15, for example, is ruled out in chain initial position by both (50ii) and (50iii), but only by (50iii) in non chain-initial position. In (51ii), the list is organized by principles.

This table therefore illustrutes both the redundancy and the independence of the principles in (50): redundancy when some possibility is ruled out by more than one of them, independence when some possibility is ruled out by only one cf them, e.g., 49.4 and 49.16 only by (50ii), 49.5 and 49.6 only by (50i) and 49.15 in non chain-initialy position only by (50iii).

Obvious questions arise with respect to such a system: is it possible to eliminate or at least limit these redundancies? As we noted earlier, the system of existing categories is not symmetrical, and in our systems (as well as that of Chomsky, 1981), this is handled by a stipulation about PRO (argumer⁺ PRO). A careful discussion of various alternatives cannot be conducted, however, without a better understanding of the Bind'ng Theory. We therefore postpone this discussion until Chapter IV.

FOOTNOTES CHAPTER II

1. Actually, Chomsky (1981) exact formulation for (5i) is

    x_1 is an NP (note the index).

    We generalize this requirement to x_i here, pending discussion of clausal arguments for which this distinction is relevant.

2. Alternatively, we could define a variable as a certain type of NP which is locally A'-bound, without being locally A-bound.

3. Recall that we assume subjects of NP's to always be θ-positions. We justify this assumption in IV.3.

4. Actually we will argue in Chapter IV that this is true at S-structure but not necessarily at LF, for control structures.

5. If that were the case, e_1 might in fact be phonetically realized as him: cf. discussion in section 6 below.

6. Cf. also the remark in footnote 5 and the discussion of section 2.3 below. Note that the same remark holds of Chomsky (1981) discussion of the verb SEEM (cf his Chapter II).

7. Note that our reasoning only needs (15) to hold at S-structure. We argue in Chapter IV that it does not at LF.

8. Chomsky (1981) and Burzio (1981) also argue for the converse. We will discuss it in section 2.3. below.

9. There are however, some underlying assumptions, namely that (i) illustrates the only type of structure with exceptional government where NP_1 is a θ-position and NP_2 can be a θ'-position. We might imagine that categories other than V enter into exceptional government, e.g. N or A as in:

    (i) ...N [$_S$ NP...

    We will argue that such cases never exist with N, and only with NP_1 a θ'-position with A (cf. Chapter IV). (The impossibility with N might be related to a requirement of non distinctness between target categories). Secondly, there might be verbs triggering S'-deletion (or taking small clauses) which are transitive, as would verbs like impress be, if Chomsky (1981, p. 19ff) is correct. We would then get structures like (ii) (which contradicts (24) of the text):

    (ii) NP_1 [$_{VP}$ V NP_2 [$_x$ NP_3

    Where VP does not θ-mark NP_1 and V governs and does not Case mark NP_3. The question would then arise of what rules out (iii), with (me, NP_3 = e_i) a chain:

> (iii) it impresses me$_i$ [e$_i$ as intelligent]

Again, it turns out that NP$_3$ is always a θ-position in such structures, and therefore such chains cannot be formed (In any case, we reject this analysis- cf. Chapter IV).

10. Recall footnote 3 of this chapter.

11. It has been noticed in Pollock (op.cit) that the locative interpretation available for the PP in: On a tiré sur le bateau is lost in Il a été tiré sur le bateau, where only the Goal reading of the PP is available. It is conceivable that an analysis of this array of data might interfere with ours here. The same point we are making can be made with English forms such as It seems to John that S or the German or Dutch equivalent to it was danced on the roof, where no such problem arises.

12. Note that Hornstein & Weinberg (op.cit) restrict the application of this filter to phonologically zero traces.

13. As Caseless head of chain it is a PRO; cf 2.6 below.

14. A plausible reason might be the following: S'-deletion can only happen if S' is governed by and adjacent to V. More generally, this condition could extend to all cases of exceptional government: P-reanalysis, small clauses, S'-deletion and Restructuring:

> (i) Exceptional government of X° into Y^n iff
>   X° governs Y^n
>   X° is adjacent to Y^n

This is particularly interesting in the case of small clauses: we will see in Chapter IV that, because we admit governed PRO's, the argument in Chomsky (1981) that small clauses are not maximal projections does not go through.
 Thus, we expect two types of small clauses: Small clauses that are in exceptional government contexts:
[$_v$ consider][small clause] permitting a lexical subject when V is a Case marker: those only occur in selected environments, adjacent to V.
Small clauses opaque to government: those can appear either in selected or unselected environments e.g:

> selected: ca rend [PRO fou]
> unselected: John left the room [PRO empty]

In the case of reanalysis of P's, (i) obviously implies that only VP-PP's can be reanalyzed. Note finally we might have to add to condition (i) above: X° θ-marks Y$_n$, depending on the grammaticality of such examples as:

> (ii) who do you believe [$_S$ [$_{NP}$ picture of e to VP]]

15. Of course, we also drop the assumption that argumenthood of non overt categories is contextually defined: we rather assume that it is arbitrarily assigned.

16. Recall that [±phonetic] is different from [±phonological] as is shown by the properties of wh-traces which are phonologically non zero but phonetically null (cf. Jaeggli 1980a; Longobardi 1978; Chomsky & Lasnik, 1977).

17. We return to this important distinction from a different point of view in Chapter III.

18. Further discussion of this topic in Chapter IV.

19. Modulo our hypothesis that only the head position may be a Case marking-position. This is equivalent to Chomsky's formulation.

20. Because of this system of assumptions, Chomsky (1981) is forced to assume Case-checking as proposed by Jaeggli (1978) rather than Case-marking, for even in D-structure, NP's need Case to be θ-marked.

21. Note that this is not an observation: although it appears to be true with respect to wh-movement – except for some cases to which we will return below – it is much less obvious with respect to other types of local A'-binding such as, perhaps, clitic binding. In French for example, the presence of a clitic prevents the NP position linked to it from being lexicalized, a fact that might be taken to indicate Caselessness of the position. Let us mention briefly the cases of wh-movement that apparently take place from Caseless positions. These are illustrated by the following (cf. also Borer, 1981, 1981a):

> (i)   *Je crois Jean être parti
> *Je crois être parti plusieurs personnes
> L'homme que je crois e_i être parti

> (ii) *I assure you John to be the best...
> John , who$_i$ I assure you e_i to be the best
> (cf. Kayne, 1981a)

If we take these examples to illustrate cases of wh-movement from Caseless positions, we would need to explain why it is not always possible to do so (cf. IV.6 for further discussion):

> (iii) *L'homme$_i$ qu'il semble e_i etre arrivé...

22. So we get a three way distinction (Chomsky, 1981, Zubizarreta, 1982): "Regular θ-role positions" allow ungoverned subjects, variable subjects and controlled subjects:

> (i)   who$_i$ e_i left
> It is unclear why e_i to leave
> John tried e_i to leave

Constant arguments do not allow vaiable subjects, ungoverned subjects but allow controlled subjects:

> (ii)   *what e_i rains
>   *It is unclear why e_i to rain
>   It may snow without e_i raining

Finally non arguments allow none of them, although it is more difficult to show that the parallel paradigm is actually telling cf. text infra, and IV.6.2.4.3.

23. Alternatively, and this is the account we finally adopt in Chapter IV, it is excluded because PRO must be bound in S, s'nce it is governed, and there is no possible binder.

24. This mechanism is assumed to be chain formation by Chomsky 1981, with co-superscripting instead of Binding between successive members of the chain.

25. Indeed, we will argue that it is not clear that (24) is correct. We will show that the work it does is redundantly done by independently needed principles, in most cases. (cf. Chapter IV)

26. French impersonal _il_ plays both the role of _there_ in ergative constructions and of expletive _it_.

27. Why should Dutch (and German and Yiddish) differ from English and French with respect to what Safir (1982) calls the "no-stripped" predicate parameter", i.e. in allowing impersonal constructions like (44iii) with intransitive verbs?   Suppose that morphological rules are constrained in the following way: they cannot have as outputs items which have properties that no non-derived predicate has. In other words, a morphological rule cannot create an item that is "new" in terms of its properties, for example, create transitive adjectives if there are no non derived transitive adjectives. If we view morphological processes as mapping sets of lexical properties onto sets of lexical properties, and if we call L the reunion of all the properties of non derived lexical items, then morphological rules are constrained to operate exclusively on P(L) i.e. on the set of all subsets of L.   In particular, no morphological rule may have as output an argumentless predicate (more precisely a predicate with no syntactically expressed argument unless there exist non derived argumentless predicates.
     If this principle is correct, we would expert to find non derived argumentless predicates in Dutch (Yiddish, German) and indeed, that seems to be the case. We have seen (cf. footnote 22 above) that weather _it_ is distinct from non arguments in that it may be controlled; this is true in French and English, but not in Dutch:

> (i)   It has rained without snowing for days
> (ii)   Il a plu sans neiger pendant des semaines
> (iii) *het heeft gesneeuwd zonder geregend te hebben

So that, it is warranted to conclude that weather _het_ in Dutch is not

a (quasi-)argument, i.e. that weather verbs are both non derived and argumentless. Modulo this lexical distinction between English-type languages and Dutch-type languages, and this constraint on morphological rules, we can deduce that there are no argumentless passives in English. The question remains of course why there is this lexical difference for weather verbs.

28. That is, assuming either that no element can be both A-bound and A'-bound locally or that, when it occurs, "only one type of binding counts".

29. Of course, we never mean reference to objects in the world but rather reference to individuals in a set of mental objects (cf. Chomsky, 1981, Chapter VI).

30. Note that allowing A'-chain formation has one desirable consequence: it predicts that a certain subclass of expletive elements, namely those not appearing in head of chain position, will have exactly the same distributional properties as regular NP-traces, except, of course, for properties related to thematic structures.

CHAPTER III: A'/A RELATIONS

1. The Map Principle

Pursuing with our basic hypotheses, we now turn to relations between an A'-position and an A-position, and more particularly relations of this type such as wh-movement that have been analysed as movement relations.

We suppose that every category appearing in an A'-position has been base generated in that position. For example, a wh-phrase may be base generated in any COMP position if either it is selected for the feature [+wh], e.g. indirect questions, or if it is not selected at all, e.g. matrix COMP's Similarly, such items as left dislocated elements may be freely generable in positions not accessible to selection, e.g. TOPIC...

In the case of A/A relations, we have been able to reduce the properties expressed by the movement relation to an interaction of properties from various other components, in particular θ-theory. Indeed, the movement relation in a way expresses the idea that some discontinuous unit at some level (e.g. John and see in John was seen at S-structure) is best analysed as a continuous unit at some other level (e.g. was [seen John] at D-structure). Chain theory can be viewed precisely as a theory dealing with the nature and properties of some type of discontinuous units.

Is such a reduction possible in the case of A'/A movement relations? The answer to this question largely depends on some further assumptions. We may however anticipate on our conclusions: we will suggest on the

basis of the discussion that follows that such a reduction is not possible but that a principled approach to this problem can be constructed, which shares with the θ-criterion the same basic idea.

If we examine the letter of the reduction we have argued for in the case of A/A relations, we see that it is both θ-criterion and the argument status of the moved category that plays a crucial role: the θ-criterion essentially establishes a one-to-one mapping between arguments and θ-positions that underlies entirely the theory of chains. In the case of A'/A relations, the same approach can be contemplated only if that part of the discontinuous unit appearing in an A'-position is an argument. Of course we must first provide criteria for determining what arguments are. If we do not want to empty θ-theory of its content, these criteria must be independent from θ-theory. In fact, the natural assumption, particularly with language acquisition in mind, is that a given NP is identified as an argument:

    (i) either by virtue of its meaning, as determined by its content

    (ii) or arbitrarily or by virtue of some universal consention (as might be the case for empty categories)[1].

In particular, it seems reasonable to require of a non-empty category that it be an argument only if it has "referential" properties, i.e. may denote one or more individuals in the set of possible mental objects. Now, if we apply this criterion to elements appearing in A'-positions, we see that they do not all fall into the category "argument".[2] For example, by the criteria given above, wh-phrases in general should not be considered arguments,[3] whereas left dislocated elements or heavy NP-shifted phrases should.[4] It thus appears that the general case of A'/A relations cannot be made to follow from θ-theory in the manner we have used for A/A relations.

However, if we keep to the spirit of this reduction, it suggests by analogy a way to formulate a principled account of A'/A relations. Recall that the basic idea for A/A relations was to establish a mapping between arguments and θ-positions, mediated by intermediate traces, which the θ-criterion forces to be a one-to-one mapping. We may therefore suppose that some principle similar to the θ-criterion is at work in this case, which is not formulated as establishing a correspondance between arguments and θ-positions but rather would establish a mapping between A'-positions and A-positions. However, because not every A-position must be linked to an A'-position we have to adopt a parallel but weaker version that we might tentatively formulate as (1) below:

(1)   Map Principle
        The set of A'-positions maps on the set of A-positions

The meaning of (1) is clear: for every A'-position, A', there is at least one A-position to which A' is linked. The parallel with the θ-criterion is apparent and the intuitive idea underlying (1) is very natural. It states that the interpretive import of categories in A'-positions must be computed through some link with a position bearing a grammatical function.[3]

We have left open the question of the form that the mapping in (1) can take. The most important case may simply involve local A'-binding. There might be cases where the situation is more complex, for example Topicalization or Relative clauses:

(2)   i. $[_{TOP}$ John $]$ $[_{S'}[_{COMP}$ O_i $]$ $[$ I know e_i $]]$

        ii. $[_{NP}$ the man $]_i$ $[_{S'}[_{COMP}$ who$_i$ $]$ $[$ I know e_i $]]$

Where the mapping is mediated by the presence of an operator in COMP, so

that in some cases, this mapping might have to be implemented through A'-binding, rather than local A'-binding. We return to this topic in more detail in section 5 below.

2. Variables and the level of application of The Map Principle

Two questions immediately arise in connection with the assumptions put forth in the preceding section. First, at what level(s) of representation must the Map principle (1) be met? Second, the elements appearing in an A-position and bound to an A'-position are usually referred to as variables: how is such a notion to be precisely defined?

2.1. Concerning the locus of application of the Map principle (1), we may assume, in the words of Chomsky (1981) about the θ-criterion, that it is a reasonable criterion of adequacy for LF-representations, which, coupled with the θ-criterion embodies the simple idea that every meaning bearing element must be assigned a semantic function by sentence grammar. In other words, we may state the Map principle as in (3):

(3)   A'-binding maps the set of A'-positions on the set of
        A-positions at LF.[6]

Note that such an assumption is by no means obviously correct. There appear to be prima facie exceptions: some languages (e.g. Japanese cf. footnote 5) fairly freely allow left dislocated structures in which the NP element in TOPIC position is related to no particular NP in the associated clause, though some weaker notion of "aboutness" must hold between the clause and the topic. For concreteness, we may assume that such forms,

although acceptable if pragmatically interpretable (hence the "aboutness" requirement) are ungrammaical or perhaps marked exceptions, requiring positive evidence in terms of language acquisition. What seems to us the important point warranting our assumption is the apparent lack of structural correlate to "weak aboutness" versus Binding, requirement.

Next arises the natural question of whether (1) holds at S-structure as well, and in which form. Chomsky (1982) suggests in effect that it should not, on the basis of the behaviour of resumptive pronoun structures with respect to parasitic gap licensing. We will investigate this question in detail in the following section. Although we will ultimately disagree with his treatment of resumptive pronoun constructions, our conclusion on this issue will be identical to his. Notice that if the hypothesis that (1) holds at LF is indeed the minimal assumption, we might find a different picture at S-structure. We might, for example, discover that (1) holds at S-structure for an arbitrary subset of A'-positions.

2.2. Consider next the question of how variables should be defined. The pre-Chomsky (1981) option of taking variables to be traces of movement from an A to an A'-position is of course not open to us. The alternative, presented in Chomsky (1981), consonant with the idea that identification of empty categories is functionally determined led him to define variables as in (4):

(4)   x is a variable iff   x is an empty category and
                                x is an A'-position and
                                x is locally A'-bound

Let us for the moment adopt a slightly modified version of (4), namely (5) below, which takes into account a further remark of Chomsky (1981)

concerning resumptive pronouns as potential variables, and which is advocated in Koopman & Sportiche (1981):

(5)   x is a variable iff   x is a $\bar{Y}$-category in an
                                   A'-position locally A'-bound

In order to get a more concrete picture of what the differences between (4) and (5) are, consider the possible A'/A configurations that might arise, where $\underline{e}$ denotes an empty $\bar{Y}$-category and P a non-empty $\bar{Y}$-category. A $\bar{Y}$-category that is not empty at S-structure will not be empty in PF, nor in LF, given the organization of the grammar. Similarly, a $\bar{Y}$-category not empty at LF cannot be empty at S-structure because of the independence of the LF and PF components which precludes the insertion of phonological material in LF. So we are left with the following three possibilities:

(6)   (i)   (ii)   (iii)

      S-Structure   A'$_i$ P$_i$ A'$_i$ e$_i$ A'$_i$ e$_i$
      PF   P$_i$ P$_i$ e$_i$
      LF   P$_i$ e$_i$ e$_i$

where we might assume that coindexing is at least present in LF (due to (3)). Koopman & Sportiche (1981) observe that (4) is different from (5) only if such cases as (6i) exist, for mere phonetic realisation of a $\bar{Y}$-category gives no indication of its status (empty or not) at S-structure and LF. Indeed, the notion of variable is potentially relevant at S-structure and at LF. Since the only difference between (4) and (5) is that (5) allows non-empty variables at these levels, the crucial case is represented by (6i).

Now it could be argued that cases like (6i) do not exist on the basis of the fact that observed cases of locally A'-bound elements which are not

empty always reduce to (6ii). Koopman & Sportiche (1981) provide the following argument that (6i) is found: the ECP is a principle applying at LF sensitive to the distinction empty/non-empty. If we find the case of a phonetically realized ?-category interpreted at LF as a variable which is not excluded by the ECP when it should be if it were empty at LF, we may conclude that it is not empty at LF, hence at S-structure. Such cases are found in Vata (cf. Koopman & Sportiche, 1981) or in classical Arabic (cf. Aoun, 1980).

We might next wonder of what use the notion of variable is in syntax, beyond mere terminological usefulness. In fact, it will play a syntactically relevant role only to the extent that it defines a set of elements which display a distinctive property or set of properties. We will argue later on this basis that (5) needs revision.

Notice in particular that the notion of variable, as characterized in (4) or (5) is only reminiscent of the notion of variable used in classical predicate calculus. For example, in a sentence like (7):

(7)   who$_i$ e$_i$ said he$_i$ was sick

Only [e] is interpreted as a variable by (4) or (5), whereas both the equivalent of [e] and of he would count as variables in a logical translation of (7), with respect to the classical predicate calculus notion of variable.

3. Properties and Parameters of A'/A relations

We know turn to the parameters and properties determining the various

a priori existing A'/A relations, restricting ourselves to the cases in which A' locally binds A.

Amongst the parameters, we have already mentioned the fact that the item in A'-position may or may not be an argument. We may furthermore take into account the nature of the elements appearing in the A'-position, which we may assume, must be Ŧ-categories. Amongst allowable locally A'-bound categories, we may find either phonetically realised or phonetically unrealised elements. Finally, given our assumption concerning the application of the Map principle in (1), we may a priori suppose that local A'-binding of a category may be established at some level (LF but perhaps non existant at the other (S-structure). We may furthermore assume that if two elements x and y are coindexed at S-structure, they are also coindexed at LF. This assumption is not necessarily correct but it is the minimal assumption since it does not require postulation of some specific process modifying an already existing index structure.

Recapitulating, we get the following set of parameters:

(8)   i.   Argument status of the local A'-binder
        ii. Phonetically realized or unrealized locally
             A'-bound category
        iii. Level at which the Binding relation is
             established (S-structure or LF)

Note that the argument status of the locally A'-bound category is insured by Principle (50i) of Chapter II, so that A'/A relations cannot be parametrized along this dimension. Nor does the empty/ non-empty distinction for A'-binders seem to play any relevant role. Are there properties attributed to A'/A relations that might correlate in some way or another with particular values of the parameters in (8). We may

consider a priori those properties that have been argued to hold of A'/A relations. Amongst those properties, consider the following three:

(9)   i.   Boundedness: the Subjacency Condition
           ii. Constraints on Crossover: Weak Crossover and
                Strong Crossover
          iii. Parasitic gap licensing.

In the following section we examine each of these, except Strong Crossover, in more detail. A given A'/A relation may be subject to any of these properties, i.e. be bounded or not, induce Weak Crossover effects or not, license parasitic gaps or not. This yields eight a priori possible combinations, if these properties are not interdependent.

As for Strong Crossover, we will provide an argument in section 8 below that it is essentially a by-product of a particular way of conceiving of A'/A relations and that, as an identifiable phenomenon, it is of very restricted scope.

4. Level of relevance of A'/A relations properties

In order to establish what the correlations might be between the various types of A'/A relations and the associated properties that we have listed in (9), we start by an investigation of the familar cases, extending our conclusions, as we proceed, to broader classes of constructions.

4.1. The Subjacency Condition

Regarding The Subjacency Condition,[7] it may be construed in any of the following three different ways or any combination of these:

      i.   it is a filter on S-structure representations
     ii. it is a filter on LF representations
   iii. it is a constraint on the application of the rules mapping S-structure onto LF (i.e. it constrains string analyzability of the structural description of the rules).

Of course, these three options do not exhaust the imaginable possibilities. For example, we might assume that the Subjacency Condition constrains only a subset of S-structure/LF mapping rules. Clearly such an alternative is less desirable than any of those we have listed. Since we strive for maximal simplicity and elegance, we will consider hypotheses different from those listed only if we fail to accomodate the relevant data within one of these alternatives.

In order to settle this issue, we consider two very closely related constructions: wh-movement constructions and wh-in-situ constructions. By a wh-movement construction, we will mean a construction in which a (perhaps empty) wh-phrase appears in COMP position and binds at LF an empty A-position. A wh-in-situ construction is a construction in which the wh-phrase appears in an A-position at S-structure; these two constructions are illustrated in (10):

(10)   i. Qui tu as vu e
         ii. Tu as vu qui

where (10i) is a wh-movement construction at S-structure, and (10ii) is a wh-in-situ construction at S-structure.

Let us start with the wh-in-situ construction. Adopting the conclusions reached in Aoun, Horstein & Sportiche (1981) and similar proposals for Chinese by Huang (1982c), we assume that the LF representation associated with (10ii) is identical to (10i) with qui coindexed with [e] and derived from (10ii) through the application of the equivalent of a rule of wh-movement applying in LF, that we call wh-raising. The rationale for postulating such a rule is twofold. First, sentences like (10ii) receive the same interpretation as their counterpart (10i), so that they may be quite naturally assumed to have identical Logical Form representations. Second, qui in (10ii) is, as (10i) shows, an element subject to scope assignment (as are QP's...). Again, it is quite natural to assume, as May (1977) proposes, that such elements must be assigned scope (his Condition on Quantifier Binding) a principle that provides independent justification for the existence of such a rule.[8]

It has been argued extensively in the literature that wh-movement constructions are subject to the Subjacency Condition (cf. Chomsky, 1976...). On the contrary, it has been argued both by Aoun, Hornstein & Sportiche (op.cit.) and by Huang (1982c) that neither LF rules nor LF representaions are subject to Subjacency requirements, precisely by looking at wh-in-situ constructions and the rule of wh-raising. We find ourselves in the following situation: wh-movement constructions obey the Subjacency Condition and wh-in-situ constructions do not. As (10) shows, two such constructions may be chosen to have identical LF representations so that this difference in behaviour can only be attributed to a difference at the level of S-structure. We are therefore driven to assume that, in wh-movement constructions, the relation between the wh-phrase and the gap is established at S-structure (otherwise, there is no level at

which we could state that the wh-phrase/trace relation obeys Subjacency).
Furthermore, given that:

    i. wh-movement in an A'/A relation established
       at S-structure obeys the Subjacency Condition
   ii. wh-in-situ constructions are constructions
       involving an A'/A relation established in LF
       do not obey that condition

we may draw the simplest conclusion from this state of affairs, namely:

(11)   The Subjacency Condition is a well-formedness
           condition on S-structure representations only.

Although in the case of wh-in-situ constructions or Quantifier Phrase
interpretation constructions, it is obvious that the A'/A relations that
they ultimately involve at LF does not exist at S-structure, the picture
is generally not that simple, in particular, in cases of constructions
involving resumptive pronouns. We can now use (11) as a probe to settle
the question of at what level some A'/A relation is established, should it
arise.

We have chosen wh-in-situ constructions as representative of A'/A
relations not existing at S-structure because their close resemblance to
wh-movement constructions makes it very easy to compare them. However,
other similar relations evidently exist. In particular, the rule of
Quantifier Raising (QR) introduced in May (1977) has similar properties.
We can briefly recall that QR is a rule of the LF component assigning
scope to Quantifier Phrases (and perhaps to other scopal elements:
negation...) by adjoining them to some S-node.' May noticed that the scope
of QP's is generally clause-bound. For example, in the following
sentences:

(12)   i. Susan forgot that Sarah liked every painting of that museum

        ii.He claimed that few people made it to the finish line

The QP's cannot be construed as having scope broader than the matrix predicate. May (1977) attempts to reduce this clause-bound restriction to the Subjacency Condition by assuming that it is a well-formedness condition on LF representations. Indeed, suppose QR assigns broad scope to the QP of (12i) yielding the LF representation (13):[10]

(13)   [Every painting of that museum]$_i$ [$_S$ Susan
            forgot [$_S'$ that [$_S$ Sarah liked e$_i$]

In (13), the local binder of [e] is the preposed QP. Because they are separated by two S-nodes and one S' node, [e] is not subjacent to the the preposed QP, May argues. Hence the ill-formedness of (13). Although it is certainly desirable to reduce this clause-boundedness restriction to other independently justified constraints, the Subjacency Condition is not a plausible candidate.

    If we want to subsume the Clause-boundedness of QR to the Subjacency Condition, we must assume that it is _precisely_ the version of the Subjacency Condition that constrains other relations, say wh-movement, that applies to QR. However, this does not seem to be the case. For in both Italian (cf. Rizzi, 1982; Chapter II) and French (Sportiche, 1981), the node S is not a bounding node. Yet, matrix scope for QP's embedded in subordinate clauses is generally impossible, contrary to the prediction made by May's proposal.

    There is moreover direct evidence that QR (and wh-R) does not obey the Subjacency Condition. Consider the following examples:

(14) i. Tu as vu la photo de la soeur de qui
   ii. Tu as vu la photo de la soeur de chaque enfant
   iii.*L'homme [dont$_i$ [tu as vu [$_{NP}$ la photo de
    [$_{NP}$ la soeur e$_i$]]]]

As (14iii) shows, wh-movement of <u>dont</u> out of two NP's yields an ill-formed structure that can be attributed to the Subjacency Condition, if NP is a Bounding node in French, as is currently assumed. However, in both (14i) and (14ii), the scopal element can take scope over the whole clause, yielding the respective LF representations:

(15) i. Qui$_i$ [tu as vu [la photo de [la soeur e$_i$]]]
   ii. [Chaque enfant$_i$] [tu as vu la photo de [la soeur e$_i$]]]

Precisely when the Subjacency Condition and the clause-boundedness restriction diverge, i.e. in the case of extraction out of NP's, do we see QR not pattern along the predictions made by the Subjacency Condition. This suggests that the explanation for the clause-bound restriction must be sought elsewhere,[11] and that our assumption (11) may stand (May, 1977, notices this problem).

  Our argument only holds if (15i) and (15ii) are the LF representations of (14i) and (14ii). There is an alternative derivation that would assume optional pied piping by QR of NP's dominating the target QP. For example, this would yield the following derivation for the LF representation of (14i):

(16) i. 1st application of wh-R with Pied piping:
    [la soeur de qui$_j$] [tu as vu [la photo e$_j$]]

   ii. 2nd application of wh-R:
    [Qui$_i$] [la soeur e$_i$]$_j$ [tu as vu [la photo e$_j$]]

which would involve no violation of the Subjacency Condition. Unless we can exhibit some case where this option is not available, our

argumentation is nullified.

We now turn to the construction of such cases. In some contexts, pronouns may be interpreted as "logical variables", ranging over some domain specified by a QP or even a wh-operator as in (17):

(17)   Everyone thinks he is a nice fellow

The pronoun he can receive what is usually called a bound interpretation whereby (17) can be paraphrased by "For all x, x believes himself to be a nice fellow", i.e. in which he functions as a variable controlled by the QP everyone. There is a general condition, whether primitive or derived, regulating the conditions under which such interpretations may arise:

(18)   A pronoun P may be interpreted as a variable
            bound by a scopal element S (or, in fact, any
            element; cf. Reinhard, 1980) only if P is in
            the scope of S at LF (i.e. C-commanded by S at LF)

Returning to our main line of concern, consider the following example (with a non echo wh-in-situ):

(19)   Il a dit [$_{s'}$ que [$_s$ chacun a vu [$_{NP**}$ une photo d'
            [$_{NP*}$ un portrait [de sa mère] [par qui]]]]

Qui takes scope over the entire sentence while chacun takes scope over the embedded S only. If wh-R of qui pied-pipes NP* (or NP**) in order to avoid a potential subjacency violation, it would pied pipe the pronoun sa contained in NP* as well. Since chacun takes scope over the embedded S only, pied piping of NP* by wh-R would remove the pronoun sa from the scope of chacun at LF. So if a violation of the Subjacency Condition is to be avoided, we expect the only available interpretation of (19) available to be one in which sa is not bound by chacun which is

wrong.[12,13] Note also that it might be that Pied Piping is always excluded in LF, as Aoun (1982) and Huang (1982a) have argued.

4.2. Parasitic Gaps

We now briefly turn to parasitic gap (PG) constructions, or rather to trying to isolate the property that certain structural configurations possess, that allow them to license Parasitic gaps. By this we mean that these constructions allow parasitic gaps to appear without there being a dramatic drop in acceptability. In this section, we start with a very sketchy account, relying primarily on Chomsky (1982). We return to these constructions in more detail in section 8 below.

Basically, PG constructions are constructions meeting the following structural description:

(20)
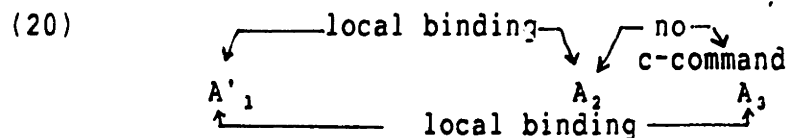
where moreover, both A_2 and A_3 are empty A-positions which are a priori acceptable A'-bindees, i.e. are arguments in Case marked positions. (Recall that locally A'-bound elements must be Case marked arguments, cf. II.2.6.)

(21)   i.   who$_i$ did your interest in e$_i$ surprise e$_i$.

       ii. It is John who$_i$ I persuaded friends of e$_i$ to visit e$_i$

       iii. they offended e$_i$ by not recognizing e$_i$
            immediately [their old friend from Texas]$_i$

       iv. [which book]$_i$ did you throw away e$_i$ without
            having read e$_i$.

which are all relatively acceptable sentences. Let 's now examine what bearing PG constructions have in telling apart the various types of A'/A relations that are in principle available.

4.2.1. Let us first take up the issue of whether PG are licensed by S-structures configurations like (20), or by LF configurations like (20). Let us consider again the minimally different constructions wh-movement and wh-in-situ. (From now on, we abstract away from the restriction imposed by English grammar on non echo wh-in-situ elements which require them to be accompanied by a wh-phrase in COMP at S-structure, i.e. we assume English is like French). The S-structures (22i) and (22ii) below are different at this level but are identical at LF (They are both identical to (22ii)):

(22)   i. who_i [your interest in e_i surprised e_i]
        ii. your interest in e_i surprised who_i

If it is at LF only that the configuration (20) must be met in order to license parasitic gaps, the pair of structural descriptions S-structure/LF (22ii/22i) should be grammatical) (and a fortiori the pair (22i/22i) should be grammatical). On the other hand, if it is at S-structure that (20) must be met, we expect (22i/22i) to be relatively acceptable, but not (22ii/22i). Of course, it is this latter alternative that turns out to be correct, since the string associated with (22ii) is totally unacceptable.

Consequently, licensing of parasitic gaps can be taken as a diagnosis for the existence of an S-structure configuration like (20). We return later to the correlated question of whether absence of parasitic gap licensing can be taken as a sufficient criterion for the non-existence of configuration (20) at S-structure.

In other words, we can state the following generalization (due to Chomsky, 1982):

(23)   If an A'/A relation licenses parasitic gaps, A'
            locally A'-binds A at S-structure.

4.2.2. There remains the question of why (22ii) is an ungrammatical S-structure configuration. Chomsky (1982) makes essentially the following argument. Consider (22ii) as representative of PG structures meeting (20) that obtain at LF but not at S-structure. How is [e]$_i$ interpreted in (22ii)? [e] in (22ii) is a free empty category, therefore, it is an empty pronominal, i.e. PRO. However, it is by assumption an acceptable variable site. By the ECP, variables must be governed. So [e] is in a governed position. On the other hand, we know that the Binding Theory, as Chomsky (1981) formulates it, has as consequence the property that PRO cannot be governed. If we assume that the Binding Theory applies at S-structure, we have an explanation for why (22ii) is ill-formed. It contains a governed PRO at S-structure.

We cannot however, accept this argument as it stands since we adopt (a slightly modified version of) Chomsky's further suggestion that empty pronominals in fact bifurcate between PRO and pro. By the assumption that [e] is an acceptable variable site, it follows that [e] in (22ii) is a Case marked empty ʏ-category, so that it is interpreted as a pro. Chomsky's argumentation therefore translates as the following conclusion:

(24)   The principles governing the distribution of pro
            (say, the ECP) apply at S-structure.

Note in particular that it is only to the extent that the principles governing the distribution of pro depend on the Binding Theory that we

have an argument that the Binding Theory applies at S-structure. Of course, this conclusion would not preclude the Binding conditions from applying at LF as well.

4.3. Weak Crossover.

Consider now the weak crossover effect (henceforth WCO), illustrated by the following examples:

(25)   i.   his mother likes John
        ii. his mother likes everyone
        iii. who does his mother like

Whereas intended coreference (or referential dependence) is possible between his and John in (25i), pronominal binding in (25ii) and (25iii) is impossible. That is, (25ii) and (25iii) cannot receive the following interpretations respectively:

(26)   i. $\forall x$, x's mother likes x
        ii. Wx, x's mother likes x   (Wx=for which x)

Postponing the discussion of what is the adequate theory of the restrictions on pronominal binding in these cases until section 8 below, let us admit for the moment that the ungrammatical configurations basically meet the following structural description at LF:

(27)
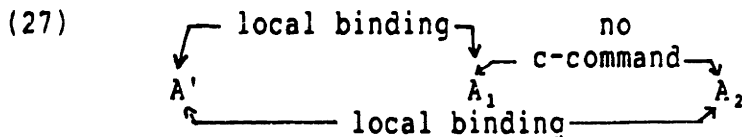
(linear order irrelevant) where A_1 is a non-empty γ-category (i.e. pronominal looking). We assume that the relevant restriction on

configurations meeting (27) holds at LF since, in the case of WCO violations involving Quantifier Phrases as in (25ii), it is only after the application of the Quantifier Raising rule to (25ii), i.e. at LF, that the configuration (27) is met.

Furthermore, we can distinguish between the constructions that meet (27) at LF but not at S-structure, and those meeting (27) at S-structure and a fortiori at LF. Consider the latter case. If the WCO effect is observed with any relation meeting (27) at LF, we should expect any relation meeting (27) at S-structure to display WCO effects. This prediction does not seem to be borne out. Indeed, consider structures involving Heavy NP-shift. Informally speaking, the process of Heavy NP-Shift can be said to relate the following pair of sentences:

(28)   i. He surprised all the representatives who were attending
            the meeting by talking about compulsory tax laws

        ii. He surprised [e] by talking about compulsory tax laws
            all the representatives who were attending the meeting

As (28ii) shows, some (heavy) NP appears in A'-position and is understood to be related to some A'-position (here the object position of surprise), that we may assume the shifted NP binds at LF. Before establishing the relevance of Heavy NP-shift to WCO, we need to establish that the Heavy-NP shift possess the properties we ascribe to it, and more specifically that the relation between the shifted category and the empty category it ultimately binds at LF is in fact established at S-structure. We must therefore establish two propositions. First that the relative ordering of the constituents in structures like (28ii) is identical at S-structure to what it is at surface structure, i.e. that Heavy-NP shift is not a stylistic rule in the technical sense. Second that the Binding relation

in question is already present at S-structure. Assuming the first proposition is established, we can provide a simple argument in favor of the second proposition of the basis of the conclusion of section 4.2.2 above. Suppose that in (28ii), the relation between the Heavy NP shifted category all the representatives who were attending the meeting and the empty A-position object of the verb surprise is not established at S-structure. This empty A-position is interpreted as containing a pro at S-structure, i.e. a Case-marked empty category which is free. Recall however, that we have show in 4.2.2 that the principles regulating the distribution of pro had to apply at S-structure. If the relation Heavy NP/empty category is not established at S-structure, the construction (28i) behaves at S-structure exactly like (29) with respect to these principles:

(29)   He surprised [e] by talking about tax laws.

i.e. should be ungrammatical. Since (28ii) is grammatical, there is at least one "derivation" in which the relation Heavy NP/empty category is established at S-structure.

An additional reason reinforcing this conclusion and also establishing the first proposition comes from the observation that Heavy NP shift constructions license Parasitic gaps, as (21iii) shows. By (23), we can conclude that the A'/A relation it involves must be established at S-structure. In particular, it shows that Heavy NP shift cannot be a stylistic rule for if it were, the parasitic gap would be unbound at S-structure and therefore ruled out as an illicit pro. We can also conclude something stronger from the fact that Heavy NP shift licenses parasitic gaps as in (21iii), namely some structural information

concerning the c-command relations between the various NP's occurring in that structure. We may infer that in (24iii) neither of the two gaps c-commands the other and that the shifted NP locally binds both of them. This leads us back directly to our present question, namely, whether or not Heavy NP-shift constructions trigger WCO effects. It is easy to see that the PG configuration (20) is structurally identical to the WCO configuration (27) except for the fact that one of the two A-positions is lexically filled by a non-empty r-category in (27).[14] Consequently, in order to find out whether Heavy-NP shift triggers WCO effects, it suffices to replace one of the gaps in (21iii) by a non-empty category as in (30) below:

(30)   He deeply offended e_i by not recognizing
          him_i immediately, [our old friend from Texas]$_i$

If Heavy NP shift triggered WCO effects, structures like (30) should be on a par with (25ii) or (25iii). However, this does not seem to be the case. (30) does not differ in acceptability from comparable Heavy NP shift structures not meeting (27) like (28ii) or like He deeply offended by not talking about the past our old friend from Texas. What are the differences between Heavy NP shift of the one hand and wh-movement on the other that could account for their unlike behaviour with respect to WCO? Chomsky (1982) makes the observation that, in Heavy NP shift constructions, the element appearing in an A'-position is argument-like. So let us assume that this is the determining factor, i.e. that WCO effects are triggered in configurations like (27) only if the A'-position hosts a non-argument, a conclusion that we might state as in (31):[15]

(31)   WCO effects are a property of A'/A relations
          involving a non-argument in A'-position.

In the case of Heavy NP-shift constructions, it is easy to indentify the element in the A'-position as an argument or not. However, there might be constructions in which such a decision is not as obvious, e.g. clitic constructions. We may hope to use (31) to settle the question, should it arise.

5. A Classification of A'/A relations

Summing up our procedure, we have classified the A'/A relations with respect to two different sets of criteria: a set of classificatory properties and a set of possible parameters. More precisely, given a LF local binding relation between an A'-position A'* and an A-position A*, we have recognized three different parametic features:

(32)   i.   A'* is an argument or not (we will note this by [±θ])
            ii. A* is a ↑-category which is empty or not ([±ph])
           iii. A'-binding of A* by A'* is established at
                S-structure or at LF.

In parallel, we have tried to partially correlate these various parameters with the following properties:

(33)   i.   the A'/A relation obeys the Subjacency Condition
                (±S): we have argued that [+S] is a property
                of S-structure binding.
            ii. the A'/A relation licenses parasitic gaps [±PG]:
                we have also argued that this is a property of
                S-structure binding (we will further qualify
                this statement in section 8 below)[16]
            iii. the A'/A relation triggers WCO effects
                [±WCO]: we have argued that [±WCO]
                held only of cases in which A* is a nonargument

Now let us run through the various possible combinations. Necessarily,

our discussion will be incomplete for lack of a systematic survey of relevant constructions across languages.

5.1. Let us consider first the cases where A* is empty. If we conjoin (32) and (33) in a single table, we get (34):

(34)   A* is empty   A'* is [-θ]   A'* is [+θ]

   S-structure binding (i) [+PG,+S,+WCO] (iii) [+PG,+S,-WCO]

   LF binding   (ii) [-PG,-S,-WCO] (iv) [-PG,-S,-WCO]

This table makes two distinct claims. First, that there are four types of A'/A relations with A* empty that should be exemplified in natural languages, and secondly, that some configurations of features is impossible, e.g., some A'/A relation which [-S, +PG] or [+S, -PG].

5.1.1. Consider first (34i). This is exemplified by wh-movement, by definition almost. We have shown that it had to be an S-structure binding relation between a wh-phrase in COMP and some A-position. It is clear that it meets the relevant properties. With respect to (34ii), the same remarks apply to wh-raising constructions and by a trivial extension to QR constructions. In both cases, the A'/A relation is obtained at LF by movement of the (quasi-)quantifier phrase from an A-position to an A'-position. Note however, that we in fact predict a different type of derivation for relations meeting (34ii). We could have a phrase in an A'-position at S-structure which gets coindexed with an empty A-position only at LF. For ease of reference, let us call this construction the empty resumptive pronoun construction. Such a construction would meet the following structural descriptions:

(35)   i.   (S-structure) ...A'...e...
          ii. (LF)   ...A'$_i$...e$_i$....

Because [e] ultimately ends up A'-bound at LF, it must be a Case-marked position and, in particular, it is interpreted as a pro at S-structure. Since we have argued that the principle(s) governing the distribution of pro must hold at S-structure (cf. 4.2.2) we expect the following generalization to hold:

(36)   Only languages allowing empty Case-marked
               pronominals at S-structure (so-called pro
               drop languages) may allow empty resumptive
               pronoun constructions.

A possible example of such a construction has been argued to exist by Taraldsen (1981) in some dialects of Italian (which is a pro-drop language) and precisely on the basis of the fact that the A'/A relation involved did not obey the Subjacency Condition. If indeed, it is an example of (34iv), two predictions that (34) makes are fulfilled. First that relations like (35) exist and second that they do not obey the Subjacency Condition. Of course, we further predict that such constructions should neither license parasitic gaps, nor trigger WCO effects. The accuracy of these predictions remains to be verified.

5.1.2. Consider next (34iii). We have already seen that this type of relation was exemplified by Heavy NP shift (in English or in French) which, as we have verified, does not trigger WCO effects and licenses parasitic gaps. We furthermore predict that it should obey the Subjacency Condition, as seems to be the case:

(37)   i.   She introduced [$_{NP}$ the man [$_{S'}$ who invented the
               first perpetual motion device]] to her brother

ii. She introduced [$_{NP}$ the man [$_{S'}$ who invented e$_i$]]
to her brother [the first perpetual motion device]$_i$

Heavy NP shift from the embedded relative clause into the matrix clause is ungrammatical. This state of affairs could be attributed to a violation of the Subjacency Condtion since the two target phrases in (37ii) are separated by two bounding nodes (S' and NP). Some caution is in order however. Because, contrary to wh-movement cases, there is no apparent violation of the Subjacency Condtion (which is usually analysed as successive COMP to COMP movement), it is quite difficult to ascertain the nature of the locality requirement imposed on Heavy NP shift. Aoun & Hornstein (1982) argue that the clause boundedness restriction on the operation of QR, which had been argued to reduce to the Subjacency Conditon, results in fact from The Binding Theory. The observationally clause-boundedness restriction on Heavy NP shift could arise because of the same kind of reason.

5.1.3. Let us turn next to (34iv). We can make the same parallel between (34iv) and (34ii) we have made between (34i) and (34ii). We predict similarly that it is only in languages allowing empty Case-marked pronominals that we should be able to find examples like (34iv).

For example, it is conceivable that we find a language with some equivalent of LF-Heavy NP shift, i.e. an A'/A relation between a shifted NP and some A-position which is only established at LF. In fact, once again, Italian may provide us with an example. Recall that Italian allows pro in subject position of tensed clauses and possesses a construction usually called Left Dislocation. When an NP is Left Dislocated from the subject position of a tensed clause, we get a configuration illustrating

(34iv) (the example comes from Cinque (1977)):

(37)   Giorgio$_i$, sapero che e$_i$ volera andare a stare in campagna
         Giorgio, I know that (he) wanted to go live in the country

As predicted, this construction, exactly as English Left Dislocation, neither obeys the Subjacency Condition nor licenses parasitic gaps, nor triggers WCO effects.

There is however a potentially more interesting candidate illustrating (34iv), namely clitic constructions. Because a discussion of these constructions would take us too far afield, let us postpone it until section 6 below.

5.2.1. Let us now turn to cases parallel to those discussed in the previous section in which the locally A'-bound category is not empty, i.e. is pronominal-like. If the parallelism between these two sets of A'/A relations is perfect, we should expect to find a construction representative of each of the following types:

(38)   A* is not empty

	A'* is [-θ]	A'* is [+θ]
S-structure binding	(i) [+PG,+S,+WCO]	(iii) [+PG,+S,-WCO]
LF-binding	(ii) [-PG,-S,+WCO]	(iv) [-PG,-S,-WCO]

These predictions are only partly fulfilled. First, there is, to my knowledge, no obvious example of constructions that would fall into category (38iii). This is significant only if the counterpart of (38iii) with A* empty were abundant, but this is not the case. However, examples of (38iv) are easy to find and are found in French or English Left (or Right) Dislocation for instance. Some category – the Left Dislocated

constituent - appears in an A'-position and is obligatorily linked to some "resumptive pronoun". Furthermore, this construction exhibits the predicted range of properties:

(39)   i.   John$_i$, I saw him$_i$
           ii. John$_i$, I met the man who taught him$_i$ how to swim
           iii. John$_i$, his$_i$ mother likes him$_i$
           iv. *John$_i$, I talked to him$_i$ without ever having seen [e]$_i$

(39i) is a simple example of the Left Dislocation construction. (39ii) illustrates the fact that the relation between the Left Dislocated constituent and the associated pronoun is not subject to the Subjacency Condition. In (39iii), the WCO configuration is met without yielding an unacceptable sentence, and in (39iv), we can observe that parasitic gaps are not licensed by Left Dislocation (Note that in English, (39iv) could be taken as a weak violation of the Subjacency Condition by Topicalization. In French, Topicalization is not possible in such cases and the equivalent of (39iv) is totally unacceptable: Jean, je lui ai parlé sans avoir vu [e]

Although potential examples of (38ii) abound in the literature, they do not seem to behave as predicted. These constructions are usually called resumptive pronoun constructions but let us call them true resumptive pronoun constructions in order to distinguish them from those of (38i). Such constructions are for example found in Spanish relatives (cf. Chomsky (1982)), Yiddish relatives (Lowenstamm p.c.), Modern Hebrew relatives (cf. Borer (1981)), Standard Arabic...[17] and all violate the Subjacency Condition. (40) illustrates this fact for Yiddish and Hebrew (the Hebrew examples come from Borer (1981) but similar examples could be constructed for other cases):

(40)   i.   (Yiddish)
           der boxer$_i$ [$_{s'}$ vos [$_s$ ix ken [$_{NP}$ dos meydl$_j$ [$_{s'}$ vos
           [$_s$ (zi$_j$) hot im$_i$ gezen
           the boy$_i$ that I know the girl that (she$_j$) has seen him$_i$

      ii.   (Modern Hebrew)
           ha' ish$_i$ [$_{s'}$she [$_s$ pagashti [$_{NP}$ et ha'isha$_j$ $_{|s'}$ she
           [$_s$ t$_j$ ra'ata 'oto$_i$
           the man that I met the woman that saw him

However, so far as I have been able to check, this type of constructions does not trigger WCO effects, contrary to the expectations of table (38). This is not in itself significant for the examples so far reviewed. The majority of true resumptive pronoun constructions are relative clauses. It has long been noted that relative clauses, even when they seem to fall under (34i) do not trigger WCO effects, as the following minimal contrast [1] demonstrates:

(41)   i. *who$_i$ does his$_i$ mother care for e$_i$
      ii. the kid$_i$ that his$_i$ mother cares for e$_i$

Clearly the account of WCO effects and the analysis of relative clauses must be so construed as to allow structures like (41ii) while rejecting structures like (41i). It is reasonable to assume that the absence of WCO effects in relative clauses will extend to cases of relative clauses falling under (38ii).

A more conclusive test for the absence of WCO effects in constructions meeting (38ii) could be constructed if we could find a case of wh-questions falling in the true resumptive pronoun construction category. This appears to be the case for Egyptian Arabic direct and indirect wh-questions as described in Kenstowicz & Wahba (to appear). In Egyptian Arabic, a wh-phrase may appear in COMP position in wh-questions in which case it is associated with a resumptive pronoun when the

wh-phrase is an NP. Furthermore, this construction may violate the Subjacency Condition:

(42)   i.   ?eeh$_i$ illi/?ayy kitaab$_i$ Fariid ishtaraa-h$_i$
              what that / which book Fariid buy it

        ii.   miin$_i$ illi/?ayy talamiiz$_i$ Fariid simi9 isaa9it inn Mona
              who that /which students Fariid heard the rumor that Mona
              yimkin titgawwiz uh$_i$/hum$_i$
              might try to marry him/them

(42i) illustrates the basic construction (if the wh-phrase is itself a wh-word it cooccurs with the complementizer illi) and (42ii) illustrates the fact that it does not obey the Subjacency Condition. Now, Kenstowicz & Wahba (op cit.) report that WCO configurations are perfectly acceptable:

(43)   i.   miin$_i$ illi marat-uh$_i$ bitbuus-uh$_i$
              who$_i$ that his$_i$ wife is kissing him$_i$

However, if the predictions concerning WCO effects for constructions that might fall in (38ii) seem falsified, those pertaining to parasitic gap licensing appear to be consistent with the available date (cf. Chomsky (1982) for some marginal English examples and some examples from Spanish due to E. Torrego).

Now we face a double problem. First, if these constructions that we have just reviewed are not examples of (38ii), this category does not seem to be exemplified, a gap that we must explain. Secondly, if the classifiction that we have proposed were correct, these constructions would not fit in it. Indeed, with respect to Subjacency, WCO and PG, these constructions behave exactly as Left Dislocation. However, in the only crucial case at our disposal, namely Egyptian Arabic wh-questions, the item in position A' seem rather to fall together with the non arguments, i.e. with elements not requiring a θ-role, so that this

construction at least should not fall under (38iv), i.e. in the same class as Left Dislocation.

There is one possible suggestion that could be made on the basis of proposals made in Chomsky (1982) that might appear to solve this problem. Suppose we postulate a further level of representations beyond LF, call it LF'. Suppose further that WCO is indeed a property of some LF configuration, as we have assumed, but that coindexing of an unindexed A'-position at S-structure with some A-position only takes place at LF' (through some rule of Predication). Then, constructions that we classified in (38iv) and constructions that we have attempted to classify in (38ii) would not fall into these categories but rather enter into some classification of LF' A'/A relations.[19] We would then get a reason why these constructions are not subject to WCO effects. This move would not get around the problem for we would need an explanation as to why no relation exemplifies the a priori possible (38ii) and (38iv) cases. Let us therefore drop this alternative and come back to our earlier hypotheses.

Suppose that it is in fact not the argument status of A'* that is relevant to the distinction between the columns of table (38), but rather some property P to be specified. For consistency, we may assume that being an argument is a sufficient (but not necessary) condition for possessing property P. Assume further that elements in A'* not having property P must bind some A-position at S-structure. In other words, let us assume that the following implication holds:

(44)   If A'* binds no A* at S-structure, A' has property P

Then we derive that there can be no examples of (38ii). Indeed, A'* has

property P or not.   If not, A'* must bind A* at S-structure, so that the relation A'*/A* does not fall under (38ii). If A'* has property P, it escapes WCO effects by assumption, so that again, it cannot fall under (38ii).

As we have seen earlier, LF-binding may arise in only two ways. Either by coindexing at LF of A'* and A*, or by LF-movement from A* to A'*. If the above argument is correct, the first option is ruled out if A'* has property not P. This argument extends of course to A'/A relations where A* is empty. It means that the only type of LF binding yielding WCO effects are those arising through QR and wh-Raising. In particular, we cannot analyze the Italian empty resumptive pronoun constructions as we suggested we could i.e. as an example of (34ii). Rather, it should fall under (34iv) and in particular, we predict that it should be exempt from WCO effects contrary to our earlier conclusions (although it is still predicted that it should not license PG).

5.2.2. Let us now turn to the last type of constructions that are predicted to exist by table (38), namely (38i). Let us start with the question of whether there are A'/A relations where A* is not empty, which obey the Subjacency Condition. In fact, such examples can be found in various languages, e.g. Relative clauses in Haitian (cf. Koopman (1982a)), Free relatives in Modern Hebrew (cf. Borer (1981)), Wh-constructions in Vata (cf. Koopman (1980, 1982)) that we illustrate below:

(45) i.   (Modern Hebrew)
          ma$_i$ she   hexlatnu 'al av$_i$
          what/that we decide on it
          'Whatever we decide on,...'

ii. (Haïtian)

fi$_i$ l ap rēmē avè l$_i$ la
girl he ASP like with her PART
'The girl he is going out with,...'

iii.(Vata)[20]

kɔ̆$_i$ mɔ̆mɔ̀ Ò$_i$ gùgù-ɓɔ́ nā Ò$_i$ kà mlī̄
man HIM-HIM he thought-REL that he FUT leave
'The man who thought that he was leaving,...'

àlɔ́$_i$ Ò$_i$ gùgù nā Ò$_i$ kà mlī là
who   he   thought that he FUT leave WH
'Who thought that he was leaving'

In order to verify whether such constructions trigger WCO effects, some
care is needed since, as we noted earlier, relative clauses in general
seem to be exempt from them, even if the relevant configuration is met.
Fortunately, we can check this with Vata wh-questions, and we find indeed
unacceptable structures in WCO configurations:

(46) i. * àlɔ́$_i$ Ò$_i$ nɔ́ gūgū nā Ò$_i$ mlī là
         who$_i$ his$_i$ mother think that he$_i$ left WH
         'Who does his mother think left'

     ii. * àlɔ́$_i$ ǹ ylà̀ Ò$_i$ nɔ́ nā Ò$_i$ mlì là
          who$_i$ you tell his$_i$ mother that he$_i$ left WH
          'Who did you tell his mother left'

Similarly, we predict that such constructions should license Parasitic
Gaps. I have not been able to test this prediction in all the languages
mentioned above, but in the case that I have been able to test, i.e.
Vata, this prediction is borne out (it is also in Welsh, if I understand
Harlow (1981) and McCloskey (1983) correctly). So consider the following
examples, where the first one is the construction without wh-movement and
the second, third and fourth are relatively acceptable parasitic gap
constructions respectively with and without a resumptive pronoun:

(47)　i.　blĭ kā mó yé yÉ , Ò dlɛ́ mó
　　　　　Ble AUX him PART see　 he beat　him
　　　　　'When Ble sees him, he beats him'

　　ii.　kò mÕmÒ blĭ kā - 6ó [$_{NP}$ e$_i$] yé yÉ lè Ò dlá-6ó [$_{NP}$_e$_i$]
　　　　　man HIM-HIM$_i$ Ble AUX-REL e$_i$ see PART he beat-REL e$_i$
　　　　　'the man that, when Ble sees, he beats'

　　iii.　kò mÕmÒ$_i$　 blĭ kā-6ó [$_{NP}$ e$_i$] yé yÉ lè Ò$_i$ guo
　　　　　 man HIM-HIM Ble AUX-REL [e] see　 PART he(resumptive) runs
　　　　　'the man$_i$ that when Ble sees, he$_i$ runs away'

　　iv.　kò　mÕmÒ$_i$　Ò$_i$ kā-6ó blĭ　 yé yÉ　 lè Ò$_i$ guo
　　　　　man HIM-HIM$_i$ he$_i$ AUX-REL Ble　 see　　 PART he$_i$ runs
　　　　　'the man$_i$ that, when he$_i$ sees Ble, he$_i$ runs away'

Note that the parasitic gap in (47ii,iii,iv) cannot be a real gap; it is not in a position accessible to movement.

So we find that the constructions with resumptive pronouns that are a priori candidates for falling under (38i) behave exactly as predicted as far as we could determine. There is one aspect of these constructions that we have so far neglected which the examples above illustrate. We see that, in Vata, relativization from subject position requires a resumptive pronoun, while relativization from other positions leaves a gap (in fact it is true for all wh-constructions in Vata).

This situation seems to obtain quite generally for the constructions with resumptive pronouns falling under (38i). This partitioning of positions between empty and non empty categories requires some explanation. Following essentially a suggestion made in Koopman (1980) let us assume that non empty categories will appear precisely in the positions in which we would otherwise get a violation of some principle of grammar governing the distribution of empty categories. For example, Koopman (1980) argues that in Vata, resumptive pronouns must appear in subject position because this position is never properly governed, so that an empty category in that position would violate the ECP. Assuming for

the moment that it is indeed to avoid ECP violations that non empty categories must appear in certain positions is very plausible given the typical distribution of these resumptive pronouns: subject (in Vata), noun complement, object of a preposition, i.e. positions that are considered not properly governed.

If this assumption provides a reason why resumptive pronouns must appear in certain positions, it does not explain why they cannot appear in the others. Here, for lack of a better alternative, we may rely on some restricted version of the Avoid pronoun principle put forth in Chomsky (1981), that we could formulate as follows:

(48)   Avoid phonological feature

(48) is taken to mean that, in a given construction, if the option exists of using a phonologically null element, this option should be used.

5.3. We can summarize our discussion in the following table:

(49)

	A* is [+ph]	
	+P (-WCO)	-P (+WCO)
S-structure: [+PG,+S]	(i) ?	(iii) Vata wh-constructions Welsh relative clauses Modern Hebrew free relatives
LF [-PG,-S]	(ii) Left dislocation Yiddish relatives Modern Hebrew relatives	(iv) *

	A* is [-ph]	
	+P (-WCO)	-P (+WCO)
S-structure [+PG,+S]	(v) Heavy NP-shift	(vii) English wh-constructions
LF [-PG,-S]	(vi) Italian Left dislocation (Clitics ?)	(viii) wh-Raising QR

Let us first introduce some terminology. Adapting a suggestion made in Zaenen, Engdahl & Mailing (1981), let us call A'/A relations established at S-structure <u>syntactic</u> and A'/A relations established at LF <u>anaphoric</u>. Similarly, we will call resumptive pronouns found in (49iii) <u>syntactic</u> <u>resumptive</u> <u>pronouns</u> and those found in (49ii) <u>true or anaphoric</u> <u>resumptive</u> <u>pronouns</u>.

5.3.1. The first remark we can make about table (49) bears on the distinction empty/non empty for A*. This distinction appears to play no role whatsoever in this classification. Furthermore, its introduction has curious results. Consider for example Vata wh-constructions, in which A* is a syntactic resumptive pronoun when it is in subject position, and is empty otherwise. If the distinction empty/non empty plays a role, we should classify Vata wh-constructions with A* a subject position in (49iii) and Vata wh-constructions with A* a non subject in (49vii). A similar classification would hold for Haitian relatives, Modern Hebrew free relatives etc... But from the point of view of the syntactic properties that these constructions display, this distinction appears arbitrary. We will therefore assume that it should be dropped entirely as irrelevant to A'/A relations classification.

It has often been noted (although it has been phrased differently) that syntactic resumptive pronouns are rare. Accordingly, some authors note that resumptive pronouns tend to be (or always are) anaphoric, while S-structure A'-bound elements (locally) are (or tend to be) non overt. This observation might very well be a fact of language to be accounted for, or it may be accidental, or even false, due to our incomplete knowledge of syntactic variation.

Assume it is a valid generalization. Is such a generalization antagonistic to our conclusions? I think not. Our arguments merely show that such cases do exist and when they do, they have the properties of non overt variables (apart of course, for properties pertaining to overt/non overt distinction). It might very well be nonetheless that such cases are rare: a fact to be accounted for, if true, but not, we show, by claiming that syntactic resumptive pronouns simply do not exist.

5.3.2. The second remark has to do with our having classified wh-R and QR constructions under (49viii). Why do they not fall under (49iv) instead? Recall that the argument we have given for the lack of A'/A relations falling in (49iv) would not exclude that possibility, namely that of postulating a "resumptive pronoun strategy" for LF-movement. One could try to argue that this is ruled out in principle by the very formulation of LF movement rules. Such rules could be formalized so that a moved category would leave no phonological feature on its trace. This assumption appears to me as arbitrary as the assumption that a moved category leaves as a trace a non empty pronominal copy. In fact, in an alternative theory in which syntactic binding arises as a result of a movement rule (e.g. Chomsky (1981)), such an option must be allowed for the cases falling under (49iii). Indeed, if the presence of a non zero element as "trace" is linked, as we have argued, to the prevention of ECP violations (or any other relevant principle), then these syntactic resumptive pronouns cannot be argued to be inserted in PF (nor in LF because of the argument of 4.2.2. above) as we have already pointed out.

In a language like English or French, LF resumptive pronouns should not be available (or only marginally) since LF movement rules do appear to

trigger LF ECP violations (cf. Kayne (1979) and Aoun, Hornstein & Sportiche (1981) for relevant examples).[21] It seems reasonable, given the nature of these phenomena, to suppose that it should be considered an unmarked option not to have LF-movement resumptive pronouns, especially for those languages not allowing resumptive pronouns at all.

That such is the case for languages allowing syntactic or anaphoric resumptive pronouns is not so obvious. It appears plausible to link differences in the functioning of the LF components of different languages to overt differences between these languages.[22] When overt differences do exist, it is natural to ask whether they are reflected in LF or not. For example, it is conceivable that languages like Vata exist in which an equivalent of French personne would be permitted to have wide scope from subject position of a tensed clause, its LF trace being treated as a resumptive pronoun.

In Vata, Koopman (1980) shows that LF movement from subject position does trigger ECP violations. If it should turn out to be systematically the case that parallelism in this respect is not found between syntactic binding and LF-movement, as we suspect will turn out to be true, it would strengthen our hypothesis that the processes involved (movement in LF, coindexing at S-structure) are of different formal nature.

5.3.3. Let us now consider the problem of what property P might be. Recall that we want arguments, i.e. θ-role bearers, to have property P. Trying to characterize not P instead of P suggests a plausible approach. The set of elements having property not P will only include non arguments. What kind of well defined class of non arguments do we find: essentially two; expletives and operators. Obviously, expletives are of no relevance

here. So let us assume that not P has to with being an operator, or more precisely that an A'/A relation between A'* and A* has property not P iff A'* is an operator. This appears very plausible if we consider for example, Left Dislocation or, say, Yiddish relative clauses, in which we can assume that the interpretation does not proceed through the intermediate coindexing with an operator. A consequence of this assumption is that we do not expect constructions involving anaphoric binding, i.e. falling under (49ii), to involve an overt operator.[23] This seems inconsistent with the existence of constructions such as Egyptian Arabic wh-questions. In the absence of a better understanding of this last construction, we can only speculate. One plausible assumption is that the wh-phrase in these cases is in fact not in COMP, but rather in the position occupied by Topics and Dislocated constituents, so that we could restrict property not P to A'/A relations where A' is an operator within the S' system (that is, excluding Topic position, Left or Right dislocated positions, heads of relatives, Focus in Cleft constructions etc...) We thus establish the following dichotomy:

    i.   A' positions within the S' system
    ii. A' positions outside the S' system

Positions in (ii) are [±P], i.e. do not trigger WCO effects. We may furthermore assume that they may bind some A-position only at LF, for we want to exclude in principle dislocated structures of the form John$_i$ I saw e$_i$ which are not mediated by an abstract operator (thereby explaining, for example, the ungrammaticality of le beurre, j'ai mis [e] au froid). Positions in (i) may be [+P] if they are not operators, e.g. in the case of Heavy NP shift or [-P] (operators). They bind some A-position at S-structure, and treigger WCO effects if they [-P].

Having partially answered the question of what the nature of property P is, we can retrun to the problem we started with. After having postulated the Map principle (1), we wondered at what level it should be postulated to apply. We have argued that, as a condition of "semantic" adequacy, it should at least be met at LF. This does not preclude that it should also be met at S-structure. In fact, given the above discussion, the hypothesis (44) put forth as a justification for the absence of constructions like (49iv) expresses that the Map principle should be met at S-structure for at least a subset of λ'-positions, namely those that are [-P], i.e. operators within the S'-system. We can thus summarize the above discussion, (44) and our original assumption concerning the locus of application of the Map principle as (50):

(50)   Map principle

i.   Every A'-position binds some A-position at LF
ii.   All and only A'-positions within the S'-system
      locally bind some A-position at every level
      of syntactic representation.

From now on, for ease of reference, we will reserve the term operator to a category meeting the adequate semantic criteria (e.g. defining a range etc...) that are in an A'-position within the S'-system. Note that (50) implies that every operator in this sense binds locally some A-position at every level of representation.

5.3.4.   Assuming that the classification of A'/A relations we have given in (49) is descriptively adequate, it is natural to wonder on what basis the language learner successfully classifies the particular A'/A relation he is exposed to.

Consider first the case a language like English or French, in which

A* is always empty in such constructions as wh-movement. Clearly, we may restrict our attention to cases when A* is a Case marked position, since, otherwise, it cannot be locally A'-bound.

We have seen in 4.2.2 that the distribution of Case-marked empty ↑-categories is regulated at S-structure (and perhaps at LF as well). In French and English, such elements are excluded when unbound (perhaps the unmarked situation in U.G.). Therefore, binding by an available overt A'-binder as in wh-questions, or by an abstract A'-binder as in purposives must be postulated in order to avoid an unlawful Case marked empty ↑-category. If we make the extra assumption that abstract A'-binders are always operators, it will follow that these constructions will be automatically be ascribed the right place in (49). Notice incidentally that to the extent that this scenario is plausible, it provides independent support for the assumption that the distribution of Case-marked empty categories must be checked at S-structure.

Consider next the case of a language like Italian. Standard Italian is identical to English in the relevant respects except for the fact that it allows null subject in tensed clauses.[24] By the same argument as above, we could conclude that:

    i.   A'-binding of empty subjects of tensed clauses
           may be postulated to hold either at S-structure
           or at LF (or, of course, not to hold at all).
    ii.   A'-binding of other Case-marked empty positions
           must be postulated to hold at S-structure.

However, this would make incorrect predictions. It would, for example, predict that "wh-movement" from subject position of a tensed clause need not obey the Subjacency Condition, contrary to fact, because in that case, A'-binding could be postulated at LF only (say, in a

relative clause, a gap in subject position could be coindexed at LF only with the head of the relative). Furthermore, the problem becomes sharper if we consider the constructions falling under (49iii) and the difference between (49iii) and (49ii).

Consider first (49iii). In all the constructions falling in this category, syntactic resumptive pronouns are not available in all the positions (subject only in Vata, oblique in Haitian and Modern Hebrew) Again, if the simple scenario we have outlined was exhaustive, we would expect a dual behaviour from these constructions. They should fall under (49vii) when A* is empty and under (49ii) otherwise. Of course this is not the case. It suggests that some notion of paradigm uniformity is involved, which, although not easy to formalize, is fairly clear. It would require that, to the extent that no principle of grammar is violated, a given syntactic construction be identified and be analysed in a uniform fashion by the language learner. So, for example, if an abstract operator must be postulated for relativization from some position, then, by paradigmatic uniformity, relativization from any position will be postulated to involve an operator, to the extent that it is possible.[25] Of course, the same reasoning would cover the Italian cases as well. Note that this account assumes a very marginal role is played by the empty/non empty distinction for A* in A'/A relations since it is crucially not taken into account by the relevant notion of uniformity. In contrast to (49iii), anaphoric resumptive pronouns in (49ii) are typically found in every A*-position. Consequently, no S-structure binding needs to be postulated (for if it could, we should expect PG licensing for example). This could be achieved for instance, by assuming that LF binding represents the unmarked option whenever possible (in particular,

only if it does not contradict paradigm uniformity requirements).

6. Clitic Constructions

Limiting ourselves to the most extensively studied cases of clitic constructions, namely clitic constructions in the Romance languages, let us investigate where they fall within our classification of A'/A relations. Let us restrict ourselves mostly to objective and dative clitics for the time being.

6.1. Chomsky's analysis

Chomsky (1982) reports an observation due to L. Rizzi according to which clitic constructions do not license parasitic gaps. A comparison of the two members of the following pairs establishes this point:

(51)   i.   quali libri$_i$ gli dobbiamo far mettere [e]$_i$ nello
           la scafale [invece di lasciare [e]$_i$ sul tavalo]

        ii.   gli li$_i$ dobbiamo far mettere [e]$_i$ nello la scaffale
           [invece di lasciare [e]$_i$ sul tavolo]
           "we must make him put them on the shelf instead
           of leaving on the table"

       iii. quel document$_i$ avez-vous fait signer [e]$_i$ par le
           président [en mettant [e]$_i$ en évidence sur son bureau]

       iv. vous l$_i$'avez fait signer [e]$_i$ par le président [en
           mettant [e]$_i$ en évidence sur son bureau]
           "you had it sign by the president by obviously
           putting (it) on his desk"

Whereas (51i) and (51iii) are relatively acceptable, thereby demonstrating that the relevant structural description may be met, i.e. that at least

one structural analysis of (51i) and (51iii) is such that the wh-phrase in COMP locally binds both empty categories, (51ii) and (51iv) are not acceptable. This does not suffice to establish that clitic constructions do not license parasitic gaps. We must first show that the clitic li (resp. l') locally binds both empty categories. Indeed, if the second [e] in (51iv) for example, is unbound, the structure is ungrammatical but for other reasons.[24] Let us reason on the pair (51iii)/(51iv). First, note that the instrumental adverbial clause hangs from VP (cf. Williams (1975), Reinhard (1976)) which, in the present case, can only be the matrix VP for obvious semantic reasons. The mean expressed by the instrumental clause bears on the causation, i.e. on the verb faire. Furthermore, the construction in (51iii)/(51iv) is a Faire-par construction, which, we may assume following Burzio (1981), crucially does not involve any alteration of the embedded VP,[27] so that the structure of (51iv) is as indicated in (51v) below (irrelevant details omitted):

(51) v.


```
                          S
              ┌───────────┴───────┐
            NP                    VP₁
           /         ┌─────────────┼──────────────────┐
          /         V             VP₂                 Adv
         /          │         ┌────┴─────┐             │
        /           │        V    NP    PP      en mettant [e]
       /            │        │    │     │        en évidence sur
   vous le₁+faire signer [e]₁ par le président   son bureau
```
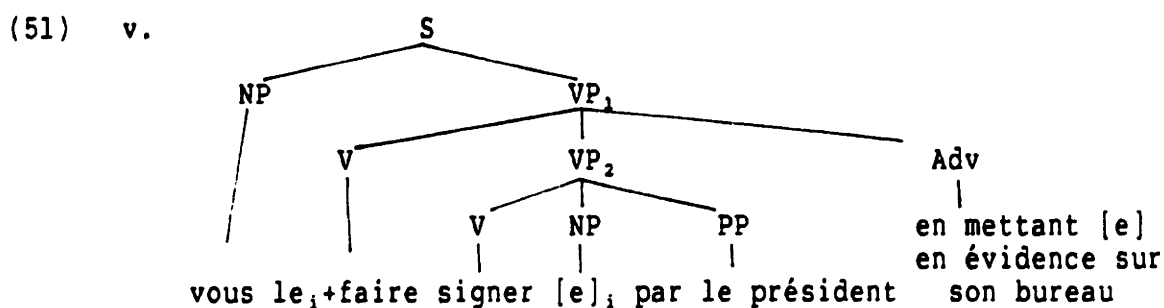
In which neither e c-commands the other (the first one does not c-command the second one only because of the intervening maximal projection of VP₂) and both are locally-bound by le.

Suppose now that, following Chomsky (1982) we take clitics to be A'-binders of the empty category they are associated with, as in (52) below:[28]

(52)   Jean   le   voit   [$_{NP}$ e]
            Ⳑ A' binding⟶

In order to explain that clitics do not license parasitic gaps, Chomsky
(1982) makes the following proposal. First, he assumes that a clitic is
an argument, and forms a chain with the empty category it locally
A'-binds. So (le,e) in (52) forms a chain. Then, because the clitic li
of (51ii) locally A'-binds two empty categories e_1 and e_2, it heads two
chains (li, e_1) and (li, e_2) each assigned a θ-role. Chomsky (1982)
concludes that this constitutes a violation of the θ-criterion, because
the argument li in (51ii) receives two θ-roles. Chomsky (1982) further
notes the problem of how to differentiate between clitic constructions on
the one hand and Heavy NP shift on the other, since, in both cases, we
find an argument in an A'-position, yet in the case of Heavy NP shift,
parasitic gaps are licensed.

In order to make the appropriate distinction, Chomsky proposes that,
in Heavy NP shift constructions, the argument in A'-position is member of
a chain since, as argument, it requires a θ-role, but contrary to the case
of clitics, it enters in a chain only with its "traces", i.e. with the
position from which it has moved. The argument thus does not form a chain
with the parasitic gap - by assumption in this system a base generated gap
- and enters in only one chain. A clitic, on the other hand, base
generated in an A'-position will form a chain with whatever empty category
it locally-binds. A number of features of this analysis are incompatible
with our earlier assumptions. First, note that the distinction between
clitic constructions and Heavy NP shift constructions is drawn on the
basis of the way in which each is derived. This crucial appeal to manner
of derivation is not formalizable within our system of assumptions.

Secondly, our system of assumptions is also incompatible with some implicit assumptions underlying Chomsky's proposal. Consider the assumption that, because an NP is an argument, it must be a member of a chain which is assigned a θ-role. Although this might follow from some version of the θ-criterion, it does not follow from the one we adopt. Recall that the θ-criterion states:

(53)   i.   Every chain containing an argument is assigned
            a unique θ-role (i.e. contains one and only
            one θ-position)
    ii.   Every chain containing a θ-position contains
            one and only one argument.

Chomsky's conclusion that an NP argument must be in a θ-marked chain only follows if it is further assumed that every argument must be a chain, and of course, this depends on the particular theory of chain formation algorithm one adopts. Besides requiring a radical reformulation of our theory of chain formation, this modification does not seem to me to be desirable. There appear to be constructions involving arguments (in A'-positions) which we certainly do not want to incorporate in a chain. Such cases are Topics in English Topicalization, Left Dislocated constituents, heads of restrictive relative clauses[29] (which partly motivated our conclusion that the Map principle does not reduce to the θ-criterion). For example, the assumption that such elements enter into chains would prevent any generalization concerning the properties of NP-traces (i.e. elements not in chain-intitial position). Suppose however, that we do modify our assumptions along the lines required by Chomsky's analysis. It still does not follow that Chomsky (1982) system of assumptions has the right consequence. For recall that at the level of S-structure and LF, the objects to which θ-roles are assigned are chains

and not particular NP's. In a construction like (51ii), even if the clitic li is assumed to head two distinct chains, no violation of the θ-criterion ensues, for it may still be true that there is a one to one correspondance between chains and θ-positions. In order for the correct conclusion to follow, we would need to reformulate the θ-theory as well and make the θ-criterion (which would entail Chomsky's implicit hypothesis stated above, that an argument in an A'-position must be linked to a θ-position in order to avoid a θ-criterion violation).

6.2. Clitic constructions as LF-A'-binding

Let us pursue along different lines. Sticking to our earlier account for Heavy NP shift constructions, whereby the argument in an A'-position at S-structure (inheriting its semantic function through this binding) and thus licenses parasitic gaps, we need to draw the required distinction between Heavy NP shift on the one hand and clitic constructions on the other. If we adopt Chomsky's suggestion that the clitic is an argument, it becomes natural to assume that clitic binding of an empty category is the LF counterpart of Heavy NP shift in table (49), i.e. falls under (49vi), since, as we have see, clitics do not license parasitic gaps. We further predict, if this assumption is correct that the A'/A relation between a clitic and its associated empty category neither obeys the Subjacency Conditon, nor triggers WCO effects. Let us begin to check the validity of these predictions by addressing the question of whether the relation clitic/associated empty category obeys the Subjacency Condition. It is fairly clear that there is some locality condition on Cl/e relations which does not reduce to the Subjacency Condition. The following paradigm illustrates this point:

(54)   i.   Jean est semblable a Pierre

          ii. Jean lui$_i$ est [semblable e$_i$]

(55)   i.   Jean considéra [$_K$ Pierre semblable a Marie]

          ii. *Jean lui$_i$ considéra [$_K$ Pierre semblable e$_i$]

          iii. A qui$_i$ Jean considéra-t-il [$_K$ Pierre semblable e$_i$]

(54) shows that the object of the Adjective _semblable_ can cliticize onto a superordinate verb. (55ii) demonstrates that this cliticization can be excluded even though [e] of (55ii) is subjacent to _lui_ of (55ii). That [e] is subjacent to _lui_ in (55ii) is shown by (55iii): the category K is not a category with COMP (it is a small clause, according to Stowell (1981) and Chomsky (1981)), so that [e] is directly subjacent to _a qui_ in (55iii), since this sentence is grammatical. This conclusion holds a fortiori of the pair (lui, e) of (55ii) since fewer nodes intervene between the two members of this pair than between the two members of the pair (a qui, e) of (55iii). The existence of this locality condition (which, in the present framework can be either some version of Government, or some version of Principle A of the Binding Theory) poses a problem. The set of configurations that it allows seems to be a strict subset of those allowed by the Subjacency Condition.[30] It therefore does not appear possible to directly test whether the Cl/[e] relation obeys the Subjacency Condition.

Let us therefore turn to the question of whether clitic constructions trigger WCO effects. We need to find a configuration in which a clitic locally binds both its associated empty category and some non-empty ϒ-category. The relevant configuration is not easy to construct (except, perhaps for "PP"-clitics as French _en_, _y_). Indeed, consider the usual Cl/[e] situation where ... contains some non empty ϒ-category P.

(56)   Jean $[_{VP}$ $[_{V}$ Cl+V] $[_{NP}$ e] ...P...

As (56) suggest, in general, the c-command domain of the clitic is identical to that of its associated empty category, so that there will always be an ambiguity as to what the local binder of P is (i.e. Cl, or [e]). We need to find a case in which the respective c-domains of Cl and [e] are distinct, which may only arise if the clitic does not appear on the verb governing the empty category, i.e. in causatives or restructuring constructions. In particular, we are led to use structures parallel to those used in examples (51).

Now, if clitics do not induce WCO effects, and if (51) meets the relevant structural desiderata, we should expect (57ii) to contrast with (57iii) and (57iv):[31]

(57)   i.   Vous avez fait signer ce document$_i$ par le président
              en le$_i$ mettant en évidence sur son bureau

         ii. Vous l$_i$'avez fait signer e$_i$ par le président en le$_i$
              mettant en évidence sur son bureau

         iii. Quel document$_i$ avez-vous fait signer e$_i$ par le
              président en le$_i$ mettant en évidence sur son bureau

         iv. Vous avez fait signer chacun des documents$_i$ par le
              président en le$_i$ mettant en évidence sur son bureau

This expectation is fulfilled. There seems to be a significant contrast between (57iv) and (57iii) on the one hand, and (57ii) and (57i) on the other. These last two examples are perfectly acceptable, while the others exhibit WCO effects. This supports the classification of clitic constructions in (49vi).

We have not used reflexive (or reciprocal) clitics in our examples. We can, however, assume that the conclusion we have reached so far may be extended to them without direct empirical problems. I have not been able

to construct examples similar to (51) and (57) supporting this extension.[32] We can also extend our conclusion to subject clitics (except, of course, for the argument character of the subject clitic, when it is not associated with a θ-position). However, because a subject position and its associated clitic have always identical c-command domains, we can neither test for parasitic gap licensing, nor for WCO effects (nor for Subjacency violations).

Summerizing, we see that the relation Cl/e should be an A'/A relation established at LF, where the clitic is considered an argument (except for some cases of Subject clitics, e.g. "[e] il faut partir"). This compatible with the observed data: directly in the case of PG or WCO effects since, in the relevant structures, neither are parasitic gaps licensed, nor WCO effects triggered; indirectly in the case of the Subjacency Condition violations since the prediction that the Cl/e relation does not obey the Subjacency Condition is compatible with the lack of observation of such violations.

Now note that this treatment of Cl/e relations is incompatible with our formulation of the Map principle (50) since it postulates the existence of an A'-position within the S' system, that does not bind some A-position at S-structure. Let us assume for the moment that clitics are exceptions to (50ii).

6.3. Problems with clitics as LF A'-binders

We can now turn to the more interesting question of what the theoretical status of the empty category associated with a clitic is. We limit ourselves to the case of non-subject clitics here. Consider again a

simple case of clitic construction:

(58)   i. $[_{VP}$ $[_V$ Cl + V $]$ ... $[_{NP}$ e$]$ ...$]$
         ii. $[_{VP}$ $[_V$ Cl$_i$ + V $]$ ...$[_{NP}$ e$]_i$ ...$]$

We have so far assumed that its S-structure representation was as in (58i), and its LF epresentation, where clitic binding is introduced as in (58ii).

6.3.1.   Note first that NP* in (58) must be the head of a chain. Indeed, if NP* is a member of an A-chain, this chain must contain some argument A at S-structure. If A is NP*, NP* is in chain initial position, since, as we have discussed in Chapter II modulo our general assumptions, arguments only appear in chain initial position. If A≠NP*, we will get a θ-criterion violation at LF. At S-structure A receives its θ-role through its comembership to a chain with NP*. At LF, because of the presence of Cl, this chain is broken (NP* becomes its head) so that A is no longer member of a θ-marked chain. Therefore, if NP* belongs to an A-chain, it must be in chain initial position.

Suppose next NP* is not a member of an A-chain at S-structure. Then we end up with a problem at LF: a clitic, as an argument (recall we are not considering subject clitics), must be linked to a θ-position. This requirement is subsumed under the more general conclusion reached in II 2.6., that locally A'-bound elements must be arguments, hence must be assigned a θ-role. In other words, because Cl is an A'-binder, the NP it locally binds, e.g. NP*, must be an argument and be member of a θ-marked chain, since it is in an A-position (arguments in A'-position do not need to be in a θ-marked chain). This option is thus excluded. Therefore, NP* must be the head of a chain.[33] Recall now that we have argued earlier (II,

6.2) that an explanation of the distribution of empty categories understood the truth of the following assertion:

(59)   Chain initial ϒ-categories are either Case-marked or PRO

From (59) applied to (58), we derive that NP* in the LF representation (58ii) is Case-marked. Indeed, it is by assumption locally A'-bound, hence not PRO. Modulo the additional, unmarked assumption that the case marked status of the position occupied by NP* does not change in the course of the derivation from S-structure to LF, we derive that, at S-structure, NP* is a Case-marked empty ϒ-category (non locally A'-bound), i.e. pro.

This conclusion is incompatible with our earlier assumptions. Recall that parasitic gap constructions provide extremely strong evidence that the principles responsible for the distribution of pro (ECP) have to hold at S-structure (cf. (24) in III.4.2.2. above). If NP* is pro at S-structure, a structure like (58i) behaves with respect to these principles exactly as (60):

(60)   i.   Jean a mis pro dehors
         ii. John put pro outside

since its relation to Cl is not established at that level. By (24), examples (60), hence structure (58i) is ruled out at S-structure since it contains an illicit pro. In other words, we are forced to conclude that our network of assumptions is inconsistent.

6.3.2.   Let us recapitulate our assumptions concerning non subject clitics.

(61)   i.   Clitics occupy A'-positions
         ii. Because clitic constructions do not license
             parasitic gaps, they involve LF A'-binding
         iii. Because clitics do not trigger WCO effects, we
             have assumed that they are arguments.[34]
         iv. Because it must be in chain initial position, an
             empty category associated with a clitic is Case
             marked.

Assumptions (61i), (61ii) and (61iii) are closely interdependent in the context of our independently justified network of hypotheses concerning A'/A relations. We cannot give one of them up, without giving up the others. Consequently, we face only two minimal alternatives: either we give up (61i,ii,iii) or we reject the implication in (61iv), i.e. (59) in its present form, since its premiss, i.e. that the empty category associated with a clitic is in chain initial position, heavily rests upon our major theoretical assumptions, made in Chapter II.

Let us start with the assumption that we give up (59). Because (59) was meant to account for the distribution of ?-categories in chain initial position, this step might seem costly. However, notice that the argument given in 6.3.1 was based on the assumption that (59) must hold at LF. Since we have not specified so far at what levels (59) was meant to apply in order to achieve the desired result, we might suppose that we restrict its relevance to S-structure configurations only. So assume that we replace (59) by the more precise (62):

(62)   At S-structure, a chain-initial ?-category is
              either Case-marked or PRO.

We must of course verify that this additional restriction imposed on the scope of (59) would have no undesirable consequences. Let us postpone this question and rather consider whether such a move will permit us to solve the inconsistency pointed out earlier.

Consider again (58ii), i.e. the structural description of a clitic construction at LF. The argumentation leading to the conclusion that NP* is in chain initial position is still valid. However, this conclusion is now compatible with treating NP* as a Caseless γ-category at LF, hence, by extension, as a PRO at S-structure.[35] In other words, we now deduce that the structural descriptions of a clitic construction at S-structure and LF are respectively as in (63i) and (63ii) below:

(63)   i. $[_{VP} [_V Cl+V]...[_{NP*} PRO]$ ...

        ii. $[_{VP} [_V Cl_i+V]...[_{NP*} e]$ ...

Consider the S-structure (63i). In the usual case of a structure containing a clitic construction, and perhaps in all of them, we observe that the empty category is in a configuration of structural government with respect to the verb on which the clitic is affixed. Take this observation in conjunction with the principles of the Binding Theory as they apply at S-structure. Consider first Chomsky (1981) version of the Binding Theory on this matter, the most widely accepted version. According to this theory, PRO cannot be governed. How is this compatible with (63i)? One possible move, taken by Jaeggli (1980) consists in assuming that the presence of the clitic morphology on the verb signals that the "governing property" of the verb is cancelled (In Jaeggli's terms, the clitic absorbs the government feature of the verb)[36]. The assumption that NP* in (63i) is not governed by $[_V Cl+V]$ entails that NP* is neither assigned Case nor θ-role from [Cl + V] since both of these properties are relational, and transferred under government only. Recall why the presence of some empty category K, argument of some predicate P at some level L is postulated. Chomsky (1981) proposes that K must be

postulated in order for predicate P to assign some θ-role T to K at L, i.e. in order for the θ-criterion to hold at level L. If we indeed assume that Cl in (63) absorbs the θ-role V would assign to NP* in the absence of NP*, the only <u>principled</u> theoretical motivation for the existence of NP* dissolves. In other words, the logic of this assumption would lead to treating predicates as in (63i) as syntactically intransitive, i.e. lacking an object position and of course, this conclusion is not acceptable.

The existence of an empty category associated with a clitic has been persuasively argued for extensively in the past few years (cf Kayne (1975); Rouveret & Vergnaud (1980); Jaeggli (1980); Burzio (1981))[37]. Getting ahead, assume the version of the Binding Theory, we will develop in more detail in Chapter IV. According to this Binding Theory, PRO may be governed. If it is, it behaves as an anaphor and must be bound in some local domain D. Consider again (63i), i.e. the S-structure representation of a clitic construction. If NP* is present, we have argued above that PRO is governed by V. It must therefore be bound in D, i.e. in the first S dominating V. Here we must distinguish two cases: non reflexive clitics and reflexive clitics. As far as non reflexive clitics are concerned, this conclusion is empirically unacceptable. It is observationally true that non reflexive clitics behave with respect to the Binding Theory exactly as if NP* were a pronouns, i.e. they must be free in the local domain D. So we get an empirically inadequate prediction.

Concerning reflexive clitics, this conclusion is a priori compatible with the data. Clitic constructions with reflexive clitics do seem to behave as if NP* were an anaphor, in that it must be bound in some local domain D. As a first conclusion, we can state that:

(64)   i.   (63), i.e. LF A'-binding of a Caseless position,
             cannot adequately represent the general case for
             clitic constructions.
        ii. (63) may be an appropriate representation for
             reflexive (or reciprocal) clitic constructions.
        iii. If (64ii) is correct, we need to investigate the
             empirical consequences of the restriction from
             (59) to (62).

In particular, expanding on (64i), we may conclude that the system of
assumption (61i), (61ii), (61iii) (and (62) replacing (61iv)) fails to
provide an adequate analysis for the general case of non-subject clitic
constructions. We will therefore assume from now on that the relation
between a (non reflexive/reciprocal) clitic and its associated empty
category is not one of A'-binding. Of course, the question now arises of
how best to characterize clitic constructions. We address this question
in the Appendix to section 6.

7. Strong Crossover

7.1. Amongst the properties that have been ascribed to A'/A relations, and
more particularly to variables, i.e. locally A'-bound elements in an
A-position, we find the so-called Strong Crossover (henceforth SCO)
phenomenon. In English, SCO can be illustrated by the following examples:

(65)   i.   who$_i$ did he$_i$ see e$_i$

        ii. who$_i$ did he$_i$ say e$_i$ left

        iii. who$_i$ did he$_i$ think you saw e$_i$

Reverting back for the moment to the usual view concerning the existence
of derivations from an independent level of D-structure to S-structure, we

can outline the usual description given for the SCO. SCO arises when wh-movement has taken place from some A-position A* asymetrically c-commanding some pronoun P in an A-position at D-structure, to an A'-position A'* which c-commands P. In a right branching language like English, this will arise only if P is in "between" A'* and A*. The movement can thus be said to "crossover" P.

In such a configuration, the pronoun P cannot be understood as coreferent with the wh-trace A* as the examples in (65) illustrate. For example, (65ii) cannot be understood as a general question meaning <u>for which</u> <u>person</u> x, x <u>said</u> <u>that</u> x <u>left</u>.

In keeping with this description of SCO, we have mentioned earlier that the notion of variable used to be defined precisely as the trace of movement to an A'-position. In (65), [e] in each case would be characterized as a variable, and the ungrammaticality of the examples in (32) can be attributed to what is referred to in Chomsky (1981, Chapter 3) as <u>Binding</u> <u>principle</u> <u>C</u>, which states:

(66)   Principle C: R-expressions must be locally A-free

where R-expressions are variables or names. Since, by assumption, A'* c-commands P, and P c-commands A*, A* is locally A-bound by P, hence not locally A-free. So are the examples in (65) ruled out. (Note that we assume throughout that there are no intermediate traces in COMP, that may act as bindees, cf. III.9 below).

7.2. Within our framework of assumptions, such an account is not possible since it in fact requires a derivational characterization of the notion of variable. Recall that we have defined a variable as being a $\bar{\gamma}$-category in

an A-position, which is locally A'-bound.   Recall also that we have shown

in III.5 above that the distinction empty/non empty for A* played no role

in the typology of A'/A relations.

Let us now examine the general case of a SCO configuration in the

light of this definition of variable. The general case of SCO meets the

structural description   (67), where P_i is a Case-marked Ɣ-category in an

A-position wh$_i$ch is an argument:

(67)   ┌─ local binding ─┐   ┌─local binding─┐
             ↓   ↘   ↙   ↓
           A'$_i$*   P$_i$ A$_i$*
                                                    |
                                                    e
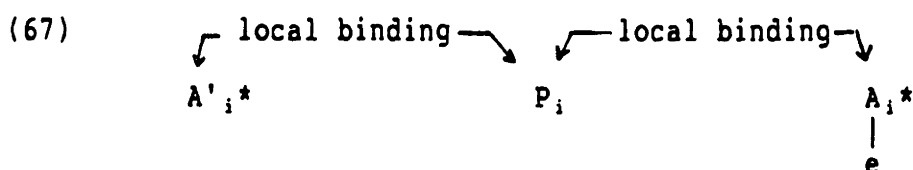
Given our definition of variable, it is P_i and not A_i which is

characterized as a variable, since P is locally A'-bound, while A* is not.

Note incidentally that, if contrary to what we assumed, P_i is not

Case-marked, or not an argument, or not a , Ɣ-category, (67) would be

excluded by independent principles, since variables must be Case-marked

arguments (and Ɣ'-categories cannot be locally A'-bound).

Moreover, A* is locally A-bound by P, so, as an empty Ɣ-category, it

may be either PRO, pro, or NP-trace.   Clearly, we may restrict our

attention to cases in which it would have been a licit target for

wh-movement, i.e. where it is a Case-marked (empty) argument (i.e.

pro).[38]

7.3.1.   Let us first assume it is at S-structure.   Then, a structure

meeting (67) will be well-formed if and only if the link A*'/P is, and the

empty category A* is licit. Recall that in English, wh-movement is

analysed as an S-structure A'-binding of an A-position, by a wh-phrase in

an A'-position. It is clear that the examples in (65) meet the structural description (67). By examining the examples below, we can see that the structures in (65) are ruled out for two reasons, independent from each other and from Principle C:

(68)   i.   *who$_i$ did he$_i$ see John
        ii. *he$_i$ saw e$_i$

(69)   i.   *who$_i$ did he$_i$ say John slept
        ii. *he$_i$ said e$_i$ left

As (68i) (resp 69i) shows, English does not allow lexicalized (i.e. non empty) ?-categories as variables. This observation explains the ungrammaticality of (65i) (resp 65ii), since it contains a lexicalized variable, namely <u>he</u>. In other words, the examples of (65) are each ruled out because English does not allow resumptive pronouns, i.e. the link A*'/P is not well-formed.

Furthermore, as is shown by (68ii) (resp 69ii), English does not tolerate pro (i.e. a Case-marked empty category in chain-initial position) in these environments, so that (65i) (and 65ii) are ruled out for the additional reason that they contain illicit pro's.

This argumentation can of course be extended to other languages which share the relevant properties with English, e.g. French, without any appeal to Principle C. We can summarize the discussion by stating that:

(70)   An S-structure configuration meeting the structural
             description (67) will be well-formed with respect
             to grammar G only if:[39]

        i.   The relation A'*/P is permitted by G: if P is not
             empty, G tolerates syntactic resumptive pronouns;
             if P is empty, G tolerates empty variables at
             S-structure.

   and ii.   G tolerates the presence of pro in position A* at
             S-structure.

The simplest case is (71i), for the discussion of 7.3.1 straightforwardly covers it. Consider (71ii). In that case, the LF representation (72ii) would correspond to an S-structure (72i):

(72)   i. A'_i* P_i A_i*

        ii. A'_i* P_i A_i*

We have described this kind of situation in section 5 above as involving anaphoric resumptive pronouns, i.e. as involving a category A'-free at S-structure, which gets interpreted as a variable at LF. We can reason on these cases in very much the same manner we have for the cases in 7.3.1 above. A sentence with the set of structural descriptions (72) will be well-formed if each of its representations is: i.e. if pro in A* is licit both at S-structure and at LF, and if anaphoric resumptive pronouns (whether empty or not) are.

The case (71iii) is more interesting. We might call it LF-SCO. According to its description, we would get the following derivation:

(73)   i. (S-structure) P_i ...A_i*

        ii. (LF) $A_i'*$ ... P_i ...A_i*

(where A'* has moved from A*). Although it is plausible to assume that there is a possible grammar in which (71i) or (71ii) would yield grammatical structures (Standard Italian is a case for (71ii)), it seems much less so for (73). Suppose it were possible in some grammar G. G might be taken to resemble English except for the fact (and correlated changes) that the following derivation would be well-formed:

(74)   i. (S-structure) you told him_i that who_i should leave

        ii. (LF) who_i [you told him_i that e_i should leave]

(Say G permits anaphoric resumptive pronouns and freely allow empty subjects). Surely we do not expect to find such a G. What could rule out (74)? Certainly not some LF restriction if we admit that cases like (72) might exist, since (72ii) is essentially identical to (74ii). Rather, it should be (74i) that should be taken as ill-formed. A natural candidate would be Principle C, as applied to who: taking who to be an R-expression at S-structure, and Principle C to apply at S-structure (as Chomsky, 1981, argues), we would derive the ill-formedness of (74i).

7.4.1. Putting aside the rather speculative discussion of cases like (74), we may conclude from the above discussion of SCO that we have been able to account for the major cases of SCO without any appeal to Principle C of the Binding Theory, for which SCO provided the strongest motivation, and more importantly, without any appeal to principles not independently necessary (Principles governing the distribution of pro, or principles governing the availablity of resumptive pronouns). To be more precise, we have argued that no appeal to the entire scope of Principle C was required. Principle C can be decomposed in the following two propositions:

(75)   i.   Names must be locally A-free
        ii. Variables must be locally A-free

The usual examples justifying the introduction of (75i) have not been affected by our argumentation. In such examples as (76):

(76)   i.   he said John left
        ii. John said John saw you

The matrix subject is usually assumed to be distinct in reference from the

:

embedded subject, as correctly predicted by (75i).[40] Furthermore, if our speculative discussion surrounding (74) has any force, it might be used as an indication that a principle as (75i) is required as a principle of grammar. In which case, (75i) should be extended to cover such expressions as who, what etc...which are not names. For example, we could straightforwardly reformulate (75i) as: $\bar{\gamma}$'-categories must be locally A-free.

7.4.2.   Let us now turn to the question whether (75ii) or some principle along its line is still needed. We have seen that the usual examples adduced in favor of it, e.g. (65), could be explained otherwise. An analysis of why this was possible will give us an indication of how to construct structures that might resist the treatment we have offered for (65). Consider a usual case of SCO illustrated in (77) below:

(77)   who$_i$ did he$_i$ say Mary kissed e$_i$

We have argued that it was not the relation between he and [e] that was impossible, as (75ii) assumes, but rather the relation between who and he on the one hand, and the presence of the empty category in an environment in which it is illicit, on the other. This was possible because the various relations of coindexation between all these elements could be analysed as a sequence of relations of local binding. In (77) we might say that he is "referentially dependent" on who because it is locally bound by it, and similarly [e] is on he for the same reason.

This situation is not necessary, however. Referential dependence (i.e. coindexation) may be reduced to binding only if the various categories involved all agree in number, gender and person ... as a

binding relation requ res. Referential dependence, however, does not require such a feature agreement. In this respect, consider the following examples:

(78)   i.   which person$_i$ do they think you saw e$_i$

         ii. Bill wonders which man$_i$ they think you saw $_e$i

         iii. I asked who$_i$ we said you should see e$_i$

In each case, we cannot have a referential overlap between the subject of the (indirect) question and the empty category. For example, (/3iii) cannot be interpreted as a general question meaning: I asked (Wx. (I and x) said you should see x). Similarly, (78ii) cannot mean: Bill wonders which man x is such that Bill and x think you saw x. Of course, this restriction bears a striking resemblance to the earlier conception of the restriction on SCO. They cannot however, receive the treatment we have offereed for the usual cases of SCO (i.e. the treatment summarized in (70) since no binding relation is possible between the wh-phrase and they (or, we), given that the first one is singular, and the second one is plural. In particular, in (76), it is [r] that is locally A'-bound by the wh-phrase and thus interpreted as a variable, contrary to what was the case in (65).

Before examining the question of what the restriction operative in (78) is, let us make some terminological adjustments and modifications of the indexing system and related concepts. We return to relevant considerations of these issues in 8.2.?.2 below. First, note that we obviously need to complexify the indexing system in order to properly represent the interpretations of phrases involving plural pronouns. Disregarding here some important problems (but cf. Chomsky, 1981, Chapter

5; Lasnik, 1981)) let us simply assume that a referential index is in fact a complex object: more specifically, assume it is a set containing one or more integers with the obvious interpretation. As illustration, consider the interpretation of (78ii) that is not available. In this system, it would be represented as (79):

(79)   Bill$_{(i)}$ wonders [[which man]$_{(j)}$ [they$_{(i,j)}$
            think you saw$_{(j)}$]]

Correlatively, we must make more explicit the notions of freedom and binding relevant to the binding theory. Altough the necessity for these adjustments is made clearer by the considerations of the referential properties of plural NP's, they are in fact necessitated even in the usual cases covered by the Binding Theory. Recall that the Binding Theory requires of certain elements (anaphors) to be bound and of others (pronominals) to be free, in some local domain D. The elements required to be bound (NP-traces, reflexives, reciprocals...) cannot have split antecedents, nor mere overlap with their antecedents (viz *John$_{(i)}$ told Bill$_{(j)}$ about themselves{i,j}, *They$_{(i,j,k)}$ like each other$_{(i,j)}$. In other words, they must have a reference identical to that of their antecedent.[40]

Without getting into details, let us acknowledge that we must therefore understand the notion **bound** accordingly:

i.   x is bound to y iff - y c-commands x and
                          - y and x have identical referential
                            indices[42]

We cannot however, define the notion **free** as being the opposite of **bound**. If that were the case, the binding theory would require of NP's that must be free (e.g. pronominals, names) not to be bound. This would not be a

strong enough requirement. We must take _free_ to mean non overlap in reference (i.e. as meaning that the -pairwise- intersection of the respective referential indices is the empty set). Cf. Lasnik (1981); Chomsky (1981) Chapter 5 for relevant disscussion. This point is illustrated by such cases as *They$_{(i,j)}$ saw him$_{(i)}$, *John$_{(i)}$ said that Bill$_{(j)}$ like them$_{(i,j)}$ ... Consequently, let us sum up the discussion as:

(80)   i.   x is bound by y iff, y c-commands x and x and y
             have identical referential index set.[42]
      ii.   x is free in D iff,$\forall$ y, y$\in$ D/y c-commands x, the
             interaction of the referential index sets of x
             and y is the empty set.

Returning now to cases (78), we see that two options are open to us in order to account for this non overlapping reference restriction. For concreteness, let us reason on example (79). We might argue that it is the relation between the wh-phrase _which_ _man_ and the pronoun _they_ that is illicit. Or we might argue that it is the relation between _they_ and [e] that must be ruled out.

Consider the first option. _They_ is not interpreted as a variable since it is not bound by the wh-phrase in the sense of (80i). It is rather identified as pronoun. In the spirit of Aoun (1982) proposal concerning the extension of the Binding Theory from a theory of A-binding to a theory of X-binding, we could argue that the relation _which_ _man/they_ is ill-formed, as follows: a slight reformulation of principle B of the Binding Theory could state that pronominals must be X-free in some local domain D including the wh-phrase. That move would rule out structures like (79).[43]

Consider however, a structure like (81):

(81)   i.   *which man$_{(i)}$ does Bill$_{(j)}$ think [COMP [they$_{(i,j)}$ saw e$_{(i)}$]]

        ii.   which man$_{(i)}$ e$_{(i)}$ told Bill$_{(j)}$ [that they$_{(i,j)}$ should leave]

In order to rule (81i) out without affecting the grammatical (81ii), we would have to appeal to an intermediate trace in the COMP of the embedded clause, so that they, bound by it, would not be free in the embedded S'. Note in particular that we could not assume that [e], a variable, falls under this revisited Principle B, for it would not only exclude the overlap between they and [e] but would exclude wh-movement in simple clauses altogether (who$_i$ e$_i$ left). This account predicts that French or Italian should allow structures parallel to (81i) since, S not being a bounding node as Rizzi (1982) and Sportiche (1981) argue, there is at least one derivation in which the embedded COMP contains no intermediate trace. Of course, this predication is incorrect: the French or Italian equivalents to (81i) are ungrammatical.

We must therefore resort to the second option. This second option would rule out as ungrammatical the relation between the variable [e] and the pronoun they in (79) or (81i). An obvious candidate for expressing this restriction is clause (75ii) of Principle C, namely variables must be locally A-free.

Notice the effect of introducing definitions (80) for free and bound. If free meant not bound, (75ii) would be trivially true since, by definition, variables are locally A'-bound, hence not locally A-bound (given that no element has two local binders). If free does not mean not bound, as we suggest, a variable may be locally A'-bound and not be locally A-free, so that (75ii) is not trivially true. Note furthermore that the notions of free and not bound will coincide when the referential

index sets of the NP's involved are singletons. When they do, we need not invoke (75ii). This is what we have done for the cases of SCO like (65).

8. Parasitic gaps and WCO

8.1. The Basic Parallelism of PG and WCO Structures.

In section 4.2 and 4.3 above, we have briefly discussed parasitc gap (PG) structures and weak crossover (WCO) configurations, in connection with our goal of establishing classificatory criteria for A'/A relations. As Chomsky (1982) amply demonstrates, the study of these constructions has proved an extremely fruitful testing ground for hypotheses about the structure of the theory of grammar. In particular, and this is why we now proceed to a more systematic investigation of these phenomena, their syntax will bear on a number of issues that we have so far discussed, and will discuss in Chapter IV.

As a cursory examination of the PG constructions and WCO configurations reveals, these structures are strikingly similar and it would be quite surprising if it turned out that their respective properties are handled by unrelated principles of grammar.

We have introduced PG constructions as meeting the S-structure (hence LF) schema (82) (recall the proviso we made concerning the non existence of intermediate traces in COMP; cf. III.9 for further discussion):

(82)   ┌──────local binding ──┐
       ↙   ↓
    A'*   A*   A   (linear order irrelevant)
      ↖   ↗
       └ normal ┘
    A'/A relation
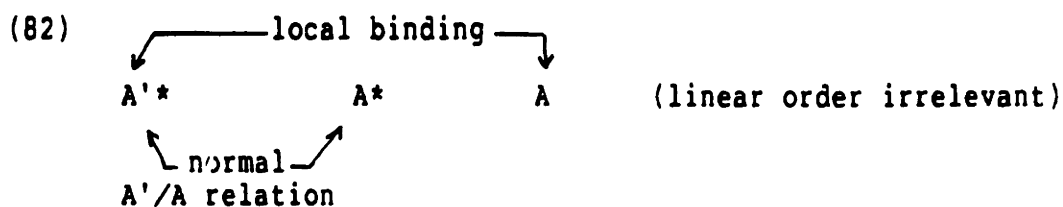
with A an empty f-category; A* is usually referred to as the real gap and A as the parasitic gap.[44]

And we have described WCO configurations as meeting the LF schema:

(83)   ┌──────local binding ──┐
       ↙   ↓
    A'*   A*   A   (linear order irrelevant)
      ↑   ↗
       └ normal ──┘
    A'/A relation
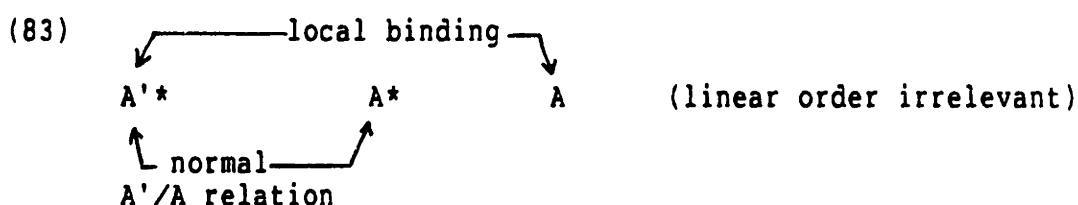
with A a non empty f-category.

In both (82) and (83), we call normal A'/A relation an A'/A relation meeting all the properties it should (with respect to some grammar G) i.e. A'* locally A'-binds A*, A* is a Case marked argument, empty (or not, depending on G and the construction)... In other words, we assume that a structure meeting (82) or (83) should be grammatical if A in it were not locally bound by A'* and filled, say, by a proper name. However, I believe that this presentation is misleading for the implicit reason why these two configurations are distinguished (besides the fact that they have been discovered at different times) does not lie in the empty versus non-empty character of A. When we have made clear what the distinguishing features of each construction are, we shall see that A of (83) does not have to be non-empty in order to be a WCO configuration.

In order to illustrate the difference between PG and WCO structures, consider the following pairs:

(84)   i.   who$_i$ did pictures of e$_i$ please e$_i$

          ii.   who$_i$ did pictures of him$_i$ please e$_i$

(85)   i.   pictures of e$_i$ pleased John$_i$

          ii.   pictures of him$_i$ pleased John$_i$

(84i) is a PG structure, (84ii) a WCO configuration. Abstracting away from Subject-Aux inversion, the structures of the type (85) are constructed from their counterparts in (84) by eliminating A'*, and replacing A* by some proper name (or pronoun...) with the same index. The results sharply differ: (85ii) is perfectly acceptable, while (85i) is totally unacceptable. I believe this is the basis for the PG/WCO distinction. For ease of reference, let us call (85i) the A'-less conjugate of (84i) and (85ii) the A'-less conjugate of (84ii). A better characterization of PG and WCO structures taking these remarks into account is given below.

In a grammar G, a structure S meeting (86)

(86)   ⌐—— local binding —⌐   (linear order irrelevant)

           A'*   A*   A K a $\bar{Y}$-category

           ⌐— normal —⌐   K
           A'/A relation
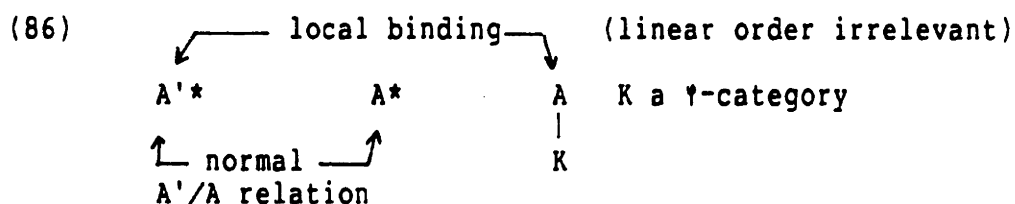
is called a:

    i. WCO configuration iff some A'-less conjugate of S is well-formed at every level of representation. ii. PG structure iff some A'-less conjugate of S is ill-formed at some level of representation.[45]

Having provided these definitions will facilitate exposition of their relevant properties. Note incidentally that, as we have noted and in accordance with (86ii), A'-binding of K must take place at S-structure.

Now suppose we abstract away from the A'-binding of K by A'* in (86), and that K is free. What can K be? Obviously, K must be in chain initial position, so it is either pro, PRO or a pronoun. We see immediately that whether some construction will be analysed as a PG structure or as a WCO structure will depend on some grammar specific properties. Suppose G never allows pro (e.g. English), A'-less conjugates of structures meeting (86), in which K is pro will always be analysed as a WCO case when K is pro and K is in P, as a PG construction otherwise. So we see, as we mentioned earlier, that the dichotomy WCO/PG does not mirror the distinction K is empty versus K is not empty.

We give below examples of WCO configurations in (86) and PG constructions in (87):

(87)   i.   Which people did they photograph e_i without ever
            having met e_i

        ii. John$_i$, I persuaded friends of e_i, to please e_i

       iii. This is a man that enemies of e_i praise e_i

       iv. Which document$_i$ should we hide e_i before someone
            steals a copy of e_i

        v. Which document$_i$ should we hide e_i before a copy of
            e_i gets stolen

       vi. Who did you hire e_i though believing e_i is incompetent

      vii. Who did you fire e_i without John's trying e_i to leave

(88)   i.   Which people did they photograph e_i without
            having met them$_i$

        ii. John , I persuaded friends of him$_i$ to please e_i

       iii. This is a man$_i$ that his$_i$ enemies praise e_i

       iv. Which document$_i$ should we hide e_i before someone
            steals a copy of it$_i$

        v. Which document$_i$ should we hide e_i before a copy of
            e_i gets stolen

vi.   Who$_i$ did you hire e$_i$ though believing he$_i$ is incompetent

vii. Who$_i$ does [e$_i$ sleeping late] bother e$_i$

These examples vary in acceptability. As we shall see, despite the fact that they respectively meet the WCO and PG structures structural descriptions, they form a heterogeneous set that we shall appropriately subdivide as we proceed. Let us, as a first approximation, assume that they are all ill-formed. What accounts for their ungrammaticality?

Consider first parasitc gap structures (37). Clearly, their ill-formedness cannot be attributed to some property of the parastic gap itself. For example, as (89i) and (89ii) show, both gaps are acceptable in a form like (87ii):

(89)   i.   Who$_i$ did you persuade friends of e$_i$ to please John

        ii. Who$_i$ did you persuade friends of John to please e$_i$

Following Koopman & Sportiche (1981) account of these violations, notice that both gaps in structures like (87) are interpreted as variables, since they are both locally A'-bound ɣ-categories. Since precisely in this situation, ungrammaticality arises, it is natural to assume that the following principle of grammar holds:

(90)   Any A'-position locally binds at most one A-position

Of course, (90) has the desired effect as far as PG constructions are concerned despite the indeterminacy as to the level of representation at which it is relevant. Remark however, that this very account extends immediately to structures in (88) and more generally to WCO configurations. The superficial difference due to the fact that, in the

WCO configurations (88), one of the ?-categories interpreted as variable is not empty, is due to the particulars of English syntax. In a language permitting syntactic resumptive pronouns (e.g. Vata) this difference disappears. It similarly disappears in languages like Italian allowing pro in some positions. Since some WCO configurations only obtain at LF, as the following examples involving QR show:

(91)   i.   (S-structure)
             Pictures of everyone$_i$ pleased him$_i$

        ii.   (LF)
             [everyone]$_i$ [pictures of e$_i$ pleased him$_i$]

It is natural to assume that (90) holds at least at LF (and perhaps at S-structure). It is worth pointing out that no particular new assumption is necessary in order to account for WCO effects, once (90) has been postulated for PG constructions. In particular, note that (90) will rule out WCO configurations redundantly, even if some other reason was shown to be relevant to their ill-formedness. One such reason may well have to do with the marginal availability of resumptive pronouns. In English, for example, resumptive pronouns are not allowed in contexts where a gap is possible. (Anaphoric) resumptive pronouns are marginally acceptable elsewhere. Note now that in WCO configuration, A is in fact analysed as a (syntactic) resumptive pronoun (e.g. "it" in (89iv)). It has been noted (cf Chomsky (1982) and references cited therein) that, ceteris paribus, a PG structure is more acceptable than a corresponding WCO structure (at least in English). For example, (87i) would be better than (88i). This might be due to the use of a resumptive pronoun in (88i). (87i) and (88i) both violate (90), but (88i) also uses an only marginally available strategy, which might make it worse. This account makes two predications.

First, it predicts that a WCO configuration in which A stands in a position "accessible to movement" should be worse than a WCO configuration in which A stands in a position not "accessible to movement". This prediction seems fulfilled (although judgments are not as clear as one would wish). (88ii), in which him stands in a position where a real gap would be possible (as (89i) shows) is worse than (88i) in which them is not a possible real gap position.

The second prediction could be verified only in a language freely allowing resumptive pronouns in positions inaccessible to gaps. In such a language, parallel WCO and PG structures should be on the same acceptability level.

8.2. Breach of Parallelism

We have so far assumed that all the structures meeting (86) were ill-formed. This was an expository simplification. We should expect PG structures and WCO structures to present exactly parallel patterns of acceptability (with WCO structures slightly worse, as we have just noted) if their behaviour was accounted for in exactly the same terms, i.e in terms of (90). We can however, make the following observations, which require some explanation:

i.   Heavy NP shift constructions do not trigger WCO effects

ii.   Relative clauses are usually considered to be exempt from WCO (e.g. 88iii is well-formed) (but cf. Higginbotham (1980) for a different
 view)

iii. Considering (87vii) and (88viii) and their A'-less conjugates as in (92) and (93) below:

(92)   i.   Who did you fire e_i without John's trying e_i to leave

ii. You fired Bill$_i$ without John's trying e$_i$ to leave

(93)   i.   Who does [e$_i$ sleeping late] bother e$_i$

ii. [e$_i$ sleeping late] bothers Bill$_i$

we see that contrary to expectation (92ii) is no better than (92i): they are both out. And (93ii) is not worse than (93i): they are both well-formed.

iv. Although the parallelism expected holds for some of the remaining PG/WCO pairs in (87), (88) not mentioned in (i) through (iii) above, it fails for others (e.g. 97v, 88v or 87vi, 88vi) where the PG structure is much worse than its parallel WCO structure. Let us consider all these questions in turn.

8.2.1. Heavy NP-shift

Consider first Heavy NP-shift constructions. Because they do not induce WCO effects, the natural step to take is to somehow make them immune to principle (90). However, because the ill-formedness of PG structures and of WCO structures is handled by (90), we are lead to assume that Heavy NP-shifted PG constructions are grammatical, a conclusion which seems to me reasonable on the basis of acceptability judgments. In other words, we are lead to assume that structures like:

(94)   i.   John offended e$_i$ by not recognizing e$_i$
             immediately his uncle from California

ii. John offended e$_i$ by not recognizing him$_i$
     immediately, his uncle from California

are both well-formed.[46] This assumption permits us to simplify the formulation of (90) and make it empirically more adequate. Recall that we

have essentially assumed that WCO effects were triggered only in the event that A'* was an operator (cf. section 5). Assuming (90) to hold at LF would have incorrectly included anaphoric resumptive pronoun constructions (such as Left Dislocation) among those triggering WCO effects. We can now more adequately reformulate (90) as:

(95)   Operators locally bind at most one A-position

Recall further that we have concluded that operators in A'-position were different from other A'-binders in that the Map principle required of them to locally A'-bind at least one A-position at S-structure (cf. section 5), hence at LF. Putting this earlier conclusion together with (95), we can state the following principle (adapted from Koopman & Sportiche (1981)):

(96)   Bijection Principle
            Every operator in A'-position locally binds one and
            only one A-position at any level of representation.

Many issues arise in connection with (96), that we will not deal with here. Some relevant discussion can be found in Koopman & Sportiche (1981).

8.2.2. Relative Clauses.

8.2.2.1. Turning now to relative clauses, we need an explanation as to why they do not trigger WCO effects. Note that we cannot adopt the same move we have for Heavy NP-shift constructions since we would have to conclude that PG constructions with relative clauses are well-formed, while those with wh-questions are not, a conclusion which seems to be unsupported by

the acceptablity judgments on such structures. Chomsky suggests that the LF representation of a relative clause such as (97) may be as indicated:

(97)   [the man]$_i$ [who$_j$ John saw e$_j$]

i.e. that at LF, the head is not necessarily coindexed with the wh-phrase in COMP. Rather, the identification of indices (i=j) is done by a rule of Predication mapping LF onto a further level of representation called LF'. Assuming (96) to apply both at LF and S-structure but crucially not at LF', we will get a well-formed derivation of a sentence like (88iii), as below:

(98)   i.   (S-structure, LF)
             The man$_i$ [who$_j$ [his$_i$ enemies praise e$_j$]]

        ii.   (LF' by Predication)
             The man$_i$ [who$_i$ [his$_i$ enemies praise$_i$]]

As can be seen, the representations at LF or S-structure do not violate the Bijection Principle.[47]

We can perhaps simplify Chomsky's proposal by assuming a particular version of the Predication rule. Suppose that the Predication rule not only identifies the indices of the head of the relative clause and of the relative clause operator but thereby makes the latter into a non operator. Clearly, the semantic function of a relative clause operator is very different from that, say, of a question operator. In a theory including a Predication rule of the sort discussed, a relative clause operator has no function whatever, once the antecedent of the relativized position has been determined by the Predication rule. It is thus plausible to assume, as we do, that it no longer is an operator at LF. Suppose next that, contrary to Chomsky's proposal, we assume that Predication is an LF rule.

We would then get the following derivation for (88iii):

(100)　　i.　(S-structure)
　　　　　　　The man$_j$ [who$_i$ [his$_j$ enemies praise e$_i$]]

　　　　ii.　(LF)
　　　　　　　The man$\underline{i}$ [who$\underline{i}$ [his$\underline{i}$ enemies praise e$\underline{i}$]]

Because the Bijection Principle only holds of operators/variables relations, it will be neither violated at S-structure, since who only binds one A-position, nor at LF since who is no longer an operator. Notice that, crucially, the same derivation could not be provided for PG structures, since, as (101) shows:

(101)　　　　(S-structure)
　　　　　　　The man$_i$ [who$_j$ [enemies of e$_i$ praise e$_j$]]

The S-structure binding of e$_i$ by the man would violate the Map principle (50): heads of relative clauses are outside the S'-system. Notice also that, in the general case of relative clause construction, the Map principle in a way predicts the existence of an LF rule of Predication. Because the head is in an A'-position (cf. footnote 48) it must, by (50), locally bind some A-position at LF. There must therefore be some process identifying its index to some A-position index (the same would apply to Left Dislocation constructions if we assume that these Left Dislocated constituents bear some index prior to LF).

8.2.2.2. Before leaving the topic of relative clauses, let us examine an alternative proposal made in Aoun (1983) as to why relative clauses are immune to WCO effects. The discussion of this proposal wil bear on some question relevant to our concerns. Aoun (1983) contains both a criticism of Chomsky's suggestion adopted and modified above, and an alternative

proposal as the why WCO effects are suppressed in relative clauses. Aoun notes that, in the following example:

(102) Mary hates the younger brother that their mother prefers

There is a reading in which the reference of _their_ includes Mary the younger brother. In the indexing system of Chomsky (1981, 1982), an element bears not more than one index, so that the representation (103) would be ill-formed:

(103) Mary$_i$ hates the younger brother$_j$ that their$_{i,j}$ mother prefers e$_j$

Consequently the account given for (98) could not carry over to (102). Of course, Aoun further argues that:

> "it goes without saying that the extension of the GB indexing possibilities to allow representations such as (103) will not solve the problem: "their" [in (103)] can be used to designate the set containing Mary and the younger brother, or a larger set properly including Mary and the younger brother...in brief, in order for a pronoun to be interpreted as coreferential with another element, we do not need to coindex this pronoun with the coreferential element. We need only a disjoint reference rule which, in certain contexts, prevents a pronoun from being construed as coreferential with another element (cf. Lasnik (1976); (1981))..."

Aoun (1983) goes on to argue for an alternative proposal to handle the lack of WCO effects in relative clauses and associated problems, which do not concern us directly here.[44] If Aoun's objection to Chomsky's proposal stands, the account given in 8.2.2.1 cannot be maintained. I believe however, that this objection is not well-founded. For not only is some extension of the GB indexing system (i.e. that of Chomsky (1981) is possible, that would permit the analysis of 8.2.2.1 to extend to cases like (102), but it is even required by some facts first pointed out by J.

Higginbotham, that we discuss below.

First consider the following representations:

(104)   Mary$_{(i)}$ hates the younger brother$_{(j)}$ that
           their$_{(i,j,k,...)}$ mother prefers e$_{(i)}$

We see that if, as we have suggested in section 7 above, we also have

referential sets, instead of indices (for plurals only, of course) we can

adequately represent the reading of (102) in which their "refers" to

S={Mary, the younger brother} or any other reading in which it "refers"

to any larger set properly including S it suffices to properly set the

content of the referential set of their, so that Aoun's criticism does not

go through. Note however that we need to reformulate the notions of free

and bound as we have done in (80) in order to avoid the probl... mentioned

in 7.4.2 (this chapter).[49] Now let us proceed to show that we in fact need

to assume referential sets instead of single referential indices.[50]

Remark first that pronominal binding by a QP or a wh-phrase (or its

trace, call it a logical variable), unlike perhaps coreference between

referring expressions, must be stipulated in the indexing structure. A

name and a pronoun may perhaps be assumed to be coreferent yet bear

different indices (cf. Lasnik (1981) for some discussion). However, if a

pronoun is not coindexed (say, at LF) with a logical variable it cannot be

understood as a logical variable itself. More precisely, pronouns may be

either coreferent with some NP (accidentally picking the same reference)

or referentially dependent on some other NP. This must be the case if NP

is not a referring expression, but as Reinhart (1980) shows, may also be

the case with referring expressions. This is why: There is a well-known

condition governing this latter possibility. When a pronoun P is

referentially dependent upon NP*, it must be c-commanded by NP* (This is

(18) of this Chapter). Clearly a non coreference rule cannot work for these cases. It is impossible to list in the indexing structure of P the (infinite and unknown) set of elements upon which P cannot be referentially dependent.

The only alternative is the opposite. If we want to be able to represent the cases of referential dependence, we must indicate in the indexing structure of P which element P referentially depends upon. Importantly, Aoun (1983) is lead by his alternative proposal to adopt this position as well. Now consider the following type of examples (due to J. Higginbotham):

(105)　i.　Everyone told someone that they should leave
　　　　ii.　Everyone believes some man to have told everyone
　　　　　　else that they should leave

Surely these sentences can receive respectively the interpretations given in (106), among others:

(106)　i.　$\forall x$, $\exists y$, x told y that x and y should leave
　　　　ii.　$\forall x$, $\exists y$, $\forall z$, $z \neq x$, x believes y to have told z
　　　　　　that x, y and z should leave.

The conclusion is inescapable. If referential dependence of a pronoun by a quantifier requires coindexing, we are forced to accept that referential indices are in fact, referential sets or arbitrary cardinality (which might require a different notation if non denumerable sets are involved) with the obvious interpretation. For example, if everyone bears index {i}, and someone index {j}, they in (105i) must bear index {i,j}, when it receives the interpretation (106).

8.2.3. PG and Subjacency

As we have mentioned, the parallelism that we expect to hold betwee PG structures and WCO configurations (except, of course, in relative clauses) fails. This suggests that the account of PG construction we have proposed is not exhaustive. We know that U.G. constrains more stringently the distribution of empty categories than that of non empty categories. It is therefore natural to attempt to link the lesser acceptability (or total unacceptability) of some PG structures to the failure of the PG to obey some constraint on empty categories. Moreover, the relative acceptability judgments can give us important clues as to how the various principles constraining the distribution of empty categories partition.

This very line of argumentation is used by Taraldsen (1981), Chomsky (1982), Kayne (1983) and Pesetsky (1982). For example, Taraldsen (1981) and Chomsky (1982) argue that the total unacceptability of (87vi) is related to an ECP violation by the PG. Kayne (1983), noting that (87v) is worse than (87iv), argues that (87v) as well violates the ECP (under the formulation of the ECP he proposes).

We will not pursue this matter here. Rather, we will consider the question of why the violation of certain constraints by the PG do not seem to affect the relative acceptability of PG-structures. More specifically, consider the relation A'*/A. This relation, we have argued, must be established at S-structure, and A is interpreted as a variable bound by A'*. While the relation A'*/A* must obey the Subjacency Condition, the relation A'*/A, i.e. A'*/PG seems to be immune to the effect of this constraint.

Consider the account of Chomsky (1982). Chomsky (1982) first assumes

that the Subjacency Condition is a condition on rule application constraining Move x: no transformation rule mapping D-structure on S-structure may move some item from position z to position y if z is not subjacent to y. It is easy to see how Chomsky can draw the required distinction between real gaps and prasitic gaps. Real gaps are created by movement, and must therefore obey the Subjacency Condition. Parasitic gaps, on the other hand, are base generated gaps so that their relation A'* is not subject to that condition.

From our point of view, the two relations A'*/A* and A'*/A are not intrinsically distinguishable as they are in Chomsky's system, since both are established at S-structure. Given our argument (cf. 4.1) to the effect that the Subjacency Condition is irrelevant to the LF component, our only option is to formulate the Subjacency Condition as an S-structure well-formedness Condition. Consequently, several questions arise.[51]

i. how can we distinguish between the pairs (x,y) where x locally binds y, subject to the Subjacency Condition from those that are not

ii. how do we formulate this condition so that it takes into account the distinction of (i).

Consider (i) first. We know that at least some A'/A relations obey the Subjacency Condition. In the absence of any convincing empirical evidence to the effect that A/A relations obey it,[52] we must resort to arguments of simplicity and elegance to settle the issue of whether they should too obey this condition or not. For example, Chomsky (1981) argues precisely on these grounds that they should.

We can characterize the relevant (x,y) by appealing to the nature of x, the nature of y, the relation between the two or a combination of all

these. A translation to S-structure of the idea that the Subjacency Condition holds of all and only movement relations would characterize the desired set of prirs by postulating that they involve a "trace" as second member, i.e. as y. It is unclear, however, how one can do it without artificiality. For as far as NP-traces are concerned, they can be defined as non-heads of chains. Non NP-traces, on the other hand, cannot be as simply characterized. Emptiness is neither a necessary criterion (because of syntactic resumptive pronouns) nor is it sufficient (because of the existence of pro) nor would Case be the relevant parameter (for the same reasons). Nor could we use local A'-binding since PG structures demonstrate that we would include PG amongst "traces".[53] So suppose rather that we appeal to the nature of x. If we suppose that both A'/A relations and A/A relations obey the Subjacency Condition, we will necessarily get a disjunctive statement to define the relevant set of x's. So suppose instead that we restrict the scope of the Subjacency Condition to A'/A relations. Then we may characterize the relevant x's simply as those A'-binders binding some A-position at S-structure. Furthermore, instead of assuming, as is usual, that the Subjacency requirement holds of every pair (x,y) such that x locally-binds y at S-structure, we may assume that it must hold of one such pair. In other words, we may answer (i) and (ii) above by stating:

(107)   If A'* locally binds some A-position at S-structure
          there must be an A* such that A'* locally binds A*
          and A* is subjacent to A'*

So, in a way, instead of being a symmetric constraint, the Subjacency Condition is viewed as a top to bottom procedure. Further discussion of the formalization of the Subjacency Condition will be undertaken in III.9

below (cf. especially III.9.2 (121)).

8.2.4. Caseless PG

Let us now turn to the observation (iii) we made in 8.2. Consider the following pairs:

(108)   i.   *who$_i$ did John fire e$_i$ without it seeming
              [[$_e$i]$_K$ to have failed]

        ii. John fired Bill without it seeming [[e$_i$]$_K$
             to have failed]

(108ii) is ungrammatical and local A'-binding of its gap, as in (108i) provokes no improvement. As we suggested in 8.2., in such cases, it is plausible to invoke a violation of some principle by the (parasitic) gap. What does K violate in (108i) and (108ii)? Consider first how K is analysed in each case. In (108ii), K is analysed as a PRO. This is ruled out by the Binding Theory. Indeed, <u>seem</u> triggers S'-deletion so that PRO is governed. In the framework of Chomsky (1981, 1982), this is excluded. In the framework we advocate for in more details in the next chapter, a PRO is allowed in a governed position provided that it is bound in its Governing category, here the without-clause. Since the Governing category of K in (108ii) contains no admissible antecedent, (108ii) is ruled out.

Turn now to K in (108i). Here, K is locally A'-bound by <u>who</u>, so that it is interpreted as a variable. Chomsky (1981, 1982) argues that this is ruled out because variables, which are arguments, must have Case in order to get θ-marked, i.e. on the very same grounds that (109) is ruled out:[54]

(109)   who$_i$ you tried [[e$_i$]$_K$ to leave]$_C$

whether (109) is a wh-movement case, or a resumptive empty pronoun case (i.e. whether who and K get coindexed at S-structure or at LF). Of course this account would understand that θ-assignment is somehow dependent on Case-marking. Because we have argued in II.2.6. (of the preceeding chapter) that this assumption (i.e. reduction of the Case Filter to the θ-criterion) was incorrect, we cannot make the same argument.

Note incidentally that we agree with Chomsky (1982) that (109) ill-formedness is not semantic. Chomsky points out that some dialects of English permit such sentences as: (i) you tried for John to leave. Another example is Kinyarwanda. Kinyarwanda has a verb try which may appear in structures like (i) or like (109) - without the equivalent of for - in Kinyarwanda try is an Exceptional Case-marking verb. We can however make an argument very similar to that of Chomsky. Recall that we have argued that chain-initial elements distinct from PRO must be Case-marked (cf. II.2.6.) It is clear that (109) would violate this requirement either at S-structure or at LF, and that (108i) would violate it both at S-structure and at LF.

However, when we consider the second facet of observation (iii) made in 8.2., we see that important difficulties arise. For (110i) below:

(110)   i. Who_i does $[[e_i]_K$ [sleeping late] bother $e_i]$

          ii. $[[e_i]_K$ sleeping late] bothers $John_i$

is a well-formed structure, as acceptable as its A'-less conjugate (110ii).[55] Yet K of (110i) is a Caseless locally A'-bound element.

A further problem arises with the ECP, a likely candidate for ruling out (109). First, it should of course be assumed that the ECP holds of empty variables, whether they are variables both at S-structure and at LF

or at LF only.[36] In (109), assuming C=S' would imply that K is not properly governed. The same account would not carry over to (108i) since K is properly governed by the verb <u>seem</u> (<u>seem</u> being an S' deletion predicate). However, the question arises as to why the ECP does not exclude (110i) in exactly the same fashion it does (109), since K in (110i) is certainly not properly governed.

In fact, all this would follow if we could somehow assume that K is immune to local A'-binding, i.e. that K behaves exactly as if it was not locally A'-bound. (108i) and (108ii) would be ruled out in exactly the same way, (109) as *you_i tried e_i to leave, and (110i) would be grammatical just as (110ii).

There are basically two ways to achieve this:

1. Require of locally A'-bound elements to be Case-marked in order to count as variables (and modify correlatively the definition of PRO...so that it may be locally A'-bound).

ii. Argue that the relations of local binding in these structures are not what they appear to be.

Clearly, each of these modifications would imply important modifications of some of our basic assumptions (e.g. classification of empty categories...) that we cannot fully discuss before we have a better understanding of how the Binding Theory functions. (ii) is too vague at this point to be evaluated but note that, as far as (i) is concerned, it starts with the inconvenience of stipulating a property that we certainly want to derive (i.e. variables must be Case-marked), if it is true. We may say for the moment that we will in fact, adopt a version of (ii) which will have the property that locally A'-bound elements need <u>not</u> have Case.

8.3. Further remarks on WCO

The presentation we have given of the WCO phenomena understands a very different conceptual view from what is usually assumed (cf. Wasow (197*) Reinhard (1976), (1980); Chomsky (1976); Higginbotham (1980), (1981); Haik (1982)). The WCO phenomenon is usually taken to be illustrative of restrictions on the referential dependency of a pronoun P to a variable V (or an indefine NP) such that neither P nor V c-commands the other. Consequently, the range of structures illustrating WCO is much wider than those meeting the structural description (86) for not all such cases apear to reduce to (86).

Notice first that if indeed variables are defined as locally A'-bound elements with no particular attention paid to whether they are empty or not, as we have argued, structures meeting (86) (e.g. who does his mother love) involve no pronoun, so that it would be incorrect to claim that such structures illustrate a referential dependency restriction of a pronoun to a variable.[57]

We may wonder however, whether all the cases that have been taken to be WCO cases can be subsumed under the Projection Principle. I think the answer is negative. We can basically distinguish two sorts of structures that are referred to as WCO configurations in the literature on the topic:

i. Those meeting the structural description (86) at S-structure, hence at LF, and which straightforwardly fall under the Bijection Principle

ii. Those that do not meet the structural description (86) at S-structure.

For those, it is natural to postulate, as far as theoretical plausibility

permits, LF mechanisms that will convert these S-structure representations nct meeting (86) into LF representations meeting (86). The most commonly accepted such mechanism[5] (originally proposed in Chomsky (1976) precisely on these grounds) is QR, which converts S-structures like (111i) into LF (111ii):

(111)   i. His_i mother likes $everyone_i$

        ii. $Everyone_i$ [his_i mother likes e_i]

More controversial are the cases involving <u>Reconstruction</u>, i.e. translation of the S-structures (112i) and (112ii) into the LF (112iii) and (112iv):

(112)   i.   [$whose_i$ book]$_j$ did his_i author sell e_j

        ii. $Whose_i$ did his_i author sell [e_i book]$_j$

        iii.   [with $whom_i$]$_j$ did his_i mother talk [e_j]

        iv. $Whom_i$ did his_i mother talk [with e_i]]

A reconstruction rule states in essentials that (a subject of) phrases in A'-positions have exactly the same c-command properties it would have if it were in the position of the variable they bind (cf. Chomsky (1976) Fourier (1980); Van Riemsdijk & Williams (1981); Higginbotham (1980) for relevant discussion, and also Belleti & Rizzi (1981) for pertinent remarks).

Now note that given a structure falling under (ii) above, we may make a further distinction. If such a structure does not meet (86), it might be reducible to (86) only if the phrases involved, i.e. A'*, A* and A all agree with respect to the features that must have identical values for the relation of binding to hold, i.e. person, number, gender... We can

however, construct cases falling under (ii) that do not meet this requirement, and are therefore not reducible tr (86). There is an important analogy between this discussion and the one we had in section 7 above on SCO. For it is clear that coindexing between two phrases without agreement is possible only if at least one of the two phrases is plural. Consider the following example:

(113)   The sultan wonders which wife$_i$ their son betrayed e$_i$

Assume the sultan has many wives and had one son with each of them. I believe (113) cannot have a reading in which their is referentially dependent on both e$_i$ and on the sultan, i.e. cannot mean "The sultan wonder Wx, the son of (x and the sultan) betrayed x. Clearly, the Bijection Principle is irrelevant to the ill-formedness of such cases, that we might call WCO with split antecedents. Rather we must invoke some principle barring referential dependence of a pronoun from a variable under certain conditions, perhaps as in Higginbotham (1981) or Haik (1982).

In conclusion, we see that WCO is analysed here as resulting from two distinct and non overlapping phenomena, accounted for by two distinct and non overlapping principles.

8.4. On the definition of Variables

Coming back rapidly to the notion of variables, we have assumed all along the definition given in (5) in section 2.2. As we have pointed out, this notion is syntactically pertinent only to the extent that the set of elements it denotes possess some distinctive property. However, no

generalization seems to hold of locally A'-bound elements and locally A'-bound elements only. For example, they are not Case-marked empty categories because of syntactic and anaphoric resumptive pronouns. Nor is it true of empty variables that they are the only Case-marked empty elements because it is also true of pro (Note that treating pro as somehow locally A'-bound would lead to the loss of the generalization according to which locally A'-bound elements must be arguments, since there are expletive pro's: this unification would therefore be illusory).

However, if we restrict the notion of variables to locally operator-bound elements, then, we do find generalization holding true of them and only them. Only variables in that sense are subject to the Bijection Principle. We therefore modify (5) and replace it by:

(114)   x is a variable iff x is a locally operator-bound
                         ?-category in an A-position

Correlatively, we can formulate the Bijection Principle as establishing a one to one correspondance between operators and variables, i.e.:

(115)   Bijection Principle
           There is a one to one correspondance between
           operators and variables at any level of
           syntactic representation.

9. Improper Movement and Successive cyclicity

Recall that we started our investigation by stating that there were four types of relations involving A and A' positions. So far, we have investigated A'/A relations and A/A relations, i.e. relations "from" an

A-position. Let us consider now relations "from" an A'-position.

9.1. Improper Movement

9.1.1. In a model with movement rules mapping Phrase markers onto Phrase markers between D-structure and S-structure, A/A' relations break down into two subclasses at S-structure:

    i. A/A' relations of coindexing
    ii. A/A' relations of movement, also called Improper Movement

Considering an S-structure representation in which some A-position A* c-commands and is coindexed with some A'-position A'*, how can we express the difference between (i) and (ii)?   Very simply as follows: movement of P from A'* to A* expresses the fact that the thematic role of P in A* is "transmitted" to P from some θ-position through A'*.

In the case of coindexing without movement no θ-transmission occurs. We want to exclude improper movement: we want each case of coindexing between A* and A'* never to involve θ-role transmission.

This in fact follows from the way we have construed the theory of chains. We have assumed that the theory of chains dealt exclusively with the network of A-positions. Since θ-role "transmission" is a property of chain structure, and since chains only contain A-positions, it follows that θ-role transmission through an A'-position is excluded, i.e. improper movement is excluded.

Of course, this is not an explanation.   Although a natural stipulation, it is still a stipulation to assume that the theory of chains deals exclusively with A-positions. Is it possible to derive it? 9.1.2. In fact, in the relevant cases, it is.

Suppose we now freely allow chain formation to partition the set of all NP's, whether in A- or A'-positions. Let C= $(x_1, x_2 \ldots x_n)$ a well-formed chain. Assume first that C contains a pair (x_i, x_{i+1}) such that x_{i+1} is an A-position locally A-bound by x_i .

Clearly, this case is more general than the improper movement situation. The θ-role, if any, transmitted from x_n to x_1 will be intermediately transmitted from x_{i+1} upward to x_i , then to x_{i-1}. Consider x_{i+1} . It is not in chain initial position. By definition, it is an NP-trace. By universal convention, justified on independent grounds, NP-traces are expletives, that is non arguments (cf. IV. 4) for further discussion).

So x_{i+1} is a locally A'-bound non argument. But we know that this must be excluded on independent grounds (cf. II, 2.6 (50)). C is in fact not well-formed. Improper movement is excluded a fortiori and we now have an explanation as to why it is.[59],[60]

The problem we now f ce is that not all chains containing an A'-position fall under the above discussion. A chain P will not if: if i is the smallest index such that x_i in C is in an A'-position $\forall j$, $j>i$, x_j in C x_j is in an A'-position.

Such chains would be of one of the forms:

i. $(A_1 \ A_2, \ldots, A_i, A'_{i+1}, \ldots A'_n)$

ii. $(A'_1, A'_2, \ldots, A'_p)$

It is easy to see that such chains are not excluded by anything. It is also not clear what roles they may play. Let us therefore exclude them by stipulation, but by a stipulation much narrower than that of requiring all e ments of a chain to be in A-positions; namely· the most deeply embedded

<u>element</u> <u>of</u> <u>a</u> <u>chain</u> <u>must</u> <u>be</u> <u>in</u> <u>an</u> <u>A-position</u>. It is clear why this excludes (i) and (ii).

Therefore, instead of II.(9), the definition of chain now reads (taking into account both II.(9) and II.(30)):

(116) $C=(x_1, x_2,...,x_n)$ is a chain iff
  i. x_i is an NP
  ii. x_i locally binds x_{i+1}
  iii. x_i is a $\bar{\Upsilon}$-category in a Caseless position $i\neq 1$
  iv. x_n is in an A-position

9.2. Successive Cyclicity and related matters.

Let us now briefly turn to A'/A' relations. We do not wish to discuss this issue in detail here, but merely to note what particular views on this question imply for the accounts we have given so far. 9.2.1. A'/A' relations have been postulated to exist to accomodate the apparent Subjacency violations of wh-movement. Indeed, in a sentence like (117i):

(117) i. Who_i do you think she believes you saw e_i

  ii. $[COMP_1[...\quad [COMP_2[...\quad [COMP_3[\quad NP$
   $Who_i\quad$ you think e she believes e you saw e_i

the relation who_i /e_i would violate the Subjacency Condition as we have formulated it (cf. the Introduction) in terms of local binding. Postulating that the relation who_i /e_i is mediated by the successive traces in $COMP_2$ and $COMP_3$ gets around the problem, as illustrated in (117ii).

Now we should distinguish two aspects of this account: First, a usually accepted view that in a synt.:tic A'/A relation (i.e. an A'/A

relation established at S-structure) between A'* and A*, all the <u>clauses</u> containing A* but not A'* (except, perhaps, for the most embedded one, if Rizzi (1982) and Sportiche (1981) are correct in their respective arguments that S is not a bounding node) have a special property, c-ll it P.

Second, a more controversial aspect: the particular notation used to represent P, e.g. in (117ii), the presence of intermediate traces in the COMP position of the clauses in question.

There is ample evidence in favor of the existence of the property P attributed to such clauses both theoretical, e.g. the fact that it permits derivation of the complex NP-constraint noticed in Ross (1967) (due to the fact that NP's may not have that property P) and perhaps the wh-island constraint,[61] or more direct, e.g. Kayne & Pollock (1978), Rizzi (1982), Torrego (1981), Sportiche (forthcoming a).

What kind of evidence do we find in favor of the particular notation used to represent P? Here, the evidence is necessarily much more theory internal, and can be classified as follows:

(118)   i.   Consistency with a particular formulation of the
             Subjacency Condition (e.g. if it is formulated in
             terms of local binding, there must be intermediate
             local binding)

       ii.   The existence of successive movement as in (117ii)
             is the null hypothesis: in a theory with movement
             such sentences are generated anyway.[62]

      iii.   Evidence based on the possibility to derive other,
             unrelated properties crucially based on the fact
             that there is an empty category in COMP. For
             example, if such a categfory acts as a binder (cf.
             the standard accounts of [that t] effects, and
             also Brody (1981); Kayne (1978); Aoun, Hornstein
             & Sportiche (1981)...)

9.2.2.   It is apparent that, so far, we have assumed that there were no

intermediate traces in COMP. Had we not, we would have had to complicate the formulation of the Map Principle and the Bijection Principle Principle, and to modify in obvious ways the argumentations developed in III.7 and III.8 above (as can easily be checked, the essential conclusions would remain valid).

Because the simplicity of the formulation of these principles speak against using the notation in terms of intermediate traces we will assume the following treatment of Subjacency. First, we adopt the mechanism, proposed for independent reasons and justified on independent grounds, in Aoun & Hornstein & Sportiche (1981) (but cf. footnote 63):

(119)   At S-structure, COMP-->COMP$_i$ i.   if COMP dominates
                                                 material with index i
                                            ii. otherwise freely and
                                                 optionally

We now understand the notion subjacent and its symmetric superjacent as in (120):

(120)   x is subjacent to y   iff   x and y are coindexed, y
      or y is superjacent to x   c-commands x and no more
                                       than one bounding node
                                       intervenes between them

Taking into account the conclusion of 8.2.3 above, we can now formulate the Subjacency Condition as in (121) (in the spirit of Bresnan & Grimshaw (1978)):

(121)   Subjacency Condition
             At S-structure, * COMP$_i$, unless it is superjacent
             to some element

It is clear that:

i.   (121) subsumes (107). If some phrase XP$_i$ is in COMP. COMP becomes

$COMP_i$ by (119). By (121), it must be linked back superjacently either to some other $COMP'_i$, or some A-position A ; in this latter case, no problem; otherwise we proceed iteratively with $COMP'_i$ ...

ii. (121) entails rightly that only "movement" to COMP is iterative, amongst "movement" to an A'-position (cf. footnote 62).

iii. It follows from (i) that indexed COMP's can only appear in chains of COMPs with subjacent links down to some A-position. In particular, they do not appear with P.G. not accessible to movement.

iv. Finally, if, as Aoun, Hornstein & Sportiche (1981) suggest, indexed COMP's are related to escape from [that t] effects,[63] this correctly accounts for the fact that subject of tensed clauses cannot be well-formed P.G.'s (if this is indeed a fact).

FOOTNOTES CHAPTER III

1. Note in particular that if arbitrariness is the rule for empty categories, the minimal assumption, the absence of say, expletive PRO requires an explanation.

2. Note that the dichotomy does not match the distinction moved phrase/base generated phrase as wh-movement and Heavy NP-shift show.

3. Nor should Quantifier phrases be considered arguments (at least when interpreted quantificationally). This remark raises important questions for the Projection Principle, (cf. footnote 8).

4. Note that these arguments do not violate the θ-criterion, since, to be more precise, it holds of A-chains and θ-positions.

5. There are well-known exceptions to that statement, e.g. Japanese "wa" phrases only require a weak link of aboutness between the "wa" phrase and its related constituent.

6. (3) might rightly remind one of May (1977) <u>Condition on Q-Binding</u> which requires every Quantifier phrase to properly bind some variable at LF. Note however, that the scope of (3) is wider in some respects and narrower in others. Wider because A'-positions do not necessarily contain Quantifier Phrases. Narrower because it does not require movement of QP's to an A'-position. This might follow from different considerations (cf. footnote 3 and footnote 8).

7. Recall that it states a restriction on certain processes or configurations involving two positions x and y, such that:

    (i)   x locally binds y
    (ii) x is not separated from y by more than one bounding node

    (iii) Bounding nodes are NP, S' (and perhaps S, PP...cf. Chomsky, 1978; Rizzi, 1982; Sportiche, 1981...for relevant discussion).

8. We have remarked above (fn.3 & fn.6) that treating QP's and wh-phrases as non arguments (when they are interpreted "quantifically") raises questions about whether the Projection Principle is correct: if it is, the θ-criterion should hold at S-structure, and would be violated by such examples as:

    (i)   everyone likes his mother,
    (ii) who saw what

    etc... since <u>everyone</u> or <u>what</u> are not arguments, yet occupy a θ-position.
    Assume therefore, that the Projection Principle does not require the θ-criterion to hold at S-structure.

If the θ-criterion holds at LF but not at S-structure, we might automatically derive the existence of May (1977) rule of QR, and of Aoun, Hornstein & Sportiche (1981) rule of wh-R. Indeed, these rules, by definition, affect non-arguments. When non-arguments occur at S-structure in A-positions (linked to some θ-role) as in (i) or (ii) above, they will have to move prior to the LF level of representation (and leave an argument "trace"), or we would end up with a θ-criterion violation. Therefore, asuming that the θ-criterion only holds at LF permits us to derive the existence of QR and wh-R, and also its obligatoriness, i.e. May Condition on Q-binding mentioned in footnote 6 above.
In order not to lose all the effects of the Projection Principle, we can reformulate it as follows:

> Chain Invariance Principle
> Chain structure is syntax invariant

This means that a given chain contains exactly the same syntactic positions at S-structure and at LF.
It is quite clear that this formulation preserves the desirable effects of having the θ-criterion hold at S-structure.

9. Or perhaps, more generally, to any node.

10. May (1977) implicitly assumes that QR does not operate successive cyclically.

11. cf. Aoun & Hornstein (1982) for a recent proposal made in terms of the Binding Theory.

12. As the reader can check, even a successive cyclic application of wh-R (from COMP to CCMP) would not avoid a subjacency violation with the reading when sa is bound by chacun.

13. Numerous arguments could be constructed in favor of the assumption of the text if there is a rule of LF reconstruction (cf. Chomsky, 1976, Fourier, 1980). R. Freidin points out the following simple case: S-structure (i) would translate as LF (ii) where e is not subjacent to the wh-phrase:

(i)   [whose book]$_j$ did you read e$_j$
(ii) whose$_i$ [$_s$ you read[e$_i$ book [$_{NP}$e]

14. We argue in section 8 below that this similarity is not accidental.

15. Note that this conclusion is strengthened by the observation that Heavy NP-shift of a QP, to the extent it is possible, is much worse than (30), viz:

(i) She deeply offended by not recognizing him immediately every former student of hers.

16. There is of course an obvious problem with the conjunction of (33i) and (33ii) since a parasitic gap need not be subjacent to its

S-structure local binder. We return to this problem in 8.2.3 below.

17. Although we say resumptive pronouns, these elements are often but not always clitics (at least in PF). We will argue later that the discontinuous constituent Cl/associated empty category _is_ a pronoun.

18. cf. Kenstowicz & Wahba (to appear) for a more detailed discussion.

19. Of course, a correlated modification of (3) would be necessary.

20. Note that, in Vata, as (45iii) shows, resumptive pronouns and pronouns "locally" bound by resumptive pronouns bear a low tone instead of the mid-high tone that regular pronouns bear (cf. Koopman&Sportiche, 1981a, for details).

21. Although it could be argued that, say, P-stranding is possible in LF because LF resumptive pronouns are available.

22. A simple example can be provided. Consider the following paradigm:

    (i)   who$_i$ did you see   a   picture of e$_i$
                                *the
                                *this
                                *his

This pattern is mirrored by wide scope possibilities for _everyone_ in:

    (ii)   you saw   a   picture of everyone
                     *the
                     *this
                     *his

In French, the equivalents of (i) are grammatical with both _a_ and _the_ but not with _this_ and _his_. _The_ _same_ _pattern_ _is_ _found_ _with_ _wide_ _scope_ possibilities for the equivalent of (ii).

23. In this connection, note that French for instance, does not allow operators in Dislocated position:

    (i) * Chacun des enfants, je l'ai vu
    (ii) * Qui, tu l'as vu
    (iii)* Il est parti hier, qui

24. If Rizzi (1982) is correct in his analysis of the Null subject Parameter, the situation is more comlex than we have described but the point we make can still be made.

25. There are few exceptions to that generalization. A possible case is relativization from subject position in infinitival relatives.

26. I suspect that this is what is happening with the French equivalent of (51ii).

27. Burzio (1981) assumes VP integrity both in Faire-infinitive (his F-S) and Faire-par (his F-VP) constructions. For reasons essentially having to do with indirect object cliticization (je lui fais envoyer la lettre par X/*je lui fais envoyer la lettre à X, where lui is an indirect object), I would disagree with this conclusion on the former construction, for which I would argue, if place permitted, that V' (≠ VP) preposing is involved, as proposed by Rouveret & Vergnaud (1980) thus destroying VP-integrity.

28. If we were to take clitic positions to be A-positions, i.e. to be potentially assigned θ-roles directly, there would be no principled motivation whatever for the existence of an associated empty category if, as Chomsky (1981) and Stowell (1981) suggest, the theory of subcategorization partially reduces to θ-theory.

29. Cf. Vergnaud (1974; 1982) for relevant discussion, and also foonote 48 below.

30. Assume, following Borer (1981) that this locality requirement is (propei) government of $\underline{e}$ by the clitic.

    In a configuration $Cl...[_x ...[_y...e$
    where x and y are bounding nodes

    (NP,S' or PP if Sportiche (1981) is correct for French, NP, S' - and perhaps PP - if Rizzi (1982) is correct for Italian). Cl will never govern (and a fortiori properly govern) $\underline{e}$ since a maximal projection blocks government as Aoun & Sportiche (1983) show (except down to the head of the maximal projection as Belleti & Rizzi (1981) argue) .

31. This expectation is not fulfilled, I believe, with the French equivalents of (51i) and (51ii). cf. Appendix for discussion.

32. If we try to construct examples parallel to (51) or (57) involving reflexive clitics, we cannot use infinitival instrumentals or manner adverbials (because the subject of the adverbial clause, which must be assumed to hang from the matrix VP as we have discussed, will be controlled by the matrix subject. This matrix subject will independently have to be the antecedent of the reflexive clitic. Since the parasitic gap (or the "pronoun" in WCO cases) is by assumption coindexed with the reflexive clitic, it will by transitivity be coindexed with the matrix subject, hence with the subject of the adverbial clause: it will not be locally-bound by the clitic.
    I could not find examples of non infinitival adverbials which can be as plausibly argued to hang from VP as instrumentals and manner adverbials: I could not thus construct relevant examples.

33. Note that this conclusion has the following interesting consequence: there cannot be a well-formed structure with clitic doubled NP-trace (since, by definition, NP-traces are not in chain initial position), i.e. the following type of structures are ruled out, where e_i is the "trace" of NP_i:

$$(i) \quad NP_i \ldots CL_i + V \ldots e_i$$
$$(ii) \quad Jean_i \; l'_i \; a \; été \; tué \; e_i$$
$$(iii) \quad Les \; enfants_i \; s'_i \; ont \; été \; présenté \quad e_i \; e_i$$

34. We could have made the more general assumption that clitics are [-P], i.e. non operators. This would not have changed our argumentation (and in particular that of 6.3.1 would have needed only trivial adjustments, as can be easily verified).

35. Jaeggli (1980) is the first to have proposed that all instances of clitic doubled empty categories be treated as PRO (both at S-structure and at LF).

36. The argument of the text is unaffected by the distinction Jaeggli (1980) makes between s-government (i.e. strict subcategorization) and c-government (structural government): Jaeggli assumes clitics absorb s-government which means that they absorb both Case and θ-role. Notice that these assumptions entail the generalization: NP-clitics may appear only on verbs strictly subcategorized by NP's, which, I believe, is falsified by verbs like <u>supposer</u> ... (viz. $*[_{VP}$ supposer NP], OK..le suppose e vrai...)

37. Notice incidentally that Jaeggli's approach can be maintained in a model with an independent level of D-structure, by assuming that (s-government absorption is an S-structure process: the existence of an empty category could be then principally justified at D-structure by the θ-criterion.

38. The other cases, i.e. PRO or NP-trace, would not really fall under the SCO category. The discussion of the text bearing on the A'*/P relevant would be relevant anyway.

39. In connection with footnote 38, it is easy to see that (70i) applies unchanged to the case when A* is PRO or NP-trace. (70ii) however, should be modified accordingly.

40. But cf. Chomsky (1981; p. 227 fn. 45, and Chapter 3), Evans (1980), Huang (1982) for relevant discussion.

41. Similarly: * John$_{(i)}$ and Bill$_{(j)}$ they$_{(i,j)}$ like himself$_{(j)}$ which would mean: John and Bill like Bill.

42. Note that it is this very notion of <u>bound</u> that is relevant to chain formation, variable definition etc...

43. Note that the incidence of such a move on our theory of resumptive pronouns is null (but cf. McCloskey (1983) for a somewhat different view). It is clear that we do not want variables to be subject to this revisited Principle B (it would exclude all types of wh-movement, e.g. who$_i$ e$_i$ left). Since syntactic resumptive pronouns are interpreted as variables at all levels, they escape the effects of this modification. As for anaphoric resumptive pronouns, we have argued that their LF local A'-binder should be considered to

fall outside the S'-system, i.e. outside the domain D of any A-position inside S: so again they would escape this principle B revisited.

44. Note that if both relations A'*/A and A'*/A* are normal, the distinction real gap/parasitic gap dissolves.

45. For a variety of reasons that would take us too far afield to expose here, we will assume, contrary to Pesetsky (1982) that Across-the-board (ATB) structures (and Right Node raising - RNR-structures) fall in an entirely different category of phenomena than PG constructions (Basically because: (i) the acceptability judgments seem to me on a different level, (ii) ATB gaps behave exactly like real gaps, unlike PG, (iii) ATB is available straightforwardly in languages not tolerating PG structures). In everything that follows, I assume that we are not dealing with conjuncts of the ATB or RNR type.

46. This conclusion is reinforced by the following observation: Dutch (H. Koopman, p.c.) does not permit P.G. structures (for some reason, (90) might hold more strongly in Dutch) except with Heavy NP shift (only possible with headless free relatives, cf. van Riemsdijk (1978)) viz):

    (i)   *wat [heb je [e gekocht] [zonder te e bekijken]
           what have you   bought   without   looking

    (ii)   Ik heb [e gekocht] [zonder te [e bekijken]]
         I have   bought   without   looking

        [wat je e gevraagt hebt]
        what   you   asked

47. Note that, because the same derivations could be assumed for the ungrammatical:

    (i)   The man$_i$ [who$_j$ [he$_i$ saw e$_j$]]   (at LF)

    (ii) The man$_i$ [who$_i$ [he$_i$ saw e$_i$]]   (at LF')

We are forced to assume that the principle ruling (ii) at S-structure or LF (i.e. no overt resumptive, or no pro in object position...) are met as well at LF'. Although not under the alternative discussed below.

48. I believe, however, that Aoun's proposal faces difficulties. He argues that a structure like (88iii) may get the following representation:

    (i)   the man$_i$ who$_i$ his$_k$ enemies praise e$_i$

in which his happens to pick up the same referent as the man, but is not coindexed with it, so that no violation of the Bijection Principle for (i).

First, I would hold, following Higginbotham (1980) that the head of a restrictive relative (contrary to appositive relatives) is not referential. The man in (i) does not refer to some particular (mental) object. Rather it sets a range for the relative clause operator, which can be even a non argument as in "the headway that he made", or "le parti qu'il a tiré" in which the head is an idiom chunk (cf. Vergnaud (1982) for relevant discussion, especially his Chapter 5). In particular, we do not consider it an A-position, i.e. a position that is attributed, or may be attributed a θ-role; rather, the whole relative clause is. If this is correct, although coindexing between a head and some pronoun may be possible, subject to interpretation, accidental coreference of a pronoun with the head as in (i) would be meaningless if we are correct; his in (i) cannot happen to refer to the same individual the man refers to because it refers to no individual at all

Secondly, for those speakers -including myself in French - who would judge (ii) and (iii) below on a par with (i), the same remarks would apply with more force since the head NP's are not referential even when they are not in relative head position:

(ii) [Chaque enfant$_i$ que ses$_i$ parents ont aidé e$_i$] a réussi

(iii) [Aucun éleve$_i$ que ses$_i$ parents ont aidé e$_i$] n'a réussi

So that accidental coreference between ses and the head is in principle impossible.

49. Note that I do not mean that this exhausts the structure of indices. It might well be the case that we need anaphoric indices in the sense of OB (=Chomsky, 1980) as Lasnik (1981) argues.

50. Note that it is misleading to call these indices or sets referential even in the restricted sense used here (i.e. "mental" reference) because the indexing or rather the coindexing plays a variety of roles some of which have little to do with reference (e.g. the index of a non-referential NP as a QP, or the coindexing of a QP and its trace under QR, or coindexing of the head of a chain and its other members).

51. There is of course the question of how the theory of Bounding nodes in a system without movement is constructed. We address it in section 9 below.

52. Chomsky (1981) discusses some evidence in this respect that we will reanalyze in Chapter IV, section 3. Cf. also Sportiche (forthcoming b).

53. Nor could we simply hold that the Subjacency Condition holds of pairs (x, y) where x locally binds y and one of x or y is in a θ'-position because of such cases as: x=John and y=his in (i) John seems to the parents of his wife e to be sick (and because of PG structures, of course).

54. Recall that such cases cannot be ruled out because of lack of case on

who because of:

(i)* the man$_i$ that you tried e$_i$ to sleep.

55. This observation is due to J. Higginbotham (1980): PRO does not give rise to WCO effects.

56. Note that this last distinction recalls the usual distinction. Empty category generated by movement (S-structure bound) versus empty resumptive at LF (LF bound), but is not identical to it because of parasitic variables which are S-structure bound but not generated by movement.

57. In fact, this depends greatly on the indexing system. If pronouns could pick the "reference" of variables without being coindexed with them, as is possible with pronouns and names, this remark would be true but incomplete for "who does his mother love" could be represented either as (i) or as (ii):

(i)   who$_i$ his$_i$ mother love e$_i$
(ii) who$_i$ his$_j$ mother love e$_i$

with (i) falling under the Bijection Principle, and (ii) under some other principle- e.g. Higginbotham (1981) Accessibility Condition. We argue later against such an interpretation of lack of coindexing.

58. Although not by authors such as Reinhart (1976) and Haik (1982).

59. May (1981) proposes a different way to exclude improper movement based on the following two assumptions:
(i)   Variables must have Case
(ii) NP-traces (i.e. traces of NP-movement) must not have Case
Although we have proceeded with the assumption (ii), based on our discussion of chain theory in Chapter II, we have also noted that it was a provisional assumption that we ultimately reject (cf. Sportiche, forthcoming b). Furthermore, we also argue against (i) in IV.6.

60. We might think of another way to proceed assuming:
(i)   NP-traces must be Caseless
(ii) Caseless elements fall under principle A
(iii) Principle A requires local A-binding
Although we do adopt (ii) and (iii) above in Chapter IV, we ultimately reject (i), as pointed out in footnote 59 above.

61. Although it is much less clear how this is achieved, since it rests on a prohibition against doubly-filled COMP's in the course of a derivation, and COMP's containing a lexical phrase and an empty category at S-structure. Cf. Chomsky (1982, p. 70) for relevant comments and also Chomsky (1981, Chapter V).

62. This remark is double edged since, of all the movements to an A'-position, only movement to COMP is freely iterative. PP extraposition, S-extraposition...are not.

63. We would have to slightly modify Aoun, Hornstein & Sportiche (1981) account of [that t] effects. They assume that an indexed COMP is a proper governor and include in (119i) a clause restricting COMP indexing to cases in which COMP only dominates phrases indexed i. The obvious modification, given our formulation of (119) is to make an indexed COMP a proper governor only if it dominates only material indexed i.

APPENDIX: THE STRUCTURE OF CLITIC CONSTRUCTIONS

0. Introduction

In this appendix, we discuss in more detail the structure of clitic constructions. We have discussed what the structure of certain clitic constructions is not in III.6. We now turn to analyses that are compatible with our general assumptions. Note that we will not make any attempt to discuss the numerous recent proposals on the topic (e.g. Aoun (1982); Bok-Bennema (1981); Borer (1981); Burzio (1981); Jaeggli (1981); Manzini (1983a); Rivas (1977); Zubizarreta (1982)...). Rather, we will explore the consequences of some very general theoretical assumptions in the domain of clitics.

We have argued in III.6 that, at least in some cases, the relation between a clitic and its associated empty category could not be one of A'-binding (even if the clitic is considered a non-operator instead of a non argument - cf. III.fn.34). For ease of reference, we will call Cl the clitic, and NP* the position (usually empty, but, perhaps, not always) associated with it.

Let us reason on French, which has a variety of clitic constructions: Cl can be a subject clitic (as il in il pleut), and object clitic (as le in Pierre le voit) a reflexive or reciprocal clitic (as se in il se rase), a "genitive" clitic (as en in il en a lu une partie), or a locative clitic (as y in il y va).

For each of these, the question arises of what the nature of the

position in which Cl appears is (A-position, A'-position, neither), its
nature (argument or not), its status with respect to Case theory (does it
need Case or not), the nature of NP* (pro, PRO, NP-trace...), the
properties of the relation between the Cl and NP* (locality conditions,
binding...).

These are questions of descriptive adequacy. Clearly, the answers
are not uniform for all the clitics involved, for, if they were, we would
expect identical clitic distribution in identical structures, which is
clearly not the case:

(1)   i.   Jean a laissé Pierre le raser
       ii. Jean a laissé Pierre se raser
       iii. Jean l'a laissé raser a Pierre
       iv. *Jean s'est laissé raser a Pierre
       v.   *Jean a laissé le raser Pierre
       vi. Jean a laissé se raser Pierre

(2)   i.   Pierre lui est semblable
       ii. Pierre en est capable
       iii. *Pierre lui croit Marie semblable
       iv. Pierre en croit Marie capable

Because there is no uniform answer to questions of descriptive adequacy,
non trivial questions of explanatory adequacy necessarily arise. On what
basis does the language learner successfully classify the various clitic
constructions encountered in a given language, given that we cannot assume
a uniform and universal analysis for all clitic constructions.
Unfortunately, we will have only scattered remarks to make in this
connection, despite its crucial importance.

Limiting ourselves to French (and perhaps, the Romance languages) let
us distinguish the various clitic constructions and examine them each in
turn. The distinctions we draw a priori here will hopefully be justified
a posteriori. We find:[1]

i.   Subject clitics, or more precisely, clitics standing for subject of tensed clauses, that we will term SCL.

ii. What we have called genitive clitics, i.e. 'en', which in fact, appears in genitive, partitive, quantitative and indefinite constructions(cf.Haik,1981a, Milner,1978)

iii. Locative clitics, i.e. 'y'

iv. Accusative and Dative clitics, which subdivide into pronominal clitics (le, la...) henceforth P clitics (or PCL) and reflexive or reciprocal clitics (me, se...), henceforth R-Clitics (or RCL)

In what follows, we consider (iv) in some detail, leaving aside (ii) (but cf Sportiche, in preparation) and (iii), and making some remarks on (i).

1. Pronominal and Reflexive/Reciprocal clitics

Let us start with clitics in (iv) above. Such clitics are found in structures like (3) below:

(3) $[_s$ NP INFL $[_{vp}$ Cl + V...NP*...$]]$

where, informally speaking, Cl forms a phonological unit with V (which says nothing about the syntactic position of Cl) and stands for some NP that would appear in the position of NP* in the absence of Cl.

Any correct theory of this category of clitics will have to account for the following observations and properties:

(4)   i.   A structure containing a pronominal clitic is interpreted exactly as if NP* were a pronoun.[3] A reflexive or reciprocal clitic is interpreted exactly as if NP* was a reflexive or a reciprocal anaphor.

        ii.   Cl may appear in some syntactic construction only if a lexical NP may appear in position NP* in the absence of Cl.

      iii.   Cl associated with NP* may only appear if associated to a θ-position through NP* (i.e. either NP* is aθ-position or NP* is in a chain containing an NP in a θ-position).

       iv.   Cl and NP* must be in a structural position meeting some some "closeness" requirement.

1.1. Clitic Constructions and A-binding

1.1.1.   Consider first the question of the nature of the Cl/NP* relation. Let us start by supposing that the Cl/NP* relation is a binding relation. Then, depending on whether Cl is in an A or an A' position, the relation will be one of A or A'-binding. We have argued in III.6 that, if it is a relation of A'-binding, it must be a relation of A'-binding established at LF, and that such a consequence was only compatiblie with the properties of R-Clitics. However, we shall see in Chapter IV, and we may simply stipulate for the moment, that such an analysis of R-Clitics violates the principles of the Binding Theory (roughly: NP*, as governed PRO must be A-bound locally, so cannot become locally A'-bound at LF by the R-Clitic). In other words, if Cl binds NP*, we may assume that Cl A-binds NP*. There are important empirical and theoretical difficulties with such an assumption, to which we now turn.

1.1.2.   If Cl A-binds NP*, Cl is by definition in an A-position. We can distinguish between two types of A-positions, and two types only, if we do

not want to empty the notion of A-position of its content.

An A-position is either lexically governed, i.e. is dominated by and only by projections of some lexical category (A,N,V,P) or it is not lexically governed.[4]

If it is lexically governed, it must be a θ-position. This property reflects the essence of the Projection Principle. An A-position governed lexically by some category L is postulated to exist only as an argument place of L. The reduction of the properties of the Phrase Structure component to other systems of grammar is rooted in this basic hypothesis (cf. Chomsky (1981); Stowell (1981)). Applied to clitics, this means that the A-position occupied by Cl is a θ-position. However, if such is the case we lose the only principled theoretical motivation, i.e. the only explanation, for the existence of NP*. Indeed, if Cl and NP* share one θ-role assigned by V, and Cl is in a θ-position, NP* is not. If NP* is lexically governed, we derive an incompatibility with the Projection principle, since NP* would be a θ'-position lexically governed by V. So NP* in fact, should not exist.

This conclusion is not problematic in itself. However, as we have already pointed out, the study of Causative constructions and more generally of the distribution of Cl (cf. e.g. Kayne (1975); Rouveret & Vergnaud (1980); Burzio (1981)...) provide a wealth of empirical evidence in favor of the existence of NP*. This conclusion therefore appears untenable. Furthermore, if NP* is not an argument of V (e.g. NP* is a subject, governed by V, of an embedded proposition as in S'-deletion constructions), we face a different sort of problem. Cl would be lexically governed by V, and in a θ-position, yet would not be an argument of V. Again, this is at odds with the Projection Principle. The

following examples illustrate this point:

(5) i. Marie considérait [$_K$ Jean triste]

   ii. Marie le + considérait [$_K$ [e]$_{NP*}$ triste]

The verb considérer takes only a small clause complement as argument, VP internally. The analysis just outlined would imply that le in (5ii) occupies a Θ-position, yieldinj a violation of the lexical properties of considérer since we would find two VP-internal Θ-positions in (5ii): le and K. (5ii) would thus be predicted ill-formed, contrary to fact. We may therefore conclude that Cl does not occupy a lexically governed position.

Suppose next that Cl occupies an A-position which is not lexically governed. In French, such a position would have to be outside VP, and a natural candidate is the node INFL. This hypothesis about the syntactic position of non subject clitics is not altogether implausible,[5] although it would become so, in my view, if such a position is further assumed to be an A-position, as we now argue.

A non-lexically governed A-position is not necessarily a Θ-position, vide subject positions. However, although it is not always a Θ-position, it has to sometimes be a Θ-position. Otherwise, nothing distinguishes it from an A'-position. If the position occupied by these clitics was an A-position, we would expect to find constructions in which Cl is an argument and bears a Θ-role, which has no equivalent with the clitic replaced by a full NP (or PP). The absence of such constructions in which Cl assumes a variety of Θ-roles as any other A-position, depending on the lexical material governing it speaks against the assumption that Cl occupies an A-position.[6]

Note finally, and this is valid whether Cl occupies a lexically governed position or not, that some a priori considerations suggest that a minimal theory should not count clitic positions as A-positions. First, because clitics are not maximal projections: they do not appear in positions that can hold NP's (or PP's) as well, so that some extension of our theory of A-positions would be needed to cover such cases. Furthermore, it is plausible to assume that the set of A-positions is not only defined once and for all in U.G. (e.g. NP, S , NP, VP ...) at some appropriate level of representation, but also that this set is fully realized in every possible grammar (except in the unlikely event of a grammar containing an extremely impoverished lexicon). For example, we do not expect to find some language with no [NP,VP] A-position. Conversely, we do not expect some language particular position to count as A-position.

1.2. Clitics and Chains

Suppose nevertheless, that these problems can somehow be circumvented and that Cl appears in an A-position, and cooccurs with an empty NP*. The question arises of whether (Cl, NP*) forms a chain or not.

1.2.1 Assume they do, i.e. assume Cl locally A-binds NP* and (Cl, NP*) is (part of) a chain. A number of consequences follow. NP* is assimilated to an NP-trace and as such is a non argument. Since some element must assume the θ-role which we know is available (cf. 4iii above), we are lead to assume that Cl is an argument.[7] Note incidentally that this assumption does not provide an answer to (4iii). It merely pushes the question to a different level. Why are clitics arguments?

Recall furthermore that we have argued that NP-traces could not appear in Case-marked position. We can reconcile this claim with the structure of clitic constructions by assuming, following Aoun (1979), that Cl absorbs the Case that would have been assigned to NP* in the absence of Cl (the existence of that Case follows from (4ii)). Because this alternative treats the relation (Cl, NP*) as being of the same nature as (antecedant/NP-trace), we expect these two relations to pattern alike in similar contexts. Consider, however, the following examples (all intended to be roughly synonymous):

(6) i.   Sarah a entendu Rosa conduire cette voiture
     ii. Sarah a entendu conduire cette voiture par Rosa
     iii. Sarah l'a entendu conduire NP* par Rosa
     iv. *[Cette voiture]$_{NP}$ a été entendue conduire e par Rosa
     v.   Rosa a été entendue conduire cette voiture

Consider the well-formed (6iii). Cl appears on the matrix verb and by hypothesis, absorbs the Case assigned by conduire to NP*. So that we may conclude first that Cl and conduire are structurally close enough for this Case-absorption to take place and secondly that Cl and NP* are structurally close enough not to violate Principle A of the Binding Theory (to which they are subject as antecedent/NP-trace type relation).

Turn now to the ungrammatical (6iv). (6v) shows that passive morphology on a verb signals, among other things the absorption of (objective - in French) Case by the passive morphology of some NP close enough to it. Unless some ad-hoc distinction is introduced,[*] the well-formed Cl/NP* relation of (6iii) indicates that the object of the embedded verb in (6iv) is close enough so that its Case can be absorbed by the Passive morphology on the higher verb. Furthermore, if the relation

Cl/NP* abides by Principle A of the Binding Theory, it means that the local domain in which NP* must be bound contains Cl, hence the whole sentence (since such domains are either NP's or S's). In particular, it indicates that the relation (NP, e) of (6iv) does not violate Principle A either.

So, as we see, the assumption that (Cl, NP*) forms a chain of A-positions makes us lose any potential basis for drawing the required distinction between (6iii) and (6iv). Their divergent status shows that the relation (Cl, NP*) cannot be analyzed as an NP-trace relation.

Consider moreover (4i). There are two ways in which (4i) could be met. It could be derived from independent, necessary principles governing the properties of the clitic constructions or it could be stipulated by ascribing P-Clitics the property of being pronominal and R-Clitics the property of being anaphors (or, more precisely, subject to Principle A). Obviously, this second alternative is less desirable than the first one. It seems, however, that it would have to be adopted of we adopted the chain-theory of clitic constructions just criticized.

So we have seen that the assumption that Cl is in an A-position and forms a chain with NP* not only contradicts some basic theoretical premises but also appears empirically inadequate, and is not minimal in that it requires stipulative answers to (4i).'

1.2.2. Suppose next that Cl appears in an A-position but does not form a chain with NP*. Clearly, NP* can be assumed to be locally A-bound by Cl (if it is not because Cl locally binds NP**, which, in turn, locally binds NP*, we choose NP** as the NP associated with Cl). Consequently, NP* can be assumed to be head of a chain. Indeed, it is not in the same chain as

its local A-binder. As we have noted in (4iii), the types of Cl we are dealing with are always associated with a Θ-position. In other words, NP* is in a chain containing a Θ-position, and, as such, must be an argument, since it is in chain initial position: Cl has thus no Θ-role to assume and must therefore be supposed to be a non argument. Now recall (4ii). Cl may appear iff NP* may be lexical in the absence of Cl. Because of the Case Filter, it is necessary for NP* to be permitted to be lexical, that NP* be Case-governed (at least in the absence of Cl). In particular, since Case is always asigned under government by some X°, NP* is in a governed position.[10] Two different situations might arise. Either NP* retains Case when Cl is present, or it does not:

i. Suppose first that NP* retains Case NP*, as an empty Case-marked head of chain (not locally A'-bound) is pro. As such, it falls under Principle B of the Binding Theory, i.e. must be A-free in S of (3) above (i.e. in its Governing Category). This, however, is incompatible with the assumption that Cl A-binds NP*, since Cl is within that S. Although we may assume that Cl A-binds NP* in a way irrelevant to Principle B, it would in essence come down to giving up the idea that Cl is in an A-position, or that Cl binds NP* (which implies that Cl is not in an A-position).[11]

ii. Suppose next that NP* does not retain Case. NP* is then PRO, since it is an empty Caseless head of chain. However, we have seen that NP* is in a governed position. If we accept Chomsky's Binding Theory, which tolerates no governed PRO, this is excluded. We, however, assume that PRO may be governed, provided that it is bound in its Governing

Category (cf. Chapter IV for justifications). Here, however, it seems that we do not want to count A-binding by Cl as fulfilling the requirements of Principle A of the Binding Theory. PRO is an argument, and such elements falling under Principle A (e.g. each other, reflexives) need to be bound by elements associated to an argument. Each other, for example, requires an antecedent ultimately assigned a reference. As illustration of this point from another angle consider the following structure:

(7)   i.   John thinks that it seems [PRO to be sick]

(7) is ill-formed (it could otherwise mean John thinks that he, John, seems to be sick). If it could be taken as valid antecedent for PRO with respect to Principle A of the Binding Theory (recall PRO is governed since seem is an S'-deletion verb), we would not be able to explain the ungrammaticality of this form. If, however, it cannot function as a proper antecedent, just as Cl for NP*, (7) will violate the Binding Theory. So again, it looks as if the binding of NP* by Cl should not count as A-binding.

We see, in other words that each alternative requires some ad-hoc adjustments. Assume, nevertheless, that we can amend the Binding Theory so that in both of these two alternatives, Cl can be taken to A-bind NP*.

How then do we explain the observation (4i)? Just as in 1.3.1., we would have to stipulate that Cl is pronominal, or anaphoric in order to get the desired result. Note however, that if, somehow, Cl were not taken to A-bind NP*, alternative (i) above would automatically yield a pronominal (it is pro). Similarly (ii) above would automatically yield an anaphoric interpretation (PRO is an anaphor, cf Chapter IV).

Again, assuming that Cl does not A-bind NP* leads to a more explanatory theory.

1.3. Clitic as Agreement marker

Let us therefore, suppose that Cl does not A-bind NP*, and, more generally, that it is not in an A-position, leaving open the question of whether Cl is VP-internal or in INFL (although, as we shall see, the fact that Cl affects syntactic properties of V suggests that it behaves as a morphological affix and should be attached to V in syntax).

As we have noted, a reasonable theory of clitics must not only ascribe each clitic construction the correct analysis, but also provide a basis for explaining how the correct analysis is imposed by U.G. on the construction in question.

In this respect, we have already noted (and cf. Bok-Bennema (1981) for similar remarks on Spanish, and conclusions close to ours) that if the notion of A-position (and, perhaps of A'-position) is universally restricted to positions that may hold maximal projections, e.g. NP's, PP's..., the hypothesis that Cl is in an A-position cannot even be made, since the position holding clitics accept only a very restricted type of nominal elements. As we have also argued that Cl do not occupy A'-positions, we must define the relation between Cl and NP* as different from a binding relation. Adapting a terminology proposed in Jaeggli (1980), let us say that Cl identifies NP* and let us represent this relation by cosuperscripting as in (8) below:

(8) $Cl^i + V...NP*^i$

Specifying further our assumptions, let us assume that Cl, which displays nominal characteristics, e.g. person, number and gender features is a [+N] affix agreeing with NP*.

Since Cl does not bind NP*, we expect NP* to behave as an autonomous category with respect to the Binding Theory, i.e. exactly as if Cl were not present.

Let us now consider how the properties of these clitics minimally follow from these assumptions. Consider (4ii) first. (4ii) states that a clitic may appear affixed to V, associated with NP*, only if NP* can be lexical in the absence of Cl. Since it is necessary for lexical NP's to be in Case-marked positions, because of the Case Filter, we can express (4ii) alternatively by saying that Cl can only be associated with a Case-marked position. This can, in turn, be reasonable derived from the Case Filter. Recall that we have argued in Chapter II that, because the Case Filter holds of non arguments, it must be stated as an independent principle. Since clitics display nominal characteristics, we have assumed that it is a nominal element and, as such, we may assume that it falls under the Case Filter.[12]

In other words, from the assumption that Cl is [+N], we can derive (4ii).

Two different situations might arise depending on whether the affix Cl absorbs the Case otherwise assigned to NP*, or merely agrees with a Case-marked NP*. The analysis of these two alternatives will enable us to provide a principled answer to observation (4i):

(i) Assume Cl simply agrees with NP*. NP*, as a Case-marked empty category must be in chain initial position (recall that NP-traces cannot be in Case-marked position) and is therefore a pro. As we have pointed

out in 1.3.2 above, in such a case, the clitic construction will behave exactly as if NP* was a lexical pronoun with respect to the Binding Theory since NP* is pronominal. So, if we assume that P-clitics are those falling in the category under discussion, i.e. agree with NP* but do not absorb its Case, we derive minimally (4i) for P-Clitics.

(ii) Suppose next that NP* does not retain Case in the presence of Cl. Must NP* be in chain initial position? The argument given above, based on the Case-marked character of NP* cannot be extended to this case. Let us therefore investigate the two possibilities.

Suppose first NP* is in chain initial position. Then, since it is Caseless, it is PRO by definition. As we have seen, since NP* is governed, this is incompatible with a Binding Theory requiring that PRO be ungoverned (cf. Chomsky (1981; 1982)). However, we assume that governed PRO's are licit and behave as anaphors, i.e. they must be bound in their governing category (cf. Chapter IV). Since here, NP* will be governed PRO, we automatically derive the anaphoric properties of R-Clitic constructions, if we assume that R-Clitics are precisely Clitics absorbing the Case of their associated NP*. We thus derive minimally the essential properties of clitic constructions.[13]

.   Suppose next NP* is not in chain initial position. Call NP the local A-binder of NP*, which is, by assumption, in the same chain as NP*. Recall that NP* is in, or is associated with a θ-position (cf. 4iii). NP must therefore be a θ'-position, hence the subject of V in (3) is the only admissible candidate (because of Principle A of the Binding Theory), since it is the only θ'-position in the Governing Category of NP*. Now, because of (4ii), in the absence of Cl, V would be a verb assigning Case to its object (or the NP it governs) but no θ-role to is subject.

If generalization (24) of Chapter II (Burzio's generalization) is correct, such verbs do not exist. On the contrary if (24) is incorrect, we would expect with such verbs alternations:

[il [V NP] / [NP [se + V [$_{NP*}$ e]]]

with il an expletive and (NP,NP*) a chain.[14] Before summing up, notice that the analysis of R-Clitics given above contradicts the generalization (15) of Chapter II (if no Case to the object, then no θ-role to the subject). However, as we have pointed out these, the argumentation relying on (15) relied in fact on a narrower generalization of (15) unaffected by this analysis of R-Clitic constructions.

1.4. Summary

Let us sum up the main features of our proposals concerning Accusative and Dative clitics:

(9)   i.   Cl not being NP's are not in A-positions and cannot enter into A-binding relations

        ii. Cl is a [+N] element subject to the Case Filter thus explaining (4ii), which identifies NP* by being cosuperscripted with it

        iii. If Cl is an R-Clitic, it absorbs the Case of NP*. The anaphoric interpretation of the R-Clitic construction follows from NP* being a PRO.[15]

        iv. If Cl is not a R-Clitic, it does not absorb the Case of NP*. The pronominal interpretation of the construction comes from NP* being pro. This, in conjunction with (iii) above explains (4i).

Two consequences also follow from this analysis. First, we derive that clitics of this sort neither trigger WCO effects, nor license parasitic gaps, as we have seen is the case,[16] since they do not involve binding at all. Secondly, and more importantly, we have seen that any analysis that we have reviewed (including the one we adopt) which observes

some standard of minimality treats NP* as a governed PRO at some level. Our binding theory must be constructed accordingly.

Consider finally (4iv), i.e. the question of the locality requirement that must hold of the pair Cl/NP*. We have so far callled this relation an identification relation. The basic assumption we have made concerning it was that it was essentially a relation of Agreement[17] between Cl and NP*. The natural assumption is then that the locality condition holding of the pair (Cl,NP*) or perhaps (Cl+V,NP*) is exactly that holding of pairs (x,y) which are in Agreement relation.

Agreement also holds of pairs (x,y) where x binds y. What is the difference between Binding and Agreement? The obvious difference is that Binding may hold only of pairs (x,y) where x and y have identical categorial status, while Agreement may hold of pairs (x,y) where x and y have (perhaps always) different categorial status Binding may hold of pairs of NP's, pairs of PP's..., Agreement relations hold of pairs (AP,NP), (Art,N), (NP,INFL), (INFL,V) etc...Notice that binding requires some kind of agreement but is distinct from Agreement in that it does not require Case-agreement whereas Agreement does, when possible (In this connection, notice that R-Clitics do not exhibit Case-agreement, as predicted if NP* lacks Case). Agreement relations are subject to the following locality constraint:

(10)   (x,y) is an Agreeing pair only if x governs y,
            or vice versa (or both)

For example, it can be shown that subject verb agreement is mediated by INFL (cf. Kayne (1972), for some arguments) and in fact decomposes into (NP,INFL) and (INFL,V) Agreement relations (recall that government of NP by INFL percolates to the head of VP). Similarly, most of

subject/predicate adjective agreement can be reduced to (10) with the help of small-clause theory e.g. they are fools is really:

they$_i$ are [t$_i$ fools] (cf. Stowell (1981)...

Applied to the (Cl/NP*) or (perhpas) to the (Cl+V/NP*) relation, we see that first it holds of pairs of element with different categorial status (one is [+N] the other is NP= [+N,-V]^2) and it requires government of one of the two members by the other. This provides an answer to (4iv).[14]

Note that it suffices for the language learner to observe feature variations amongst clitics to deduce that it Agrees with NP* and thus that the relation (Cl/NP*) is sujbect to the locality Condition (10).

2. Subject clitics

Of all the observations listed in (4), we have said nothing of (4iii). The reason is that the rationale we may provide for (4iii) depends on the analysis we adopt for Subject clitic (SCL). In fact, a full discussion of (4iii) (and (12iii)) below will also necessitate a deeper understanding of the workings of the Binding Theory and the ECP. Let us therefore postpone it until the next chapter and, in the meantime discuss the analysis of SCL. Let us recall some basic assumptions about SCL.[19] Consider a SCL construction as in (11):

(11)   [$_S$ NP*   INFL   [SCL] [$_{VP}$...
                       +tense
                       +AGR

Kayne (1972) provides ample evidence that SCL do not behave as NP's.

Following Jaeggli (1980), who bases his conclusions on Kayne's extensive discussion, let us suppose that SCL is a property of the INFL node.[20] As we remarked, SCL can only appear in tensed clauses. Since, in French, [+tense] and [+ AGR] always cooccur, it is not clear to which of these features SCL are actually linked.

2.1. Analysis of SCL Constructions

Let us now turn to the analysis of SCL constructions, i.e. the nature of NP*, and of the relation (NP*/SCL). Again here, I will not review recent proposals on the topic but rather examine the consequences of our theoretical assumptions on the analysis of these constructions (but cf. Jaeggli (1980); Safir (1982)).

Except for complex inversion constructions, which we will discuss in 2.2 below, NP* is empty when a SCL appears in INFL. The fact that a lexical NP may sometimes appear supports Chomsky (1982) Extended Projection Principle or, more precisely, that part of this principle stating the obligatory character of subject positions. We will provide some more direct empirical evidence for this assumption in 2.2. below. There are a number of similarities but also some differences between SCL and object clitics.

(12) i.   A structure containing a SCL is interpreted as
           if NP* were a pronoun
      ii. A SCL may appear only if NP* may be lexical
           in the absence of SCL
      iii. SCL may be associated to any type of NP*,
           arguments, quasi-arguments or non-arguments,
           contrary to object clitics
      iv. The relation between NP* and SCL meets some
           locality condition

As we can see, (i), (ii) and (iv) parallel similar observations for object clitics, while (iii) differs.

These similarities a priori speak in favor of simply extending the analysis we have proposed for P-Clitics to SCL. This analysis would explain immediately the similarities, i.e. properties (12i), (12ii) and (12iv) above in exactly the same fashion it did for P-Clitics. It would furthermore complete the casual paradigm of NP* positions linked to a Clitic. Since Agreement relations must meet a government requirement and given that clitics appear in INFL and affixed to V (in French), the set of NP positions that may be linked to a clitic is narrowed down to Nominative, Accusative and Dative. Postulating an identity in the relations (SCL/NP*), (P-Cl/NP*) would fill a gap in the Case paradigm of possible P-clitics. We would simply hhave SCL = nominative P-Cl. Finally, this would explain why these SCL parallel all the additional characteristic properties of Object P-Clitics as Kayne (1975, pp. 84-90) shows.

In other words, the a priori derivable conclusion is a uniform analysis of SCL and P-Clitics.

Let us now proceed more systematically and examine to which analysis our theoretical assumptions lead us. Could SCL be in an A-position binding NP*? Some of the argumentation adduced against such an assumption in the Case of object clitics carries over. First, SCL are not NP's. This assumption would require an implausible extension of the inventory of A-positions (Note however, that the Extended Projection Principle guarantees the existence of a position NP*, so that that part of the parallel argumentation on object clitics does not carry over).

If SCL forms a chain with NP*, it must be considered an argument and

NP* an NP-trace. We will see in 2.2 below that the Complex Inversion facts suggest that such an analysis is implausible: if SCL were an argument, it should enter into coreference relations as any other pronominal argument. But it in fact does not.

Now, if SCL does not form a chain with NP*, either SCL absorbs the subject θ-role if any (hence NP* is an expletive element) and it is an argument and we get the same problem as above. Or SCL does not absorb the subject θ-role and (as we have seen in the Case of P-Clitics) there is no reason left and every reason to assume that SCL binds NP*.

So SCL is not in an A-position. Could SCL be in an A'-position, binding NP*? Here, the argumentation given for object clitics based on WCO and PG cannot be extended at all since, due to the way the locality condition mentioned in (12iv) can be realized, the relevant situations can never be constructed. We can however, take advantage of the observation (12iii). Recall that, quite generally, locally A'-bound positions must be (linked to) θ-positions. As (12iii) indicates, the assumption that SCL A'-binds locally NP* would lead to the loss of this generalization. We would have to distinguish two types of A'-binding; SCL binding of NP*, where NP* is not necessarily an argument and every other type of A'-binding in which the locally A'-bound element has to be an argument. This assumption thus appears artificial.

Consequently, this argumentation substantiates what seemed a priori desirable. The relation between SCL and NP* is not a relation of binding but can be assimilated to other (Cl/NP*) relations. It is an Agreement relation as appeared a priori plausible from the fact that SCL and NP* have different categorial status.

Is Case still assigned to NP* as in the case of pronominal Object

Clitics or is it not - SCL absorbing it - as in the case of R-Clitics.

Suppose NP* receives no Case. Since it is in subject of tensed clause position (and does not form a chain with INFL), it is Caseless and in chain initial position, i.e. is PRO. Moreover, it is governed by INFL:[21] this is excluded. Either because we have a governed PRO, or because we have a governed PRO not bound in its governing category. It must then be that Case is assigned to NP*, so that NP* is pro.

Summing up, we analyze the (NP*/SCL) relation of (11) as in (13) below:

(13) $[_s [_{NP} \text{pro}]^i [_{INFL} \text{SCL}]^i [_{VP}...]$

where the subject NP* is pro and agrees with a nominal SCL in INFL, an Agreement relation that we will represent by cosuperscripting. Notice incidentally that this analysis makes French into a Null Subject language. Of course, it would remain to explain why French does not have all the properties usually attributed to the Null Subject languages (cf. Sportiche, in preparation, for a discussion). We mentioned earlier that it was unclear whether the presense of SCL was dependent on [tense] or AGR. Note that a very natural assumption that would answer this question is that SCL is what is referred to as AGR (or more precisely, represent it phonetically).

2.2. Complex Inversion

Very interesting (partially) confirming evidence for the above analysis comes from observations and analyses made in Kayne (1972) (and taken up in Jaeggli (1980) and Safir (1982)) that we almost exactly adopt,

and to which we refer the reader for more details. Kayne (1972) remarks the following three sets of facts. First, the complex inversion construction, which is essentially a main clause phenomenon, shows a SCL cooccurring with any non empty NP*, as (14) illustrates:[22]

(14)   i.   Quand Jean a- t-il mangé
        ii. Depuis quand tout est-il en ordre

(14) further indicates that this clitic doubling situation may occur even with a class of NP's such as tout, rien, cela, sentential subjects, ...that we will call H'-NP's for ease of reference.

Second, although a SCL can double any NP whatever, a SCL cannot "refer back" to H'-NP's:

(15)   i.   Jean$_i$ est parti parce qu'il$_i$ avait fini de manger
        ii.* Tout$_i$ est en ordre maintenant, mais il$_i$ sera en
             désordre demain

Thirdly, the same pattern holds of the relation between an NP and a full pronoun: coreference is not permitted with H'-NP's:[23]

(16)   i.   Jean$_i$, je pense souvent a lui$_i$
        ii.* Tout$_i$ est tombé parce qu'elle s'est appuyée sur lui$_i$

As Kayne (1972) notes, we get a unified explanation for the last two sets of facts, which is furthermore compatible with the first one, if we assume that the structure of these SCL constructions is:

[$_S$ [$_{NP^*}$ pronominal] [$_{INFL}$ SCL] ...],

i.e. one in which the subject position is itself a pronominal element. The ungrammaticality of (15ii) and (16ii) result from the impossibility for a pronominal element to "refer" to H'-NP (whatever the reason for

that may be).[24] In particular, the SCL does not enter itself in coreference relations and structures like that of (14ii) are thus permitted.

Clearly, this analysis supports our conclusions. It argues both for the existence of a position NP* (thus providing independent evidence for the Extended Projection Principle), for its analysis as a pronominal, as in (13), and for the fact that SCL does not participate in Binding relations. The only point of difference with Kayne's analysis is that we assume this pronominal NP* to be empty all along, whereas he assumed that it was not empty and deleted later. Also, note that the grammaticality of the following sentences:

(17)   i.   [Que Jean ait dit cela]$_i$ indique que S, sans
              PRO$_i$ indiquer que S'

         ii. Tout$_i$ se disait dans le but PRO$_i$ d'être répété

in which a PRO refers back to H'-NP's, provides additional support for our treating NP* as Case-marked, and not as Caseless, ie. as pro instead of PRO.[25]

As a final point, the fact that the pattern found in (15) above is reproduced with P-Clitics in place of SCL suggests that it was correct to analyse P-clitics and SCL alike.[26]

2.3. Clitic Doubling

Our analysis has some consequences with respect to the analysis of Clitic Doubling that we now review briefly (but cf. Aoun (1982); Borer (1981); Jaeggli (1980); Ingria (1981); for recent discussion).

By clitic-doubling, we mean a construction as in (3) or (11) above, in which Cl or SCL cooccur with a non empty NP*. If the question of Clitic Doubling arises, it is because it is generally assumed that the unmarked situation is one in which NP* is empty. So, we may ask the following two questions:

i. why is it that NP* must be empty in some clitic constructions (e.g. P-Clitics in Standard French) while it seems not to have to be in others (objects in River Plate Spanish, French Complex Inversion)?

ii. if the following generalization attributed to R. Kayne in Jaeggli (1980) is correct: if NP* is Clitic Doubled, it is flanked by a P(reposition). Why is it true?

Let us note that although some clitic Doubling constructions seem to accord with Kayne's generalization, some others do not seem to (e.g. Algonkian object agreement markers, French Complex Inversion). Of course mere observation of a cooccurrence of a clitic and a non empty NP "associated" with it does not suffice to establish the truth or falsity of Kayne's generalization, it must also be decided whether that NP is in position NP*...

Consider the cases where K's generalization holds. Aoun (1979) has made a suggestion tying together (i) and (ii) above. Aoun suggests that Cl absorbs the Case that would otherwise be assigned to NP*. If follows by the Case Filter that NP* must be empty, unless, Aoun argues, some rescuing device, e.g. insertion of a Case-marker like a P, Case-marks NP*.

Now it is clear that this suggestion is not directly compatible with our analysis (as far as French, and perhaps, the Romance languages are concerned), as we have concluded that NP* is a Case-marked position even in the presence of Cl so that we must suppose that the explanation for (i)

does not lie straightforwardly in Case Theory as Aoun (1979) suggests. Perhaps clitic constructions parallel wh-constructions in a sense. In both cases, we find a Case-marked position (NP* here, variables in wh-constructions) which does not tolerate phonetic material systematically. Clitic doubling is thus parallel to resumptive pronouns. Some common explanation might be found. We have none to offer.

Before leaving the topic, it is worth pointing out that the non universal validity of Aoun (1979) suggestion - which is discussed in Aoun (1982, Chapter III) - shows that even in the grammars where it might be taken to hold, this suggestion displaces the question, admittedly to a higher level, for it is now the fact that certain languages use a rescuing device, while others do not, or need not, that requires explanation.

FOOTNOTES APPENDIX

1. Appart from scattered remarks, we will be ignoring inherent-Clitics and middle (or ergative clitics (se).

2. The following table lists the various clitics of (iv) found in French.

	Singular		Plural		
	Accus.	Dative	Accus.	Dative	
1st pers.	me	me	nous	nous	PCL
	me	me	nous	nous	RCL
2nd pers.	te	te	vous	vous	PCL
	te	te	vous	vous	RCL
3rd pers.	le/la	lui	les	leur	PCL
	se	se	se	se	

3. For some speakers, in French, overt pronouns (strong forms) tend to pick [+human] referents (cf. Kayne (1972) for relevant remarks).

4. Note that a position may be governed, without being lexically governed: subjects by INFL, NP's by V under exceptional government (S'-deletion).

5. This has been proposed by Belletti (1980) for Italian, J.C Milner (GLOW Colloquim, Pisa), for French and Koopman (1983) for Gbadi.

6. Note however, the existence of a class of such constructions with very specific semantic interpretation rules, the so-called ethical Dative construction (and perhaps, although it is not as clear, the Clitics of unalienable possession again – if in this category of constructions – with very specific semantics). Cf. Kayne (1975); Jaeggli (1980) for some discussion of these constructions.

7. Of course, this follows only if Cl is the head of the chain containing NP*. This must be insured some way since, Cl being ≠e cannot be an NP-trace.

8. Whose formulation would not be easy given that the analogues of (6) are grammatical in Italian (cf. Burzio (1981, p. 37, p. 576, for examples with Faire constructions). Note that the French Faire-constructions are not telling since they do not allow passive morphology.

9. Notice that the argumentation of this last section carries over to a chain-theory of clitic constructions in which Cl is in an A'-position (with all the obvious extensions that it requires of the theory of

chains) of Chomsky (1982).

10. Recall that we have argued against the possibility that Cl may prevent (or absorb) governement, as Jaeggli (1980) proposes (cf. Chapter III, section 6).

11. Remark that it would not be possible to restrict Principle B to A-binding in the following way. Pronominals must be argument A-free..., because of ill-formed examples like

    "John was believed t_i to have seen him_i "

    in which it is the A-binding of $\underline{him}$ by the non argument t_i which is ruled out by Principle B. Similarly, it is not possible to restrict the scope of Principle B to freedom from A-binding by a θ-position because of ill-formed examples as:

    " * $John_i$ seems to him_i t_i to be sick"

    where it is the A-binding of $\underline{him}$ by $\underline{John}$ in a θ'-position that Principle B excludes (Note incidentally that Principle B says nothing of the relation him/t_i , since t_i is not a pronominal.

12. Note that this requires a generalization of the Case Filter to:

    $*X=[+N]^n$ (or perhaps $*[+N,-V]^n$) if X is phonlogical and Caseless

    (instead of *[NP], NP phonological with no Case). Note that, this reformulation is compatible with the analysis of gerunds adopted in Chapter IV. Note also that, if we postulate zero clitics, they would have to be assumed [+N] as well, unless they can be associated with Caseless NP*.

13. But not all their properties. First, some additional mechanism must be introduced - in any theory - to account for the reciprocal interpretaion (cf. Chapter IV, section 2, for relevant discussion). Also, note that we must prevent the local antecedent of an NP* associated with an R-Clitic to be a non subject:

    * Tu s'est présenté les $enfants_j$ e_j

    (cf. Huybregts (1979); Rizzi (1983) for some proposals, incompatible with our analysis).

14. There are no clear examples of such verbs. Notice, however, that if we assume that processes of Case absorption may always be accompanied with subject θ-role absorption as well either obligatorily, as Chomsky (1981) suggests (Passive morphology) or optionally (R-Clitic morphology) as we might suggest here, we derive the syntax of ergative (or middle, anticausative) R-Clitic constructions.

15. This treatment of R-Clitics is also suggested in Bouchard (1982).

16. Cf. section 6 where examples are provided for P-Clitics. I have been

unable to construct relevant examples with R-Clitics.

17. We use Agreement with Capital A to designate the formal binary relation, as distinct from agreement which is the informal way of talking about the facts that certain features must be identical in certain binary relations (e.g. binding).

18. We return in Chapter IV to the identification relation performed by the clitic when we discuss pro and the ECP.

19. Here is a table of SCL in French.

	Singular	Plural
1st pers.	je	nous
2nd pers.	tu	vous
3rd pers.	il/elle/on/ce	ils/elles

Besides on and ce and 3rd person expletives, SCL are all definite in interpretaion. On is either arbitrary or specific indefinite or an equivalent to nous. We will in fact do not analyze on and ce but cf. Kayne (1972) on on and ce, Jaeggli (1980) and Safir (1982) on ce.

20. But cf. Kayne (1983) for recent discussion of this point.

21. Recall that we have rejected an analysis like Jaeggli's in which government may be absorbed (cf. also Rizzi (1982, Chapter IV, Appendix, for relevant discussion).

22. The precise conditions governing this Doubling are complex. Cf. Kayne (1972); Safir & Pesetsky (1981); Safir (1982a), for some discussion).

23. Recall (cf. footnote 3), that for some speakers, full pronouns tend to be preferred with [+human] referents. (16ii) is uniformly unacceptable, however, as Kayne (1972) notes.

24. Kayne (1972) suggests that pronouns cannot refer to headless NP's, i.e. NP's with no head nouns (hence our notation: H'-NP's). However, as he and Jaeggli (1980) note, it is not clear how this could be extended to such H'-NP's as rien or idiom chunks like (porter-) secours,...

25. It also suggests that PRO may not be pronominal at all, as we argue in Chapter IV, at least in these contexts.

26. Note however, that contrary to the Case of Object clitics, no assignment of θ-role to NP* in the presence of a SCL is not excluded. It was with object clitics because it would have implied the non existence of NP*. Not so for SCL given the obligatoriness of subject positions. Safir (1982) suggests (roughly) that the SCL ce is of this type. [ce displays some idiosyncratic properties, as Kayne

(1972) shows, e.g. it may be a SCL only if it is immediately string adjacent to the verb <u>etre</u> (or in certain fixed expressions with another Cl intervening which begins with a vowel as <u>c'en est trop</u>, but not always * <u>c'y était</u>)]. As Safir (1982) notes, analyzing <u>ce</u> in this fashion would explain why it does not permit Doubling *<u>cela est-ce vrai</u>. In our terms, the subject position NP* would be analyzed as an expletive <u>pro</u> (an empty version of the English <u>it</u> of <u>it</u> seems...) found in

[_S [_{NP} pro] [_{INFL} il] semble ...]

CHAPTER IV: †-CATEGORIES AND BINDING THEORY

1. The Nature of the Problem

1.1. Introduction

In the preceding chapters, we have dealt in some detail with several problems related to NP-distribution, with some emphasis on the import of θ-theory and Case theory to that question. We have however, frequently called upon the principles of the Binding Theory and the ECP throughout, without ever discussing their form and structure. We now turn to such a discussion, exploring also some consequences of the theoretical choices we favor.

So far, we have been basically assuming the theory of Binding presented in Chomsky (1981), although we have repeatedly hinted at some modifications we wanted to introduce.

Because of the very intricate nature of the relations between the various subcomponents of grammar, some of the arguments presented in the previous chapters had to rely on some consequences of the structure of the Binding Theory presented in Chomsky (1981). We have to make sure that whatever modifications we introduce do not alter the conclusions of the arguments which depend on some specific property of the Binding Theory. Some other arguments lead to conclusions incompatible with that version of the Binding Theory. We have argued that a principled approach to R-Clitic constructions analyzes the empty category it contains as a governed PRO. For consistency, we must introduce modifications of the Binding Theory

that will have the effect of permitting instances of governed PRO.

Before we begin, let us make a note about the ECP. We will from now on designate under this name whatever principle (or sets of principles) that imposes specific restrictions on the distribution of non overt elements (and more particularly of pro) without prejudging its content. In order to use a neutral terminology, we will say of (certain) non overt elements that they need to be identified, in a sense to be made precise as we proceed.

1.2. Questions on the Binding Theory

Putting aside Principle C of the Binding Theory, which we have already discussed in Chapter III, section 7, and will be of no relevance to our present concerns, we can say that the Binding Principles reduce to the conjunction of statements in (1):[1]

(1)   i. $x \in A$ must be A-bound in $D(x)$
          ii. $x \in B$ must be A-free in $D'(x)$

where A and B are non necessarily disjoint sets and $D(x)$ and $D'(x)$ are syntactic domains, which must at least contain a governor of x, and whose value may be fixed independently of the nature of x, or not.

Such a formulation raises a number of closely interrelated questions that we will try to address each in turn:

(2)   i.   What kind of elements belong to sets A and B.

         ii. How is membership to A and B determined

        iii. How are these elements interpreted

        iv. How is $D(x)$ (or $D'(x)$) computed for a given x
              and do we have $D = D'$ for any x

v.   At what level of representation must the principles
     in (1) be met

vi. Coupling the principles in (1) with the consequences
     of the other subtheories that we have reviewed
     (Case-theory, and θ-theory), do we account
     exhaustively for the distributional properties
     of the members of A and B

Considering (2i), we may provisionally adopt the assumptions shared by all the recent formulations of these principles, namely that the elements concerned by (1) are lexical anaphors (X-self, each other...), pronouns, and non overt categories (PRO, pro, NP-trace...), or, to put it in a different perspective, lexical anaphors and ϒ-categories.

Before deciding whether such and such a ϒ-category belongs to set A or set B, we must have a better idea of which ϒ-categories there are. Obviously, this knowledge will crucially influence the way the principles in (1) (and the ECP, if necessary) should be formulated, so that we can maximize, as far as naturalness and simplicity permits, the work that they will do in the area of NP distribution. Note also that, reciprocally, a choice among different sets of predicted ϒ-categories may be motivated on the grounds that it permits a formulation of the Binding Theory which naturally predicts why we observe the ϒ-categories we do and only those.

1.3. Are lexical anaphors ϒ-categories?

Note first that we have assumed throughout that lexical anaphors (e.g. X-self and each other) are not ϒ-categories. Recall, as Chomsky (1981) points out, that these items should somehow be distinguished from pronouns by some grammatical feature, whatever it is, which will fall outside the set of ϒ-features. So that lexical anaphors are not

ϒ-categories. This is no argument, however, but rather is true by definition since ϒ-categories, we have assumed, are categories bearing no other grammatical features than those of pronouns. There is a more interesting way to look at ϒ-features, implicit in Chomsky's definition. Assume that, given a ϒ-category in some context in a string, the values of the ϒ-features composing it may be chosen arbitrarily. We may define the set of ϒ-features so that only the correct combination of values will be allowed to occur by the various principles of grammar in that context (Of course, we wish to eventually exclude from consideration language particular rules that may affect ϒ-categories, which specify the way it is spelled out phonetically, and which are obviously idiosyncratic).

In this light, we may conclude that a theory a priori maximizing the size of the set of ϒ-features will have a higher predictive value, hence, is a priori preferable. This desirable increase is moderated by the possibilities to attain the goal just outlined: to provide an account for the distributional properties of any ϒ-category on the basis of well-motivated and explanatory principles.

Applying these remarks to lexical anaphors, we see that it is a priori desirable to consider them to be ϒ-categories. If they are, we are in effect claiming that there is no difference between lexical anaphors and empty anaphors apart from those following from some difference in value of some ϒ-feature composing them, the differences here being either the presence or absence of phonetic features, or their status as argument or non argument. However, the following observation seems to me to support the idea that some idiosyncratic feature distinguishes lexical anaphors from their empty counterparts. Postponing the question of how we decide it, which will become clear later, assume that, on the hypothesis

that lexical anaphors are ↑-categories, their empty counterparts are what have called PRO and NP-trace. Clearly, NP-trace being a non argument, will systematically differ in its properties from lexical anaphors, which are arguments. (Note that we should however, expect anaphoric phonetic non arguments to exist, paralleling NP-traces: we return to this below in section 2). This is not so for PRO however. In particular, we should expect that the range of interpretation of PRO exactly matches that of lexical anaphors since, whatever features make reflexive a reflexive or reciprocal a reciprocal, they should be, on our assumptions, able to freely be assigned to PRO as well. This prediction does not seem to be borne out, as the following examples illustrate:

(3)   i.   They expect [PRO to leave]
        ii. They expect [themselves to leave]
        iii. They expect [each other to leave]

A sentence like (3i) in which the non overt subject of to leave is identified as PRO cannot be freely interpreted either as (3ii) or (3iii) as we would expect, were both types of lexical anaphors ↑-categories. In fact (3i) is interpreted exactly as (3ii) a significant fact to which we will return. More generally, it seems that no PRO can ever be interpreted as a reciprocal.[2]

Two possible conclusions follow from this observation. First, we might conclude that lexical anaphors are ↑-categories but some yet to be discovered principle prevents an empty anaphor from being treated as a reciprocal - the hidden variable theory. For lack of directions to follow towards the discovery of some non ad-hoc principle having this effect (as, e.g. [+reciprocal] ---> [+Case]) we will not pursue this alternative. Rather, we will consider the second possible conclusion, namely that

reciprocal phrases are not ϒ-categories. This leaves open the question of whether reflexives should be considered ϒ-categories or not. We will explore this issue in the following sections.

1.4. The Problem of ϒ-Categories Classification

We can now return to the questions listed in (2) and more specifically (2i), (2ii) and (2iii). We can address these questions in a slightly different way. Categories subject to the Binding Conditions exhibit three different sets of properties:

(4)   i.   Properties of interpretation: some are
          necessarily referentially dependent (e.g.
          lexical anaphors), while some others may possess
          independent specific reference (e.ç. pronouns),
          some may be arguments, while some others are not.

       ii. Behaviour with respect to the Binding principles:
          some belong to set A, some belong to set B.

      iii. What we might call structural and functional
          properties: being Case-marked, appearance in
          chain initial or internal position, being
          phonetically null, being locally A'-bound...

Given the set of all elements falling under the Binding Principles in (1), i.e. the elements of the set A U B, it might be that we have to stipulate for each element what its properties are with respect to each of (4i, ii,iii). However, we surely expect there to be implicational relations between all these properties, as e.g. being necessarily referentially dependent and belonging to set A etc...

Whatever the interconnections between these three sets of properties (total independence at one extreme, one to one correspondance at the other) it will induce a classification of categories along the various

independent dimensions necessary to postulate in order to attribute each expression its correct properties. In particular, it will induce a classification of ?-categories and thereby predict the existence of a certain number of ?-categories. The adequacy of these predictions will be evaluated on two different grounds:

(5)   i.   Exhaustivity: no needed type of ?-category has
              to be postulated beyond those predicted by this
              classification

        ii.   Principled reasons (Binding, ECP...) can be
              constructed that explain the non occurrence of
              categories which are predicted to exist but do
              not surface

2. The classification of ?-categories.

Let us now turn to the parameters involved in the problem of ?-categories classification.

2.1. Classificatory Dimensions

2.1.1. Let us start with the Binding principles themselves. Categories falling under the Binding principles subdivide into 3 subclasses depending on whether they must obey (1i), (1ii) or both. Adopting the usual terminology, let us call (1i) Principle A of the Binding Theory, and (1ii) Principle B. Say that we attribute to x the feature [+A] if x obeys Principle A (i.e. belongs to set A), [-A] otherwise, similarly for [+B] and [-B].

2.1.2. Turning now to interpretive properties of these expressions, let us follow Chomsky (1982) in assuming that expressions are assigned to semantic classes defined by the features [± anaphoric], [± pronominal]. These distinctions have to do with the way in which these expressions select some element in D, where D is some postulated domain of mental entities accorded no ontological status apart from mental representation. Limiting ourselves to elements falling under the Binding conditions, we can define the content of these features by examining the behaviour of overt elements, i.e. pronouns and lexical anaphors.

Simplifying somewhat for the moment,[3] we may say that a pronoun P may always select some element in D (or rather in subset of D determined by P's grammatical features) whose identity is not determined by P (but may be left unknown, or be determined by contextualy indications). Let us assign the elements under discussion sharing this mode selection in D the feature [+pronominal]. Consider next lexical anaphors. A lexical anaphor always fails to denote. Rather, it is in effect interpreted as a variable assigned its value by virtue of its relation to its antecedent. Let us assign elements sharing this behaviour the feature [+anaphoric]. In a sense, we may say that the difference between pronominal elements and anaphoric elements is that pronominals may select directly in D while anaphoric elements can select only indirectly in D.

Note that in general we cannot take [+anaphoric] to be the opposite of [+pronominal] (but it may be true over some subset of expressions as we mention in section 2.4 below) for there exist linguistic expressions that behave differently from both. For example, names do select directly some element in D as pronouns, but may identify it by virture or their inherent properties contrary to pronominals.[4]

Again, we expect a priori 4 categories of expressions exemplifying the four possible combinations of the features [±anaphoric], [±pronominal].

Furthermore, categories can be semantically distinguished as being arguments or non arguments (e.g. pronouns and expletive pronouns) as we discussed already, which adds another feature: [±argument]. In sum, this predict 8 different types of categories along semantic dimensions.

2.1.3. Finally, putting aside for the moment functional properties (i.e. position inside a chain and A'-binding), and looking at structural properties that we have already mentioned, we get the following parameters. A category may be Case-marked or not and may be phonetically realized, i.e. overt, or not.

Again, this gives 4 a priori combinations.

2.1.4. If all these features that we may assume to be ↑-features could vary independently, we would predict the existence of 4 x 8 x 4 = 128 distinct ↑-categories. However, we do not expect total independence between all these dimensions, and we find specific proposals in the literature proposing interconnections between them.

For example, we have argued in II.2.6 that the Case Filter had to stated as an independent principle of grammar. One way of viewing the Case Filter is as a redundancy rule between the feature [+phonetic] and the feature [+Case]. Such a proposal eliminates Caseless phonetic categories and thus reduces the number of possible ↑-categories (in fact, it eliminates 32 possibilities).

Obviously, an exhaustive discussion of all these possible

combinations would be cumbersome and unnecessary, for we start with some idea of which categories we need and what their properties should be.

Let us instead start the discussion with a modified version of a proposal made in Chomsky (1982). Chomsky proposes a certain classification of categories relevant to the Binding Principles. We will in fact examine and evaluate the saturation of Chomsky's classification (i.e. the entire space generated by the dimensions Chomsky implicitly or explicitly considers).[5]

In section 2.2 below, we will apply the same procedures to the preliminary classification presented in Chapter II, which differs in important ways from Chomsky's and is underlying to many specific analyses presented in Chapter III.

2.2. The Semantically based Inventory

We can describe Chomsky's proposal as one in which the features [+A] and [+anaphoric] are identified, as well [+B] and [+pronominal]. This divides the number of possible Y-categories by 4: we are thus left with 32.

Concomitantly, the Binding Principles now read as in (6):

(6)   i. if x is [+anaphoric], x must be A-bound in D(x)
        ii. if x is [+pronominal], x must be A-free in D'(x)

We will furthermore make our discussion easier by assuming for the moment that D(x)=D'(x), as Chomsky (1982) assumes.[6] As we said, the resulting inventory is subject to the criteria given in (5i,ii) but also, here, to the additional requirement that some element falling, say, under (6i), does behave semantically as [+anaphoric] element...

2.2.1. Arguments

Exploring the full range of possibilities, we would get a table of ↑-categories with 32 entries. In order to facilitate the presentation let us first limit ourselves to those ↑-categories that are arguments, which are given in table (7):

Arguments

+Anaphoric				−Anaphoric				
+Case		−Case		+Case		−Case		
S1 ?	S2 ?	S3 *	S4 ?	S5 pronouns	S6 pro	S7 *	S8 ?	+ Pron.
S9 ?	S10?	S11*	S12?	S13 overt vrbls	S14 non overt vrbls	S15*	S16?	− Pron.
+Phon.	−Phon.	+Phon.	−Phon.	+Phon.	−Phon.	+Phon.	−Phon.	

The slots in this table (indicated as S) contain either names of categories, stars (*) or question marks.

Names stand for categories (feature bundles) which are attested and fit the feature composition of the slot they occur in. Stars are found in slots of unattested feature bundles, whose absence is predicted by principles we have already adopted in II.2.6 at the time of our preliminary classification, principles whose formulation might have to be adapted to these different classificatory features. Question marks stand in slots requiring more elaborate discussion.

Starting with stars, recall that we have argued that (8) was an independent principle of grammar:

(8)   Case Filter: Only Case-marked NP's are PF-visible

The effect of (8) is to forbid [+phonetic,-Case] categories. This explains the presence of stars in the slots (S3), (S7), (S11) and (S15).

Turning now to attested categories, we find it uncontroversial that pronouns fit slot (S5), and pro -which is the exact non overt counterpart of a regular pronoun- in slot (S6). Chomsky (1982) assumes that variables fit in (S14). Hopefully, some relation must be derived that links the following two properties:

(9) i. being a locally A'-bound (or operator-bound) Ɣ-category
    ii. being a [-anaphoric,-pronominal] Ɣ-category

If not, it would mean that there may be [-anaphoric,-pronominal] categories not interpreted as variables (what might they be?), so that this slot would require closer scrutiny, and the term variable used in (S13) and (S14) would be misleading. Similary (S13) contains overt variables, a better name than resumptive pronouns which misleadingly suggests that these elements have some property in common with pronouns beyond more physical appearance (non overt categories also all have the same physical appearance)

Let us now turn to question marks. Starting with (S16),which would be a Caseless variable in the usual terminology. We have assumed all along that such categories are barred, perhaps by some extension of a principle like the one we have adopted in II.2.6 which read:[7]

(10)   Heads of chains are Case-marked or PRO

As we noted in Chapter II, such a statement introduces a suspicious asymmetry in whatever classification has to adopt it. We will return to this question later in this Chapter (cf. section 6).

More interesting is the case of (S8) which would be a Caseless

pronominal argument, i.e. Caseless pro. I believe no specific claim has ever been made in the literature for (or against) the existence of such an item. Consider however the analysis we have proposed for subject clitic constructions in the Appendix on clitic constuctions. We have concluded that the structure of a SCL construction is as in (11):

(11) $[_{s'}$ COMP $[_s$ NP* $[_{INFL}$ SCL] ... where NP* is pro

Recall however, that the main argumentation only showed that NP* could neither be PRO, nor a variable, nor an NP-trace, i.e. could neither be [+anaphoric] nor [-anaphoric,-pronominal]. Since it must be a [-anaphoric,+pronominal] element (which, incidentally, matches the way it is interpreted) which is non overt, it may fall either under (S6) as we concluded, or fall under (S8), i.e. be Caseless (in which case, we may assume that SCL retains the Case assigned by INFL). Some evidence could be taken to suggest that this might be the correct analysis for NP*. It is widely assumed that some phonological processes analyzing two terms X and Y are sensitive to the adjacency of X and Y and, in particular, are blocked by a non overt category when it intervenes between X and Y and is Case-marked. Amongst such processes, we find English contraction (want + to ---> wanna), French liaison, certain Filters.* The fact is that NP* does not block certaim contraction processes between COMP and INFL in structures like (11). This observation (made in exactly the same context of discussion in Rizzi (1982, p. 176, footnote 17, but valid also for French cf. [que+il parte] --> [qu'il parte]) would be consistent with an analysis of NP* in SCL constructions as Caseless pro. This argument is moderated by the fact that some phonological processes are oblivious to the Case/Caseless distinctions of non overt categories.

Suppose nevertheless that Caseless pro exists. We must account for its restricted distribution. As a [+pronominal] element, it may freely denote some element in D and by assumption may appear in Caseless positions linked (perhaps through a chain) to a θ-position. The problem, of course, is that it does not seem to occur anywhere else than in subject of tensed clause. Indeed, if we examine the Caseless positions, we find governed Caseless NPs governed by a verb not assigning Case (e.g. with Passive morphology). In such positions, Caseless pro is impossible:

(12)   i.   *it was killed pro
         ii.   *There was pro on the roof

Neither can it appear in ungoverned Caseless positions, as, e.g., subject of an infinitival, as the following examples illustrate, where the relevant structures are grammatical, but not with the intended reading: (Bresnan (1982), exhibit some counterexamples - only for structures like (13i). cf. section 5.3 below for further discussion):

(13)   i.   *it is unclear how pro to leave (≠ it is
             unclear how he should leave)
         ii.   *I expect [pro to sleep] (≠ I expect him to sleep)

Because the distribution of pro in fact matches that of Caseless pro, it is plausible to assume that both are subject to the ECP, so that it may account for the restricted distribution of Caseless pro.

Let us now consider anaphoric elements. We can reasonably eliminate (S1), which is not attested, and which we expect not to be if we assume, as is optional, that Case is always assigned under government (i.e. that the set of Case assigners is a subset of the set of governors). Indeed, if some element E is a pronominal anaphor, it falls both under Principle A - i.e. (6i) - and Principle B - i.e. (6ii) - of the Binding Theory. Given

that we have assumed that D(x)=D'(x) (but cf. Footnote 6) E will have to be both A-free and A-bound in D(E) if D(E) exists. D(E) is supposed to be a syntactic domain containing a governor of E. The apparent contradiction is resolved if E is not governed. If E is ungoverned, D(E) does not exist and neither (6i) nor (6ii) apply to E. But if E is ungoverned, it cannot be Case-marked since Case-marking requires government. So (S1) does not exist. The same reasoning extends unchanged to (S2).

Consider next (S4). Chomsky (1982) argues that (S4) is PRO, i.e. that element appearing as subject of infinitival clauses. By the same reasoning as above, if such an element exists, it must be ungoverned. Is treating PRO as a [+anaphoric,+pronominal] consistent with its interpretation? Chomsky (1982) argues that it is. He writes (p. 83-84):

> "it [PRO] may denote, as determined by context (bound PRO),
> or it may function as a free variable lacking an operator
> (free PRO) with arbitrary interpretation..."

Although we might grant that these behaviours illustrate one the [+anaphoric] properties of PRO, the other the [+pronominal] semantic *they do not show that PRO exhibit both properties at the same time* properties of PRO. Consequently, these remarks might suggest instead that we are not dealing with one single type of element, but rather with two, one being perhaps [+anaphoric,-pronominal], the others perhaps [-anaphoric,+pronominal].'

Note furthermore that if indeed the features anaphoric and pronominal have the semantic content we have attributed to them, this last conclusion seems inescapable, for it appears that no element should be able to be both [+pronominal] and [+anaphoric]. Indeed, a pronominal may always directly select some element in D, an anaphor never can.

This therefore, suggests that no element may fit in slot (S4) (or for

that matter (S1) through (S4) for the same reasons).

We might envision reconsidering the way the semantic properties [pronominal] and [anaphoric] are defined so as to permit such an element as PRO in (S4). For example, we could assume, and it is implicit in the quote excerpted from Chomsky (1982) above, that it is a distinguishing property of pronominals that they may either be free, or bound to an element with an independent θ-role. As we have seen, this is simply another way of stating that pronominals would only appear in chain initial position and may be either free or bound. However, this property is not particular to pronominals as Chomsky (p.c.) points out given the well-formedness of the following sentences:

(14)   i. lies about each other trigger the fight
        ii. They$_i$ lied about each other$_i$

in which the anaphor each other is in chain initial position and either free -(14i)- or bound -(14ii).[10] In any case, we shall see directly that some more independent considerations suggest that it would be redundant to postulate the existence of a pronominal anaphor such as PRO.

Consider (S8). The natural candidates are lexical reflexives. As we pointed out, whether lexical reflexives are treated as ϒ-categories will depend on extraneous considerations. In this system, it is natural to assume they are.

Let us now consider (S10), the non overt counterpart of a lexical anaphor. Such an element would be an anaphoric analogue of pro. One plausible structure in which such an element might appear is as the empty category of a R-Clitic construction. Recall that our argumentation led us to analyze these constructions as:

(15)   i.   ...[$_{VP}$ [$_V$ R-Cl+V]...NP*...] with NP* non overt

Recall further that we had assumed that the R-Clitic absorbs Case because
we were working under the assumption that a Case marked non overt category
is not anaphoric. However, if table (7) is on the right track, (S10)
might represent an example of a Case-marked non overt category which is
anaphoric. So that NP* in an R-Clitic construction could be taken as an
example of (S10), especially if it could be shown to act as a terminal
element with respect to the phonological processes mentioned above (in
connection with Caseless pro) which pay attention to the adjacency of the
terms they analyze. If we admit that (S10) is exemplified, the problem
becomes then how to account for its very restricted distribution, for it
certainly cannot appear in any Case-marked position (where it would find a
binder fulfilling the requirements of (6i):

(16)   i. *John's$_i$ pictures of NP*$_i$ (=John's pictures of himself)
         ii. *John$_i$ saw NP* (=John saw himself)

With NP* the element filling the slot (S10), the forms in (16) should, if
well-formed, respectively mean what is in parentheses. Clearly, the
ungrammaticality of these forms can neither be attributed to Case Theory
θ-theory or the Binding Theory, as the well-formedness of the intended
meanings show. Some other reason must then be brought into play to rule
them out.

Again, a plausible candidate is the ECP, that we have already invoked
to regulate the distribution of pro -(S6)- and Caseless pro -(S8). If
correct, we expect (S10) to have a distribution even more constrained than
pro for they must be both in identified positions but (S10) must further
be bound in D(S10) as (6i) requires of anaphoric elements, and, in

particular, will not be permitted in subject of tensed clause position. However, if NP* in (15) is an instance of (S10), we can take the R-Clitic to fulfill the identification relation required by the ECP.

Consider finally (S12). This element would be a non overt Caseless anaphoric argument. Call it PRO'. We therefore expect it to appear in Caseless positions and to be subject to Principle A of the Binding Theory. Consequently, if PRO' is in a governed Caseless position, it will have to be bound in D(PRO'). However, if it is ungoverned, Principle A requires nothing of it. Of course, it is expected that some other subsystem of grammar will provide PRO' with a reference in D, since PRO', an anaphoric argument, would otherwise be uninterpretable.

So the system of principles we have developed will permit PRO' to appear in all the contexts PRO appears in, and in governed Caseless contexts as well. If the distribution of PRO' properly includes that of PRO, it is natural to wonder whether we do not have an unnecessary duplication of category types. We could in fact show that this duplication is justified if there were some cases in which the interpretation of PRO and PRO' differed (or perhaps, if we found some other distinguishing feature). We would therefore expect structures permitting both PRO and PRO' to appear in some position to be ambiguous semantically (or otherwise), depending on whether it is PRO or PRO' that is appearing.

In fact, such structures permitting both PRO and PRO' are not systematically ambiguous, as we would expect. The non overt category is interpreted in one of two ways. It may be arbitrary in reference, compatible with the nature of both pronominals and anaphors: since pronouns do function as variables in the appropriate conditions (e.g. when

bound by a quantifier) and argument anaphoric elements function also as variables, assuming the referential value of their antecedent, the default procedure applying to PRO or PRO' in these contexts seems compatible with their nature.

When not arbitrary in reference, the non overt category appearing in these structures displays anaphoric properties uniquely.[11] So, consider the following examples:

(17)   i.   Only John expects he will win
        ii. Only John expects NP* to win
        iii. Only John expects himself to win

(18)   i.   John knows how NP* to solve this problem
        ii. John knows how NP* to solve this problem and Bill too

Take (17ii) (and (18i)). If NP* is [+pronominal], it could be taken to select some element of D, which, because of some other subtheory of grammar, say Control theory, must corefer with John. In other words, if NP* was pronominal, we would expect the possibility of coreference between NP* and John without referential dependence of the first on the second. But clearly this is not the case as the non ambiguity of (17ii) (or (18ii)) shows. (17ii) means exactly (17iii) (and we get obligatory non sloppy identity in (18ii)), a result we would not expect if coreference were possible as in (17i), since (17ii) does not mean (17i) (with the coreference reading).

So we see that there are some redundancies between PRO and PRO', and some reasons to believe that PRO does not exist. Let us therefore assume that there is no category exemplifying (S4) and let us call that exemplifying (S12) PRO instead of PRO'. Summing up, we get the following table for argument †-categories:

(19)

+Anaphoric				-Anaphoric				
+Case		-Case		+Case		-Case		
+Ph	-Ph	+Ph	-Ph	+Ph	-Ph	+Ph	-Ph	
*	*	*	*	pron.	pro	*	Caseless pro	+Pron.
lexical reflex.	Case PRO	*	PRO	overt vbls	non overt vbls	*	*	-Pron.

Recall also that we have postulated that Case marked PRO, pro, Caseless pro and perhaps non overt variables (although we have not discussed it) are subject to the ECP.

2.2.2. Non Arguments

Consider now the table equivalent for (19) or (7) for non arguments:

(20)   Non arguments γ-categories

+Anaphoric				-Anaphoric				
+Case		-Case		+Case		-Case		
+Ph	-Ph	+Ph	-Ph	+Ph	-Ph	+Ph	-Ph	
t1	t2	t3 *	t4	t5	t6	t7 *	t8	+Pron.
t9	t10	t11*	t12	t13	t14	t15*	t16	-Pron.

A number of stars immediately follow from the Case Filter: (t3), (t7), (t11) and (t15).

Instead of proceeding in the same way we did for arguments, let us instead try to fit the elements we know exist in the various slots available.

Consider first NP-traces. NP-traces are the elements appearing in

non chain initial position so that they are non arguments. Furthermore, we know that we can partially account for the distribution of NP-traces under the assumption that they fall under Principle A of the Binding Theory. NP-traces are therefore [+anaphoric]. Moreover, in order to explain why Raising is not possible with Control predicates, it is usually assumed that NP-traces are subject to the ECP. Whether this is correct or not does not matter at this point. The important point is the consequence of this assumption, namely that NP-traces must be governed. How is this derived? Recall that the ECP took identification to be proper government. One can easily see that, for NP-traces, properly governed positions are positions which are governed by a lexical category or positions belonging to a subset of subject postion of tensed clauses. Now, NP-traces cannot appear in subject of tensed clauses position (for various reasons: Case, binding). So that this version of the ECP applied to NP-traces reduces to the requirement that NP-traces must be governed by a lexical category, hence governed. From this, it follows that an NP-trace is an example of one of (t9), (t10), (t11) or (t12) (it could not be [+pronominal] since it would imply it is ungoverned). (t11) is excluded as we have just seen. (t9) and (t10) do not seem to exist, an unexplained gap within this classification, that we might describe as:

(21)   Case-marked non arguments are not anaphoric

From this, we conclude that NP-trace should fall in (t12). If it makes sense to say that an NP-trace must fall under Principle A, does it make sense to say it is [+anaphoric]? If it does make sense to say that an NP-trace fails to denote, it makes little sense, I think, to say that an NP-trace behaves as a variable assigned a value by its antecedent. As a

non argument, an NP-trace can hardly be said to assume a referential value in D. Rather, an NP-trace plays a formal role of θ-role transmission subject to locality requirements. We must therefore slightly modify our definition of [+anaphoric]. We will now hold that an element is [+anaphoric] iff it fails to inherently denote in D.[12] This raises an immediate problem with such elements as pleonastic (or expletive) _it_ and _there_ of (22):

(22)   i.   it seemed that S
       ii.   it turned out that S
      iii.   there was a man on the roof

If we assume, as appears a priori most plausible, that these elements fail to denote in D, they should be considered anaphoric. Obviously, this is not a welcome result. Expletive _it_ may be governed (so it could not be a pronominal anaphor) but does not have to be bound. We thus do not want such elements to count as [+anaphoric]. This can be achieved in different ways.

One way is suggested in Chomsky (1982). It consists in assuming that expletives, contrary to appearances, do denote in D, say, some designated element. Some empirical reasons suggest however, that this is not desirable. Recall that we have distinguished between three types of elements: regular (or true) arguments (such as names...), quasi-arguments or constants (idiom chunks, weather it...) and non arguments (or pleonastic, or expletive elements). Positions demanding regular arguments (as determined by the θ-role assigned, or transmitted to that position) may be wh-questioned, receive arbitrary interpretation or be controlled (i.e., informally speaking, appear in control structures). Quasi-arguments may not be wh-questioned, may not receive arbitrary

interpretation but may be controlled. Non arguments can be neither of these (cf. Chapter II, footnote 22, for examples).

This differentiated behaviour can be easily explained under the natural assumption that true arguments denote or range in a subset (usually infinite) of D as determined by the particular semantic restrictions (selectional restrictions) imposed by the θ-role they receive, while a quasi-argument Q denotes a designated element of D E_0; that is to say, a "quasi θ-role" is compatible only with some particular element of D. $_{13}$ Finally, the "θ-role zero" -i.e. no θ-role- is compatible with no element of D, i.e. non arguments fail to denote altogether. Now, wh-questionning or assigning arbitrary interpretation to some position only makes sense if the θ-role assigned to that position defines a non-trivial set of compatible arguments that could rightly fit that position: so that they are permitted with arguments only, not with quasi arguments or non arguments.

Control on the other hand means identity of denotation in D so that any element denoting in D may be potentially controlled. A constant Q denoting some designated element E_0 may be controlled only by some other constant Q' denoting E_0 as well. Non arguments, clearly, cannot be controlled (cf.IV.6.2.4 for further remarks).

Now this pattern of explanation is lost if we assume that non arguments do denote some designated element in D, just like quasi arguments.

Another way to achieve the treatment of expletives as [-anaphoric] is to drop the equivalence between failure to denote and being anaphoric. Instead we make it a simple implication:

(22)   iv. if x is anaphoric, it fails to denote, but nct conversely

If this move is correct, we expect to find two categories of elements that fail to denote: [+anaphoric] elements, such as NP-trace, and [-anaphoric] elements. And we can take expletives as precisely exemplifying this type of [-anaphoric] element failing to denote.

Pursuing this line of reasoning, we can see immediately that no non argument should be taken to be [+pronominal] as we have assumed that a [+pronominal] element may always be chosen to denote some element in D, and expletives do not denote.

We thus derive the non existence of elements fitting the slots (t1) through (t8). Furthermore, we are led to conclude that expletive _it_ is [-anaphoric,-pronominal], and thus exemplify (t13). This conclusion might seem surprising: expletive elements now seem to fall together with variables and also names, but, recall names are not ʔ-categories. Notice that the very treatment of variables as [-anaphoric] implied the adoption of (22iv) over the alternative since although they have "referential" properties, they fail to denote in D.

The only slots that remain are (t14) and (t16). Here, we can proceed by analogy with the case of arguments. In (t14), we should find the expletive analogue of _pro_, i.e. of (S6). As we have argued, _pro_ appears in P-Clitic constructions, but P-Clitics do not double non-arguments,[14] so that expletive _pro_ cannot appear there. Depending of the analysis of empty subjects in tensed clauses, we might argue that expletive _pro_ appears in subject position of tensed clauses. If an empty subject of tensed clasues is always Caseless, i.e. is Caseless _pro_ when it is an argument, then empty expletive subjects of tensed clauses will always exemplify (t16), i.e. be expletive Caseless _pro_. If, as would seem more likely, these empty subjects are sometimes Case-marked, sometimes

Caseless, depending on particular properties of INFL, empty expletive subjects of tensed clauses will either exemplify (tl4) or (tl6).[15] (For example, our analysis of SCL constructions would provide an example of (tl4) with expletive subjects. The analysis of empty subjects by Rizzi (1982, Chapter IV) could be easily adapted to provide an example of (tl6), with expletive subjects).

Summing up, we end with the following table for non argument Ʈ-categories:

(23)   Non argument Ʈ-categories

| +Anaphoric | | | | -Anaphoric | | | | |
| +Case | | -Case | | +Case | | -Case | | |
+Ph	-Ph	+Ph	-Ph	+Ph	-Ph	+Ph	-Ph	
*	*	*	*	*	*	*	*	+Pron.
*	*	*	NP trace	expl. "pron."	expl. pro	*	Caseless expl. pro	-Pron.

2.2.3. Consequences and Problems

The consideration of tables (19) and (23) suggests a number of remarks and raises a number of questions.

Let us first make a terminological adjustment. We have referred to the element appearing in (tl2) as NP-trace. This is the usual denomination but I believe it is in fact misleading. A comparison of tables (19) and (23) suggest that we should rather call it expletive PRO: indeed, it is the non argument counterpart of PRO. The notion of NP-trace belongs in fact to an altogether different register: that of functional

notions pertaining to Chain-Theory: an NP-trace is simply an NP appearing in chain-internal position. Whether the set of NP-traces and that of expletive PRO's are in fact identical sets is an empirical question.[16]

Consider next the following problems about this inventory of ↑-categories:

   i.   we have provided no explanation for the gaps (t9)
       and (t10) or, equivalently, for the stipulation (21).

   ii.   we have so far, paid no attention to functional
       properties of these ↑-categories. The following
       questions arise: can we predict on the basis of its
       feature composition:
       (a) in which position in a chain a given
           ↑-category may appear
       (b) whether a given ↑-category may be locally
           A'-bound or not

   iii. We have to verify that this classification meets
       the adequacy criteria given in (5) and in
       particular (5i). As we see, we find both
       horizontal and vertical asymmetries within each
       table (19) and (23) and between tables as well.
       How are they explained?

   iv.   Finally, as we have indicated, the principles
       already discussed are insufficient to account for
       the distribution (along non functional criteria) of
       non overt items. We have thus, invoked the ECP to
       do the required additional work. How do we
       formulate it and does it apply to a natural class?

2.2.3.1. Let us start with (ii) above. Clearly, because arguments may only appear chain initially, the elements in table (19) may only appear as heads of chains. This follows from chain theory. The question is not as easily settled for the elements in table (23). A priori, non arguments may appear either in chain initial possition, or chain internally, i.e. as NP-trace.

Consider first what can be an NP-trace. Clearly, expletive PRO can be an NP-trace. This is the usual case. NP-trace as in Passive or

Raising constructions are Caseless anaphoric elements. Nothing else can be. As we have argued had to be the case, and as we have stipulated amongs the properties chains must have, Case-marked elements, and, in particular, Case marked anaphoric elements are excluded from chain internal positions. So that (t9) and (t10) cannot be NP-traces, as well as (t13) and (t14). This leaves out (t16). Nothing so far excludes Caseless expletive pro from appearing chain internally. Note first that this option could not be excluded by an argument based on the ECP even if the only identified positions are subjects of tensed clauses. This is because such a position could be an NP-trace if it is filled by (t16) given that (t16) is not anaphoric and thus does not have to be bound at all. As illustration, consider the structure:

(24)   *Jean semble que [$_{NP}$ e] [$_{INFL}$ il] a entendu
            (=il semble que Jean a entendu)

We have argued that NP* in SCL constructions might be taken as examples of Caseless pro (in French or, perhaps, some other language). Suppose we form the chain (Jean, NP*). NP* is a Caseless pro, identified by the AGR element il. It is non anaphoric so that no binding violations occur, and it is expletive so that the chain (Jean, NP*) meets the θ-criterion.

Clearly, we want to exclude such a possibility. If (t16) exists, as this classification predicts, it seems that we have to stipulate that:

(25)   NP-traces are anaphoric

which means that we cannot derive certain functional properties, we have to stipulate them. Note also that, if (t16) does not exist, we must then introduce some ad hoc statement excluding it, instead of (25).

Consider now the inverse question. Which of these expletive elements

can appear as head of chain? Non anaphoric elements all may. No anaphoric element may (we qualify this in 4.2. below). We may account for this observation in different ways. One is to adopt stipulation (21), which rules out Case marked anaphoric expletives both as heads of chains and redundantly as NP-traces, and to provide a different explanation for the non occurrence of expletive PRO as head of chain. Another is to erect the observation above as a principle of grammar:

(26)   Anaphoric expletives may not be heads of chains

But we see that (26) is the converse of (25). (25) states that NP-traces which are expletives by definition are anaphoric. (26) expresses that anaphoric expletives are NP-traces: they cannot be heads of chains. Since every NP belongs to some chain, they must appear chain internally. So we get (27):

(27)   A ϔ-category is an anaphoric expletive
            if it is an NP-trace

which answers remark (i) of 2.2.3. This shows that this classification provides no explanation for the gaps (t9) and (t10), nor for the fact that expletive PRO appears in chain internal position only. (We will see in 4.2. below that (26), hence (27) is not quite descriptively adequate).

Turning now to the second part of remark (ii), consider whether we can predict which elements may or may not be locally A'-bound. First, we may restrict ourselves to arguments: as we have shown in II.2.6, some principle like (28) below:

(28)   locally A'-bound elements are arguments

is fairly natural, given that A'-binders define some domain or set some

referential value that the elements they bind must respectively range over or assume. Consequently, consider the elements of table (19) and, more specifically [-anaphoric,-pronominal] elements. Following Chomsky's proposal,[17] we have supposed that these elements are variables, i.e. locally A'-bound. If this assumption is correct, U.G. should contain a statement of the form:

(29)   for x, x a ɣ-category, x is locally A'-bound
         iff it is a [-anaphoric, -pronominal] argument

Can (29) be derived? It seems not. It appears that it must be stipulated for, except for one subclass of elements in table (19), namely governed anaphors, which must be, by Principle A, locally A-bound, (cf. below) nothing prevents any other ɣ-category of table (19) from being locally A'-bound. Conversely nothing implies that [-anaphoric, -pronominal] elements must be locally A'-bound.

Consider this last point first. Why could there not be a [-anaphoric, -pronominal] ɣ-category which is free. It would not fall under the Binding Theory and would never inherently select any element of D (arbitrary PRO would fit this description exactly).

So we must stipulate (29) from right to left. Stipulation of (29) from left to right is also necessary. Consider governed anaphors. We might fairly plausibly, as we have just mentioned strengthen Principle A to a requirement that anaphors must be locally A-bound within their domain, if they have one, as was, I believe, always implicitly intended. From this, it follows that governed anaphors cannot be locally A'-bound.

Let us now consider the elements that are neither governed anaphors nor [-anaphoric, -pronominal] elements. For these, there does not appear to be any semantic contradiction whatever in their being locally A'-bound.

This is particularly striking for [+pronominal] elements since the way their referential import would be computed if they were locally A'-bound would be identical to the way it is when they act as semantic variables, i.e. when they are A-bound by quantified phrases and other non referential expressions. This last remark seems to me to cast very serious doubts on the well-foundedness of the type distinction between pronominals on the one hand, and variables on the other.

2.2.3.2. Let us now turn to problem (iii).

The vertical asymmetry in (23), i.e. the asymmetry between the two lines of table (23) follow directly from straightforward semantic considerations. Expletives cannot be [+pronominal]. And the horizontal asymmetry had to be stipulated. This is statement (27).

Consider now table (19). It is clear that there is only one asymmetry that we have to account for: the non-existence of Caseless [-anaphoric, -pronominal] elements. A visibility account like the one proposed in Chomksy (1981, Chapter 6), that we have discussed and rejected in II.2.6, would be even harder to make feasible here for we would have to stipulate that not only PRO is intrinsically visible, but also what we have called Caseless pro.

Note that we have so far admitted that there were no Caseless [-anaphoric,-pronominal] elements on the basis of the fact that local A'-binding of a Caseless position is not always permitted. We have just noted however, that there was no principled bar to having ungoverned PRO locally A'-bound. If that was indeed permitted, we could not assume that it would be an example of locally A'-bound PRO properly speaking because of (29): but nothing would prevent us from assuming that it is instead a

[-anaphoric,-pronominal] ungoverned (hence Caseless) locally A'-bound
element, thus filling a gap in table (19) and making the asymmetries
entirely predictable:[14]

    i.   No Caseless phonetic elements because of the Case Filter
   ii.   No Pronominal anaphor for semantic reasons

Note also that this would also render entirely predictable the asymmetries
between the tables (19) and (23).

    We return, in somewhat different terms, to that suggestion in section
6 below.

2.2.3.3. Consider finally the ECP. We have subsumed under this name the
additional requirements that some non overt categories must meet. In
fact, if we put together our informal suggestions with the usual
assumptions concerning the scope of the ECP which takes NP-trace and non
overt variable to be subject to it, we get that every non overt $\acute{1}$-category
is subject to the ECP except PRO.

    Without going into the question of how such a principle should be
formulated, let us simply ask whether it is a priori plausible that some
principle apply to this class of elements, namely:

    - Expletive PRO chain internally
    - pro, Caseless pro
    - Case-marked PRO
    - non overt variables

This class looks rather heterogeneous. There is no simple way to
characterize them. This suggests that at least for some elements, some
other principle than the ECP is involved. Three of these elements would
appear to go together since they have identical distributions namely:

subject of tensed clauses (independently excluded, for Case-marked PRO, by the Binding Theory) and clitic constructions.

So perhaps expletive PRO on the one hand, and non overt variables on the other, are subject to some other type of restriction (e.g. S' breaks a chain as Aoun (1982) suggests for NP-traces).

However, even if this is the case, we still do not end up with a natural class, for Case-marked PRO and Caseless pro, for example, differ in every feature value (except argumenthood). Of course, we could further dissociate Case-marked PRO from Caseless pro and pro, since neither of these two resulting set has a distribution resembling that of expletive PRO or non overt variables, it would lead to a proliferation of very specific principles, i.e. a series of stipulative statements. This state of affairs seems to indicate that this classification misses some generalization, or simply overgenerates.

2.3. The formally based Inventory

Although, as we have seen, the classification that we have just discussed meets with a certain degree of success, its very richness raises new problems whose solutions are not obvious. Suppose that we try to reduce the number of independent f-features by simply eliminating the features [anaphoric] and [pronominal] from amongst the primitive features, allowing ourselves to state entailment relations by using functional notions instead (a use that we have seen was necessary also in the previous classification). That is, instead of using intrinsic features and semantic features and trying to deduce from them functional properties, let us try to do the opposite. For ease of reference, let us

call the inventory that we have just discussed, the <u>S-inventory</u>, and the alternative that we will now present, the <u>F-inventory</u>.

In fact, it is an inventory of this last type that we had provisionally adopted in II.2.6 and on which parts of Chapter III are crucially dependent.

According to this system of assumptions, the predicted inventory of ʔ-categories is given in table (30) below:

(30)

	+Argument		-Argument		
+Case	-Case	+Case	-Case	
pronouns	*	expl. pronouns	*	+Ph
pro	PRO	expl. pro	expl. PRO	-Ph

Notice first that lexical anaphors do not fit in this inventory. We thus have to assume that they are lexically specified as [+anaphoric]. This is a negative result since we replace the predictive value of the S-inventory by a stipulation. However, the overall simpler system provided by (30) largely outweighs, I think, this inconvenient.

Let us now apply to (30) questions parallel to those we have applied to the S-inventory:

    i.   Exhaustiveness
    ii. Explanatory reasons for existing asymmetries, relation to
         functional properties
    iii. Prediction of the Binding behaviour
    iv. Prediction of the semantic behaviour
    v.   Formulation and scope of the ECP

2.3.1. The issue of exhaustiveness must remain open at this point (as it was, in fact, for the S-inventory). The F-inventory lacks Case-marked PRO, Caseless pro and Caseless expletive pro as compared to the

S-inventory. While discussing table (19), we have given possible examples of such cases. In each instance, the category [xCase] in question could just as well have been analyzed as [-xCase]. We had suggested that [xCase] and [-xCase] would differ with respect to their analizability by certain phonological rules sensitive to the distinction [±Case] for non overt categories. If it can be convincingly shown that there exist some phonological processes which not only belong to the right category of rules (certain phonological processes pay no attention to this distinction) and also draw the distinction predicted by the S-inventory, the F-inventory would have to be given up in favor of (19) and (23). At this point, such an argument is lacking.

2.3.2. Consider next questions of functional properties. As far as NP-traces are concerned, the answer is straightforward. By definition, only Caseless non arguments may appear as NP-traces. It follows that the only candidate is expletive PRO.

Consider possible A'-binding: we may a priori exclude non-arguments for the reason put forth and semantically justified in II.2.6, and needed also for the S-inventory:

(31)   locally A'-bound t-categories are arguments

As far as arguments are concerned, pronouns and pro may be freely locally A'-bound. When they are, they are interpreted as variables, overt or not. As we have mentioned, there is no semantic oddity in this assumption. Variables and pronouns as bound variables appear to have exactly the same semantic properties. In other words, we are suggesting that there is no type distinction between pronouns and pro on the one hand and variables

overt or not on the other.

As for PRO, we still face the same problem of the apparent non existence of Caseless locally A'-bound elements. If it is true that variables must be Case-marked, it must be stipulated one way or another, for example, as we have suggested in II.2.6, i.e. by stipulating some version of (10) above.[19]

2.3.3. Consider next the Binding properties of ↑-categories. Recall that the Binding Theory is formulated as:

(32)   i.   if x∈A, x must be A-bound in D(x)
         ii. if x∈B, x must be A-free in D'(x)

We want to achieve the following result. Pronouns and pro (and perhaps PRO) should fall under (32ii), expletive PRO and PRO under (32i). Locally A'-bound should not fall under (32i). They are usually assumed not to fall under (32ii) either, but it does not matter whether they do or not.[20]

The most straightforward algorithm that would yield this result is (33):

(33)   i.   Caseless ↑-categories belong to A
         ii. Case-marked ↑-categories belong to B

According to (33), a Case-marked category, whether locally A'-bound or not falls under (32ii).

Note that PRO, as a Caseless element, falls under (32i) only; this conclusion, based on a simplicity argument agrees with the conclusions we had reached while discussing the S-inventory. We will see that the very same considerations apply to this classification when we consider the interpretative properties of the elements in (30).

2.3.4. Turning now to the interpretative properties of ?-categories, we still adopt the terminology [± anaphoric], [± pronominal]. First, note that we may adopt a more natural approach than in the S-inventory by restricting these semantic properties to arguments. There is no need for example, to postulate that expletive PRO is anaphoric under certain conditions. If we want to keep at least a three way distinction variables/anaphors/pronominals, as in Chomsky (1982), we may proceed as follows. We may adopt the following algorithm:

(34)   Let x be a ?-category argument

i.   if x is Caseless, it is [+anaphoric]
ii. if x is Case-marked and locally A'-bound, it is neither
     [+anaphoric], nor [+pronominal]
iii. if x is Case-marked and not locally A'-bound, it is a
     pronominal

However, if we take into account the full range of properties that "pronouns" may have (where, informally speaking, we mean those elements traditionally identified as pronouns), it suggests a slight reformulation of the semantic notion [+pronominal].

A pronoun may either select independently some element in D as its referential value, whose identity may be contextually determined (coreference) or left unknown (free pronoun) or it may act as a variable ranging over some domain, as determined by its antecedent (bound pronoun) be it an A'-binder, a non referential expression in an A-position (Quantified Phrase) or a referential expression. In this latter case, that of a pronoun referentially dependent upon its antecedent (a situation often called pronominal binding),[21] there is no semantic difference in the various ways the referential function of the pronoun is determined. In particular, this suggests that locally A'-bound elements should count as

[+pronominal], if "pronouns" are taken to be archetypes of this type of behaviour.

In that case, we can modify the rules of entailment given in (34) determining the referential properties of f-categories from their structural content and replace them by the simpler (35):[22,23]

(35)   Let x be a f-category argument

    i. if x is Caseless it is [+anaphoric]
    ii. if x is Case-marked, it is [+pronominal]

If (35) is on the right track, it means that <u>there</u> <u>is</u> <u>no</u> <u>type</u> <u>distinction</u> <u>between</u> <u>pronouns</u> <u>and</u> <u>variables</u> but simply functional differences.

If we now construct the table of f-categories that this system predicts, using the same classificatory features as in the S-inventory so as to compare them more easily, we get:

(36)

+Anaphoric				-Anaphoric				
+Case		-Case		+Case		-Case		
+Ph	-Ph	+Ph	-Ph	+Ph	-Ph	+Ph	-Ph	
1. *	2. *	3. *	4. *	9.pron, overt vbls	10.pro, empty vbls	11. *	12. *	+Pron
5. *	6. *	7. *	8. PRO	13.expl.	14.expl.	15. *	16.expl.	-Pron

The structure and asymmetries of this table follow immediately from the assumptions made: (36.1) through (36.4) are exluded by (35) since no element is both Caseless and Case-marked. Note that if we had taken PRO to illustrate (36.4), we would have no explanation for the gap in (36.8) (or equivalently for the redundancy with (36.8), if some PRO' were to illustrate it). (36.3), (36.7), (36.11) and (36.15) are excluded by the

Case Filter. (36.5), (36.6), (36.11) and (36.12) are excluded by the rules (35).

Finally, on the side of the functional properties, we get: pronominals may be freely locally A'-bound. Following standard usage, we give them a different name when they are (bottom line in (36)) but no difference of nature is intended. Expletives cannot be bound due to (31).

As for PRO, it is an open question at this point whether it has a locally A'-bound analogue. If it may, the only thing left to investigate is precisely under what conditions and for which (hopefully) principled reasons.

If it may not, ?-categories behave asymmetrically with respect to A'-binding. We would then need to state the relevant stipulation. As for chain internal positions, the only question is: why can't expletive PRO's head a chain (but cf. section 4 below).

2.3.4. Let us now briefly consider the question of which set of categories the ECP should be taken to apply to. If we put aside NP-traces, as we have argued in 2.2.3.3, was plausible, given that they would be be the only elements whose relevance to the ECP is functionally rather than intrinsically determined, we are left with the non overt categories pro/variable and expletive PRO perhaps. As can be easily checked, any other plausible subset of this set of elements is easily and naturally characterizable.

2.4. Summary

Summing up the discussion, we have explored two different theories of ϒ-categories leading respectively to the S-inventory and the F-inventory. The S-inventory takes intrinsic and referential properties as primitives and tries to derive their binding and functional properties. The F-inventory takes intrinsic properties as primitives and tries to derive their binding and referential properties, arguing that functional properties belong to a different dimension.

These two approaches share some basic assumptions:

    i.   The Case Filter
    ii. Locally A'-bound elements are arguments

Some problems appear in both:

    i.   If Caseless locally A'-bound elements are not permitted,
        neither offers an explanation for this gap
    ii. The fact that expletive PRO's cannot head chains requires
        stipulation in the two systems.[24]

The S-inventory includes a wider array of ϒ-categories, including lexical reflexives, whose grammatical behaviour, it claims, is as predictable as that of non overt categories, or pronouns.

The richness, however, leads to a number of stipulations or problems that we list below:

    x a ϒ-category
    i.   x is anaphoric expletive iff x is an NP-trace
    ii. x is locally A'-bound iff it is [-anaphoric, -pronominal]
    iii. [+A] = [+anaphoric] [+B] = [+pronominal]

The F-inventory is more restrictive and must stipulate the binding behaviour of lexical reflexives (i.e. [+anaphoric] = [+A])[25] but offers a

more natural system of correlations between the various properties:

x a ?-category (let [-A] = [+B] for ?-categories)
i.   x is [yCase] iff x is [yB]
ii. x an argument, if x is [-Case], it is anaphoric
                    if x is [+Case], it is pronominal

Overall, I believe, the second system is simpler and more natural. We will therefore adopt it.

Note finally that some conclusions are common to both systems, namely that PRO is an anaphor subject to Principle A. This conclusion has rather deep consequences for the theory of grammar, some of which we will explore in subsequent sections.

3. Binding Principles and the Distribution of PRO

Now that we have made some of our assumptions concerning the inventory, classification and interpretation of ?-categories more precise, thereby answering the questions (2i), (2ii) and (2iii), we can turn to issues surrounding the actual formulation of the Binding principles, and, in particular, how the Binding Domains D and D' apppearing in (1) are determined, i.e. question (2iv).

For convenience, let us repeat the basic form of these principles:

(37)   i.   if x∈A, x must be A-bound in D(x)
        ii. if x∈B, x must be A-free in D'(x)

We have concluded that B consisted of Case marked ?-categories, and that A consisted of Caseless ?-categories and lexical anaphors (i.e. reflexives and reciprocals).

As compared to the assumptions of Chomsky (1981; 1982), the only major difference is our assumption that what we have called PRO belongs to set A only and not to both A and B.

This conclusion has also been reached by several authors - e.g. Koster (1981), Bouchard (1982), Manzini (1983) and Sportiche (1982a), all, I believe, from different points of view and is in fact evoked and briefly criticized in Chomsky (1982, p. 104). Conceptually, very similar hypotheses taking Case to play a major role in the distribution of PRO can be found in the early proposals by Vergnaud concerning the relation between Case Theory and NP distribution.

This conclusion leads to a number of obvious questions, especially when it is taken in comparison with the alternative theory of Chomsky (1981; 1982) according to which PRO must be ungoverned. If we take Chomsky's formulation as point of departure, we must:

  i.   verify that the different status we attribute to PRO
       does not run into empirical problems
  ii. introduce appropriate modifications correlated with this
       different status of PRO

In order to proceed more easily with the discussion, let us assume provisionally, as we had done in section 2 above, that in (37) $D(x)==D'(x)=$the Governing Category of $\underline{x}$, i.e. the first NP or S containing x and a governor of x.

3.1. Government, Case-marking and PRO

Let us start with (i). According to our assumptions, the distribution of PRO is determined by two distinct factors:

First, it may only appear in Caseless positions
Second, it may only appear in positions in which it meets
Principle A (i.e. 37i)).

Recall that the alternative is to assume that PRO may appear in all and only ungoverned positions.

As Chomsky (1982) notes, to perform a comparison of these two alternatives, we must consider constructions in which there is Case but no government, and conversely, constructions in which there is government but no Case. Let us consider them in turn.

3.1.1. Case-marked ungoverned Positions

We are now considering structures containing an ungoverned Case-marked position.

Note first that some care is needed to establish why such structures are relevant at all, for, strictly speaking, if a non overt category were to appear in such a position, we would not analyze it as PRO, but rather as pro. What is really intended is this: assuming the existence of such positions, the occurrence of a non overt category K otherwise displaying properties identical to that of what we call PRO (e.g. referential, binding properties...) would seem to indicate that our classification of ʔ-categories is incorrect (Note that this does not apply to the S-inventory, which tolerates Case-marked PRO's).

Consider now such structures and assume that the non overt category appearing in them displays the usual properties of PRO. Call P the ungoverned Case-marked position they contain. We can superficially distinguish two classes of structures that have been argued in the literature to contain such a position P:

i.   structures in which P may freely host a lexical NP
ii. structures in which P does not (freely) host a lexical NP

The structures presented in Neidle (1982), Simpson (1982), for example, are, if my understanding is correct, of the second type. Accepting the analysis of these authors leads to the postulation of a Case-marked PRO in position P (at some level of representation) and given that positions such as P do not freely permit the occurrence of overt NP's, we are driven to a reconsideration of the theoretical notion of Case that we have used so far.

Indeed, following the ideas of the theory of Abstract Case originally introduced by Vergnaud, and in Rouveret & Vergnaud (1980) and Chomsky (1980), we have assumed that the following proposition lies at the core of Case Theory:

(38)   A position P may contain (or be linked to - as in wh- or
          clitic constructions) overt material, iff it is Case-marked

if we grant the existence of structures as in (ii) and if we assume that observed (i.e. morphological) Case features are always indicative of Abstract Case structure, (38) is no longer true. However, it is clear that the existence of structures like (ii) requires the introduction of a distinction between a property, call it Kase, which will meet (38), i.e. which will be a necessary and sufficient condition for overt material to be able to appear, and another property, Case which would not meet (38). It is worth pointing out that this Kase/Case distinction, however it is drawn, is necessary regardless of the theoretical status of PRO and other elements. For the purpose of characterizing PRO and more generally for our classification of †-categories, it would suffice to replace everywhere Case by Kase so that such structures as in (ii) are not directly relevant

to the issue.

Let us pursue with structures of the type described in (i). Candidates for such structures are illustrated by English poss-ing or NP-ing constructions (where NP≠genitive).[26] Consider the following pairs:

(39) i. I like John's singing
      ii. I like   singing

(40) i. John having left,...
      ii. [$_{NP}$ e] having left,...

Consider first, the absolutive constructions in (40) (cf. Reuland (1983) and the references cited therein for extensive discussion of these constructions). It seems quite clear that they have a clausal structure. All the arguments presented in Chomsky (1981, Chapter 2) to the effects that clauses have obligatory subjects carry over to such structures. Consequently, we must accept the presence of a non overt subject in (40ii). This subject has none of the referential properties of pro. It is therefore plausible to take it to be PRO. However, Case seems to be a property of the subject position of these adjuncts as (40i) shows.

Is it necessarily a property of these positions. This of course, depends on the analysis of such structures. One possibility that we reject for reasons that will become clear in the next subsection is that Case assignment is optional.[27] However, the same effects can be obtained if the Case assigner is optionally present. Precisely such an analysis has been argued for in Reuland (1983). Reuland argues that the clausal structure of adjuncts like those in (40) is:

(41)   [$\underline{S'}$ [$\underline{S}$ NP* INFL VP]]

where INFL contains the Affix -ing, an element that governs, and

Case-marks (or transmits Case) to the NP subject it governs. This affix, Reuland argues paralleling Chomsky (1981) treatment of the "Pro-drop parameter", is subject to the rule of Affix Hopping (the rule R of Chomsky (1981)), which is freely ordered with Case-marking (or Case-checking) rules. If Affix Hopping applies first, INFL neither governs nor Case-marks NP* which is therefore PRO.[28] Otherwise, INFL both governs and Case-marks NP* as in (40i).

Consider next, the poss-ing constructions (39). A number of alternatives can be explored. One possibility would be to extend Reuland's analysis to those cases. We could make exactly the same assumptions taking poss-ing structures to be clausal structures with an _ing_ INFL subject to an unordered rule of Affix Hopping. We would need however, to introduce one substantial difference given that NP-ing structures and Poss-ing structures do not have identical distributions. Without going into detail, we could assume that poss-ing structures have a nominal character, or perhaps a stronger nominal character than those NP-ing structures in which -ing is nominal (cf. Reuland, op.cit. for discussion), so that in effect, poss-ing structures are literally clausal NP's. Their head, being noun-like, assigns genitive Case in the same condition the INFL of NP-ing structures does, or does not assign or govern the subject position if Affix Hopping precedes Case-marking, as Reuland (op.cit) suggests for NP-ing.[29]

Clearly, a full investigation of these suggestions would be necessary in order to determine whether they are empirically adequate but, in the context of our discussion, this is not the essential point. We are trying to argue that there is no situation in which Case (or, more precisely, Kase) is assigned without government, as certain analyses of structures

like (39) or (40) would have it. Surely, some analysis, perhaps similar to those we have described above can fulfill this task, however natural or plausible they might be.

We must not lose sight of the fact that we are trying to compare two alternative theories of PRO, and it is important to note that the failure or success of the analyses we have discussed cannot establish the superiority of one theory over the other given that <u>both</u> require such analyses. For ours, it is clear why. For the alternative theory taking PRO to be ungoverned as a consequence of its being a pronominal anaphor, this is required for a different reason. If there are contexts in which Case is assigned without government, we would expect to find in these contexts and only in these contexts overt pronominal anaphors. The systematic absence of such elements strongly suggests that no such context exists, and thus implies the adoption of analyses for structures apparently falling under (i) that would have the right properties, namely, that the binary relation involved in Case-marking, just like that involved in θ-marking, or subcategorization, is always a narrower relation than the government relation.

Furthermore, this assumption will yield exactly the correct result in the structures falling under (ii) that we have reviewed earlier. In those structures, the common assumption of advocates of any theory of PRO is that morphological Case appears without government. The above assumption concerning Case-marking will draw the required distinction between Case and Kase.

3.1.2. Governed Caseless Positions

Consider now positions that are governed but not Case-marked. Here the claims made by the respective approaches can be phrased as follows:

    i.   if PRO must be ungoverned, it should not appear in such positions.

    ii.   if PRO may be governed, it should be able to appear in such positions provided that it obeys Binding Principle A.

Before assessing each position, we have to make explicit one assumption that is clearly required. We have to suppose that Case-assignment (or Case-checking) is an obligatory process. Exactly as for θ-assignment (for which it is also required) whenever two appropriate categories - a Case assigner and an NP - are in the appropriate structural configuration - government - Case-marking takes place. Otherwise, we would have no explanation for the ill-formedness of such forms as:

(42)   *John saw $[e]_{NP}$.

If NP* was not obligatorily Case-marked by <u>see</u>, it would be analyzed as a governed PRO bound by <u>John</u>, i.e. would be licit. (42) however, is ill-formed. If Case is assigned, NP* is analyzed as <u>pro</u> and thus ruled out by the ECP for lack of identification. In fact, the obligatoriness of Case-marking is not required only by such cases. Recall that we have argued in Chapter II that a stipulation had to be included in the definition of a chain forbidding NP-traces to be Case-marked. Although this stipulation might not be the only way to handle such a restriction, it would remain, I believe, that the particular characteristics of ill-formed chains illustrated in Chapter II and excluded by this stipulation have to do with the Case marked character of NP-traces. It

was implicit in this account that Case-marking was obligatory.

Consider now the question of where we find governed Caseless positions. We have assumed that:

> Governors are X° in the X' system, i.e. A,V,N,P, INFL
> with some restrictions for INFL(e.g. INFL of infinitivals
> is not)[30]

In the unmarked situation, we can take Case assigners to be those X° that are transitive, i.e. that may appear in a structure governing some NP only dominated by projections of that X°. Consequently, a X° is a Case assigner unless:

> i. it is not transitive
> or ii. it is subject to some lexical (morphological)process
> removing its Case-marking ability (e.g. Passive
> morphology, reflexive clitics)
> or iii. by Universal convention, it is considered not a
> Case-marker (as, perhaps, INFL [-tense])

So a priori, we expect A,N,V,P and INFL to be Case markers. Note incidentally that we take A and N to be Case-markers. A transitive A or N Case-marks an NP it governs. It is equivalent for our purposes to say that A and N are not directly Case-assigners but rather trigger the insertion of a Case-marker when they govern some NP (e.g. insertion of of or 's).

3.1.2.1. Consider first prepositions. Putting aside the Case of reanalyzed prepositions, which may be considered to form part of a verb, prepositions taking NP-complements, i.e. transitive prepositions are always Case-assigners, at least in the languages in which they are not subject to the equivalent of Passive morphology.

The same conclusion applies to INFL. We have assumed that it is

either both a governor and a Case marker (as in tensed clauses, -ing clauses prior to Affix Hopping) or neither (as infinitival INFL). The one exception is INFL of tensed clauses which we consider a governor regardless of Affix Hopping. If Affix Hopping removes its Case-marking properties, we would find a governed PRO in subject position which would be left unbound in its Governing Category (here the clause it is subject of) and thus excluded. Consequently, we never find PRO governed but not Case-marked by INFL.

3.1.2.2. Consider now nouns and let NP* be some NP governed by a noun, say N. We can basically distinguish two cases depending on whether NP* is assigned a thematic role by N or not.

Suppose NP* is assigned a thematic role by N. Then, it is governed by N, which is, by definition, transitive, and by the assumption of obligatoriness of Case-marking, it cannot be PRO. Therefore, in such structures as:

(43)   i. $[_{NP}$ NP* $[_{N'}$ N...$]]$

        ii. $[_{NP}...[_{N'}$ N NP*...$]]$

NP* is obligatorily Case-marked and cannot be PRO.[31]

Before examining the situations in which NP* is not θ-marked by N, let us note that, by the same argument, NP* cannot be an NP-trace, as we have argued that NP-traces cannot be Case-marked.[32] From this, it follows immediately that there cannot be "NP-movement" in NP's. A phrase like Rome's destruction must be assumed to be base-generated as such.

It is worth pointing out that this consequence of our analysis is independently justified by the Projection Principle in its maximal

interpretation and therefore a priori appears preferable over the alternative. Recall that, in essence, the Projection Principle states that positions governed by lexical heads (except for subject of clauses, which are obligatory) are projected from the lexicon in that they are postulated to exist so that the lexical properties of lexical heads be met. Thinking of a lexical head as a n-places function, each place being assigned a θ-role, syntactic positions are postulated so that that function may actually have n arguments.

Consider the case of subjects of NP's ([NP,NP]). First, the argumentation showing the obligatoriness of subject of clauses does not extend to NP's (recall we analyze gerunds as clausal internally): pleonastic elements never appear as subjects of NP's (*there's being of a man, *its tendency that S...). Thus, the subject position of NP's is not an obligatory position. It is therefore predicted by the Projection Principle that occurrence of a subject position in an NP (a position governed by the head noun) follows from the thematic properties of the head noun. Accordingly, a subject position will appear in some NP only if it is a thematic position, i.e. is projected to fulfill the thematic properties of the head noun. However, because a θ-position can only be the most deeply embedded position in a chain, there can be no NP-movement in NP's.[33] The alternative, consisting in permitting a θ'-position as subject of NP's requires some complications of the theory of θ-marking (cf. Chomsky, 1981, p.40-41, for an explicit attempt).

Let us now return to the configuration N/NP* with N governing NP*, NP* not an argument of N. This may occur only in Exceptional Government situations, e.g. an N triggering S'-deletion or selecting a small clause.

We may however, simply assume that nouns are never exceptional

governors (cf. Kayne (1981b) for a detailed argumentation). Indeed, exceptional government is postulated (in syntax) for one of two reasons: either for reasons of Exceptional Case-marking, or because of the ECP. Consider the following structures:

(44)   i.   expect [NP* to VP]
        ii. seem [NP* to VP]
        iii. [$_{NP}$...N [$_{NP}$* to VP]]

verbs like, _expect_, _believe_, etc... may be followed by an infinitival clause with an overt subject. Since INFL [-tense] does not assign Case, it is plausible to assume that NP* is somehow Case-marked by _expect_ (an assumption corroborated by the Binding behaviour of NP*). Similarly _seem_ in (44ii) allows NP* to be an NP-trace. NP-traces must be governed. However, INFL [-tense] is not a governor, so that it is plausible to assume that _seem_ somehow governs NP* in (44ii). Neither of these situations arise in NP's. Thus, it is a well-known observation that NP* in (44iii) can never be an NP-trace (no Raising in NP's) nor can NP* be lexical. The resulting string N [_of_ NP* _to_ VP] is always ill-formed.

There is therefore no reason to assume that N can be an Exceptional Governor and/or Case assigner. It follows that PRO is never governed by N.[34]

Let us now turn to adjectives. Contrary to nouns, adjectives must be assumed to be sometimes, at least, exceptional governors, e.g. in raising constructions:

(45)   i.   Advantage$_i$ is likely [t$_i$ to be taken of John]

        ii. John$_i$ is liable [t$_i$ to succeed]

Indeed, _t_ in these structures is an NP-trace and must be governed. It

must be assumed under any analysis that Raising Adjectives do not assign Case to t_i. This follows immediately from the fact that Raising adjectives are never transitive. They never appear in $[_{AP}$ A NP] constructions contrary, to say, _proud_. Furthermore, they do not have an external argument, that is to say, no θ-role is assigned to their subject (viz. it is likely that S...), thereby permitting raising. Can PRO appear in the position of $\underline{t}$ in (45). Do we find structures like (46) in which NP* is PRO?

(46) $\left\{ \begin{array}{l} [_{small\_clause} \text{ NP}] \\ [_s \text{ NP be} \end{array} \right\}$ $[_{AP} \text{ A } [_s \text{ NP* to VP }]]]$

If A is a raising adjective, the answer is negative. NP* being governed by A, must be bound in accordance with Principle A, i.e. in a domain at most equal to the c-command domain of NP of (46), i.e. by NP of (46), the only available binder. As we have just seen, NP is not an argument as a possible binder for PRO. (46) would therefore violate the Binding Theory. However, if there are exceptionnally governing adjectives with a θ-subject, such structures as (46) should be possible, with NP* = PRO. Such adjectives would of course, have to be intransitive so that we would have to assume without explanation that only intransitive adjectives may be Exceptional Governors (which is not true for verbs), given that NP* can never be Case-marked by A in structures like (46). This indicates that the minimal assumption appears to be that syntactic S'-deletion or Exceptional Government is postulated only in view of positive evidence to that effect.

Consider finally verbs. Contrary to what we find with other categories, we do find structures like (47):

(47)   [$_s$ NP INFL [$_{VP}$ V...NP*...

In which NP* is governed by V, but not assigned Case by it. Given our assumptions about Case-marking, this may arise in one of two ways. Either NP* is not an argument of V, as in Exceptional Government configurations. In that case, V must be intransitive, or it must be morphologically marked as assigning no Case. Or NP* is an argument of V. Only the morphological option is available. Furthermore, NP in (47) may be assigned a θ-role by NP or not. Consider the possible combinations of these situations.

Assume first, that NP* is an argument of V. V is therefore morphologically marked as assigning no Case and may or may not assign a θ-role to its subject. We find examples of both:

(48)   i.   [$_s$ NP [$_{VP}$ [se+rasé] [PRO]$_{NP*}$]]

         ii. *[NP [$_{VP}$ [$_V$ a été rasé] [PRO]$_{NP*}$]]

In (48i), NP* receives no Case as we have argued was the R-Clitic effect on the verb, and NP receives a θ-role. NP*=PRO, must be bound in S, i.e. by NP. (48ii) is an example of such a structure in which NP does not receive a θ-role from V (or more precisely from VP). This is the standard analysis for passive morphology. In that case, PRO governed by V must be bound. NP being a non argument (or rather, in a θ-less chain) does not qualify. (48ii) is therefore ruled out.

Consider next, the case of NP* not an argument of V. By the same reasoning, NP*=PRO must be bound within S in (47). If it cannot be bound, the structures violate Principle A. Here are some examples:

(49)   i.   [$_s$ NP [$_{VP}$ se+considérait [PRO intelligent]]]

          ii. *[$_s$ NP [$_{VP}$ was believed [PRO to VP]]]

iii.*[_s there was [PRO sick]]

iv. *[_s NP [_vp happened [PRO to VP]]]

v. [_s NP [_vp semble [_NP** a Jean] [PRO to VP]]]

So (49i) and (49ii) parallel (48i) and (48ii). (49iii) is essentially identical to (49ii). In (49iii) through (49v), we have intransitive verbs not assigning a θ-role to their subject. (49iii) and (49iv) are ungrammatical for no Binder is available for PRO. The case of (49v) noticed in Rouveret & Vergnaud (1980) is more interesting. As Bouchard (1982) points out, the indicated analysis must be preferred to an analysis in which NP** has raised from embedded subject position (as Rouveret & Vergnaud, op.cit., argue, because of the grammaticality of il semble à Jean que S showing clearly that NP** may be associated with a θ-role independent of the embedded clause. (Notice the raising analysis is also excluded by the Projection Principle). Bouchard (1982) further notices that a theory allowing governed PRO predicts correctly the grammaticality of such structures. PRO governed by sembler must be bound within S and can in fact be by NP** (Recall that in French, contrary to English, indirect objects behave as NP's in terms of c-command properties). Furthermore, the well-formedness of (49v) is problematic for a theory assuming that PRO must be ungoverned for it would have to assume that sembler obligatorily triggers S'-deletion unless it takes an indirect object, in which case S'-deletion is optional. Indeed, if S'-deletion does not take place with bare sembler we would expect il semble [PRO to VP] to be well-formed with il expletive and PRO ungoverned (which it is not).

These examples provide further evidence in favor of the theory presented here.

As a final note on tne topic of Exceptional Government, notice that there is a gap in the paradigm in (49). There do not seem to be intransitive verbs which are both exceptional governors and assign a θ-role to their subject. Furthermore, the existence of structures like (48i) and (49i) indicate that the generalizaton (15) of Chapter II is too strong. Putting these two remarks together, we can formulate a weaker version of (15) as (cf. Chapter II. footnote 9)

(51)   If an intransitive verb V governs some NP (hence does
         not Case-mark it) V^{max} assigns no θ-role to its
         subject.

We can even generalize to (52), if we take into account the preceding discussion on Adjectives:

(52)   If an intransitive [+V]° governs some NP, [+V]^{max}
         assigns no θ-role to its subject.

This concludes the discussion of governed Caseless contexts, whose properties, we have seen, are entirely compatible with our assumptions.

3.2. Formulation of the Binding Principles

We now discuss some necessary modifications to the formulation of the Binding Theory given in Chomsky (1981; 1982).

The Binding Theory imposes distributional constraints on ¥-categories -i.e. pronouns, pro, PRO, NP-trace (=expletive PRO)...- and on lexical anaphors. We have assumed so far that, as far the distribution of PRO was concerned, the relevant binding restriction was as in (53):

(53)   x=PRO must be A-bound the first NP or S (or small clause)
         containing x and a governor of x

We now proceed to justify this formulation, showing that it in fact extends to all Caseless empty categories, and that the relevant principle for Case-marked Î-categories, i.e. Principle B, is its exact counterpart.

3.2.1. On SUBJECTS and Binding Domains

First ~~let us sketch~~ some observations and arguments made in Huang (1981; 1982b) showing that the Binding Domains of lexical anaphors and pronouns is distinct (cf. Huang, op. cit., for a more detailed exposition). The general form of the Binding Principles is as in (54):

(54)   i.   Principle A : x∈A must be A-bound in D(x)
        ii. Principle B : x∈B must be A-free in D'(x)

Where we will call D (or D') indifferently Binding Domain or Governing Category. Consider the following examples:

(55)   i.   They saw their/each other's pictures

        ii. They saw pictures of them/each other

       iii. They expected that pictures of them/each other be on sale

       iv. They expected that for their/each other's pictures to be on sale would be possible

       v. They expected that for them/each other to come would be possible

       vi. They expected that it would be possible for (friends of) them/each other to come

As Huang notes, although taking D=D' is largely correct, the forms in (55) constitute a systematic set of counterexamples in English (similar facts are found in other languages, e.g. Chinese). Indeed, if D=D', pronouns and lexical anaphors should not be able to appear in the same contexts bound

by the same NP. In (55) however, both <u>them</u> and <u>each other</u> can be bound by <u>they</u>. Huang concludes that D(lexical anaphors)≠D'(pronouns). On that basis, and on the basis of some conceptual considerations as well, Huang (op.cit) argues for a characterization of D and D' distinct from that of Chomsky (op.cit).

Now, both theory and Chomsky take PRO to belong to both A and B and wish to derive from this assumption that it must be ungoverned. Chomsky (1981; 1982) obtains this result by making D=D': if governed, PRO would have to be both bound and free in D.[35] However, as Huang shows this is inconsistent with the observations in (55). Huang therefore proposes the following formulation,[36] in which D(lexical anaphors) ≠ D'(pronouns):[37]

(56)   i.   Principle A : x∈A must be A-bound in D(x)

        ii.   Principle B : x∈B must be A-free in D(x)

        iii. D(x) is the minimal category containing x, a governor of x
              and a SUBJECT c-commanding x which, if x∈A, must be
              accessible to x

Where (56) must be interpreted in the context of the auxiliary assumptions (57) taken from Chomsky (1981) - except for (57i), due to Huang, and a slight reformulation of the notion "accessible" where we make coindexing asymmetrical to avoid minor technical problems of Chomsky's formulation:

(57)   i.   A SUBJECT of a phrase P is the subject of P or the
              nominal head of P (i.e. AGR in S or N, head of NP)

        ii.   AGR is coindexed (or cosuperscripted) with the subject NP
              it governs

        iii. The index of a phrase XP percolates to its head X

        iv.   y is accessible to x iff attribution of y's index to x
              does not yield a configuration meeting the i-within-i
              Filter: *[K...C...] where K and C share an index and C
              is not the head of K

Note that the only (and crucial) differences between Chomsky's formulation and Huang's are:

- The inclusion of the phrase: if x∈A in (56iii)
- Counting N as a SUBJECT of NP

Now it is clear that Huang's theory has the desired effects concerning PRO and also for pronouns and lexical anaphors. For these categories of expressions, we can formulate the constraints equivalently as:

(58)   i.   A lexical anaphor L, must be A-bound in the first NP or S (or small clause) containing L, a governor of L and a SUBJECT accessible to L.

          ii.   A pronoun P must be A-free in the first NP, S (or small clause) containing P and a governor or P.[38,39]

For lexical anaphors, (58i) is a conflation of (56i) and (56iii). For pronouns, (58ii) follows from the fact that, by (57i), any S, NP (or small clause) has a SUBJECT.

Now let us briefly examine the examples of (55) in the light of (56) or (58). As far as pronouns are concerned, it suffices that they be free in the first NP or S containing them. This is clearly the case in all of the forms in (55). Consider next lexical anaphors.

In (55i) and (55ii), the first SUBJECT accessible to each other is they. Indeed, the noun pictures is not. It is coindexed with the NP it is the head of (by 57iii). Coindexing of that NP with each other would yield a violation of the i-within-i Filter in (57iv). The Binding domain of each other is the whole clause and each other is bound in it. The same reasoning carries over to (55iii), (55iv), (55v) and (55vi). In each case the first accessible SUBJECT is they so that they are all well-formed. A comment is in order concerning (55vi), since the above conclusion rests on

an assumption we have not so far made explicit, namely that a clause "extraposed from subject position NP*" is coindexed (or cosuperscripted) with the pleonastic element filling the position NP*. Justification for this assumption comes from θ-theory and Case theory applied to clausal arguments, a matter we have not discussed here.[40]

If some such coindexing holds, the structure of (55vi) is actually:

(59)   They expected [it^i AGR^i would be possible
            [for [((friends of) each other] to come]^i

Coindexing of __each__ __other__ with either __friends__ or __AGR__ would yield a violation of (57iv).

3.2.2. Binding Domains

The arguments presented so far only dealt with lexical anaphors and pronouns. For Huang (op.cit), (56) is also intended to extend to other ϒ-categories. We are now going to argue that a further modification is required for two reasons:

      i.   first, because we admit governed PRO's, (56) would give incorrect results

      ii.   second, because the behaviour of NP-traces seems to require some further modifications independently of the behaviour of PRO

More specifically, we are going to argue for the following formulation:

(60)   i.   Principle A : x∈A must be locally A-bound in D(x)

        ii. Principle B : x∈B must be locally A-free in D(x)

        iii. D(x) is the minimal category containing x, a governor of x and a SUBJECT which, if x is a lexical anaphor, must be accessible to x (where these notions are defined as in 57)

Recall that we have argued that A is the set of Caseless Y-categories and lexical anaphors, while B is the set of Case-marked Y-categories. If (60) is correct, as we claim, it means that the Binding Theory is totally symmetrical, as far as Y-categories are concerned. Lexical anaphors deviate from this symmetry: a not unlikely behaviour for lexical elements whose properties can be stipulated.

Notice next that we have formulated Principle A and Principle B as involving local A-binding and local A-freedom respectively. This makes (60i) a slightly stronger principle. It rules out a situation in which x∈A is locally A'-bound and A-bound within D(x). Conversely, it makes (60ii) a slightly weaker principle. It permits a locally A'-bound element to be A-bound within D(x) (i.e. permits the situation described in footnote 20). We will see in section 6 below that this change may have desirable consequences.

Let us now examine the empirical differences between (56) and (60). Of course, these differences will materialize only in the predicted behaviour of Y-categories falling under Principle A, i.e. PRO and NP-trace (=expletive PRO). This is because the only difference between (56) and (60) is the replacement of if x∈A in (56iii) by if x is a lexical anaphor in (60iii). More precisely, for such categories, the difference will show up in a context in which the minimal category containing a SUBJECT c-commanding some Caseless Y-category K and containing a governor of K is strictly included in the minimal category containing an accessible SUBJECT c-commanding K and a governor of K (Note that the reverse situation is impossible). Furthermore, by considerations of Case and government -the only contexts relevant to our discussion are those in which PRO and NP-trace appear in governed Caseless contexts- we can narrow down the

relevant structures to such cases as (61) below, having the abstract structure (61i) where X is an exceptional governor:

(61)   i.   NP [$_{XP}$ V [$_S$ NP [$_{XP}$ X*   [$_S$ NP*...]]]

        ii. *They think [it^i [is expected [^{i} PRO to leave]]
            (they think it is expected they will leave)

        iii.*They seem [it^i [is expected [t^i to leave]]
            (=it seems it is expected they will leave)

Both (61ii) and (61iii) are ill-formed. In a theory where PRO must be ungoverned, the ungrammaticality of (61ii) follows: PRO = NP* is governed by X*=expected. For us, however this is not true. This means that we must prefer (60) over (56). According to (56), (61ii) would be well-formed if PRO is considered to belong to set A only (which is not the case for Chomsky, op.cit.,; Huang, op.cit.) since its Binding Domain is S*, containing no possible binder. Thus, for internal consistency, we must prefer (60) over (56).   .

(61iii) is more interesting however, because it suggests that (60) must be preferred over (56) independently of the status of PRO. According to (60), (61iii) is ruled out exactly for the same reason (61ii) is. There is no antecedent in the Binding Domain S* of t.

For (56) (and Chomsky's formulation as well) (61iii) is well-formed as far as the Binding Theory is concerned since the first SUBJECT accessible to t is they. What then rules (61iii) out? This question is discussed at length in Chomsky (1981, p.58 and p. 306ff). Note first that there are clearly no θ-theory nor Case theory violations. Chomsky (op.cit) argues it is Bounding Theory. He assumes that the antecedent/NP-trace relation is subject to the Subjacency Condition and that t is not subjacent to they. There are a number of reasons to doubt

this explanation (cf. III 8.2.3. and III 9.2.) the most simple of all being the well-foi..edness of structures parallel to (61i) in the relevant respects involving wh-movement. Such cases can be found in those dialects of French and Italian (cf. Rizzi (1982); Sportiche (1981)) in which S does not count as Bounding. In such dialects, the equivalent of (61iii) are of course ill-formed, while such examples as those ʲn (62) are grammatical:

(62)   i. $[_{s'}$ wh$_i$ $[_s$ NP $[_{vp}$ V $[_{s'}$ wh$_j$ $[_s$ NP*$_j$ $[_{vp}$ V $[_s...t_i...$

        ii.   voila des gens [que$_i$ [l'on ne [sait pas [qui$_j$ [e$_j$
              [a vu [t$_i$ sortir

As the comparison between (62) and (61) makes clear, if the relation wh$_i$ /t$_i$ of (62) obeys the Subjacency Condition, as is clear from the grammaticality of (62ii), the relation NP/NP* of (61) must do so aь well since they are identical in the relevant respects.[41] This suggests that (61iii) is not a Subjacency violation. Therefore, if (56) is correct, we are left with no explanation for the ill-formedness of (61iii). This clearly favors (60).

Why do lexical anaphors behave differently from Caseless ϒ-categories in terms of Binding? This behaviour is not systematically found. For example, in French, in which the relevant data do not seem to exist, or, in Vata, a West African language of the Kru family, in which lexical anaphors (i.e. reflexives) behave exactly as NP-traces do in English (cf. Koopman, 1983). In that language, the equivalents of (55) are ill-formed with a reflexive in place of each other. The behaviour of PRO and of NP-trace (or more precisely, expletive PRO) does not seem to vary in this fashion from grammar to grammar. These observations do suggest some type of parametric or markedness approach to the English data. Huang (1983) argues that this is suspicious because facts similar to the English facts

are found in Chinese (and perhaps other languages). But I do not see why. We could assume that marked options may be structured only in very specific ways: e.g. upon presentation of the relevant data (e.g. 55i), the language learner has only the option of adding the extra condition: "if x is a lexical anaphor, consider accessible SUBJECTS instead of SUBJECTS" in (60iii). There are many implicit assumptions in such an account, but it does not appear a priori implausible. If correct, it means that the core of Binding Theory is represented by the behaviour of f-categories, a fairly natural conclusion.

3.2.3. Some further remarks and outstanding problems

Consider finally the following examples:

(63)   i.   *each other's pictures are on sale
         ii.   *for each other to leave would be premature
         iii. Some lies about each other have triggered the fight
         iv. There^i has been [a number of games against each other]^i

(64)   i.   *PRO's pictures are on sale
         ii.   *for PRO to leave would b. a mistake
         iii. *[for it^i to be expected [PRO to leave]^i] would be
            premature

Consider first (64). In (64) we find governed PRO's with no accessible SUBJECT. Both for Chomsky's theory and for Huang's (i.e. (56)), some stipulation has to be added to rule out these ungrammatical cases. Indeed, for these theories, PRO does not have a Binding Domain in such cases, since a Binding Domain must include an accessible SUBJECT. Consequently, no binding constraint is imposed on such PRO's. Chomsky (1981) adopts the following suggestion due to N. Hornstein:

(65)   The main clause is a Binding Domain for a governed element

Adopting (65) permits the assignment of a Binding Domain to these PROs, thereby ruling out the forms in (64).

Under (60), note that no such stipulation is necessary. In (64) we find a free governed PRO, which is ruled out by Principle A (note that (64i) and (64ii) are redundantly ruled out by the fact that the PRO's are in a Case context).

Now the consideration of (63) shows that (65) is too strong. As we have already pointed out, Chomsky (class lectures, fall 82) has remarked that lexical anaphors do not always have to be bound. This is illustrated by the acceptablity of such examples as (63iii) and (63iv). The conditions under which this may occur are rather obscure.[42] However, these facts show that (65) is too strong. The lexical anaphors in (63) are all governed and are all assigned a binding domain because of (65). They should therefore have to be bound in this domain, contrary to fact. Again, these data seem to support a differentiated treatment for lexical anaphors and Caseless Y-categories by the Birding Theory as (60) claims. If we compare the two approaches, we see that the ones we just reviewed face the problem of explaining why (63iii) and (63iv) are well-formed. Our assumptions face the problem of why (63i) and (63ii) are ill-formed: since the lexical anaphors are assigned Binding Domains, the Binding theory (60) requires nothing of them.

There is thus a trade-off of problems between the two approaches, with the difference that we have no need of postulating (65): again (60) appears slighly superior over the alternatives.

4. ECP and Expletive PRO

Of all the questions we started with in (2), only (2v) and (2vi) remain unanswered.

4.1. Where dces the Binding Theory apply?

(2v) concerned the level at which the Binding principles must be met. We will not investigate this question here. The question has been discussed in many recent works, e.g. Aoun (1982), Chomsky (1981), Chomsky (1982), Fourier (1980), Van Riemsdijk & Williams (1981)... The evidence presented ir most cases is very highly theory internal, tied to specific assumptions these authors make, usually different from ours. No conclusion can thus be drawn from it here. One exception is the evidence presented in Chomsky (1981, p. 196ff). However, it only deals with so-called Principle C of the Binding Theory, that we have discussed in III.7, and which is of a very different nature from Principles A and B. Furthermore, as Chomsky (1981) remarks, this evidence is only of limited weight. If correct, it would require Principle C to apply at S-structure. There is little reason, however to assume that this conclusion extends to Principles A and B (in fact, given our formulation, nothing requires Principle A and B to constrain the same level(S)).

For concreteness, we will assume that Principles A and B apply both at S-structure and at LF:

4.2. What does the ECP apply to?

Let us now turn to (2vi). This is more easily answered. It is quite clear that the theories of Binding, Case, Bounding and Thematic structure as we have formulated them are not sufficient to account for the distributional properties of non overt categories, in particular, pro, whether locally A'-bound (i.e. "wh-trace") or not and expletive PRO (and in particular NP-trace). The required additional principle has been so far referred to as the ECP.

4.2.1. ECP for pro≠ECP for NP-traces

Prior to the work of Chomsky (1982), the ECP was thought to require of "traces" to be properly governed, a notion slightly weaker than lexical government. Because Chomsky (1982) has proposed the existence of the category type pro, which appears to be subject to the same kind of restrictions as as "wh-traces", the set of elements to which the ECP is supposed to apply cannot be characterized as "traces". Furthermore, within the framework of assumptions developed here, there is no notion of "trace" applying both to NP-traces and "wh-traces". We have thus two reasons to doubt that a unique principle, the ECP, governs the distribution of NP-traces, pro, and "wh-traces" (more accurately, locally A'-bound pro). The most natural break seems to be between pro (whether locally A'-bound or not) and NP-traces.[43] From now on, we will reserve the term ECP to whatever Identification Principle pro is subject to. Furthermore, we will see immediately that treating NP-traces apart is a posteriori justified by the fact that the independently needed principle

in (67) below will subsume the effects of the ECP for NP-traces.

4.2.2. The distribution of Expletive PRO

Let us first examine NP-traces in more detail. Recall that we have noted that an NP-trace is nothing else than an expletive PRO in non chain initial position. Furthermore, because we have assumed that expletive PRO was barred from chain initial position (i.e. if x is expletive PRO, then it is an NP-trace) and reciprocally that NP-traces are Caseless expletives (i.e. if x is an NP-trace, it is a Caseless expletive, that is, expletive PRO) it made no difference to talk about NP-traces or expletive PRO (x is an NP-trace iff it is an expletive PRO).

This assumption was not entirely accurate, however. We have shown in Chapter II (cf. in 2.6.1 the discussion around the examples (44)) that Caseless expletives (i.e. expletive PRO's) were excluded from chain initial position. The argument was based on the behaviour of impersonal passive constructions in Dutch (or German) or, more generally, on the behaviour of predicates with no syntactically expressed arguments (as **weather** verbs in Dutch, German, Yiddish, which, contrary to their English or French counterparts, do not take quasi-argument subjects, but expletive subjects. cf. II.fn.27).[44] As can be easily checked, it turns out that the empirical evidence presented in Chapter II only showed that expletive PRO are barred from ungoverned chain initial position.

Could expletive PRO appear in governed chain initial position? In fact, the answer to this question is positive. Recall that we have also argued in Chapter II that the chain formation algorithm simply stated: Partition the set of NP's into chains. Other conditions, we argued, would

insure that each chain meets the θ-criterion appropriately. In most cases, we derive the result that there is only one way to carry out this partition for a given sentence S, and end up with well-formed chains only. There is one exception: chains containing only expletives. Consider (66):

(66)   [$_{NP}$ het] schijnt [$_{NP*}$ e] te regenen
              it   seems   to be raining

It is easy to check that the theory as it stands permits two different chain structures for (66): in one, (NP,NP*) forms a chain, NP* is an expletive PRO in NP-trace position. In the other, each of NP and NP* forms its own chain; NP* in this case is expletive PRO in chain initial position. Of course, the reason why such ambiguities are possible comes from the fact that such chains do not involve any θ-role transmission. More generally, given a well-formed expletive chain (NP$_1$,..., NP$_n$) any partition of the set {NP$_1$ NP$_2$,..., NP$_n$} into continuous subparts can be taken as a partition in well-formed chains.

So in fact, expletive PRO may appear in chain-initial position provided that it is governed.[45] The most general way to state this observation is as in:

(67)   The Constraint on Expletive PRO: Expletive PRO must be
                                          governed

(67) a priori appears too strong. It seems that it should be qualified to apply to expletive PRO's in chain initial position only. However, (67) as stated has the interesting property of being able to derive the effects of the ECP for NP-traces.

Indeed, as we have already noted in IV.2.2.2, the work done by the ECP for NP-traces comes down to requiring of NP-traces to be governed

[Recall that the ECP required of NP-traces to be properly governed. Properly governed positions are governed positions and subject positions of tensed clauses. But, for independent reasons, -Case, Binding-NP-traces are barred from subject of tensed clauses position: so, in fact, the ECP required of NP-traces to be governed].

Furthermore, we have accepted the generalization that NP-traces are always expletive PRO's: it follows that (67) requires of NP-traces to be governed, i.e. subsumes the effect of the ECP. Let us therefore accept (67) as a valid generalization. In 6.2.3. below, we will propose a possible way of <u>deriving</u> it.

5. Remarks on Control Theory

We now sketch some remarks on Control Theory suggested by the framework we have developed. By Control Theory, we mean, as is usual, the theory dealing with the referential properties of PRO.

5.1. The Obligatory/Non-Obligatory Control Distinction

5.1.1. All recent theories of control assume a distinction between Obligatory Control (henceforth OC) Constructions and non Obligatory Control (henceforth NOC) Constructions, either explicitly as in Williams (1980), Koster (1981), Bresnan (1982), Bouchard (1982) (of course, each author has a specific way of implementing this distinction) or implicitly, as in Manzini (1983a).

We wish to argue here that the OC/NOC distinction is necessary -

although, perhaps not sufficient -- and propose particular ways to derive the properties of OC inspired by Koster (op.cit.).

Manzini (1983a) is the only recent theory in which this basic distinction is not explicitly accepted. We will see below that this theory in fact does contain provisions having the same effect. It should be noted however, that none of the theories mentioned above seems totally satisfactory. I accept here some of Manzini's criticisms of Williams (op.cit.) and Bresnan (op.cit.), which carry over to Koster (op.cit.) and Bouchard (op.cit). Also, it is clear, as we shall see in 5.3. below, that more than a binary distinction OC/NOC is necessary.

Before examining Manzini's theory, let us define what we will mean by OC. We will take the essential differences between OC and NOC to be as follows:[46]

Given PRO*:   in OC construction, there must be a syntactically
                   expressed antecedent for PRO*, which binds PRO*
                   (i.e. c-commands PRO*, has identical f-features
                   and referential index as PRO*)

                   in NOC construction, there need not be a
                   syntactically expressed antecedent for PRO* and,
                   if there is one, it need not bind PRO*.

5.1.2. In Manzini (1983a), the referential properties of PRO are determined as follows. First, given a PRO, call it PRO*, a certain algorithm computes whether PRO* possesses what Manzini calls a Domain Governing Category (henceforth DGC). If PRO* possesses one, it must be bound in it. Otherwise, it freely (co)refers. Manzini defines DGC as in (68) (cf. Manzini, op.cit., for the definitions of the technical terms used, some having a slightly different meaning from the one we assume):

(68)   K is a DGC for PRO iff:

    i.   K is the minimal category with a subject containing PRO,
        the c-domain of PRO (i.e. the minimal maximal projection
        dominating PRO) and a governor of the c-domain of PRO.
    ii. K contains a subject (or SUBJECT) accessible to PRO.

There are two reasons why a PRO might lack a DGC. Its c-domain might be ungoverned (as, e.g., if the clause containing the PRO has a non-null COMP); or, the category K of (68i) contains no subject (or SUBJECT) accessible to PRO (as, e.g., if the c-domain of PRO is coindexed with the subject of K. This leads to the following relevant consequences:

(69)   i.   a PRO subject of an object sentence S with null COMP is
        obligatorily bound within the first category with a
        subject containing S.
    ii. a PRO subject of an extraposed clause S* freely refers.

(69i) is transparent. (69ii) follows because, either S* has a non null COMP or it is coindexed with the subject of K (or both). This is what it means to be extraposed - so that the subject (SUBJECT) of K is not accessible to the PRO S* contains.[47] We shall now see how Manzini uses the OC/NOC distinction. Consider the following examples:

(70)   i.   John attempted PRO to shave himself/*oneself/*herself
    ii. John decided PRO to shave himself/*oneself/*herself

Both sentences in (70) fall under (69i). PRO must be bound by <u>John</u>. Consider now the passive counterparts of (70):

(71)   i.   *It was attempted PRO to shave oneself/himself
    ii. It was decided PRO to shave oneself

(71i) is ill-formed, while (71ii) is not. this shows that the PRO in the object sentence of <u>attempt</u> requires a syntactically expressed antecedent, while the PRO in the object sentence of <u>decide</u> does not. Given the way we

have defined OC and NOC, <u>attempt</u> would be an OC predicate, while <u>decide</u> would be an NOC predicate.

Manzini achieves the correct result by taking the S of (71ii) to be extraposed so that its PRO freeely refers by (69ii), while she takes the S of (69i) not to be extraposed, so that it would fall under, and thus violate (69i).[48]

Since Manzini provides no criteria independent of her theory of control to decide whether a clause is extraposed or not,[49] this stipulated distinction is simply another way to implement the OC/NOC distinction.

5.2. Obligatory Control

We have argued that governed PRO's should be permitted. When a PRO is governed, it must be bound. We see immediately that the properties of PRO in OC constructions would immediately follow if such PRO's were governed. One very natural way to implement this idea is to assume that OC predicates are in fact exceptional government predicates, i.e. S'-deletion predicates. Consider the following examples:

(72)   i.   John believes [Bill to have left]
        ii. John attempted [PRO to leave]

Suppose that, exactly as the verb <u>believe</u> is assumed to exceptionally govern the subject NP <u>Bill</u> of its complement clause, the PRO subject of <u>leave</u> in (72ii) is exceptionally governed by the OC predicate <u>attempt</u>. Following usual assumptions, we may assume that Exceptional Government is rendered possible because some process renders the S' bracket of the complement clause transparent, say S'-deletion, for concreteness.[50]

It is clear how the properties of OC PRO follow. In particular,

notice that (71i) is excluded by the Binding Principle A. PRO, governed by attempted, must be bound within its Binding Domain (here the main clause). Since no possible binder is availabe, (71i) is excluded.

This analysis, first proposed by Koster (1981) (and adopted and defended in Bouchard (1982)) raises some questions. Consider the following examples:

(73)   i. John attempted [NP* to leave]
       ii. *John was attempted [NP* to leave]

If indeed attempt triggers S'-deletion, why can't NP* be lexical in (73i)? The verb attempt is a transitive verb which, by assumption, governs NP*. Therefore NP* should be a Case position (recall obligatoriness of Case-marking) allowing NP* to be lexical and thereby excluding PRO. Another problem arises with (73ii). In (73ii) we must form a chain (John, NP*). Otherwise John, an argument, would be in a θ-less chain. However, (73ii) is ill-formed. The usual explanation in terms of the ECP, (which would have come from the CEP (67) requiring NP* to be governed is not available since NP* is governed What is then the explanation for the ungrammaticality of (73ii)? Suppose that we construe the lexical property of OC predicates:

(74)   OC predicates trigger S'-deletion in LF.

We can see immediately that (74) solves both problems. First, NP* in (73i) cannot recieve Case. Case-marking (or Case checking) is an S-structure property. At S-structure NP* is ungoverned since the S' boundary is present at this level. As maximal projection, it blocks government, hence Case-assignment.

Secondly, (74) allows us to provide a simple explanation for the

ill-formedness of (73ii). We have argued that the requirement that NP-traces be governed was best stated as in the CEP (676). We have not yet considered the question of what level of representation (67) applies at. It is easy to see that taking it to apply at S-structure (and, perhaps, at LF) will cause the exclusion of (73ii). Let us therefore assume that the CEP applies at S-structure.

At S-structure, the NP-trace or expletive PRO NP* of (73ii) will not be governed, due to the presence of the S'-boundary, thus violating th CEP.

Summing up, we get the folowing paradigm in which (75) are S-structure representations and (76) corresponding LF representations:

(75)   i.   John$_i$ attempted [$_s$· [$_s$ PRO$_i$ to leave]]
        ii. *John$_i$ was attempted [$_s$· [$_s$ t$_i$ to leave]]
        iii. *it was attempted [$_s$· [$_s$ PRO to leave]]

(76)   i.   John$_i$ attempted [$_s$ PRO$_i$ to leave]
        ii. John$_i$ was attempted [$_s$ t$_i$ to leave]
        iii. *it was attempted [$_s$ PRO to leave]

The representations (75i) and (76i) are both well-formed. PRO is not in a Case position because it is ungoverned at S-structure as (75i) chows. It must be bound by John because it is governed at LF and falls under Principle A of the Binding Theory at this level. (73ii) is ill-formed because its S-structure representation (75ii) is. It contains an ungoverned expletive PRO. Finally, (71i) is ill-formed because its LF representation (76iii) is. It contains an unbound governed PRO.

Before pursuing, notice that our account of sentences like (71i) is incomplete. We attribute its ungrammaticality to the ill-formedness of its LF-representaion (76iii). However, the validity of our argument rests on there not being a possible antecedent for PRO. Suppose, however, that

this PRO is expletive. Then _it_ would become an acceptable antecedent and the corresponding sentence should be well-formed. We should therefore expect (77ii) to be well-formed:

(77) i.   es regnet
           'it rains'

      ii.* es wurde versucht zu regnen
           'it was attempted to rain'

just like in Dutch, weather verbs in German are argumentless.[51] As (77i) shows, expletive _es_ is a possible subject for the verb regnen. If we take the subject of _regnen_ in (77ii) to be the non overt counterpart of _es_ -i.e. expletive PRO- the subject of the main clause should be able to act as its antecedent. (77ii) contrary to (76iii) thus involves no Binding Theory violation. The LF representation of (77ii) is therefore well-formed.

However, its S-structure representation is not. At S-structure (77ii) contains an ungoverned expletive PRO, namely the subject of the infinitival clause. This example is important because it shows that the exclusion of examples like (75ii/76ii) on semantic grounds (as in Koster, op.cit) would have no explanation for (77ii).

5.3. Non Obligatory Control

5.3.1. Let us now examine the referential behaviour of PRO in NOC constructions. Consider the following structures:[52]

(78) i.   John attempted [NP* to leave]
      ii.   John proposed to Bill [NP* to leave]
      iii. John described to Bill [how [NP* to solve the problem
            (together)]]
      iv.   It is unclear [what [NP* to do]]

In each case, NP* must be analyzed as a Caseless empty argument, i.e. as a PRO. Since it is Caseless, we have argued that it should be [+anaphoric], i.e. lacks inherent reference. Superficially however, NP* in (78) displays various referential properties.

In (78i), NP* is in an OC construction. It must have a unique antecedent binding it. In (78ii), NP* does not have to be bound. Recall that we have argued in III.7 that binding requires identity of referential indices and, in particular, excluded split antecedents. In (78ii), NP* can have John, Bill or both as antecedent(s). We are therefore dealing with coreference (or referential dependency, cf. below) rather than with binding. Notice that we cannot assume that (78i) is similar to (78ii) in this respect but that lack of more than one possible antecedent gives the illusion that (78i) involves binding. The difference between the two is clearly brought out by their respective passive counterparts. No antecedent need be expressed in that of (78ii), contrary to that of (78i) (viz. it has been proposed to leave versus (71i)).

In (78iii), NP* can have either John, Bill or both as antecedent. There is however, an additional possibility. NP* can also have an arbitrary reading which is impossible in (78ii). Finally, in (78iv), NP* can be eitner arbitrary in reference, or as Bresnan (1982, p.381) remarks, it may also pick a "specific extrasentential referent" in an appropriate discourse context:

(79)   she sighed and looked around the empty room. It was
            unclear what to do with herself now that Molly was gone
            (Bresnan (1982, p. 381)).

5.3.2. These observations raise two questions.

(i) Is this behaviour compatible with the [+anaphoric] status we
     attribute to Caseless elements and to PRO in particular.
(ii) How do we account for the referential properties
     of these NOC PRO's.

Obviously, these two questions are related. However, we will focus
here on the first question, postponing some suggestions about possible
answers to the second until 6.2.4. below.

Let us nonetheless note two conceptually very similar proposals made
in Bouchard (1982) and Bresnan (1982) according to which these PRO should
be treated as "pure" pronominals (i.e. what we would describe as
[-anaphoric,+pronominal]). Bresnan (1982) argues that the ability to have
"specific extrasentential referents", as PRO does in (78iv), is
illustrative of pronominal properties and unlike those of (lexical)
anaphors. Assuming that we are dealing with the same element in (78ii),
(78iii) and (78iv), we would be led to postulate that NOC PRO is in fact a
pronominal element.

Note first that there seems to be a difference between pronouns (and
pro) on the one hand, and PRO as in (78iv) on other other. In order for
this PRO to get specific extrasentential reference, the discourse context
must be appropriately constructed. For example, it seems to me difficult
to get such an interpretation if (78iv) is in isolation. This is not at
all the case for pronouns (or pro); specific extrasentential reference is
always readily available. This might suggest that such interpretations
for PRO do not reflect grammatical properties of the elements involved.

Secondly, notice that if indeed the PRO's of (78ii), (78iii) and
(78iv) (or even only the last two) are treated as pronominals, it raises
the question of why selection of specific extrasentential referent is not
possible in (78ii) and (78iii), since pronouns (or pro) in comparable

positions always freely have this option.

Thirdly, notice that it is not quite accurate to say that lexical anaphors lack the capacity to select specific extrasentential referents. We have seen in 3.2.3. above that, given appropriate structural conditions (and, perhaps, appropriate discourse context) we do find examples of lexical anaphors with extrasentential referents.

Finally, consider in more details how pronouns refer (in D). We have followed Chomsky (1982) in proposing that pronouns may always have inherent (or independent) reference, while anaphors may not. We could phrase the difference in other terms by saying that anaphors are necessarily referentially dependent. The identity of the element they select in D must be determined through coindexation (sometimes extrasentential) with an expression selecting some element or set in D.

Pronouns may also be referentially dependent but they do not have to. For example, the following sentence in which John and he refer to the same individual is (at least) two way ambiguous:

(80)   I told John he is sick

Roughly speaking, it can mean either, I said to John: "you are sick", a reading we might represent as: x=John, I told x that x is sick, in which he is referentially dependent upon John (a relation represented formally by coindexing). Or it might mean, I said to John: he is sick in which he - that I identified some way other than by his name - happens to be John (I might not even be aware of it). In this case, John and he simply happen to pick the same referent. This is a case of coreference which, of course, requires no coindexing. This distinction is the long acknowledged distinction between the referential interpretation and the "bound"

variable interpretation for pronouns (cf. for example Sag (1976); Williams (1977) or Reinhart (1980)). As Reinhart (1980) discusses at length, the difference between the two clearly appears in VP deletion contexts:

(81)   Bill likes his home town and John does too

Suppose his and Bill refer to the same individual. The second conjunct of (81) is ambiguous meaning either that John likes his own home town (a reading often called, after Ross (1967), sloppy identity) or that John likes Bill's home town (non sloppy identity). This observation relates straightforwardly to the coreference/referential dependence distinction. If his in the first conjunct corefers with Bill, its reference is fixed independently from that of Bill and we get the non sloppy reading (cf. the references cited for why). If his is referentially dependent upon the subject of its clause, i.e. Bill (which is possible only under certain specific conditions, cf. Reinhart (1980) and Chapter III condition (18)), we get the sloppy reading.

Again PRO behaves in these constructions unlike pronouns (or pro. Consider for example, (78ii) or (78iii) embedded in a VP deletion context (which is not possible for (78iv)):

(82)   i.   John proposed to Bill to leave, and Harry did too
         ii. John described to Bill how to solve the problem and Mary
              did too

Neither of these sentences is ambiguous (in the relevant ways). For example, (82i) cannot mean "John proposed to Bill that they two should leave" and "Harry proposed to Bill that John and Bill should leave".

In summary, all these considerations cast serious doubts on an analysis of PRO in (78ii) and (78iii) (and by extension, in (78iv)) as

pure pronominals. If they were, they should be able to select their referent and "keep it constant across conjuncts".

The non ambiguity of (82) indicates that these NOC PRO's, unlike pronominals, are necessarily referentially dependent.

Concluding these remarks, we see that control constructions fall into two categories: OC and NOC. Control Theory for OC, we have argued, reduces in essentials to Binding Theory. Let us therefore reserve the term Control Theory for NOC constructions. We have also argued that PRO in NOC constructions was not pronominal. Its referential properties remain to be accounted for. One conclusion is clear however. They cannot be reduced to Binding Theory as the notion involved - referential dependence - is not binding, since binding does not allow for split antecedents. One possible way is to adapt Manzini's theory to our present hypothesis, so that it would apply only to NOC predicates.[53] We present a possible alternative below.

6. Caseless variables

We have noted in several occasions that all the classifications of non overt categories reviewed (Chomsky, 1981, 1982, our proposal in IV.2) suffered from a basic unexplained asymmetry with respect to local A'-binding. This asymmetry arises from the apparent truth of the following proposition:

(83)   Locally A'-bound elements must be Case-marked

If (83) is true, it seems to have to be stipulated one way or the other.

We have shown this to be the case for the systems of Chomsky (1982), our own (cf. 2.6. above, for a summary). Similarly, we have shown in Chapter II (section 2.6) that in a system like that of Chomsky (1981, Chapter 6), which attempts to reduce the Case Filter to θ-theory through the notion of visibility, the stipulation is displaced to: PRO is intrinsically visible, contrary to other categories.

In the rest of this section, we will explore what it would mean if (83) was false. We will argue that, if it is false, we can derive a number of properties and solve some problems that we have noted as we proceeded. Before doing so, recall that (83) is not the only statement that U.G. has to contain about locally A'-bound elements. We have also seen that the truth of the following proposition must be assumed:

(84)   Locally A'-bound elements are arguments

We have argued that (84), contrary to (83), derives from simple considerations about what it means to be a local A'-binder and what it means to be an argument (cf. II.2.6.1.2).

6.1. Must Variables have Case

Let us now investigate in more detail what it means to stipulate proposition (83), in the context of our assumptions concerning the inventory of non overt categories and their behaviour with respect to the Binding Principles.

Clearly, (83) is trivially true for overt elements. If they may be A'-bound, overt elements require Case. This is the content of the Case Filter, which, we have argued (II.2.6) is an autonomous principle. Let us

therefore turn to non overt elements. A non overt element e* can be either governed or not, and, if governed, can be Case-marked or not (since we assume Case requires government).

Suppose first e* is governed and Case-marked. This is the usual case for locally A'-bound category. For such cases, (83) is of course, unnecessary.

Suppose next that e* is ungoverned. Again, it must be Caseless, so that e*=PRO, but it does not have a Binding Domain. There is no Binding requirement imposed on it. Now recall that we have shown in III 8.2.4. that one could not assume that Caseless locally A'-bound elements were subject to what we then referred to as the ECP. This conclusion is consistent with the argument, made in IV.4.2. on independent grounds, that the scope of the ECP reduces to some _identification_ .function of Case-marked non overt categories (i.e. _pro_). So it appears that if such structures must be ruled out, they are by (83) and by (83) only.

Clearly, a theory not including (83) as underived statement is preferable to one that does, if dropping (83) has no negative effects. Similarly, a theory not including (83), but including some other stipulation of more explanatory force is superior to one including (83). We will argue that such a theory can be constructed. Assume we drop (83). The above discussion indicates that the following conclusion holds:

(85)   Locally A'-bound ungoverned elements (i.e. PRO) are
            permitted at any syntactic level of representation

Suppose e*=PRO is a locally A'-bound ungoverned, hence Caseless, element. What can its local A'-binder A* be? Let us restrict our attention to cases in which A'* does not locally bind anything else than e* (if it does, we get the PG structures discussed in III, 8.2.4., to which we

return in 6.2.5. below). Can A'* be overt? If it is, it must have Case, because of the Case Filter, and it cannot get Case through its being coindexed with e*, since e* is Caseless.

Could it get Case some other way. We will assume not, in the unmarked case by assuming that A'-positions are invisible to Case assignment (or Case checking).[54] From this it follows that A'* must be empty.

The possibility described in (85) is simply the consequence of rejecting the stipulation (83). In other words, it constitutes the null hypothesis. Of course, admitting this possibility will create new problems. For example, what are the properties of these non overt A'-binders? Is the A'*/e* relation subject to locality requirements. We will not explore the possible answers to these questions here. Rather, we will replace the stipulation (83) by another one, derived from (85) which, as we will see, will provide solutions to some problems we have noted along the way.

(85) allows ungoverned elements to be locally A'-bound. Where do we find ungoverned elements? Essentially in subject of infinitival position (also gerunds). So suppose we strengthen (85) to (86):

(86)   An ungoverned subject (i.e. a PRO) is always bound by an
          empty operator in the adjacent COMP

(or adjoined to S, in the case of gerunds).

It is clear that (82) is no longer the null hypothesis. It is also clear that it will create no new problems with respect to (85). Let us explore some consequences of (86).

6.2. Some consequences

6.2.1. Symmetry of ϒ-categories classification

The first obvious consequence of (86) is that the unexplained asymmetry of ϒ-categories with respect to A'-binding disappears. Any ϒ-category (which is an argument) can be locally A'-bound.

6.2.2. On deriving the CEP

Recall that we have argued that a principle had to be adopted, the CEP (67), to the effect that <u>expletive</u> <u>PRO</u> <u>had</u> <u>to</u> <u>be</u> <u>governed</u> (cf. IV.4.2.2.). The CEP subsumed both the ECP for NP-traces and derived the observation that expletives cannot appear as ungoverned subject of infinitivals. Adopting (86) permits us to <u>derive</u> the CEP.

Suppose indeed that some PRO is ungoverned. It must, by (85) be locally A'-bound. However, as we know by (83), locally A'-bound elements must be arguments, so that expletives can never be ungoverned. We now see one advantage of trading stipulation (83) for (86). We can eliminate the otherwise necessary CEP (67).

6.2.3. Obligatory Control

It is clear that we now have to slightly revise our analysis of OC constructions. Recall that we have argued in 5.2. above that OC predicates trigger S'-deletion in LF. This means that, at S-structure, the PRO appearing in an OC construction is in fact locally A'-bound by an

empty operator in the adjacent COMP. In other words, the respective structures of a form like (87i) are (87ii) and (87iii):

(87) i.   John attempted [to leave]
      ii. John$_i$ attempted [$_{s'}$ O$_i$ [$_s$ PRO$_i$ to leave]] (S-structure)

      iii. John$_i$ attempted [$_s$ PRO$_i$ to leave]   (LF)

It is quite clear however, that the argumentation given in 5.2. is unaffected by this modification: The argumentation depending on LF representations remains unchanged, since the LF representations of OC structures remain unchanged; The argumentation depending on S-structure representations rested on the CEP which we now see derives from (86) and thus remains valid as well.

6.2.4. Non obligatory control

6.2.4.1. By (86), we must now analyze all NOC PRO's as variables bound by an adjacent non overt operator both at S-structure and at LF. In 5.3. above, we have shown that Control theory did not involve binding. We further assumed that Control Theory was the theory dealing with the referential properties of NOC PRO's. We can now be more explicit. The reference of a PRO, or the values it may take will be determined by the range of the empty operator binding it. We derive the following characterization of Control Theory:

(88)   Control Theory:
           Control Theory is the theory of the range of non overt
           operators

A non overt operator is simply a non overt category in an A'-position. We thus see that non overt categories bifurcate in two sets:

i.   non overt categories in A-positions are subject to
     Binding Theory
ii. non overt categories in A'-positions are subject to
     Control Theory

6.2.4.2 What is the content of Control theory? To answer this question we must investigate the distribution and properties of non overt operators, a non trivial matter.[55] We will not pursue this question here. Rather, we will mention some qualitative observations. Basically, the idea is to require of non overt operators to have their range set within a certain syntactic domain containing them. To have their range set means to be coindexed with one or more phrases within that syntactic domain, call it the Control domain.

If a non overt operator has no Control domain, or if its control domain contains no possible antecedent, it gets assigned an arbitrary range. This corresponds to the arbitrary interpretation for PRO. Furthermore, under well-designed discourse conditions, it may also select an extra-sententially defined range (as, e.g. in (79)). These remarks must of course, be technically implemented so as to explain the observations made in 5.3. I know of no theory able to achieve this result.

6.2.4.3. Finally, note that this approach to NOC entails a different explanation of the following facts that that presented in II.2.6.

Recall that a distinction is necessary between argument' quasi-arguments and expletives (cf. II.2.6 and II.fn.22). Arguments can be questioned, arbitrary and controlled. Expletives can be none of them (and we have explained this). Quasi-arguments can be controlled, but neither questioned nor appear in a position where an argument gets an

arbitrary reading:

(89)   i. *It is unclear why e to rain
        ii. It rained after e having snowed for days

In particular, note that the ungrammaticality of (89i) would be surprising if e was analyzed as a pronominal. Why could it not pick the correct designated element in D (i.e. weather it) as extrasentential referent.

For us, e in (89) is bound by a non overt operator O. In (89i), this operator has no available antecedent (or no Control domain) so it is assigned an arbitrary range. It must therefore bind an element whose semantics allows to range. (89i) is therefore excluded. In (89ii), O can set its range by being coindexed with (weather-) it. (89ii) is well-formed.

In other words, the fact that both questioning and "non-arbitrariness" of quasi-arguments are ruled out receives a natural explanation here. Of course, these remarks imply that (84) must be modified as:

(90)   Locally A'-bound elements must be able to select some
        element in D

so as to allow, a priori, locally A'-bound quasi-arguments.

6.2.5. PG/WCO and PRO

Returning now to the examples discussed in III.8.2.4, we see that the adoption of (85), coupled with our theory of OC and NOC permits a simple resolution of the problem noted there. Recall that we discussed the following two structures:

(91) i. *who$_i$ did John fire e$_i$ without it seeming [e*$_i$ to have failed]

ii. *who$_i$ did John fire e$_i$ without Bill trying [e*i be replaced]

(92) who$_i$ does [e*$_i$ sleeping late] bother e$_i$

In both cases, we have a PG or WCO structure in which one of the two gaps is a PRO. Yet, one is excluded, the other is not. Two questions arise: why does (92) escape the Bijection Principle requirements? Why is (91) much worse than a Bijection Principle violation?

Consider now the analysis we would give of each example. In (81), e*$_i$ will be governed (both at S-structure and at LF in (91i), at LF only in (91ii)). It will have to be locally A-bound in the adverbial clause, but cannot (no available antecedent). We get a violation of the Binding Theory. This answers the second question above. In (92), e*$_i$ is ungoverned. The analysis of the subject constituent is in fact [O$_i$ [[e]*$_i$ sleeping late]] bothers No violation of the Bijection Principle occurs since who locally A'-binds e$_i$ only. In other words, ungoverned PRO never triggers WCO effects. (Note that the range of O$_i$ is determined by coindexation with e$_i$ as in [O$_i$ [e*$_i$ sleeping late]] bothers John$_i$.

6.2.6. Summary

We have traded stipulation (83) for stipulation (85). This was noted to raise some problems, although most of them (related to the syntax of non overt operators) are not specific to (85). However, (85) allowed us to:

i.   restore the symmetry of our classification of
     ʔ-categories
ii. derive the ECP for NP-traces
iii. derive the prohibition against expletive PRO's in
     ungoverned chain initial position
iv. explain why ungoverned PRO never trigger WCO effects
v.   explain the distribution of quasi-arguments

FOOTNOTES CHAPTER IV

1. But cf. Brody (1981), and Aoun (1982), for versions of the Binding Theory extending it to A'-binding.

2. Unless, of course, some specific rule introduces this reading as we would argue is the case in Gbadi (reanalyzing the proposals of Sportiche (1982) slightly).

3. I am ignoring here expletive elements and pronouns with bound variable interpretations (i.e. pronouns bound by quantified expressions).

4. Note that some obvious adjustment has to be made for 1st and 2nd person singular pronouns. Note also that expressions like "that male person" resemble pronouns -he- in their referential behaviour - but only for the aspect discussed here (cf. fn. 3)

5. It is in fact implicit throughout Chomsky 's work that there is a fundamental difference between the ways the binding properties of overt categories and those of non overt categories are determined. Binding properties for overt elements are stipulated. Binding properties for non overt elements are either derivative, or arbitrarily assigned. This seems to us to be an arbitrary decision which leaves unexplained why, as far as γ-categories are concerned, their behaviour can be entirely predicted whether overt or not, unexplained.

6. It is crucial that D and D' meet some property like identity if the distribution of PRO is to follow as desired in Chomsky's sytem. If D(x) contains D'(x), x could be A-free in D'(x) and A-bound in D(x). It would not follow that x obeying both Principle A and Principle B does not have a governing category. Huang (1982) reformulation overcomes this problem by having D(x) = D'(x) (although dD/dx≠0).

7. We will see when the table is completed that a visibility theory as the one reducing the Case Filter to the θ-criterion would be unfeasible here. We have shown in Chapter II that it failed also in principle. This does not preclude taking Case as a necessary feature for θ-role visibility, although the resulting redundancies are suspicious.

8. e.g. Longobardi (1978) *VV Filter. Cf., Jaeggli (1980a), Aoun & Lightfoot (1982) for some recent discussions.

9. As, in fact, Bouchard (1982) proposes. His conclusions, reached from a rather different point of view than ours would not be extensionally not identical to that of the text due to the distinction bound PRO ≠ obligatorily bound PRO cf. his chapter 5.

10. Obviously, there is another course of action, if one believes that "PRO is ungoverned" is a desirable theorem of grammar. It would consist in taking the feature [±A], [±B] as primitives, reformulate (6) in terms of them and derive the [±anaphoric], [±pronominal] properties somehow. The semantic criticisms against PRO as a pronominal anaphor would not extend to the existence of an element [+A,+B]. The discussion that follows in the text suggests that it would not be sufficient either.

11. There is some simplification here which does not affect the basic point: there is no ambiguity between PRO and PRO'.

12. Another possibility is to proceed as in footnote 10 above. Take [±A, ±B] as primitives instead of [±anaphoric], [±pronominal] and derive these last properties from the others. Notice however, that such a move would leave no explanation for the non existence of a [+A,+B] Caseless non argument (i.e. the equivalent of (t4) of (20)).

13. As Chomsky (1981; Chapter 6) first suggested.

14. An observation ((4iii) of the Appendix to section 6, Chapter III) that remains unexplained.

15. If the first hypothesis mentioned turned out to be correct, some principled reason would have to be found as to why INFL always absorbs Case. Otherwise, the gap that it would create in (t14) would remain mysterious.

16. In fact, here (but cf. Sportiche, forthcoming b) these sets are identical. NP-traces are non arguments by definition, Caseless by stipulation hence [-phon], and expletive PRO, we have remarked in Chapter II cannot appear in chain initial position, a fact requiring explanation (cf. 4.2. below).

17. Chomsky (1982) restricts this assumption to non overt elements.

18. Of course, if such an element exists, its distribution should be adequately constrained.

19. (10) as such is insufficient for we now call PRO a given feature bundle, regardless of the nature of its local binder. Presumably, local A'-binding should be brought into play.

20. There might be situations in which it would matter. Suppose x is some element, that is locally A'-bound and both A-bound and A'-bound in D(x):

$$[_{D(x)} \quad A_i \qquad A'_i \qquad x_i \ldots]$$

binding (over A_i ... x_i), local binding (from A'_i)[5]
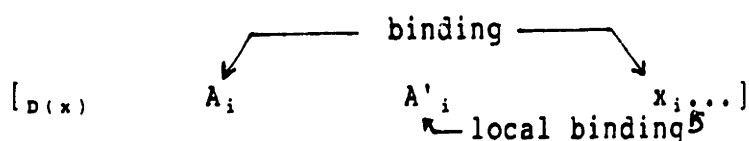
if locally A'-bound elements fall under (32ii), this should be impossible. Otherwise, it should be well-formed.

21. Note that this terminology might be misleading for a pronoun referentially dependent upon its antecedent is not technically bound. For example, it may be referentially dependent upon several antecedents at the same time (split pronominal binding). Binding does not permit that (cf. the discussion in Chapter III, section 7: in terms of that discussion, it is rather "anti-free")

22. Perhaps a better terminology than anaphoric and pronominal is anaphoric, i.e. necessarily referentially dependent and non anaphoric, i.e. non necessarily referentially dependent.

23. There is no footnote 23.

24. A careful examination of the argument, leading to this conclusion, given in Chapter II shows in fact a narrower property. Expletive PRO's cannot appear in ungoverned chain initial position. We will return to this in 4.2. below.

25. Notice incidentally that this entailment is not necessary. Some anaphoric elements must be bound but are not necessarily [+A].

26. Cf. also Mohanan (1982) for a discussion of some structures which have the same abstract structure as English gerunds (+ "Quirky Case").

27. Cf IV.3.1.2. where we argue Case assignment must be obligatory.

28. Note that Reuland takes INFL after Affix Hopping not to be a governor in order to avoid a governed PRO. We also need this assumption for different reasons. Cf. next subsection.

29. Alternatively, we could argue that the structures (39i) and (39ii) are distinct, despite appearances. One such suggestion is put forth in Bouchard (1982) who investigates the same questions for the same reasons. Essentially, Bouchard (1982) proposes that we take into account the nominal/verbal distinction between gerunds due to Warsow & Roeper (1972) Nominal gerunds have the internal structure of NP's. They assign genitive Case to their subject when they have one. Like other NP's they do not have to have a subject (I am ignoring some complications here. Some nominal gerunds might have verbal properties, e.g. assigning objective Case to their complements. Cf. Bouchard (1982) for details) Verbal gerunds behave like infinitival clauses, they do not assign Case to their obligatorily present subject. Adopting this dichotomy, we can adapt its consequences to our framework of assumptions. If a gerund is nominal and has no subject, there is no PRO involved. If a gerund is nominal and has an overt subject, we get structures like (39i). If it has a non overt subject, this non overt subject is pro, since it is in a Case-marked position. This, however, would be excluded since such a position is not identified (ECP). Finally, if a gerund is verbal, it must have a subject, which is Caseless by assumption, so it is PRO. A structure like (39ii) is therefore ambiguous. Either it is nominal and subjectless or it is verbal with a PRO subject. Note however, that the examples in (39) are perhaps misleading. The problem is why we

do not have "*the reading books" which, one would assume, should be possible with nominal gerunds. The matter is obscured by the acceptability of such examples as "all that studying French did not do me any good".

30. The reasons for these restrictions have become different, however, since PRO may appear as subject of infinitivals, it has been assumed that INFL of such clauses was not a governor, if PRO must be ungoverned. Our reason is different. It has to do with binding. Assuming every governed element has a governing category, if INFL of infinitivals was a governor, it would imply that all PRO's must be A-bound. If arbitrary PRO is a PRO in our sense, this would be false. Note also that it would disallow split antecedents for non arbitrary PRO's (cf. the discussion of bind, in III.7.)

31. Note that, as Chomsky (1982) does, if Case-marking requires adjacency, this conclusion is not obvious. In fact, it would force us to take the view that N triggers the insertion of a Case-marker under government (which does not require adjacency) but is not a Case-marker itself. Similar remarks would apply to A.

32. An additional argument for this conclusion comes from the ill-formedness of such structures as [Rome$_i$'s destruction of t$_i$] (which would mean Rome's destruction of itself) which must be somehow excluded - e.g. because $\underline{t}$ is Case-marked - where Case-marking has applied to both the subject and the object (or, also, only to the subject) of the head noun. Notice that this problem is independent of the obligatoriness of Case-marking.

33. Similar conclusions have been reached by several authors recently, sometimes from different points of view: Williams (1982); Higginbotham (1982a); Rappaport (1982). Obviously, as all these authors note, it follows immediately that there can be no Raising in NP's. However, it does not automatically exclude Tough-movement (say if the analysis of Chomsky (1981, Chapter V), is correct, Tough-movement involves no NP movement).

34. At least at S-structure. We will argue that it may be at LF. Let us rapidly examine a potential counterargument to our conclusions, put forth in Chomsky (1982). Consider the following paradigm:

    i.   The belief [PRO to be sick]
    ii. John's belief [PRO to be sick]
    iii. John's belief [t to be sick] by Mary
    iv. belief [NP* to be sick]

None of these structures is well-formed. Given that believe triggers S'-deletion in syntax, it might be tempting to try to explain the ill-formedness of (i) and (ii) with the following assumptions:

    i.   belief triggers S'-deletion
    ii. PRO must be ungoverned

since in both (i) and (ii), PRO would be governed. However, putting

aside the ungrammaticality of (iv) - arguably, <u>belief</u> might be taken as an intransitive noun - the ill-formedness of (ii) would remain unexplained. The case of NP-movement involved would be structurally quite parallel to admitted cases of NP-movement. There are two additional comments First, as OC structures show, verbs and nouns systematicaly differ in that an OC verb requires that an antecedent be syntactically expressed (*it has been attempted to leave), but the corresponding nominal does not (the attempt to leave). We provide in section 5 below some evidence that OC verbs trigger S' deletion precisely for that reason. Obviously, corresponding nominals do not. Secondly, in French although <u>croire</u> is not an S'-deletion predicate, the equivalent of (i) and (ii) above are ill-formed. Whatever the account for the French case is likely to extend to the English case.

35. I am ignoring some refinements in which PRO is governed but no domain D or D' exist, because no SUBJECT is accessible to it. Cf. Chomsky (1981, 219f).

36. I am modifying Huang's formulation slightly to avoid some minor technical problems. If c-command is included in the definition of accessibility, it would yield the wrong result for pronouns. D' (pronouns) could be computed by considering - wrongly - a non c-commanding SUBJECT.

37. Notice that by making D and D' category dependent instead of Principle A or B dependent, Huang avoids the problems mentioned in footnote 6, this chapter.

38. Huang considers the question of whether (58) is not a better formulation than (56). He argues not because:

> - it loses the PRO theorem - but this is irrelevant to us
> - it loses the collapses of NIC and SSC (but I do not see why)
> - it requires stipulating NP's,S...are binding domains. Here we agree. His notion of SUBJECT permits derivation of that and that small clauses, which presumably are neither S's nor NP's are also binding domains.

39. This formulation raises the question of whether the requirement that the binding domain of x contain a governor of x is necessary at all. We will see later (section 4, below) that it is.

40. But cf. Chomsky (1981, Chapter VI) and Stowell (1981, Chapter III), for discussion. Note that although the conclusion - coindexing or cosuperscripting of a pleonastic subject and a clause is compatible with our assumptions, some particular arguments leading to it (e.g. in terms of visibility) are not. Cf. also Safir (1982), Freidin & Harbert (1982) for some opposing views.

41. Notice that the facts are as predicted by the theory of bounding nodes in French. Notice further that the presence of wh_j in the intermediate COMP of (61) prevents analyzing the relation wh_i /t_i is a two step relation.

42. It is clear that lack of accessible SUBJECT is not a necessary and sufficient condition, but it might be a necessary condition.

43. Furthermore, if it is not true that NP-traces are necessarily expletive PRO's, the ECP could not be taken to apply to "traces" (Cf. Sportiche, forthcoming b, for discussion)

44. In Dutch, however, there is a difference between impersonal passives and weather verbs. Although they both take expletive subjects, weather verbs select het (equivalent of English it) while impersonal passives select er (equivalent of English there). Let us assume for concreteness some feature difference between these two expletives. Chomsky (1981, p. 87ff) makes a concrete proposal about this difference based on English. English there always cooccurs with some coindexed NP. Not so for Dutch, to which Chomsky's proposal thus cannot extend.

45. Safir (1982) comes to a very similar conclusion (although not termed as concerning expletive PRO (cf. his Chapter II, section 2.4.2, especially (57)).

46. These observations are due to Williams (1980). Note however, that unlike Williams, we do not require that OC positions, unlike NOC positions, tolerate PRO but no lexical NP's.

47. Notice, incidentally, that Manzini's theory, as she formulates it, fails for nominals, as it predicts that forms such as: Yesterday's attempt PRO to leave are either ill-formed (Yesterday having to bind PRO) or are predicted wrongly to contain a freely referring PRO.

48. Note that the distinction could not be argued to be the verbal/adjectival passive of Williams (1981) (with extraposition in the case of verbal passive only) viz: (i) It has been recommended to Bill by John [to shave oneself in such circumstances] which must be a verbal passive because of the by-Phrase thus not extraposed, but allows arbitrary PRO.

49. As is well known, allowed occurrence of the complement S in subject position as a test for extraposition would fail, viz: il faut [PRO partir] in which PRO "freely" refers, thus seems to fall under (69ii) despite * [PRO partir] faut.

50. Note that if S'-deletion requires a phonologically null COMP, the following examples:

> i. Jean a essayé de partir (OC)
> ii. Le probleme mérite d'etre étudié (Raising)

show either that S'-deletion is not the appropriate device, or that de is not a complementizer in French.

51. This argument cannot be made in Dutch. As pointed out in footnote 44 above, Dutch, unlike German (or French) takes different expletive subjects in impersonal passives (er) and weather verbs (het). *het

werd geprobee:d te regenen (it has been tried to rain) is arguable related to the fact that the subject of the passive must be het to bind that of regenen, violating the requirement of impersonal passives. Similarly for "er werd geprobeerd te regenen in which the subject of regenen is governed at LF but lacks an acceptable binder.

52. (78ii) is due to Koster & May (1980). For some English speakers, it is awkward. The same argument can be as easily made in French (proposer,...).

53. Essentially, we would have to modify (68) sc that it require of PRO to have its reference fixed (instead of being bound) in DGC.

54. Several authors have proposed Case-marking in COMP, in contradiction of the assumption of the text, e.g. Borer (1981a); Groos & Van Riemsdijk (1979); Kayne (1981a); Pollock (1982). We might perhaps accept these analyses as marked options if they are compatible with the necessity for e* to be ungoverned (that is, only Kayne, op. cit).

55. Cf. Chomsky (1982, p. 70-71) and Levin (1983) for some discussion.

REFERENCES

Anderson,S. and P.Kiparsky (1973), A Festschrift for Morris Halle, Holt, Rinehart and Winston, New York.

Aoun,J. (1980), "ECP, Move x and Subjacency", in Linguistic Inquiry 12.4.

Aoun, J. (1979), "On Government, Case marking and Clitic placement" mimeographed, MIT

Aoun, J. (1982), The Formal Nature of Anaphoric Relations, MIT PhD dissertation.

Aoun, J. (1983), "Logical Forms", in Linguistic Inquiry 14,2.

Aoun, J and N. Hornstein (1982), "Binding and Quantifier Scope" paper presented at NELS, UQAM, Montréal.

Aoun, J., N.Hornstein, & D.Sportiche (1981), "Some aspects of Wide Scope Quantification" in Journal of Linguistic Research 1.3.

Aoun, J. and D.Lightfoot (1982), "Government and Contraction", to appear in Linguistic Inquiry.

Aoun, J. and D.Sportiche (1981) "On the Formal Theory of Government", to appear in The Linguistic Review

Baltin, M. (1978), Towards a Theory of Movement Rules, MIT PhD dissertation.

Baltin, M. (1982), "A Landing site Theory of Movement Rules", in Linguistic Inquiry 13.1.

Belletti, A. (1980), "Morphological Passive and Pro Drop", manuscript, MIT and Scuola Normale Superiore.

Belletti, A., L.Brandi and L.Rizzi, eds, (1981) A Theory of Markedness in Generative Grammar, Proceedings of the 1979 GLOW conference, Pisa, Scuola Normale Superiore.

Belletti, A., & L.Rizzi (1981), "The Syntax of 'ne': Some Theoretical Implications" in The Linguistic Review 1.2.

Bok-Bennema, R. (1981), "Clitic and Binding in Spanish", in May & Koster eds.

Borer, H. (1981), Parametric Variations in Clitic Constructions, unpublished MIT PhD dissertation.

Borer, H. (1981a), "On the Definition of Variables", in Journal of Linguistic Research 1.3.

Bouchard, D. (1982), On the Content of Empty Categories, MIT PhD dissertation.

Bresnan, J. (1982), "Control and Complementation" in Linguistic Inquiry 13.3.

Bresnan, J. (1982a), The Mental Representation of Grammatical Relations, MIT Press.

Bresnan,J & J.Grimshaw (1978), "The syntax of Free Relatives in English", in Linguistic Inquiry 9.3.

Brody, M (1981), "Deriving ECP from the θ-theory", manuscript MIT, paper presented at NELS XII, MIT.

Brody, M. (1983), "On Categorial Definitions and the Theory of Chains", manuscript, University of London.

Burzio, L.(1981), Intransitive verbs and Italian auxiliaries, MIT PhD dissertation.

Chomsky, N. (1955), The Logical Structure of Linguistic Theory

mimeographed (Harvard); Plenum, 1975 (in part).

Chomsky, N. (1965), Aspects of the Theory of Syntax, MIT Press.

Chomsky, N. (1970), "Remarks on Nominalizations", in Chomsky (1972).

Chomsky, N. (1972), Studies on Semantics in Generative Grammar, Mouton, The Hague.

Chomsky, N. (1973), "Conditions on Transformations" in S.Anderson and P.Kiparsky (eds)

Chomsky, N. (1975), Reflections on Language, Pantheon.

Chomsky, N. (1976), "Conditions on Rules of Grammar", in Linguistic Analysis 2.

Chomsky, N. (1980), "On Binding", in Linguistic Inquiry 11.1.

Chomsky, N. (1980a), Rules and Representations, Columbia.

Chomsky, N. (1981), Lectures on Government and Binding: The Pisa Lectures, Foris Publications Dordrecht.

Chomsky, N. (1982), Some Concepts and Consequences of the Theory of Government and Binding. MIT press.

Chomsky, N. & H.Lasnik (1977), "Filters and Control", in Linguistic Inquiry 8.3.

Cinque, G. (1977), "The Movement Nature of Left Dislocation", in Linguistic Inquiry 8.2.

Evans, G. (1980), "Pronouns", in Linguistic Inquiry 11.2.

Freidin, R. (1978), "Cyclicity and the Theory of Grammar", in Linguistic Inquiry 9.4.

Freidin, R. and W. Harbert (1982), "On the Fine Structure of the Binding Theory", in the proceedings of NELS XIII, Amherst.

Fourier, Z. de (1980), "Remarks on the Extended Linear Model", manuscript, MIT.

Groos, A. & H.van Riemsdijk (1981), "Matching Effects in Free Relatives", in Belletti et al.

Halk I. (1981), "Indirect Binding", paper presented at GLOW 6, to appear in Linguistic Inquiry.

Halk I. (1981a), "On Clitic en in French", manuscript, MIT.

Harlow, S. (1981), "Government and Relativisation in Celtic", in F. Heny (ed).

Heny, F. (ed) (1981), Binding and Filtering, MIT Press, Cambridge, Mass.

Higginbotham, J. (1980), "Pronouns and Bound Variables", in Linguistic Inquiry 11.

Higginbotham, J. (1981), "Anaphora and GB", in Cahiers linguistiques d'Ottawa, 9.4, Université d'Ottawa.

Higginbotham, J. (1982), "The Logic of Perceptual reports: an Extentional Alternative to Situation Semantics"

Higginbotham, J. (1982a), "LF, Binding and Nominals", in Linguistic Inquiry 14.3.

Hornstein, N. & D. Lightfoot (1981), Explanation in Linguistics, Longman.

Hornstein, N. and A. Weinberg (1981), "Case Theory and Preposition Stranding", in Linguistic Inquiry 12.1.

Huang, C.-T.J (1982) Logical relations in Chinese and the Theory of Grammar", unpublished MIT PdD dissertation.

Huang, C.-T.J (1982a), "ECP and Preposition Stranding", manuscript, MIT.

Huang, C.-T.J (1982b), "A Note on the Binding Theory", in Linguistic Inquiry 14.3.

Huang, C.-T.J (1982c), "Move wh in a Language without Wh-movement", in The Linguistic Review 1.4.

Huybregts, R. (1979), "Clitic Binding as Variable Binding" paper presented

at MIT.

Ingria, B. (1981), Greek Complementation, unpublished MIT dissertation.

Jackendoff, R. (1972), Semantic Interpretation in Generative Grammar, MIT Press.

Jackendoff, R. (1977), X-bar Syntax, MIT Press.

Jaeggli, O. (1978), "On Abstract Case", manuscript MIT.

Jaeggli, O .(1980), On Some Phonologicallly-null Elements in Syntax, MIT PhD dissertation.

Jaeggli, O. (1980a), "Remarks on to -Contraction", in Linguistic Inquiry 11.1.

Katz, J. and P. Postal (1964), An integrated Theory of Linguistic Description, MIT Press

Kaye, J., H. Koopman, D.Sportiche and A.Dugas (eds), (1983), Current Approaches to African Linguistics vol 2, Foris Publications, Dordrecht.

Kayne, R. (1975), French Syntax: The Transformational Cycle MIT Press, Cambridge, Mass.

Kayne, R. (1978), "Extensions of Binding and Case marking", in Linguistic Inquiry 11.1.

Kayne, R. (1979), "Two notes on the NIC", in A.Belletti et al.

Kayne, R. (1981), "ECP extensions" in Linguistic Inquiry 12.1.

Kayne, R. (1981a), "On certain Differences between French and English", in Linguistic Inquiry 12.3.

Kayne, R. (1981b), "Unambiguous Paths", in May and Koster, eds.

Kayne, R. (1982), "Complex Inversion Chains in French", in Wiener Linguistische Gazette 27/28.

Kayne R. (1983), "Connectedness", in Linguistic Inquiry 14.2.

Kayne, R. & J.-Y Pollock (1978), "Stylistic Inversion, Successive Cyclicity and Move NP in French", in Linguistic Inquiry 9.4.

Kenstowicz, M. and W. Wahba (to appear), "Wh-in-situ in Egyptian Arabic", in Kaye et al. eds.

Koopman, H. (1980) "Subject Object Asymmetries in Vata", in Publications de l'ILA, Université d'Abidjan.

Koopman, H. (1982), "Control from COMP and Comparative Syntax", to appear in The Linguistic Review.

Koopman H. (1982a), "Les constructions relatives" in C.Lefebvre, H.Magloire-Holly, N.Piou (eds) La syntaxe de l'haïtien, Karoma.

Koopman, H. (1983), The Syntax of Verbs, to appear by Foris Publications.

Koopman, H. & D.Sportiche (1981), "Variables and the Bijection Principle", to appear in The Linguistic Review.

Koopman, H. & D.Sportiche (1981a), "Binding and Coreference of pronouns in Vata", manuscript UQAM, presented at CLA 1982, Ottawa.

Koster, J. (1981), "Binding and Control", mimeographed, University of Tilburg.

Koster, J. and R. May (1980), "On the Sentential Status of Infinitives and Gerunds", paper presented at GLOW 1980, Nijmegen.

Lasnik, H (1976), "Remarks on Coreference", in Linguistic Analysis 2.1.

Lasnik, H. (1981), "On two Recent Treatments of Disjoint Reference", in Journal of Linguistic Research 1.4.

Levin, J. (1983), "Government Relations and the Structure of INFL", unpublished manuscript, MIT.

Longobardi, G. (1978), "Double-Inf", in Revista di Grammatica generativa 3.2.

Manzini, M.-R. (1983), "On Control and Control Theory", in Linguistic

Inquiry 14.3.

Manzini, M.-R. (1983a), MIT PhD dissertation.

May, R. (1977), The _Grammar of Quantification_", MIT PhD dissertation.

May, R. (1981), "Movement and Binding", in _Linguistic Inquiry_ 12.2.

May R. and Koster J. (1981), _Levels of Syntactic Representations_, Foris Publications, Dordrecht.

Mc Closkey J. (1983), "On the Binding of Resumptive Pronouns in Modern Irish", paper presented at GLOW VII, York.

Milner, J.-C (1978), _De la Syntaxe à l'Interprétation_, Le Seuil, Paris

Mohanan, K.P (1982), "Infinitival Subjects, Government and Abstract Case", in _Linguistic Inquiry_ 13.2.

Neidle, C. (1982), The _Role of Case in Russian Syntax_, unpublished MIT PhD dissertation.

Pesetsky, D. (1982), _Paths and Categories_, unpublished MIT PhD dissertation.

Pollock, J.-Y (1979), "Réanalyse et constructions impersonelles", in _Recherches Linguistiques_ 8, Paris VIII.

Pollock, J.-Y. (1981), "On Case and Impersonal Constructions", in May & Koster (eds).

Rappaport, M. (1982), "Further Remarks on Nominals", manuscript, MIT.

Reinhart, T. (1976), The _Syntactic Domain of Anaphora_, unpublished MIT dissertation.

Reinhart, T. (1980), "Coreference and Bound Anaphora", manuscript, Tel Aviv University.

Reuland, E. (1983), "Governing -ing", in _Linguistic Inquiry_ 14.1.

Riemsdijk, H.van (1978), _A Case Study in Syntactic Markedness_, Foris Publications, Dordrecht

Riemsdijk, H.van & E. Williams (1981), "NP-structure" in The _Linguistic Review_ 1.3.

Rivas, A. (1977) _A Theory of Clitics_, unpublished MIT PhD Dissertation.

Rizzi, L. (1982), _Issues in Italian Syntax_, Foris Publications, Dordrecht.

Rizzi, L. (1983), "On Chain Formation", paper presented at GLOW VII, York.

Ross, R. (1967), "Constraints on Variables in Syntax", MIT PhD dissertation.

Rouveret, A. and J.-R.Vergnaud (1980), "Specifying Reference to the Subject", in _Linguistic Inquiry_ 11.1.

Safir, K. (1982), _Syntactic Chains and the Definiteness Effect_, unpublished MIT PhD dissertation.

Safir, K. (1982a), "Inflection-Government and Inversion", in The _Linguistic Review_ 1.4.

Safir, K. & D. Pesetsky (1981), "Inflection, Innversion and Subject Clitics", in _Nels XI_.

Sag, Y (1976), "Deletion and Logical Form", unpublished MIT PhD dissertation.

Schein, B (1981), "Small Clauses and Predication", unpublished manuscript, MIT.

Simpson, J. (1982), _Control and Predication in Warlpiri_, unpublished MIT PhD dissertation.

Sportiche, D. (1981), "Bounding nodes in French", in The _Linguistic Review_ 1.2.

Sportiche, D. (1982), "Bete Reciprocals and Clitic Binding", to appear in Kaye et al. (eds).

Sportiche, D. (1982a), "Speculations on the Binding of PRO", paper presented at NELS XIII, UQAM, Montréal.

Sportiche, D. (forth.a), "Successive Cyclicity in Dida", to appear in Rapport de Recherche sur les langues Kru II.

Sportiche, D. (forth.b), "Vata Raising and Case marked Traces", to appear in Rapport de Recherche sur les langues Kru II.

Sportiche, D. (in preparation)

Stowell, T. (1981), Origins of Phrase Structure, unpublished MIT dissertation.

Taraldsen, T. (1981), "On the Theoretical Interpretation of a Class of Marked Exceptions", in Belletti et al eds.

Torrego, E. (1981), "On Inversion in Spanish and Some of its Effects", unpublished paper, University of Massachusets, Boston.

Vergnaud, J.-R (1974), French Relative clauses, MIT PhD dissertation.

Vergnaud, J.-R.(1982), Dépendances et niveaux de représentation en syntaxe, thèse de doctorat d'Etat, Université de Paris VII.

Wasow, T. (1972), Anaphoric Relations in English, unpublished MIT dissertation.

Wasow, T.& T.Roeper, (1972), "On the Subject of gerunds", Foundations of Language 8.1.

Williams, E. (1975), "Small Clauses in English", in J. Kimball ed. Syntax and Semantics vol 4, Academic Press.

Williams, E. (1977), "Discourse and Logical Form", in Linguistic Inquiry 8.

Williams, E. (1980), "Predication", in Linguistic Inquiry 11.1.

Williams, E. (1981), "Argument Structure and Morphology", in The Linguistic Review 1.2.

Williams, E. (1982), "The NP-cycle", in Linguistic Inquiry 13.2.

Zaenen A., E. Engdahl, and J. Maling (1981), "Resumptive pronouns can be syntactically bound", in Linguistic Inquiry 12.4.

Zubizarreta, M.-L (1982), On the Relationship of the Lexicon to Syntax, unpublished MIT PhD dissertation.