

**Jaume Solà i Pujols**

**Agreement and Subjects**

**directora: Gemma Rigau i Oliver**

**Departament de Filologia Catalana  
Facultat de Lletres  
UNIVERSITAT AUTONOMA DE BARCELONA**

**7 de setembro de 1992**

## List of errata

### 1. Most significant errata:

- p. 33, 4th line: ...a Null case PRO-infinitives-> a kind of weak Nominative Case in PRO-infinitives.
- p. 103, ex. 34: (, que plogui) -> (, que plou)
- p. 132, 1st line: co-reference -> co-denotation
- p. 132, par. before (69), last line: Greek) -> Greek and German)
- p. 183, 2nd par., 7th line: is at odds -> is not at odds
- p. 209, last line: (24), (40) -> (24)
- p. 210, par. after (25), 1st line: (25), (41) -> (25)
- p. 226, 2nd line: (24), (40) and (25), (41). -> (24) and (25) (repeated here as (40) and (41)).
- p. 234, 2nd par., last-but-1 line: very unnatural -> very natural
- p. 255, note 39, 2nd line: are comparable to ECM verbs like -> do not behave like complements of ECM verbs such as
- p. 270, ex. (6).b): Una cotxe -> Un cotxe  
gloss and translation: snake -> car
- p. 274, 2nd par., last-but-1 line: 'Everybody likes... -> 'Everybody loves...
- p. 296, 6th line: not host of -> not hosted by
- p. p. 297, 16th line: interrogative Wh-phrases are sentence initial the... -> since interrogative Wh-phrases are sentence initial, the...
- p. 297, last-but-1 line: mutual distributional complementarity -> complementary distribution
- p. 304, section 1.4., lines 2, 3, 5, 9 and 11: host(s) -> guest(s)
- p. 305, par. d):
  - 1st line: non-NSLs use Spec of AGR (i.e., Spec of the FC containing AGR) as the AGR-identifier. -> non-NSLs use option (34).a).
  - 4th line: to be quantifier -> to be expressible as a quantifier
  - last two lines: so that the AGR-identifier can be a quantifier of any type. -> so that the Specifier of (the host of) AGR can be a quantifier of any type.

### 2. Typos and errata not impairing interpretation:

- p. 21, 2nd line: what we he terms -> what he terms
- p. 24, penult. line: very thin -> very feeble
- p. 25, 2nd line: pragmatiscal -> pragmatic
- p. 25, 1st line after ex. (24): poin -> point
- p. 45, gloss in ex. (21).b): redundancy -> unemployment
- p. 45, translation in ex. (21).b): redundant -> unemployed
- p. 48, last but 2 line: such a kind of -> such an
- p. 85, 1st line: coindexed the -> coindexed with the
- p. 94, end of big paragraeph: otherwise would be nothing but a theoretical artifact used to our convenience -> otherwise they would be nothing but a theoretical artifact used for our

convenience

- p. 94, next par., 1st line: The solution we propose to this puzzle -> The solution we propose for this puzzle

- p. 94, same paragraph, last word: distinctions -> predictions
- p. 96, par. after ex. (19), 5th line: in an relation of subset -> in a subset relation
- p. 106, 1st par., after (38), last line: conceived. -> conceived of.
- p. 114, last par. last but 2 line: A-movement or (... -> A-movement (...)
- p. 117, note 21, 3rd line: agreement-less -> agreementless
- p. 118, 1st par., 2nd line: nun-null-subject -> non-null-subject
- p. 119, last-but-2 par., last line: that PRO also has Case -> that also PRO has Case
- p. 119, last-but-1 line: let us take as -> let us take it as
- p. 129, ex. (65).a) & b), gloss: you-were -> you-are
- p. 133, 2nd par., 7th line: of unusual -> of the unusual
- p. 140, ex. (78), translation: Being tired went... -> Being tired, Gianni...
- p. 154, last-but-1 par., 2nd line: a dislocated -> dislocated
- p. 171, note 53, last line: expeltive -> expletive
- p. 186, last line: concerned (for instance -> concerned. For instance
- p. 213, last-but-1 line: English case construction -> English construction
- p. 222, 2nd par., 3rd line: case French -> case of French
- p. 223, central par., 9th line: enclisis because (... -> enclisis (...)
- p. 263, 11th line: gives reasonable -> provides a reasonable
- p. 267, 4th line: So in a -> So a
- p. 267, 1st line after ex. (2): would ambiguous -> would be ambiguous
- p. 268, 1st line after ex.(3): have an only optional pause -> involve a pause only optionally
- p. 275, last par., Negative quantifiers and cannot -> Negative quantifiers cannot
- p. 297, 13th line: basis two -> basis of two
- p. 309, 1st par., last-but-1 line: exclusive of -> restricted to

## Table of contents

Acknowledgements.....	ix
List of Abbreviations and Glosses.....	x
Chapter 1 - <b>Introduction</b> .....	1
1. Aims and Scope of this Thesis.....	3
2. Some Basic Premises.....	6
2.1. The Internal Subject Hypothesis.....	6
2.2. Subject Inversion.....	7
2.2.1. Transitive/Unergative Constructions.....	11
2.2.2. Unaccusative/Passive Constructions.....	15
2.2.3. Copulative Constructions.....	19
2.2.4. On the Availability of Subject Inversion.....	23
2.2.5. Floating Quantifiers.....	26
2.3. The Projection of Arguments.....	28
2.4. The Split INFL Hypothesis and Verb Movement.....	30
2.5. Case Theory.....	33
Chapter 2 - <b>Burzio's Generalization and EPP</b> .....	35
1. Introduction.....	35
2. Possible Solutions Based on CHAIN Theory.....	39
2.1. The Extended Projection Principle.....	39
2.2. CHAINS.....	40
2.3. Conditions on Expletive CHAINS.....	43
2.4. Problems.....	48
2.4.1. Empirical Problems.....	49
2.4.2. Theoretical Problems.....	51

3. Binding Theory.....	52
3.1. Binding Theory and Expletives.....	52
3.2. A Binding Theoretical Approach to the Nature of Inverted Subjects.....	55
4. Summary.....	78
 Chapter 3 - <b>AGR and Subjects</b> .....	81
1. On the Nature of the AGR-Subject Dependency.....	81
1.1. A Reformulation of the EPP.....	81
1.2. Two Parameters.....	87
1.2.1. Impersonal Constructions.....	87
1.2.2. Ergative Languages.....	89
1.3. A Problematic Case.....	92
1.4. The EPP and the ECP.....	107
2. AGR-Identifiers and Binding Theory.....	116
2.1. The Notion of AGR-identifier.....	116
2.2. Case Theory.....	118
2.3. Binding Theory.....	124
2.3.1. A Definition of Binding Domain.....	124
2.3.2. Anaphoric Copulative Constructions.....	126
2.3.3. On the Status of Binding Theory.....	137
2.4. The Position of pro and Floating Quantifiers.....	140
2.5. On the Nature of AGR Requirements.....	150
2.6. Inverted Copulative Constructions.....	154
3. Indefinite I-subjects.....	156
3.1. Partitive Case.....	157
3.2. Indefinite DPs and phi-features.....	162

4. CP I-subjects and Small Clauses.....	166
5. French Stylistic Inversion and Northern Italian Dialects.....	173
6. A-dependencies and Minimality.....	181
7. Summary.....	187
 Chapter 4 - <b>Infinitival Constructions</b> .....	189
1. Some Properties of Infinitival Constructions.....	190
1.1. Parallelism between Finite and Infinitival Constructions.....	191
1.2. V-movement in Infinitives.....	194
1.3. Kayne's Proposal.....	196
2. The AGR-identifier of Infinitives.....	212
2.1. Control.....	213
2.1.1. Enclisis.....	217
2.1.2. The PRO Theorem.....	225
2.2. Raising.....	231
2.3. ECM and 'for' Infinitives.....	241
2.4. Infinitives with an Overt Subject.....	244
2.4.1. IOS with an Inverted Subject.....	248
2.4.2. IOS with a Subject in Specifier of AGR.....	257
3. Summary.....	263
 Chapter 5 - <b>Preverbal Subjects in NSLs</b> .....	265
1. Finite sentences.....	265
1.1. Dislocation vs. Specifier (of AGR).....	266
1.2. On the Nature of Preverbal Specifier in NSLs.....	291

1.3. Interrogative WH-movement.....	296
1.4. A parameter for the FC hierarchy.....	304
2. Infinitives.....	306
3. Summary.....	309
References.....	310



I wish to express my gratitude to the following people (lists are in alphabetical order -first names count):

Gemma Rigau, who as the director has always supported and encouraged me even when I felt like throwing my thesis into the basket of hopeless linguistic vocations.

Joana Rosselló, who first aroused my interest for generative linguistics and has always been open to discussion and help.

Linguists from the Grup de Gramàtica Teòrica in the Universitat Autònoma de Barcelona, who have offered comments, data, information and support in seminars and corridors: Albert Branchadell, Anna Bartra, Carme Picallo, Enric Vallduví, Francesc Roca, Jeff Runner, Joan Mascaró, Josep Maria Brucart, Josep Quer, Lluïsa Hernanz, Mireia Llinàs, Núria Martí.

Linguists from the Un. of Geneva, where I wrote a pre-version of this thesis in an excellent atmosphere of discussion: Anna Cardinaletti, Cristina Figueiredo, Ian Roberts, Luigi Rizzi, Robin Clark, Tali Siloni, Teresa Guasti.

Linguists who had the patience to answer letters, E-mail messages or rush annoying questions: Alain Rouveret, Anne Zribi-Hertz, Elisabet Nebot, Giovanni Albertocchi, Giuseppe Longobardi, Hagit Borer, Háldor A. Sigurdhsson, Isabel Herrero, Joana Louro, Kjel Åke Gunnarson, Lea Nash-Haran, Liliane Haegeman, Patrick Sauzet, Sabine Iatridou, Sten Vikner, Tali Siloni, Uri Shlonsky, Virginia Motapayane.

Friends I tortured by abusing them as subjects with unbelievable questions about what they never would have thought was possibly their language: Luigi Giuliani, Michael Laurence, Pilar Pedraza.

The following are uncommon abbreviations and glosses we will use for the sake of convenience. In the text most of these abbreviations are first introduced in full form, but they might be hard to identify in consulting or reading through:

#### Abbreviations

BD	<b>B</b> inding <b>D</b> omain
BG	<b>B</b> urzio's <b>G</b> eneralization
BT	<b>B</b> inding <b>T</b> heory
CLLD	<b>C</b> litic <b>L</b> eft <b>D</b> islocation
DP(s)	<b>D</b> eterminer <b>P</b> hrase(s)
EA(s)	<b>E</b> xternal <b>A</b> rgument(s)
EL(s)	<b>E</b> ragtive <b>L</b> anguage(s)
EPP	<b>E</b> xtended <b>P</b> rojection <b>P</b> rinciple
FC(s)	<b>F</b> unctional <b>C</b> ategory/ies
IOS(s)	<b>I</b> nfinite(s) with an <b>O</b> vert <b>S</b> ubject
IOS-INV(s)	<b>I</b> OS(s) with an <b>I</b> nverted subject
ISH	<b>I</b> nternal <b>S</b> ubject <b>H</b> ypothesis
non-NSL(s)	<b>non</b> <b>N</b> ull <b>S</b> ubject <b>L</b> anguage(s)
NSL(s)	<b>N</b> ull <b>S</b> ubject <b>L</b> anguage(s)

#### Glosses:

SE	Reflexive particle not being a clitic (at least not a clitic of the Romance type) (e.g., Italian <b>se</b> , German <b>sich</b> , etc.).
SELF	A type of crosslinguistically recurrent of word which, among other uses, is typically a member of reflexive DPs (German <b>selbst</b> , Italian <b>stesso</b> , Catalan <b>mateix</b> , etc.).
SI-	Any Romance reflexive clitic (which has a wider range of use than mere reflexivity).

## Chapter 1

### Introduction

The present thesis is conceived within the Principles and Parameters framework.<sup>1</sup> It is an attempt to explore an alternative hypothesis to some standard assumptions within this framework, basically concerning sentence subjecthood and the Null Subject phenomenon.

Since a good deal of hypotheses in recent years are based on assumptions we challenge here (especially concerning null expletives in subject inversion structures), and since the consequences of the alternative view we will propose are far-reaching, I cannot by any means commit myself to provide an adequate answer to many of the questions that could naturally be posed to the proposals I will advance. Ideally, a new hypothesis has to cover all the relevant data covered by previous hypotheses and possibly some more, and has to face less problems. But it is also legitimate to draw back to challenge some basic premises, even if, by doing so, some data that previous theories could handle fall now out of the predictions. This is what happened, for instance, in generative linguistics when specific rules were abandoned in favor more principled accounts during the 1970's.

So, for instance, the present thesis has little to say about Nominative assignment in contexts of INFL-to-COMP.<sup>2</sup> My opinion is that ideally the theory should say little about it, because INFL-to-COMP should be orthogonal to Nominative assignment. But actually some facts (especially concerning the V-2 phenomenon) seem to suggest that subjects have a special

---

<sup>1</sup> See Chomsky (1986-b) for a review of some essential contentions specific to the framework, although the spirit of the framework can be traced back to the early 1980's (Chomsky (1981)).

<sup>2</sup> This has been a fruitful topic of debate in recent proposals by Rizzi & Roberts (1989), Rizzi (1991-a), and Roberts (1991-a).

behavior in INFL-to-COMP contexts.<sup>3</sup> The essential proposal in this thesis is silent (and to some extent neutral) about these facts.

We will introduce some descriptive generalizations that to my knowledge had not been considered and even less accounted for thus far. And, in addition, we will provide a means of deriving Burzio's Generalization, which, as we will see, can hardly be obtained under standard hypotheses. I think, therefore, that the present thesis is more than an ingenious variant of standard theories.

The presentation and style we will use are exploratory: although some of the ramifications in the argumentation will lead us to more specific assumptions, we will often go back to the initial, more general formulation for expository purposes, especially when the ramification is intended to cover limited and language specific data. On the other hand, some of the proposals are clearly subsidiary to the main hypothesis: the point at stake is often not whether they are the best of hypotheses, but rather whether the main hypothesis can be extended to cover a certain theoretical field in a reasonable and even plausible way.

Finally, I ask the reader to be patient: some crucial proposals cannot work without each other but obviously have to be introduced one after the other in due time. I will anyway try to point out, as I proceed, where any momentary potential problem or apparent inconsistency will be properly addressed.

---

<sup>3</sup> See Adams (1987) for evidence that Old French is a V-2 language having null subjects only in V-2 contexts (i.e., contexts of verb movement to COMP). In Germanic V-2 languages, on the other hand, subjects in Spec of COMP seem to have a different behavior from other constituents in that position (see, e.g., Cardinaletti (1990)).

## 1. Aims and Scope of this Thesis

The aim of the present thesis is to explore and work out an alternative to standard theories in the GB framework on the nature of AGR and subjects, as a basis for an account of the standard cluster of properties which hold for most Null Subject Languages (NSLs):

- null subjects
- subject inversion
- absence of that-t effects

Although these correlations may not be universal, I think they are not just a matter of coincidence, even if they hold only of a certain subset of languages. Our account of the facts essentially consists in a reinterpretation of the theory by which: a) all languages have subject inversion in an abstract sense: non-NSLs (which have been traditionally assumed not to display such a possibility) have anaphoric inverted subjects, whereas NSLs have [-anaphoric] I-subjects; b) all languages have null internal subjects (I-subjects), the contrast being between the ones having anaphoric null I-subjects (i.e., non-NSLs) and the ones having [-anaphoric] null I-subjects (i.e., NSLs).

We will also provide an account for another fact that holds for most Romance NSLs, namely infinitival long head movement.

In addition, the theory is conceived as a way of deriving a classical descriptive generalization, namely Burzio's Generalization (BG). The two facts correlating under this generalization (presence of an external Argument and availability of Accusative Case) are not easily derivable in a straightforward way from a single syntactic premise, since they look very different in nature. In the present account, BG is reduced to a specific version of the Extended Projection Principle, ultimately derived from the licensing of the AGR morphology in a given language.

In the remainder of this chapter, I will set out some theoretical premises that I will

crucially rely on or refer to in the following chapters: the Internal Subject Hypothesis, Subject Inversion, the Projection Principle, the Split INFL Hypothesis, and Case Theory.

In Chapter 2, I develop a possible account of BG based on standard notions of subject inversion and expletives. Thereafter I argue, on the basis of Binding Theory considerations, that this account is not appropriate, in that it misses some generalizations on subject inversion and expletives. Finally, I present a binding-theoretical approach to the nature of inverted subjects which will provide the basic facts to be explained in the following chapters. Throughout the chapters, the contrast between NSLs and non-NSLs languages will be at stake.

In Chapter 3, I develop the central proposals, which are essentially two: one on the way the dependency between AGR and the sentence subject is established (from which BG can be derived), the other on the way AGR identifies its subject across languages (from which the cluster of properties of NSLs vs. non-NSLs languages is derived). The first proposal is in fact an alternative to the Extended Projection Principle. The second proposal is based on Binding Theory. It involves a reformulation of the notion of Binding Domain, in a way that covers the classical facts in Binding Theory as well as the new facts presented in Chapter 2, and some others. Case Theory is also reformulated in a way consistent with our version of Binding Theory and Nominative Case assignment. This theoretical apparatus allows for a simple characterization of the contrast NSLs vs. non-NSLs, the former having the classical cluster of properties (null subjects, subject inversion, absence of that-t effects). In fact, however, we will argue that all languages have a more subtle form of null-subjecthood and subject inversion.

We also consider some independent motivation for our reformulation of BT, concerning a special type of copulative constructions ('John is not himself anymore'). We devote special attention to some specific types of subjects: 'indefinite subjects', namely indefinite NP's in existential and presentational constructions; CP subjects, and inverted subjects in French Stylistic Inversion. Finally, we propose a solution to a problem for Relativized Minimality concerning A-dependencies.

Chapter 4 is devoted to infinitival constructions. On the one hand, it aims at explaining

why infinitives in NSLs allow long head-movement. On the other, it presents an alternative conception of the notion of infinitival sentence, according to which infinitives in a given language are a 'simplified' version of the finite sentence structure in that language: the missing AGR-morphology is recovered in content in some way (control being the most typical strategy), or it is entirely absent.

Finally, Chapter 5 is devoted to preverbal subjects in NSLs. If the theories in the preceding chapters are on the right track, the question arises as to whether preverbal subjects in NSLs are of the same nature as in non-NSLs (for the latter we assume that (preverbal) subjects occupy the Specifier of AGRP). We will show that the classical test for subjecthood vs. dislocation (namely, only true subjects can be quantified) is not as clear cut as has often been claimed. Our conclusions point to the idea that preverbal subjects in NSLs can (besides being dislocated) be in a functional specifier which is not exclusively occupied by subjects.

## **2. Some Basic Premises**

### **2.1. The Internal Subject Hypothesis**

We assume some version of the Internal-Subject Hypothesis (ISH) as advanced in Koopman and Sportiche (1988)/(1990). This hypothesis has also been developed in some way in Zagana (1982), Kitagawa (1986), Speas (1986), Contreras (1987), Kuroda (1988), and, specifically for Catalan, in Bonet (1989).

The essential idea of this hypothesis is that the External Argument (EA) is generated inside the VP (as in 0.a) or as an adjunct to the VP (as in 0.b)):

- (1) a.  $[_{VP} \text{ John}_{Ag} [_{V'} \text{ put the book}_{Th} \text{ on the table}_{Goal} ] ]$
- b.  $[_{VP} \text{ John}_{Ag} [_{VP} \text{ put the book}_{Th} \text{ on the table}_{Goal} ] ]$

Among the advantages and consequences of such an analysis we have the following:

a) locality in  $\theta$ -marking: all Arguments of a verb, even the EA, are projected and  $\theta$ -marked in the local domain of the VP. I think that if one adopts some version of the split-INFL hypothesis (as we will) it is hard not to adopt some version of the ISH in order to avoid an extremely long-distance verb-subject predication relation, specifically if the surface subject position is several functional categories up from the VP.<sup>4</sup>

b) a more straightforward notion of  $\theta$ -position: there are no A-positions being optional  $\theta$ -positions (as the specifier of INFL was in previous hypotheses).  $\theta$ -positions are all and only those positions which are projected as a consequence of the Projection Principle. The notion of A-position may still be required (see Rizzi (1991-b) for an attempt to define A-positions as a super-set of the set of  $\theta$ -positions), but even so, there is no A-position being an optional  $\theta$ -position.

c) a more uniform characterization of the source for sentence subjects: all of them are generated in the local domain of VP. Thus, a theory trying to derive which Argument will become the sentence subject has more explanatory power than one simply stipulating that the EA is directly generated in its surface position. The relevant proposal will be introduced in Chapter 3, in connection with Burzio's Generalization.

The ISH will be crucial for our account of subject inversion, of which we advance some basic assumptions in the next subsection.

---

<sup>4</sup> Chomsky's (1986-b) proposal that the sisterhood condition on predication overlooks FCs is, I think, too powerful.



## 2.2. Subject Inversion

Subject inversion is closely related to the ISH in that it is quite tempting to assume that inverted subjects (in languages such as Romance NSLs) are in fact occupying their basic position according to the ISH. The classical account of subject inversion (as in Chomsky (1981)/(1982), Rizzi (1982-b)) claims that inverted subjects are (right) adjoined to the VP, which is not at odds with the ISH in 0.b), where the EA also occupies an adjoined position. For Catalan, the idea that inverted subjects occupy their basic position has been proposed by Bonet (1989) and Solà i Pujols (1989).<sup>5</sup> We will assume the essentials of Bonet's hypothesis, with some qualifications.

The idea that inverted subjects occupy their basic position raises many questions. Basically it predicts that inverted subjects which are EAs will occupy a position peripheral to VP, while inverted subjects which are internal Arguments will appear inside the VP.<sup>6</sup> To test these predictions is not easy, as there are many theoretical variables and obscuring facts:

a) the basic position for the EA could in principle be left-adjoined or right-adjoined to the VP. Notice that the usual subject-predicate word order for Small Clauses (if the EA forms a SC with the VP) is a consequence of the requirement of adjacency for Case-marking, but in the present case this requirement is possibly not relevant, for the EA will receive Case by becoming a sentence subject, not by being governed by an adjacent head.

---

<sup>5</sup> Rosselló (1986) proposes that the inverted subject position is the A-position for subjects, preverbal subjects being left dislocated elements, but she assumes that inverted subjects are in a rather high position, outside the VP. We will adhere to some of her arguments concerning the status of preverbal subjects in NSLs.

<sup>6</sup> I think that assuming that internal Arguments may raise to the same position as the basic position for EAs is not an appealing hypothesis: it brings back the dubious notion of optional  $\theta$ -position which can be overcome by the ISH. So I will not adhere to Bonet's (1989) proposal that internal Arguments in unaccusative verb constructions raise to Spec of VP, i.e., to the position where, according to her, EAs are generated in transitive and unergative verb constructions.

b) Word order with respect to the  $V^0$  cannot be telling, since  $V^0$  in null subject languages undergoes head movement (see Belletti (1991)), so that the verb will always appear to the left of the inverted subject whether the latter is left or right-adjoined to the VP.

c) Word order between the inverted subject and other arguments is difficult to test because in many cases (but not all) the co-occurrence of an inverted subject and a non-clitic/non-dislocated argument gives unnatural results (so in the paired examples below I will be using relative judgements, not absolute ones, since sometimes neither example is perfect).

d) When this co-occurrence gives good results, the facts might be obscured by extraposition of VP-internal constituents, which is relatively free in Romance languages.

e) Sometimes two alternative word-orders are not equivalent from the point of view of the Topic/Focus distinction, so that to compare its acceptability can be misleading.

In spite of all these problems, I will try to establish some minimal characterization of the facts. In the following sections, we will consider subject inversion in transitive/unergative, unaccusative/passive and copulative constructions.

Only for the sake of simplicity, I will be assuming throughout this thesis that the EA is (left- or right-) adjoined to VP. The alternative hypothesis, namely that the EA is in Spec of VP, raises some questions: when the inverted subject (assumed here to be in the subject basic position) occurs to the right of VP, then:

a) either VP has a right branching specifier, as Bonet (1989) assumes.

b) or we assume that lexical categories are projected unordered, as opposed to FCs (this is proposed in Ouhalla (1991)).

c) or inverted subjects appearing to the right of the VP are not in the subject basic position.

Option a) is merely stipulative. Option b) does not explain why, as we will see, some languages have obligatory right adjunction (the inverted subject always follows the complements: Catalan, Italian), others obligatory left adjunction (Romanian) and others either optionally (Spanish). Option c) is appealing, for right-adjoined inverted subjects are usually

Focus, which suggests this is a Focus position. It is, however, difficult to imagine what kind of position this could be: by X-bar theory constraints it can only be a right branching specifier or the lowest complement in the structure (thus a complement of V, which makes little sense). On the other hand, that Focus comes rather to the end of the sentence seems to be a widespread situation across languages, so that a strictly structural account for the Focus status of inverted subjects might be redundant.

The idea that the inverted subject position is the specifier of VP could raise a potential problem: verb complements do not seem to c-command this position as far bound variable binding is concerned. So (2) shows WCO effects:

- (2) \*No<sub>i</sub> condueix cap cotxe<sub>i</sub> el seu<sub>i</sub> propietari  
 Not drives<sub>i</sub> no car<sub>i</sub> the his<sub>i</sub> owner  
 intended reading: 'No car<sub>i</sub> is driven by his<sub>i</sub> owner'

Therefore strict c-command, and not m-command, should be postulated for bound variable binding if one adopts the hypothesis that inverted subjects occupy the specifier of VP. To stay on the safe side, I prefer to adopt the view that the basic position for EAs is one of adjunction to VP. Perhaps this idea, together with the hypothesis that Small Clauses are adjunction structures, has the advantage of characterizing all predicative nuclei in a sentence as Small Clauses (which can be of any of the four lexical categories):

- (3) a. John<sub>i</sub> [VP e<sub>i</sub> [VP eats bananas ]]  
 b. John<sub>i</sub> is [AP e<sub>i</sub> [AP apish ]]  
 c. John<sub>i</sub> is [NP e<sub>i</sub> [NP a monkey ]]  
 d. John<sub>i</sub> is [PP e<sub>i</sub> [PP in the jungle ]]

The only gap in this paradigm would be unaccusative verb clauses, where there would be

no EA adjoined to the VP.

### 2.2.1. Transitive/Unergative Constructions

Transitive and unergative verbs are by hypothesis the ones having an EA. If this Argument is adjoined to the VP, the predictions are:

- the inverted subject will occur to the right of the other Arguments (if there are any) if right-adjoined to the VP.
- it will appear to the left of the other Arguments if left adjoined to the VP (still to the right of the verb because of  $V^0$  movement).

Let us consider the following examples from Catalan:

- (4)
- a. Avui farà el dinar en Joan  
Today will-make the lunch the Joan  
'Today JOAN will cook the lunch'
  - b. ???Avui farà en Joan el dinar
  - c. Aquest mes pagarà les factures en Joan  
This month will-pay-for the bills the Joan  
'This month JOAN will pay the bills'
  - d. ???Aquest mes pagarà en Joan les factures

Notice that inverted subjects are (generally) interpreted as Focus when they appear after the object or an obligatory oblique Argument (see below). When they precede the object they are not, but the result (in Catalan) is unnatural. So I think the contrasts are genuine in spite of the fact that they are not interpretatively neutral. This does not hold when the inverted subject precedes an adjunct or optional Argument (see below): in this case the inverted subject need not exhaustively be Focus (it may rather be part of the a larger Focus constituent, namely the VP).

Thus the examples in 0 seem to indicate that the inverted subject is right adjoined to the VP, as it more naturally follows the object. This is true for Catalan. For Spanish the b) and d) examples would be as natural as the a) and c) examples, the difference being that in the latter the inverted subject is interpreted as Focus, as in Catalan.<sup>7</sup>

If now we shift to (obligatory) oblique arguments, the facts are similar:

- (5) a. Avui s'encarregarà dels nens en Joan  
 Today will-be-in-charge of-the children the Joan  
 'Today JOAN will be in charge of the children'
- b. ??Avui s'encarregarà en Joan dels nens
- c. Parlarà d'aquest tema en Joan  
 Will-speak of this subject the Joan  
 'JOAN will talk about this subject'
- d. ??Parlarà en Joan d'aquest tema

Again, Spanish does not show any contrast in acceptability. So objects and obligatory oblique Arguments seem to suggest that, at least in Catalan, the EA inverted subject is preferred as right-adjoined to the VP.

Optional Arguments (such as Datives and Obliques, see 0) and Locatives not being the first Argument -0- seem to allow free word order w.r.t. the inverted subject (perhaps with a slight preference for the PP-inverted-subject order):

---

<sup>7</sup> There seems to be some V-2-like phenomena in Spanish VSO sentences, in that they are preferred when there is some preverbal constituent (including Wh-constituents, Negation and some adverbs). But we cannot claim this is a case of subject-verb inversion, for the position of adverbs may show that the subject is lower down in the structure:

(i) No está **todavía** Juan en casa  
 Not is still J. at home  
 'Juan is not home yet'

- (6) a. Aquest llibre el va regalar a la Maria en Joan

This book it-gave to the M. the J.

'Joan gave this book to Maria'

- b. ?Aquest llibre el va regalar en Joan a la Maria

- c. D' això, en parlarà amb la Maria en Joan

Of this of-it-will-talk with the M. the J.

'About this, Joan will talk with Maria'

- d. ?D'això, en parlarà en Joan amb la Maria.

- (7) a. El nen, l'ha portat a casa en Joan

The boy, him-has taken to house the J.

'Joan has taken the boy home'

- b. El nen, l'ha portat en Joan a casa

- c. El llibre, l'ha posat al prestatge en Joan

The book it-has put on-the shelf the J.

'The book, Joan put on the shelf'

- d. El llibre, l'ha posat en Joan al prestatge

Adjuncts of time, place and manner -0- also allow free word order, although perhaps the order inverted subject-adjunct is slightly preferred:<sup>8</sup>

- (8) a. Aquest llibre, el va comprar en Joan a Londres

This book it-bought the J. in London

'This book was bought by Joan in London'

- b. (?)Aquest llibre, el va comprar a Londres en Joan

---

<sup>8</sup> In the following examples, none of the adjuncts is intended to be interpreted as dislocated (which would render the examples irrelevant for the discussion): actually, the non-dislocated construction is the most natural one in a neutral context.

- c. Comprarà el menjar en Joan aquesta tarda  
Will-buy the food the J. this afternoon  
'Joan will buy the food this afternoon'
- d. ?Comprarà el menjar aquesta tarda en Joan
- 0 e. Ho escriurà en Joan amb l'ordinador  
It-will-write the J. with the-computer  
'Joan will write it with the computer'
- f. (?)Ho escriurà amb l'ordinador en Joan

We will interpret the preceding facts in the following way:

a) In Catalan, the EA inverted subject is right-adjoined to the VP. It can be extraposed to the right of the time/place/manner adjuncts, giving less natural sentences.

b) Therefore the other Arguments precede the inverted subject, unless they are extraposed to its right: since extraposition is easier for optional and locative Arguments in general, the latter will more easily appear to the right of the inverted subject than objects and other oblique Arguments, which cannot be extraposed so easily.

c) Spanish, unlike Catalan, allows the inverted subject to be left or right adjoined to the VP freely (*modulo* Focus interpretation), so that no contrast in acceptability appears in the word order w.r.t. arguments. I assume that other languages cluster together with either Catalan or Spanish. Italian is like Catalan. Portuguese is, as far as I know, like Spanish. Some dialects or varieties of Catalan (Valencian, speakers with strong Spanish interference in immigration areas) are probably like Spanish. In Romanian left adjunction (the inverted subject preceding the object) is not only possible, but in fact obligatory (see Motapanyane (1989)).

d) Right VP-adjoined inverted subjects usually have Focus interpretation for some reason.

The last point may suggest that we are missing something in simply saying that we are dealing with VP-adjunction. Perhaps there is a right-branching Focus Specifier. However, right

adjoined inverted subjects do not always have Focus interpretation. Specifically, embedded clauses do not necessarily involve Focus interpretation for the inverted subject exhaustively:

(9) Allò que no sé és QUAN PRESENTARA LA TESI EN JAMES

That that not know-I is when will-submit the thesis the J.

'What I don't know is WHEN JAMES WILL SUBMIT HIS THESIS'

where the whole embedded sentence is Focus and the inverted subject *en James* is not (necessarily) Focus: the sentence can be used even if we are not contrastively considering when either *James* or *Albert* is submitting his thesis. So I prefer to keep to the idea, however vague it may be, that VP-final inverted subjects are usually Focus because final (non-dislocated) constituents tend to be Focus in general.

### 2.2.2. Unaccusative/Passive Constructions

The question we want to answer is the following: what is the position for an inverted subject originally being an internal Argument? If the internal Argument is in the underlying object position, we would expect it to precede the other Arguments (if any).<sup>9</sup>

With unaccusative verbs, the prediction is not clearly fulfilled. When we are dealing with an obligatory Argument (such as the obligatory locative Argument of *anar* 'to go' in 0), the preferred word order is locative-subject, contrary to what we would expect:

---

<sup>9</sup> For the moment, let us abstract away from indefinite inverted subjects (which seem to occupy the object position, not only in languages allowing subject inversion, but in all languages in general).



- (10) a. Ha anat a casa en Joan  
 Has gone to house the J.  
 'JOAN went home'
- b. ??Ha anat en Joan a casa  
 Has gone the J. to house  
 'Joan has gone home'

With optional Arguments, the order inverted subject-Argument seems preferable:

- (11) a. M'ha caigut l'agulla al forat  
 Me-has fallen the-needle in-the hole  
 'I dropped my needle into the hole'
- b. ??M'ha caigut al forat l'agulla
- c. Ha entrat el Zorro al palau  
 Has gone-in el Zorro in-the palace  
 'El Zorro has broken into the palace'
- d. ?Ha entrat al palau el Zorro
- 0 e. Ha vingut en Joan a casa  
 Has come the J. to house  
 'Joan has come home'
- f. ?Ha vingut a casa en Joan
- g. ?S'ha transformat l'àcid en sal  
 SE-has transformed the acid in salt  
 'The acid changed into salt'
- h. ??S'ha transformat en sal l'àcid

With adjuncts, the clearly preferred order is inverted subject-adjunct (although there is

some variation):

- (12) a. S'ha fos la bombeta al menjador  
SI-has fused the bulb in-the dining-room  
'The bulb blew in the dining room'
- b. ?\*S'ha fos al menjador la bombeta
- c. Ha vingut en Joan aquesta tarda  
Has come the Joan this afternoon  
'Joan has come this afternoon'
- d. ??Ha vingut aquesta tarda en Joan
- e. Hi ha anat en Joan amb cotxe  
There-has gone the Joan with car  
'Joan has driven there'
- f. ?Hi ha anat amb cotxe en Joan

In the examples in both 0 and 0 it is worth noticing that the order subject-PP does not entail contrastive Focus on the subject (rather the whole sentence minus the preverbal elements is 'new information'). This further confirms our position that inverted subjects do not occupy a Focus position. The order PP-subject (or object-subject) does involve a Focus interpretation for the subject, but the sentence is not very natural.

If we consider passives, the facts are more clear cut: (standard) passives and SE impersonal passives seem to allow for the subject to appear in its basic object position (i.e., preceding all other Arguments) as the preferred word order, while the PP-subject order involves contrastive Focus on the subject and is not very natural:

(13) Passives:

- a. Han estat posats els llibres al prestatge  
Have been put the books in-the shelf  
'The books were put on the shelf'
- b. ??Han estat posats al prestatge els llibres
- c. Ha estat aprovada la llei al parlament  
Has been passed the law in-the parliament  
'The law has been passed in the parliament'
- d. ??Ha estat aprovada al parlament la llei

(14) SE impersonal passives:

- a. S'han posat els llibres al prestatge  
SE have put the books in-the shelf  
'The books were put on the shelf'
- b. ??S'han posat al prestatge els llibres
- c. S'ha aprovat la llei al parlament  
SE has passed the law in-the parliament  
'The law has been passed in the parliament'
- d. ??S'ha aprovat al parlament la llei

Even if contrasts are not strong and clear cut, I think it is reasonable to conclude that:

a) Internal-Argument inverted subjects are subjects in their basic position (object position).

b) The prediction is not problematic for passives. As for unaccusative verbs, the problematic behavior of verbs like *anar* 'to go' could be due to the unclear status of this verb (and possibly others) as a truly unaccusative verb. At least from the point of view of  $\theta$ -theory, this verb could perfectly pattern with agentive verbs, as has been argued in Gràcia i Solé (1986). We could claim that 'to go' is systematically ambiguous between an unaccusative verb (which it

would be in cases like: 'The document went to the chairman's hands') and an agentive verb (as in 'John went to see Mary full of passion'). In the non-agentive usage, the order V-subject-PP is more readily acceptable:

- (15)    Han anat totes les bales a la paret  
           Have gone all the bullets to the wall  
           'All the bullets hit the wall'

### 2.2.3. Copulative Constructions

The standard analysis for copulative constructions (since Couquaux (1981)) assumes that they are raising constructions, their subject being originated as the subject of a small clause. If so, subject inversion would consist in having the subject in this basic position.

Consider the following cases:

- (16)    a.    Està cansat en Joan (no en Pere)  
               Is tired the Joan (not the Pere)  
           b.    ?\*Està en Joan cansat  
           c.    Ara és de vacances en Joan  
               Now is on vacation the Joan  
           d.    ?\*Ara és en Joan de vacances  
           e.    Serà campió el Barça  
               Will-be champion the Barça  
           f.    ?\*Serà el Barça campió

The pattern in 0 is similar to the other constructions considered above in that the inverted subject is 'VP-final', the difference being that here we are not dealing with a VP, but rather with

an AP, PP or NP predicate. As before, Spanish does not forbid the b., d. and f. constructions.

If we assume, as in Koopman & Sportiche (1988), that the EA forms a Small Clause with its VP in Agentive constructions, and that for some reason the subject of the Small clause follows the predicate in Catalan (optionally in Spanish), then this idea extends naturally to non-VP small clauses.

It has been contended in several papers (Bonet (1989), Saccon (1991)) that subject inversion is not possible with copulative constructions involving an individual-level predicate (in Kratzer's (1988) terminology).<sup>10</sup> The evidence is based on examples like the following:

- (17) a. \*És intel.ligent en Joan  
Is intelligent the Joan  
b. \*És de Barcelona la Maria  
Is from Barcelona the Maria

The question is: are these examples unacceptable because individual-level predicates do not admit subject inversion at all, or because there is some restriction which excludes these examples without excluding all cases of subject inversion with individual-level predicates? I will assume that the latter idea is on the right track.

First of all, it is a general fact that the inverted subject, when sentence-final is Focus. Then it might be the case that in the preceding examples there is some problem concerning the

---

<sup>10</sup> In fact, several authors have contended that individual-level-predicate copulative constructions the subject is directly generated in Spec of INFL (Kratzer (1988), Torrego (1989)). I think this is hardly tenable if the split INFL hypothesis is to be kept, unless it is somehow reinterpreted (e.g., by assuming that the Argument of an individual level predicate is generated higher up than the Argument of a Stage level predicate). I will abstract away from the issue. I think that, being equipped with the ISH and empty categories, we should not give up strict locality constraints on predication, unless semanticists themselves were to say otherwise for strictly semantic reasons, which is unlikely.

interpretation of the Focus element.

As argued in Kuno (1972), the element interpreted as Focus has what he terms an 'exhaustive reading' interpretation. Consider the following sentences:

- (18) a. *En Joan* es presentarà en aquesta plaça  
The J. SI-will-present in this post  
'Joan will apply for this post'
- b. En aquesta plaça, s'hi presentarà *en Joan*  
In this post SI-there-will-present the J.  
'It is John who will apply for this post'

Sentences a. and b. differ in that *en Joan* is Focus in b. but not in a. Unlike the sentence in a., the sentence in b. naturally suggests that *Joan* is the only applicant, or at least the only one out of a discourse-determined set of people. This is what we mean by 'exhaustive reading' for Focus.<sup>11</sup>

Now consider the sentences in 0 again. We could argue that what makes these statements odd is the fact that the predicates involved do not naturally admit an exhaustive reading, especially if they are stated out of the blue: it is odd to say that 'someone is the one who is intelligent', implying that the others simply are not. If this line of reasoning is correct, the prediction is that individual level predicates will not allow (focused) inverted subjects as far as there is some conflict with Focus interpretation. If some individual level predicates are not liable

---

<sup>11</sup> Perhaps the 'exhaustive reading' belongs to the presupposition, if it is the case it can be cancelled, as the following dialogue could suggest:

A: JOHN will apply for the post!

B: And Mary?

A: Well, yes, Mary too.

It is not clear, though, if speaker A is correcting his previous assertion in his reply, or perhaps resetting the range of the discourse presupposed set, of which the 'exhaustive reading' holds.

to such a conflict, subject inversion will be alright with them. I think this prediction is fulfilled. First of all, suppose we add a degree adverb to a predicate such as 'intelligent' ('more/less intelligent'): the modified predicate will more easily allow a Focus interpretation for its subject, because it is the case that somebody is exhaustively the person who 'is more intelligent than others belonging to some discourse-determined set of people'. So subject inversion is quite natural in this case:

- (19) a. És més intel·ligent en Joan  
           Is more intelligent the Joan  
       b. És menys perillosa la dinamita  
           Is less dangerous the dynamite

On the other hand, individual-level predicate copulative constructions with inverted subject improve if we make the exhaustive reading linguistically more explicit, as in 0:

- (20) a. Només és intel·ligent en Joan  
           Only is intelligent the Joan  
           'Only Joan is intelligent'  
       b. De nosaltres, només és de Barcelona en Joan  
           Of us only is from Barcelona the Joan  
           'Out of us, only Joan is from Barcelona'  
       c. En aquest llibre, només és interessant el pròleg  
           In this book only is interesting the preface  
           'In this book, only the preface is interesting'

So I will assume that individual level predicates do not exclude subject inversion in principle. What excludes some of the constructions is the independent fact that the inverted

subject is interpreted as Focus and Focus requires an exhaustive reading.

Similar considerations are valid for non-copulative individual level predicates (such as 'to know how', 'to hate'): they easily allow an inverted subject as far as an exhaustive reading interpretation is pragmatically available or linguistically emphasized:

- (21) a. De nosaltres, només sap cantar bé en Joan  
Of us only knows to-sing well the Joan  
'Out of us, only Joan can sing well'
- b. Jo odio les cerimònies, però encara les odia més en Joan  
I hate the ceremonies but even them-hates more the J.  
'I hate ceremonies, but Joan hates them even more'

#### **2.2.4. On the Availability of Subject Inversion**

It has often been claimed that subject inversion is a highly restricted option in languages allowing it. Since we are going to argue that the inverted subject position is a Case position available in principle for any sentence in a NSL, we should say something in this connection. Let us revise two recent proposals on the issue.

Rizzi (1991-a) suggests that subject inversion is only possible if no overt complements intervene between the verb and the inverted subject. This would be due to a requirement of adjacency for Nominative assignment. I think this suggestion is problematic for at least two reasons:

- a) We would be dealing with merely phonological adjacency, for when the complements are absent, they must be realized as empty categories. This is not in accordance with traditional proposals on adjacency requirements for Case marking.
- b) If the adjacency requirement was due to Case assignment, we would expect sharply



degraded constructions when the requirement is violated (comparable to *\*John put on the table the book*). In fact, there are many examples of subject inversion with the order V O S being perfectly acceptable. I think that the varying degrees of acceptability of the word order V O S rather depend on the contextual or pragmatic plausibility for the V O sequence to be interpreted as a topic of conversation (so V O sequences like 'cook the meal', 'pay the bill', etc. readily accept a postverbal subject, whereas sequences like 'find a curious worm' do not). So, both in Catalan and, as far as I know, in Italian, sentences like the following are perfectly acceptable:

- (22) a. Avui farà el dinar en Joan (Catalan)  
 Today will-make the lunch the J.  
 'Today Joan will cook the lunch'
- b. Questo mese, pagherà le fatture Gianni (Italian)  
 This month will-pay the bills G.  
 'This month Gianni will pay the bills'

Delfitto & Pinto (1992) make the surprising claim that inverted subjects are subject to the Definiteness Effect, and that apparent counterexamples should be analyzed as involving a Small Clause structure. They present acceptability contrasts as the following:

- (23) Ha recensito il libro un professore/\*Gianni  
 Has reviewed the book a professor / G.

From the empirical point of view, if there is a contrast in 0, it is very feeble:<sup>12</sup> in fact neither of the options in 0 should be very natural if asserted out of a context. No such contrast will appear in 0, which is a more felicitous sentence from the pragmatic point of view:

---

<sup>12</sup> The Italian speakers I asked found no such contrast.

(24) Oggi, ha condotto la macchina un ragazzo/Gianni

Today has driven the car a boy G.

'Today A BOY/GIANNI has driven the car'

From the theoretical point of view, assuming that a Small Clause structure is an available alternative option for verbs being agentive predicates looks problematic w.r.t. the Projection Principle. On the other hand, if we assume that Focus is an operator that creates a lambda predicate in its scope at LF, the Small Clause structure is in some sense guaranteed without having to revise the underlying projection of predicates, which should optimally be uniform: a lambda predicate for Focus would create a structure roughly represented like 0.b) (for 0.a)):

(25) a. I saw JOHN

b. John<sub>x</sub> [ I saw x ]

roughly read as: JOHN is the x I saw

In fact, lambda operators are required for the interpretation of any quantifier (or operator, more generally) if we adopt a Quantifier Raising approach (see Heim (1989)). So for independently motivated reasons a structure being roughly a Small Clause is obtained for free without having to revise the underlying projection of predicates at D-structure, which should optimally be uniform.

In conclusion, although subject inversion is not apparently a freely available option, I assume that this is not due to grammatical factors, but rather to the fact that inverted subjects are usually Focus and Focus is an operator that has strong pragmatic and discursive interactions.

### 2.2.5. Floating Quantifiers

Koopman & Sportiche (1988) argue that FQs may be used as a diagnostic for detecting the internal subject position even in languages where there is no free subject inversion. For English, this position would be pre-VP (or pre-XP, where XP is an adjectival, nominal or prepositional predicate, in copulative sentences), as the following examples suggest:

- (26) a. The boys have all/both understood  
b. The boys are all/both intelligent/students/in the kitchen

For Catalan (Spanish, and Italian), FQs can appear both in pre-VP and post-VP position:

- (27) a. Els nois faran (tots) la feina (tots)  
The boys will-do (all) the work (all)  
b. Els nois estan (tots) cansats (tots)  
The boys are (all) tired (all)

If Koopman & Sportiche's hypothesis is correct, then our idea that the subject basic position is one of right-adjunction to the VP (or XP predicate) in Catalan will have to be qualified.

Our suggestion is that, since the subject basic position is one of right-adjunction to VP in Catalan, FQs appearing to the left of the object actually occupy a higher position in the structure (a specifier of some FC). This is the hypothesis adopted in Bonet (1989).

On the other hand, for Catalan it is not clear that FQ are material left by NP-movement as is assumed in Koopman & Sportiche. Consider the following examples:

- (28) a. Els nois han presentat *tots dos* la sol.licitud  
 The boys have submitted *all two* the application  
 'The boys have both submitted the application'
- b. Ja ha presentat *tothom* la sol.licitud  
 Already has submitted *everybody* the application  
 'Everybody has submitted the application now'
- c. Encara no ha presentat *ningú* la sol.licitud  
 Yet not has submitted *nobody* the application  
 'Nobody has submitted the application yet'

In 0.a) *tots dos* 'both' is not likely to be an element left by movement, in that it cannot form a constituent with the preverbal subject (*\*tots dos els nois*). In 0.b/.c) *ningú* 'nobody' and *tothom* 'everybody' cannot be considered floating elements in any reasonable sense: nevertheless, they have a distribution similar to that of FQ.

In conclusion, we will assume that the basic position for subject is post-VP (or XP) in Catalan, and that a certain type of elements (FQs and certain single-word quantifiers) can raise to a higher pre-VP position.

On the other hand, any theory of FQs has to admit that there are several positions for FQs, as can be seen in 0:

- (29) They might (all) have (all) been (all) chatting her up.

Obviously, not all these positions can be in the subject basic position, and possibly none of them are, inasmuch as FQs cannot occur in VP-final position, which we claim is the subject basic position:

- (30) \*They might have been chatting her up all.

I will adopt the view that FQs undergo leftward movement, which is a typical situation for light quantifier particles (e.g., French *rien* 'nothing', *tout* 'everything' as objects -see Pollock (1989)).

\* \* \*

These assumptions will be crucial in our characterization of subject inversion in the following chapters. In fact, we already considered the alternative possibility that inverted subjects are not in their basic position, but rather in some derived (Focus) position. To the extent that either alternative is generalizable through the range of languages we will be considering, this will not be a problem.

### 2.3. The Projection of Arguments

The ISH does not challenge the hypothesis that there is, in some sense, an external Argument (EA) in the cases it had been traditionally postulated (transitive and unergative verbs). This hypothesis makes sense for both empirical reasons (singling out EAs provides a basis for explaining their peculiar behavior) and theoretical reasons. Concerning the latter, it is desirable to preserve some version of the Uniformity of Theta Assignment Hypothesis (UTAH), as defined in Baker (1988:46):

(31) *Uniformity of Theta Assignment Hypothesis*

Identical thematic relationships between items are represented by identical structural relationships between those items at the level of D-structure.

It might turn out that 0 is too strong: it is a well-known fact that different languages allow for different realizations of Arguments which are nonetheless thematically identical:

- (32) a. English: to look at DP (prepositional complement)  
b. Catalan: mirar DP (DO complement)

Even so, some implementation of a weakened version of the UTAH could be stated as follows (in the spirit of Belletti & Rizzi (1988)):

- (33) a. There is a thematic hierarchy which has to be mirrored by the structural hierarchy when the Arguments are projected.  
b. The Agent Argument (which is at the top of the thematic hierarchy) not only has to be projected as the highest Argument in the structure, but also as the EA.  
c. The EA is projected in a position external to VP, specifically, adjoined to VP (as a DP or, possibly, as a by-phrase, in passives).

The Theme Argument, on the other hand, would be the lowest one in the Thematic hierarchy and therefore it would be projected as the closest Argument to the verb, as a DP in the unmarked case. Or alternatively, as Larson (1988) proposes, the Theme could be the highest Argument inside the strict VP, and then possibly the second Argument in the hierarchy.

In any case, what will be crucial for our characterization of the facts concerning Burzio's Generalization is that the EA is structurally higher than the object Argument: the presence of the EA will prevent any other Argument (and specifically the object) from becoming a subject, while the absence of the EA will force some other Argument (usually the object) to become a subject in some specific sense we will make precise.

## 2.4. The Split INFL Hypothesis and Verb Movement

We will be assuming a version of the split-INFL hypothesis, specifically one in which AGRP dominates TP, as in Belletti (1991). The structure of the sentence is as in 0:

(34) [<sub>AGRP</sub> ... [<sub>TP</sub> ... [<sub>VP</sub> ] ] ]

where other possible FCs (such as NEGP and Object-AGRP) are omitted.<sup>13</sup> For the contention that AGR is higher than T in the hierarchy, I adhere to the motivation presented in Belletti (1991) (the Mirror Principle effects for verbal tense and AGR affixes, the plausibility that the subject occupies Spec of AGRP and not that of TP, etc.).

My working hypothesis throughout Chapters 2 to 4 will be that the hierarchy in 0 is uniform across languages (or at least the languages I will be considering): this will facilitate the discussion of the contrast between NSL and non-NSLs by minimizing the parametric factors of variation. In Chapter 5, however, we will explore the alternative view that the FC hierarchy is parameterized in a minimal way (concerning only two FCs), this parameterization being tightly related to the Null Subject parameter.

One of the facts our theory will be concerned with is  $V^0$ -movement. The general fact is that NSLs seem to exhibit longer  $V^0$ -movement than some non-NSLs. This will be a natural consequence of our theory, although we will not be able to predict the exact details of  $V^0$ -movement.

Specifically, two facts will naturally follow from our theory:

---

<sup>13</sup> See Pollock (1989), Chomsky (1988), Belletti (1991), Laka (1990) for some well-known proposals on the issue. See Chomsky (1992) for a proposal on object AGR.

- (35) a. Finite sentences in NSL always undergo long verb-movement (while they do not always in non-NSLs, e.g. in English). See Pollock (1989) and Belletti (1991) for a characterization of the facts.
- b. Infinitival sentences in NSLs allow long verb-movement, while they do not in non-NSLs (see Pollock (1989) and Belletti (1991)).

The following examples from Italian show that there is no asymmetry between finite and infinitival sentences concerning the word order w.r.t. negative elements:

- (36) a. Non ha *mai* detto la verità (finite)  
Not has never told the truth  
'(S)he has *never* told the truth'
- b. Non dire *mai* la verità... (infinitival)  
Not to-tell never the truth  
'(*Never*) to (*never*) tell the truth'

In Catalan and Spanish the word-order facts suggest, if anything, that infinitives undergo longer head movement, as some adverbs that can precede the finite verb cannot precede the infinitival (I give examples from Catalan):

- (37) Finite sentences:
- a. (Sempre) diu (sempre) la veritat  
(Always) tells (always) the truth  
'(S)he always tells the truth'
- b. (Ja) té (ja) el permís  
(Already) has (already) the permission  
'(S)he already has the permission'



(38) Infinitives:

- a. No cal (?\*sempre) dir (sempre) la veritat  
Not must (always) to-tell (always) the truth  
'It is not necessary to always tell the truth'
- b. M'agradaria (\*ja) tenir (ja) el permís  
Me-would-please (already) to-have (alr.) the permission  
'I would like to already have the permission'

In fact, in Chapter 4 we will make crucial use of the idea that in some NSLs V-movement in infinitives is a slightly longer than in finite clauses. But whether or not the above examples indicate longer verb movement for infinitives than for finite sentences, it seems clear that verb movement in Catalan (and Spanish) infinitives is no shorter than in the corresponding finite clauses.

For simplicity, we will assume that long verb-movement is movement up to AGR<sup>o</sup>, which is the highest FC in the hierarchy. The facts in both 0.a) and 0.b) will follow from the way Nominative Case is assigned in NSLs. In fact, the essential of Pollock's (1989) initial idea that V is allowed to move to a functional head only if that head is 'rich' can be kept under the approach we will propose. But here we will use the clearer notions of 'contentful'/'empty' instead of 'rich'/'poor'. In fact 'contentful' does not imply morphologically realized, as we will see (the opposite is not true: when a FC has some distinctive morphology -e.g., agreement in English present tense-, it obviously has some content).

## 2.5. Case Theory

As for Case theory, our main concern will be the Case which is assigned to the sentence subject, namely Nominative in the unmarked case, Accusative in ECM constructions, Oblique in *for*-infinitives, and, we will assume, a kind of weak Nominative Case in PRO-infinitives.

In the GB tradition, Nominative is assigned by INFL (AGR) to the specifier of INFL (AGR). In Koopman & Sportiche (1988), a more complex view is proposed, according to which there are two ways in which Nominative Case can be assigned:

- a) Government: INFL assigns Nominative to the subject basic position, which it governs.
- b) Agreement: INFL assigns Nominative to the subject in Spec-INFL, with which it agrees.

While the second option takes place only when the Argument becoming subject raises to Spec-INFL, the first one involves no raising of the subject (or at least no raising up to Spec-INFL). According to these linguists, these two possibilities are not exclusive, so that languages can choose either or both.

In the present thesis, we will instead adopt the view that these two options are disjunctive, so that languages can use either but not both. Specifically, we will contend that NSL (which allow subject inversion) are languages exclusively using the government strategy, while non-NSL exclusively use the agreement strategy. On the other hand, the option chosen by a language will not be directly set as a parameter value, but rather will be derived from constraints on how AGR is licenced in a given language.

If they are exclusive, then NSLs having subject inversion (then government-Nominative) do not have agreement-Nominative. This in turn implies that preverbal subjects are not in a Case position, since Nominative is assigned to the (empty) DP in subject basic position. Therefore the dependency between the preverbal subject and the (empty) position in the subject basic position will consist in resumptiveness. In Chapter 5 we will argue for this view.

Within the Split INF Hypothesis, we should address the question of which FC is responsible for Nominative Case assignment. In Chapters 3 and 4 we will adopt the working hypothesis that it is  $AGR^0$  which assigns Nominative in both the agreement and government strategy. In Chapter 5, however, we will propose that  $T^0$  is the basic Nominative-assigner by government, and that  $AGR^0$  assigns government-Nominative only by combining with  $T^0$ .  $AGR^0$ , on the other hand, is the only Nominative-assigner by agreement.

## Chapter 2

### Burzio's Generalization and EPP

#### 0. Introduction

The so-called Burzio's Generalization (from now on BG, see Burzio (1981)/(1986)) appears to be an accurate characterization of the distribution of Accusative Case, at least for a wide range of languages:

- (1) (Structural) Accusative Case is available iff there is an external  $\theta$ -role.

The biconditional in 0 can be split into two different implications:

- (2) a) External  $\theta$ -role  $\rightarrow$  Accusative  
b) Accusative  $\rightarrow$  external  $\theta$ -role

In fact, 0.a) is not difficult to derive for verbs having an external Argument (EA) and an internal one realized as object: if there is no Accusative case, a verb having these two (not inherently Case-marked) Arguments will not be able to appear in a well-formed structure, since either the EA or the object will remain Caseless.

Still, 0.a) makes a prediction that is not trivial: intransitive (verbs having an EA but no object) are Accusative assigners, so that they are able to assign Accusative to some DP. Consider:

- (3) a. He talked my head off (see Burzio (1986))  
 b. Vuit hores, no les treballa /dorm (Catalan)  
 Eight hours not them<sub>Acc</sub>-works/sleeps  
 'He does not work/sleep for eight hours'  
 c. Els cent metres els corre fàcilment  
 The hundred meters them<sub>Acc</sub>-runs easily  
 'He easily runs the hundred meters'

0.b) is, however, the most difficult part to derive and the one we will concentrate on: it is not obvious why a verb cannot exist, in English as in many other languages, being like *to fall* (i.e., an unaccusative verb) but assigning Accusative. On the other hand, 0.b) strongly undermines the standard Case-theoretical account for DP-movement, namely, that a DP moves to obtain Case: it follows from 0.b) that a DP will fail to have Accusative Case precisely in the cases where DP-movement to obtain another Case is available,<sup>14</sup> the EA being absent. Consider the following D-structures:

- (4) a. John [<sub>VP</sub> melt the ice ] (transitive)  
 b. - [<sub>VP</sub> be melted the ice ] (passive)  
 c. - [<sub>VP</sub> melt the ice ] (unaccusative)  
 (5) a. J. [<sub>VP</sub> believe [<sub>IP</sub> the ice to VP ] ] (ECM)  
 b. - [<sub>VP</sub> be believed [<sub>IP</sub> the ice to VP ] ] (ECM-passive)  
 c. - [<sub>VP</sub> seem [<sub>IP</sub> the ice to VP ] ] (raising)

---

<sup>14</sup> Abstracting away from PRO, if it does not obtain Case. In fact we will assume it does.

In the a. examples we have verbs assigning an external  $\theta$ -role: as predicted by BG, Accusative Case is available. In the c. examples there is no external  $\theta$ -role, while in the b. examples the external  $\theta$ -role is not projected (at least as a visible DP). In both b. and c., Accusative is not assigned, and consequently the Argument *the ice* has no option but DP-movement to the subject position. These facts strongly suggest that DP-movement is not due to lack of Case, but rather to the possibility for DP-movement to take place. So, 0.b) could be replaced by:

- (6) Accusative is assigned only if DP-movement is not an alternative available option for the DP to obtain Case (or be realized as PRO).

Put another way, it can be easily shown that BG stands in a relation of theoretical circularity with Case Theory and Theta theory. In fact, for any well-formed structure involving an unaccusative verb or a raising verb, the following implications hold:

- Case theory: [-Accusative]  $\rightarrow$  DP-movement
- $\theta$ -criterion: DP-movement  $\rightarrow$  [-Ext.  $\theta$ -role]<sup>15</sup>
- Burzio's Gen.: [-Ext.  $\theta$ -role]  $\rightarrow$  [-Accusative]

So, BG is problematic in two respects:

- it establishes a correlation between two facts that do not look akin in nature: the existence of the EA and the availability of Accusative Case.
- it undermines the classical Case account for DP-movement as a last resort device to obtain Case, since it leads to circularity in this connection.

---

<sup>15</sup> Within the Internal subject hypothesis, NP-movement requires absence of external  $\theta$ -role not because of the Theta criterion, but because of Case theory: there would be two Arguments competing for a Subject Case-position. In any event, the result is the same.

Another question that should be addressed concerning BG is its universality. Are all languages subject to BG? If we take the so-called ergative languages (ELs) into consideration, it seems it is not. Let us consider a transitive/unaccusative verb like *to melt* again and let us compare the Case array in ELs with that of non-ELs:

(7)		<u>External Argument</u>	<u>Internal Argument</u>
ELs	Trans.	Ergative Case	Absolutive Case
	Unacc.	-	Absolutive Case
<u>non-ELs</u>	Trans.	Nominative Case	Accusative Case
	Unacc.	-	Nominative Case

If we analyze ELs as involving obligatory passive for transitive verbs (as in some traditional analyses), then BG is falsified in one sense: if obligatory passive is interpreted as systematic lack of Accusative Case, then it is not the case that the presence of an EA implies Accusative, as BG predicts in one sense.

Another analysis is possible for ELs in which Ergative Case is equated with Nominative and Absolutive with Accusative.<sup>16</sup> With such an analysis, BG is falsified in the other sense: Absolutive (=Accusative) is available independently of the presence of an EA. In the next chapter we will propose a parameter which accounts for the contrast between ELs and non-ELs, so that BG will be derived from the [-Ergative] value of that parameter, while another the [+Ergative] value will give rise to another generalization holding of ELs.

Before making our proposal in the next chapter, we will explore the possibility of deriving BG from more or less standard assumptions. The conclusion will be that a different approach is necessary.

---

<sup>16</sup> See Levin (1983-a/b) for such an analysis applied to Basque.

## 2. Possible Solutions Based on CHAIN Theory

In this section we will try to provide a way of deriving BG which would be based on Chomsky's (1986-b) theory of CHAINS, together with a version of the Extended Projection Principle and other more or less independently motivated assumptions. Later we will argue that this solution is not adequate, and argue for a radically different hypothesis. Some version of the EPP, whether a primitive or a theorem, will be required in any case. Let us start by a brief review of how the EPP can be used to derive BG.

### 2.1. The Extended Projection Principle

As far as the EPP requires some DP to appear in the sentence subject position, it can be conceived as a way to derive BG, specifically 0.b) above: what we want to explain is why Accusative cannot be assigned when there is no EA. Roughly stated the answer would be: when the EA is not present, the EPP requires the internal Argument to raise to subject position. So, supposing there were verbs not having an EA and at the same time assigning Accusative, and *to sink* was one of them, a sentence like 0 could be excluded because of the EPP, since no DP is filling the subject position:

(8) \*Sank [the boat]-acc

Since *the boat* has to raise to subject position, where (in a finite sentence) it will receive Nominative, then Accusative cannot be assigned, because if it were the Chain (*the boat*, *t*) would have two Cases:

(9) \*The boat<sub>Nom</sub> sank t<sub>Acc</sub>

This first approach cannot work without further developments. First of all, we want to exclude cases such as 0:

(10) \*The boat<sub>i</sub> to sink t<sub>i-acc</sub> would be terrible.

where the EPP is fulfilled and *the boat* inherits Accusative Case from its trace: there is no Case-conflict, because the head of the Chain has no Case, and still the structure is ill-formed. Here we cannot appeal to a Last Resort principle for DP-movement in search for Case, since we are crucially assuming that DP-movement is not triggered by Case requirements, but rather by the EPP.

There is another problem having to do with the existence of expletive constructions. Since expletives seem to be able to satisfy the EPP, we must ask ourselves why we cannot have structures like 0, where the expletive satisfies the EPP and the internal Argument remains in situ and receives Accusative:

(11) \*There developed [ the problem ]<sub>Acc</sub>

(Cf. There developed a problem)

In order to overcome these problems, a more sophisticated theory is required. We will see how the notion of CHAIN can be useful in this connection.



## 2.2. CHAINS

A solution for the above problems can be based on Chomsky's (1986-b) theory of CHAINS, together with a version of the Extended Projection Principle. CHAIN can be defined as the unification of two traditionally different syntactic concepts: A-Chains (ordered sets of A-positions linked by an antecedent/trace dependency) and expletive/Argument dependencies. CHAINS are the entities that receive  $\theta$ -roles, subject to the condition they have only one Case. So both (*a man*, *t*) in 0.a) and (*there*, *a man*) in 0.b) would be instances of CHAINS:

- (12) a. *A man came t*  
b. *There came a man*

In both CHAINS in 0 there is a single  $\theta$ -role (namely the one assigned by *to come* to its internal Argument) and a single Case (namely Nominative). Since in both cases the internal Argument ends up being part of a CHAIN, we could try to exploit the idea that it is the obligatoriness of CHAIN formation that prevents this Argument from receiving Accusative, even if Accusative is available in principle. Suppose, then, we assume the following principles, one of which is a special version of the Extended Projection Principle (EPP):

- (13) Accusative assignment is always available but *optional*.

- (14) *Extended Projection Principle*:

Spec of INFL has to be a member of a CHAIN.<sup>17</sup>

---

<sup>17</sup> We could alternatively define 0 as:  
(i) Spec of AGR has to head a CHAIN  
were it not for the case of infinitives in raising constructions, where Spec of AGR in the infinitive does not head the (maximal) CHAIN. (i) can work if we are able to define the sub-CHAIN headed by Spec of AGR as a CHAIN in an interesting way. If we simply state that any sub-CHAIN is a CHAIN, the

These two principles are at the basis of an explanation for the BG facts: if there is an EA, this Argument can (and probably must) obtain Case by becoming member of a CHAIN (let us assume that a situation where the EA receives Accusative and the internal Argument enters an INFL-CHAIN is excluded in some way<sup>18</sup>); if there is no EA, the object has to become member of a CHAIN. This is so because, as is essential in a theory of CHAINS:

(15) A CHAIN must contain one and only one  $\theta$ -role.

Still, we want to exclude cases like the following, where there is a CHAIN fulfilling the above conditions but where Accusative has been assigned:

- (16) a. \*The problem<sub>Nom</sub> developed t<sub>Acc</sub>  
 b. \*There<sub>Nom</sub> developed the problem<sub>Acc</sub>

These cases are naturally excluded within the spirit of a theory of CHAINS: since CHAINS are the expression of an Argument, we can assume that a CHAIN cannot have more than one Case:

(17) A CHAIN must contain one and only one Case.

---

notion of 'heading a CHAIN' is not different from the notion of 'being member of a CHAIN'.

<sup>18</sup> Since the EA is not governed by V even in its basic position, Accusative will not be available for it. Let us assume that V-raising cannot widen the Case assigning scope of the verb. In the theory we will develop in the next chapter, Accusative assignment to the EA is further blocked because AGR coindexation overrides this possibility, in the same way it prevents Accusative assignment to the object in unaccusative or passive constructions.

Both examples in 0 involve CHAINS containing two Cases: Accusative (assigned to the foot) and Nominative (assigned to the head).

There are still some cases to be excluded if we want the present theory to be minimally accurate. All of the following examples fulfil the preceding principles and still they are ungrammatical:

- (18)
- a. \*It is strange there to develop the problem<sub>Acc</sub>
  - b. \*It is strange the problem to develop t<sub>Acc</sub>
  - c. \*It seems there to have developed the problem<sub>Acc</sub>
  - d. \*It seems the problem to have developed t<sub>Acc</sub>

In 0 the CHAINS ((*there, the problem*) and (*the problem, t*)) contain one and only one Argument and one and only one Case, and still they are strongly ungrammatical. In a standard theory CHAIN formation is triggered as a last resort for Case marking. So the above examples could be excluded because of a last resort principle that would prohibit forming a CHAIN longer than Case requirements demand. But we cannot adopt a last resort principle based on Case marking here because it is essential to the present theory that CHAIN formation is triggered by our version of the EPP in 0. So we have to stipulate that Case can only be assigned to the head of a CHAIN.

Summarizing, the following set of principles can derive the BG facts in a rough way:

- 0 Accusative assignment is always available but *optional*.
- 0 *Extended Projection Principle*:  
Spec of INFL has to be a member of a CHAIN.
- 0 A CHAIN must contain one and only one  $\theta$ -role.
- 0 A CHAIN must contain one and only one Case.
- (19) Only the head of a maximal CHAIN can be Case marked.

Notice that 0 and 0 are standard assumptions for Chains (therefore they should also be for CHAINS). 0 is a theoretically desirable result on the grounds of simplicity and generality.<sup>19</sup> The extended projection principle in 0 is also a natural development once we adopt a theory of CHAINS.

Notice also that 0 in conjunction with 0 implies that expletives are always linked to an Argument. So, the subject of a weather verb like *to rain* has to be an Argument itself, since there is no other Argument in the sentence for an expletive-Argument CHAIN to be formed. Let us assume, then, that what Chomsky (1981) dubbed a quasi-Argument (i.e., the subject of weather verb) is an Argument as far as 0 is concerned.

For the above theory to work, we have to assume that PRO has some Case. This is in the spirit of the theory of Visibility (see Chomsky (1986-b)), by which Case is necessary for a Chain (CHAIN) to be assigned a  $\theta$ -role: a CHAIN headed by PRO needs a  $\theta$ -role as any other, so some kind of Case should be available for PRO. Let us therefore assume PRO bears some (possibly intrinsic) Case.

### 2.3. Conditions on Expletive CHAINS

The theory developed thus far does not say anything about the wellformedness conditions on expletive CHAINS. In fact, there are only three types of well-formed expletive CHAINS, according to standard accounts:

- expletive-CP CHAINS (*It strikes me that...*);
- expletive-indefinite CHAINS (*There came a man*); and

---

<sup>19</sup> In fact, we could expect that Accusative, like Nominative, is not always available: still, like Nominative, its distribution would not be subject to lexical idiosyncrasies, but rather to syntactic constraints (e.g., passives could possibly be non Accusative-assigning contexts).

- null expletive-inverted subject CHAINS (Italian: *pro* ha telefonato *Gianni* 'pro has telephoned Gianni').

The third type seems to be restricted to NSLs. A theory about CHAINS has to be able to predict why these (and only these) are well-formed expletive CHAINS.

Concerning expletive-indefinite CHAINS, let us consider Belletti's (1988) theory of Partitive Case. According to Belletti, the indefinite DP receives Partitive Case, which is an inherent Case.

It predicts, on the basis of the assumption that Partitive is an inherent Case, and that inherent Case is always assigned to an Argument by a head that  $\theta$ -marks it, that indefinite DP's in these constructions are restricted to  $\theta$ -positions, (see 0.a)), and specifically to  $\theta$ -positions which are  $\theta$ -marked by the verb which assigns partitive (see 0.b)):

- (20) a. \*There seem men to have come  
b. \*There were considered men intelligent

In fact, Belletti's theory should be qualified in order to deal with certain types of Small Clause subjects. Specifically, Small Clauses which are complements of causative verbs allow Partitive subjects (examples from Catalan):

- (21) a. No deixeu [<sub>SC</sub> llibres oberts ]  
Not leave books open  
'Don't leave (any) books open'  
b. Aquesta política manté [<sub>SC</sub> gent en atur ]  
This policy keeps people in unemployment  
'This policy keeps some people unemployed'<sup>20</sup>

---

<sup>20</sup> The English glosses might be misleading because English bare NPs can be both indefinite and generic. Catalan bare NPs are only indefinite.

We could deal with these and similar cases by resorting to Rizzi's (1986) idea of Affectedness: in  $\theta$  the subject of the Small Clause is 'affected' by the process expressed by the main verb. We could in fact go further to say that the main verb participates in  $\theta$ -marking the SC subject (i.e., the SC subject receives a compositional  $\theta$ -role). If so, Belletti's contention that Partitive Case is an inherent (hence  $\theta$ -associated) Case is not falsified.

There are two other cases where the indefinite subject of a Small Clause could possibly not be  $\theta$ -marked by the verb governing it: these are the complement SC of *to have* and the SC in existential constructions (if they involve a SC) (ex.s from Catalan):<sup>21</sup>

- (22) a. Tinc [SC llibres a la nevera ]  
           I-have books in the refrigerator  
       b. Hi ha [SC MITWPLs al congelador ]  
           There are MITWPLs in-the freezer

For *to have*, one could in fact argue that there is affectedness w.r.t. the SC subject. In the case of existential constructions (which are a central case for the study of Partitive), their internal predicative Structure is a subject of debate: Moro (in several papers: (1988), (1991), (1992)) claims that the predicate in existential constructions is the raised locative (*there* in English). So it might turn out that the problem disappears once we have a better understanding of the construction.

I think Belletti's theory is appealing on empirical grounds. From the theoretical point of view, however, it is not clear that Partitive should be an inherent Case: in fact it does not look like a Case at all, if Case is conceived as a merely formal entity of no interpretative import:

---

<sup>21</sup> For these cases another constraint seems to be at work: individual-level SCs do not allow a partitive.

Partitive forces the DP receiving it to be indefinite.<sup>22</sup> For convenience, however, we shall adopt the essentials of Belletti's theory in the next chapter, leaving the question open of whether it could be reinterpreted in a different way still compatible with our theory.

In any event, if we adopt Belletti's theory together with the above proposal on CHAINS, the account for BG does not work as it stands, since an expletive-indefinite CHAIN would contain two Cases (Nominative and Partitive).

Another problem the above theory does not account for is subject inversion in languages such as Italian and its absence in languages like English. If subject inversion is analyzed as involving an expletive/Nominative CHAIN, then why are overt-expletive/inverted-subject CHAINS not possible?

A solution to both problems can be based on the following principles (which replace 0 above):

- (23) A CHAIN can not contain two structural Cases.
- (24) Null expletives do not require Case, while overt ones do.

0 and 0 together allow the three only cases of expletive CHAINS which are attested:

- expletive/CP (assuming that the CP does not require -or even accept- Case): there is one Case, which is assigned to and retained by the expletive:

- (25) *It strikes me that...*

- expletive/indefinite: the expletive is assigned Nominative (or whatever Case is assigned

---

<sup>22</sup> See Pesetsky (1982) for the contention that Partitive is restricted to certain positions because of its quantificational nature. See also Reuland & ter Meulen (1987) for discussion on the semantic/formal nature of the (In)definiteness Constraint.

in Spec of INFL) and the indefinite is assigned Partitive. The two Cases do not conflict according to 0 if Partitive is an inherent Case:

(26) *There* arrived *a man*

- null- expletive/Nominative (free subject inversion): the null expletive is assigned Nominative, but, since it does not require Case, Nominative can be transmitted to the inverted Subject:

(27) *pro* Lo farà *Gianni*  
it-will-do Gianni

## 2.4. Problems

The above theory crucially relies on a theory of expletives. In the next section, I will consider Binding Theory facts which appear to undermine such an approach. I will contend that some generalizations are not expressed by the preceding theory. Let us see, for the moment, some more immediate problems.



### 2.4.1. Empirical Problems

The theory presented in the preceding sections essentially tries to derive Burzio's Generalization by assuming that Spec of INFL always has to contain a member of a CHAIN, so that the absence of an EA will force the object to become a member of such a CHAIN. Assuming that in the absence of a projected EA the object is always the candidate chosen, this theory covers all cases that BG is intended to cover. This theory, however, makes a prediction which goes beyond BG: even in cases where BG is not at stake, the theory advanced above predicts that an INFL-CHAIN must be formed. Now consider a case where there is no EA and none of the internal Arguments apparently enters a CHAIN. Such cases are found in German with verbs like *schwindeln* ('to feel dizzy') or *grauen* ('to be afraid'), in which AGR is not coindexed with any Argument (see 0; the examples are taken from Cardinaletti (1990)):

- (28) a. dass (es) mir schwindelt  
that (it) me-DAT is-dizzy  
'that I feel dizzy'  
b. dass (es) mir davor graut  
that (it) me-DAT it-of fears  
'that I am afraid of it'

These constructions do not (necessarily) violate BG, since there is neither EA nor Accusative.<sup>23</sup> But they seem to falsify the claim, essential to the theory above, that there is always an INFL-CHAIN containing an Argument, unless we assume that oblique Arguments

---

<sup>23</sup> Some verbs of this kind subcategorize for an Accusative:

(i) dass (es) mich dürstet  
that (it) me-ACC is-thirsty

This is not necessarily in violation of BG, if we assume that this Accusative is inherent and, on the other hand, BG is only concerned with structural Case. German abides by BG in the general case.

(such as the Dative in 0 can be members of an INFL-CHAIN. Another case where our version of the EPP is violated is found in the German impersonal passive:

- (29) dass gestern getanzt wurde  
that yesterday danced was

Cardinaletti (1990) assumes that the kind of verbs in 0 involve a quasi-Argument. I think this assumption is problematic: we should expect that quasi-Arguments are the manifestation of some semantic property of a class of verbs (weather verbs, time verbs), and not a free option for other classes of verbs. So, for instance, with weather verbs the quasi-Argument roughly represents some atmospheric Cause Argument. No such abstract entity can be understood in the case of the verbs in 0. Since quasi-Arguments fall under the poverty-of-stimulus learning problem, it is reasonable to assume that they cannot vary from language to language, but rather that they are projected because of the semantics of the verb.

One possible way of handling the above problematic Cases is to loosen the conditions on CHAIN formation, so that we could claim that, in spite of appearances, there are INFL-CHAINS in the examples in 0 and 0. This, however, amounts to admitting that expletive-Argument CHAINS do not have a uniform pattern. In the case of 0, it we would have to admit that a CHAIN can be formed consisting of an expletive and an oblique Argument (for instance the Dative in these examples). In the case of impersonal passives -0-, it is not clear at all what the Argument in the CHAIN should be, for there appears to be none available. So, trying to maintain that expletives always involve a CHAIN in these cases leads to a weakening of the notion of CHAIN.

It seems therefore preferable to abandon the EPP as defined above.

### 2.4.2. Theoretical Problems

Another intrinsic problem for the hypothesis above is the requirement in 0, repeated here:

- 0 Case can be assigned only to the head of a maximal CHAIN.

In more classical accounts this requirement can be derived from the last resort character of DP-movement or, more generally, of CHAIN-formation: a CHAIN cannot be longer than necessary for it to obtain Case. But, as we pointed out, in the above account it is crucial that CHAINS are not formed because of Case requirements (Accusative is available in principle), but because of the EPP. So the requirement in 0 has to be merely stipulated.

A conceivable way to derive 0 would be the following: there is a kind of requirement to the effect that Nominative is 'preferred' to Accusative. So, to exclude 0 above, repeated here:

- 0     a. \*The problem developed  $t_{Acc}$   
      b. \*There developed the problem<sub>Acc</sub>

we could argue that within a CHAIN, whenever we can choose between Nominative and Accusative, Nominative has to be chosen. This strategy, however, does not work for infinitive sentences, where the head of the CHAIN is not assigned Nominative, but Accusative or Oblique (in ECM or 'for'-infinitive constructions) or no Case (in PRO-infinitives, if PRO does not bear any Case). So, the fact that the head of a CHAIN is not assigned a uniform Case makes it impossible to derive 0 from a Case-hierarchy strategy.

Another set of theoretical problems for the theory above come from Binding Theory. Since it is a crucial problem, we will treat it in a separate section.

### 3. Binding Theory

#### 3.1. Binding Theory and expletives

It has traditionally been noticed that a theory of expletives poses an immediate problem for Binding Theory (BT): if the expletive is coindexed with the Argument it is linked to, there should be a BT violation. There are several solutions that have been proposed for this problem. The classical solution (basically stemming from Chomsky's (1982) account for subject inversion) is co-superscripting: The link between the expletive and the Argument would not be co-subscripting, which is the device used for binding dependencies, but co-superscripting, which is, by assumption, irrelevant for BT.

It is clear that a theory of co-superscripting does not fit into a theory of CHAINS, if we reasonably assume that members of CHAINS should be uniformly coindexed in the same way: otherwise the concept of CHAIN would hardly be a unitary concept. On the other hand, as argued in Borer (1986) a theory having a single coindexing device is preferable, if tenable, to one using two different types of indexes.

In Chomsky (1986-b), a solution is suggested with a single indexing device, consisting in simply stating that Expletive/Argument binding relations do not count for BT:

(30) Binding of an Argument by a non-Argument is not subject to Binding Theory.

Obviously, this is a mere stipulation of no independent interest. Later in the same book, another solution is suggested: expletives are replaced at LF by the Argument they are linked to: if BT holds at LF, violations of BT by expletives are overridden:

(31) SS: *There<sub>i</sub> came a man<sub>i</sub>*

LF: *A man<sub>i</sub> came t<sub>i</sub>*

This solution does not account for the obligatory narrow scope reading of the indefinite w.r.t. negation in cases like:

(32) There *aren't many* linguistics freaks here

where *many linguistics freaks* cannot have wide scope w.r.t. negation. Since the LF representation Chomsky proposes is essentially the same as the SS in 0:

(33) Many linguistic students aren't here

where the wide scope interpretation is possible and preferred (or, for some speakers, exclusive), the theory of LF expletive-replacement makes wrong predictions in this connection.

To overcome this problem, Chomsky (1988) proposes an alternative approach: at LF the Argument does not replace the expletive, but rather adjoins to it. The theory of LF adjunction, Chomsky argues, would be a solution for the interpretation problem concerning the relative scope between Negation and the indefinite: if the Argument adjoins to the expletive as in 0 for an S-Structure like 0, there is no scope relation between the negative particle and the quantifier *many*, so that *many* can be assumed to have narrow scope, as desired:

(34) [<sub>DP</sub> there [ *many* l. students ] ] aren't t here

It is a mystery why *many* has to have narrow scope w.r.t. negation when there is no structural scope relation between them. In addition, it is doubtful that this is so in cases where there is no scope relation at SS:

(35) Pictures of *many* students aren't here

For many speakers 0 cannot be interpreted with *many* having narrow scope, while the narrow scope interpretation in 0 is clear and exclusive.

In conclusion, a theory LF expletive-replacement is problematic for scope interpretation, and a theory of adjunction to the expletive, conceived as a solution to this problem, does not work much better.

Both the theory of co-superscripting and the theories of LF-replacement/adjunction share a common idea: expletives would pose a problem for BT if some grammatical principle or process did not neutralize their BT effects. I think that this approach is suspect in the following sense. As noted by Borer (1986):

(36) Overt expletives never agree with the Argument they are linked to.

Typical overt expletives are singular (or adverbial, as English *there*) independently of the number feature of their Argument. We can conceive expletives as essentially uninflected elements, and then it is quite plausible that this fact alone is sufficient to exclude them from the scope of BT. If we assume that Binding involves sharing of phi-features, then expletives cannot bind, and no further stipulation is required. In other words, a theory which neutralizes BT effects in expletive constructions appears to be spurious because there is nothing to neutralize.

On the other hand, if expletives are not coindexed with the Argument in the same way as in Chains, the notion of CHAIN is considerably weakened: either we allow some other linking device (such as co-superscripting) for expletive CHAINS or we give up any linking device: in both cases, there is no unitary linking device for CHAINS.

In conclusion there seems to be a tension between a theory of CHAINS, which conceives

expletive/Argument links as having essentially the same nature as Chain links, and the generalization in 0, which rather suggests no linking between expletives and Arguments.

The theory we will introduce in the next chapter avoids these problems by:

- reducing the range of expletives to expletive-indefinite cases.
- assuming indefinites are defective in phi-features.<sup>24</sup>

### 3.2. A Binding Theoretical Approach to the Nature of Inverted Subjects

Now consider subject free-inversion. The standard analysis since Chomsky (1982) is that subject free-inversion involves an expletive *pro*. We will be considering an array of BT facts that appears to undermine the idea that expletives are involved in subject free inversion.

Let us assume that all Arguments in a sentence (including the EA) are generated inside (or close to) VP, as in Koopman & Sportiche's papers (1988)/(1990). Let us call I-subject (suggesting internal subject) the Argument in this basic position which is coindexed with AGR. We also assume that inverted subjects in Italian are Arguments in this basic position, as we proposed in Chapter 1, hence they are I-subjects.

Then we can pose the following question: What is the nature of I-subjects with respect to the features [ $\pm$ pronominal] and [ $\pm$ anaphoric]?

Within standard assumptions, at S-structure, I-subjects in English are always DP-traces (abstracting away from indefinite and CP Arguments), in other words they are null anaphors. In Italian, I-subjects can be DP-traces too, but, since null expletives licence inverted subjects, they can also be R-expressions (full DP's or variables) and pronominals (specifically overt pronominals). So, concerning the status of I-subjects, English would be a subset of Italian. Within standard assumptions, the status of I-subjects would be, abstracting away from indefinite

---

<sup>24</sup> It is probably not the case the approach I will take is essentially better fit than the hypothesis sketched above for dealing with this issue: this above hypothesis on CHAINS, which I reject for other reasons, is mainly conceived to set out the problems to be solved.

and sentential subjects:

(37) BT-status of I-subjects (standard assumptions):

		<u>English</u>	<u>Italian</u>
- anaphors	null	+	+
	overt	-	-
- pronominals	null	-	-
	overt	-	+
- R-expressions	null	-	+
	overt	-	+

Let us assume the contrast English/Italian is representative of the contrast NSLs vs. non-NSLs. The distribution of values in 0 is derived from the standard theory on expletives: Italian would be like English if null expletives did not allow a wider range of I-subjects. Suppose we try to redefine the distribution in 0 by tentatively assuming the following generalization:

(38) I-subjects are [-anaphoric] in NSLs and [+anaphoric] in non-NSLs.

What are the consequences of this assumption? The new picture that emerges is expressed in the following table:



(39) BT-status of I-subjects (according to 0):

		English	Italian
- anaphors	null	+	-
	overt	+	-
- pronominals	null	-	+
	overt	-	+
- R-expressions	null	-	+
	overt	-	+

0 differs only in three respects from the table of values in 0:

- (40)
- a. overt anaphors are possible I-subjects in non-NSLs.
  - b. (null) anaphors are not possible I-subjects in NSLs.
  - c. null pronominals are possible I-subjects in NSLs.

Is there any evidence for these predictions? We start with prediction 0.a). Consider a sentence like 0:

(41) John has done it himself

We could take what is traditionally called emphatic anaphors as an instance of I-subject overt anaphor. Emphatic anaphors have been traditionally considered non-Arguments.<sup>25</sup> But in fact, constructions such as 0 share some properties with inverted subjects in Italian. 0 has an

---

<sup>25</sup> As in Burzio (1986:102). In the next chapter (section 2.4.) we will in fact consider the possibility these emphatic elements are not really in an A-position. But this does not affect the argumentation that follows, even if the idea that these elements are in the same position as inverted subjects is only an approximation.

interpretation similar to the Catalan sentence:

(42) Ho ha fet ell

It-has done he ('HE has done it')

in that both *himself* and *ell* have an emphatic (probably Focus) interpretation. On the other hand, both occur in a position that can be roughly characterized as VP-final.<sup>26</sup>

The obvious objection one can pose to the contention that *himself* is the I-subject in 0 is that this element is likely to occupy an A'-position simply because it has a kind of adverbial interpretation, roughly paraphrasable as 'in the flesh', 'by himself' or 'alone'. However, if we consider the Catalan equivalent to 0, namely 0:

(43) En Joan ho ha fet ell (*mateix*)

The Joan it-has done he (SELF)

we can see that the element that is used instead of *himself*, namely *ell* (*mateix*) is in complementary distribution with the inverted subject:

---

<sup>26</sup> American speakers seem to allow empathic anaphors in non-final position (thanks to B. Schwartz and E. Pierce for pointing this out to me), :

(i) John has himself done it.  
while British speakers seem not to allow this word order.

Parallellly, some Romance languages (Spanish, Romanian) allow inverted subjects in non sentence-final position (the order being VSO), while others do not (Italian, Catalan). Since I am going to assimilate empathic anaphors to inverted subjects (both being I-subjects), I think that this kind of variation is of the same nature: some languages allow I-subjects only as right VP-adjuncts and others allow them as left VP-adjuncts. I will not provide any explanation for this contrast. However, see below for Italian.

(44) a. \*Ho ha fet en Joan *ell* (*mateix*)

It-has done the J. he (SELF)

b. \*Ho ha fet *ell* (*mateix*) en Joan

If instead of *ell* (*mateix*) we use an adverbial expression like *en persona* 'in person', which has roughly the same interpretation, no such complementary distribution arises:<sup>27</sup>

(45) Ho ha fet en Joan *en persona*

It-has done the J. in person

This strongly suggests that emphatic elements such as *ell* (*mateix*) (and, likely, its English construction-mate *himself*) do not have the same distribution of an adverbial, but rather that of an inverted subject. One conclusion of this thesis will be that all languages have 'subject inversion', if overt anaphoric I-subjects were to be termed inverted subjects. To avoid confusion, we will keep to the traditional terminology, and call inverted subjects only overt [-anaphoric] I-subjects. This is an informal term, the actual theoretical term being I-subject which is neutral w.r.t. both the features [ $\pm$ anaphoric] and [ $\pm$ overt].

Now let us consider 0-b). The prediction is that anaphoric I-subjects are not possible in Catalan. This prediction is also apparently fulfilled:

---

<sup>27</sup> **en Joan** and **en persona** do not necessarily form a constituent:

(i) Ha vingut el ministre a dinar EN PERSONA

Has come the minister to lunch in person

'The minister has come to have lunch IN PERSON'

but they undoubtedly can:

(ii) [**Error! Marcador no definido.DP;Error! Marcador no definido.** El ministre en persona ] i [**Error! Marcador no definido.DP;Error! Marcador no definido.** la seva dona ]

The minister in person and his wife

The latter fact does not falsify the contention that **in person** is in a A'-position: a subconstituent of an Argument in an A-position is not (necessarily) an Argument.

- (46) \*En Joan ho ha fet si mateix  
 The J. it-has done SE SELF  
 'Joan did it himself'

As we have seen, only a pronominal (*ell (mateix)*) can be used in this construction.

Is the contrast 0/0 really significant? According to 0, this contrast should generalize to a contrast between NSLs/non-NSLs. The prediction is that, for this type of constructions, non-NSLs will have emphatic anaphors, while NSLs will not: instead, they have emphatic pronominals. We will see directly that this prediction is borne out in a wide variety of languages, although some qualifications will be required. Concerning NSLs, all of the following use emphatic pronominals instead of anaphors.<sup>28</sup>

- (47) Peter had this work done by a lawyer, but...
- a. en Joan l'ha fet *ell (mateix)*/\**si mateix* (Catalan)  
 the J. it-has done he (SELF) / SE SELF
  - b. Joan l'a fait *el (mateis)*/\**se mateis* (Occit.)<sup>29</sup>  
 J. it-has done he (SELF) / SE SELF
  - c. Juan lo ha hecho *él (mismo)*/\**si mismo* (Spanish)  
 J. it-has done he (SELF)/ SE SELF
  - d. Gianni l'ha fatto *lui (stesso)*/\**se stesso* (Ital.)<sup>30</sup>

---

<sup>28</sup> Thanks to Patrick Sauzet (Occitan), Luigi Giuliani (Italian), Joana Louro (Portuguese), Virginia Motapayane (Romanian), Josep Quer (Greek) and Itziar Laka (Basque) for the data and comments.

<sup>29</sup> In modern colloquial Occitan **mateis** has been replaced in most dialects by the French borrowing **même**. This is a merely lexical fact.

<sup>30</sup> Some speakers reported to me that in Italian this construction with emphatic **lui (stesso)** is much more emphatic than in Catalan, and that the alternative use of **da solo** 'by himself' is more natural. In fact **da solo** is not, I think,

- G. it-has done he (SELF) / SE SELF
- e. O Jo~ao fê-lo *ele (mesmo)/\*si mesmo* (Portugu.)  
The J. did-it he (SELF) SE SELF
- f. Ion a scris *el însuși* acest proiect (Romanian)  
I. has written he SELF this project
- g. o jánis to káni o *ídjos / \*o eafτός* tu (Greek)  
the J. it-did the HE-SELF/ the self of-his )
- h. Jon, ordea, *berak / \*bere buruak* egin du (Basque)<sup>31</sup>  
J. instead HE-SELF/ his self done has

In all the above examples, the reflexive forms are unacceptable as emphatic I-subjects.<sup>32</sup>

The acceptable emphatic I-subjects in 0 are possible in contexts where they are not bound

---

intensionally equivalent to the emphatic I-subject (although it is extensionally equivalent in most pragmatic situations): **da solo** implies 'with no help', while the emphatic subject does not necessarily exclude 'help'. It rather stresses that the action has not been delegated to someone else.

These speakers also report that adding **stesso** to the emphatic subject **lui** is not very natural (at least in their dialect). See below for discussion on the position of **lui** and **lui stesso** as emphatic elements.

<sup>31</sup> **Jon berak** is actually a possible constituent, but not necessarily: I put the **ordea** 'instead' element in between to clarify the example. **Berak**, on the other hand, is not an anaphoric element in modern Basque, but rather a logophoric / emphatic expression, as is usual with emphatic I-subjects.

<sup>32</sup> Hebrew allows neither emphatic anaphors nor emphatic pronominals as I-subjects. In fact, this language does not allow pronominals as inverted subjects (thanks to Tali Siloni for the data and comments). Perhaps the intermediate status of this language as a NSL (however it should be properly characterized) is the reason for this situation.

Hungarian does not provide clear examples relevant for the theory either, perhaps because emphatic I-subjects, as Focus elements, should occupy the obligatory Focus position in this language, which is preverbal and, likely, not an I-subject (it would be the specifier of some (Focus) FC).

I think that a more detailed study would be necessary to extend the present theory to these languages.

(preverbal subjects, dislocated positions, complement positions).<sup>33</sup> I give examples to show that in Romance the *mateix/mismo/stesso/mesmo* element does not turn the preceding pronoun into any kind of anaphor, but simply into an emphatic or logophoric pronominal. From now on we will use the term SELF (in capitals) for these elements. In some of the examples I provide some possible context in the translations only to suggest what SELF adds to the meaning:

### Catalan

- (48) a. *Ell mateix* no ho farà  
 He SELF not it will-do  
 'He himself will not do it' (-> His lawyer will)
- b. *A ell mateix*, no l'he vist  
 To him SELF not him-have-I seen  
 'Him (himself) I haven't seen' (->I saw his lawyer)
- c. No he parlat amb *ell mateix*  
 Not have-I spoken with him SELF  
 'I haven't spoken to him himself' (->but actually to his wife)

---

<sup>33</sup> Italian seems to allow **lui stesso** in non-final position (thanks to L. Rizzi for this remark), i.e., in a position where inverted subjects are not possible:

(i) Gianni ha lui stesso fatto questo  
 G. has he SELF done this

This seems to suggest that the emphatic I-subject **lui stesso**, beyond occupying the subject basic position, can raise to some intermediate FC Specifier. We can still assume, however, that, whenever it appears in final position, it does occupy the same position as inverted subjects since they cannot cooccur:

(ii) \*Lo ha fatto Gianni lui stesso  
 It-has done G. he SELF

### Spanish

- (49) Ella misma estaba allí  
She SELF was there  
'She herself was there'

### Italian

- (50) Lui stesso era nella riunione  
He SELF was at-the meeting  
'He himself was at the meeting'

### Portuguese

- (51) Ele mesmo pode venir  
He SELF can come  
'He himself can come'

What SELF adds to a pronominal form is either emphasis or logophoricity. To the extent that emphatic or logophoric elements are not required to be bound (but rather to have a discourse prominent antecedent) they are pronominals in the technical sense of BT. The view that such elements are actually pronominals in the technical sense of BT ([+pronominal,-anaphoric,]) appears to be often challenged in the literature. Some authors contend that they are a sort of long distance anaphors or at least suggest that binding by an antecedent is a licensing condition.<sup>34</sup> I think this a misleading tack to take. These authors start by assuming or suggesting that these elements have to be bound by an antecedent outside their Binding Domain, on the basis of examples like:

---

<sup>34</sup> See Bickerton (1987) for English **he himself**; Iatridou (1986) for Greek **o idhios**.

- (52) a. *O Yanis<sub>i</sub> theli i Maria na voithisi ton idhio<sub>i/\*j</sub>*  
 The J. wants the M. that helps him-himself  
*'Yanis wants Maria to help him himself'*  
 (Greek, Iatridou (1986:767))
- b. *John<sub>i</sub> told Bill<sub>j</sub>'s sister that he himself<sub>i/\*j</sub> had been arrested*  
 (Bickerton (1987:346))

Then, they point out that actually there are exceptions, (Iatridou (1986) treats them as only apparent). I think that the fact that these elements often appear in constructions where they are bound is only an epiphenomenon which should not be granted theoretical status. If these elements are logophoric or emphatic (see Zribi-Hertz (1990-a/b)), they require a discourse-prominent antecedent (a subject of conscience, when logophoric). Of course, if one introduces examples out of the blue, with no context, as in 0, then the most prominent element in the discourse will be the main clause subject, or at least a preceding DP, for there is no other available antecedent unless one makes up a plausible context having one. So I think the optimal theory is one treating logophors and emphatic pronouns as pronominals in the BT technical sense, leaving the account for their often bound status to discourse grammar (i.e., prominent discourse antecedents can happen to be present in the sentence and even c-commanding the emphatic/logophoric element).<sup>35</sup>

There is in addition a strong theoretical argument against the claim that logophoric/emphatic elements consisting of a pronoun and an adjoined SELF element are (long distance) anaphors. It appears that the distribution of SELF is neutral w.r.t. the BT status of the host. So in German we have:

---

<sup>35</sup> I think that the distinction between logophoricity and (referential) emphasis could easily be reduced to a single concept, the distinction being then a matter of meaning nuance or vagueness.



- (53) a. sich selbst (SE SELF) (emphatic anaphor)  
 b. er selbst (he SELF) (emphatic pronoun)  
 c. Hans selbst (Hans SELF) (emphatic R-expression)

In other languages (English, French) the facts are less clear because the SELF element has become an affix and the whole pronoun-SELF element has become both a SELF element (*John himself*) and a not necessarily emphatic reflexive (*John shaves himself*), and even a logophor (*John thought that Peter would take a picture of himself!*).

In the above paradigm we have overlooked an important fact: in some of the languages (Catalan, Occitan, Romanian), the colloquial forms used as reflexives are the emphatic/logophoric elements (*ell mateix*, *el mateis/même*, *el însusi*) or even the bare pronominals (Catalan *ell*, Occitan *el*). Therefore the colloquial versions of these languages are neutral w.r.t. our prediction that I-subjects have to be pronominal and not anaphoric, because the (emphatic) pronominals are used as neutral forms (pronominal/anaphoric). But in all three languages speakers using the unambiguously reflexive forms have the clear intuition that these forms are completely excluded as I-subjects.

Non-NSLs should instead have emphatic anaphors in constructions equivalent to the ones in 0, as predicted by the generalization in 0. Let us consider the following examples:<sup>36</sup>

---

<sup>36</sup> Thanks to Michael Laurence (German), Liliane Haegeman (West Flemish and Dutch), Sten Vikner (Danish), Kjell-Åke Gunnarson (Swedish), Haróld A. Sigurdhsson (Icelandic) and Alain Rouveret (French) for their data and comments.

(54) Peter had this work done by a lawyer, but...

- a. John did it *himself* (English)
- b. Hans hat es *selbst* gemacht (German)<sup>37</sup>  
H. has it SELF done
- c. Hans har gjort det *selv* (Danish)<sup>38</sup>  
Hans has done it SELF
- d. John har gjort det *själv* (Swedish)<sup>39</sup>  
Hans has done it SELF
- e. Jón gerðhi þetta *sjálfur* (Icelandic)  
J. did that SELF<sub>Nom</sub>
- f. Jean l'a fait *lui(-même)* (French)  
J. it-has done he (himself)

The analysis for the above examples is far from trivial. We will see, however, that under a reasonable interpretation it supports the generalization in 0: specifically, I-subjects are [+anaphoric] in non-NSLs.

The French example might at first glance look like a counterexample, since the emphatic element looks like a pronominal (*lui*), or like an element similar to the emphatic/logophoric

---

<sup>37</sup> Dutch has essentially the same behavior as German:

(i) Ik heb het eten nie zelf gekookt  
I have the meal not self cooked  
'I haven't cooked the meal myself'

<sup>38</sup> Danish allows emphatic **selv** not only in final position, but also in pre-VP position, as the following example shows:

(i) ...at Hans måske selv har gjort det  
that H. maybe SELF has done it

The same happens with other Germanic languages. As far as this position can be argued to be a low ('internal subject') position, this is not problematic. Otherwise, it can be that floating can take place at intermediate positions in DP-raising to Spec of AGR.

<sup>39</sup> Norwegian behaves exactly like Swedish.

element in the Romance NSLs (*lui-même*). But French strong pronominals function as anaphors,<sup>40</sup> as is clear from the following example:

- (55) Jean parle de *lui* (-*même*)  
 J. speaks about *lui* (-*même*)  
 'J. speaks about himself/him (himself)'

Therefore, we can argue, our prediction that French uses anaphoric (reflexive) elements in these constructions is not falsified, it is only vacuously fulfilled (obviously, *emphatic* I-subjects cannot be reflexive *clitics*, because clitics cannot be emphatic). However, there is positive evidence that our claim is accurate: in French, when the subject is of a certain quantificational type (*tout le monde* 'everybody', *chacun* 'each person', *on* 'one', *personne* 'nobody', *nul* 'no', PRO<sub>Arb</sub>, etc.), the strong reflexive is *soi(-même)* (see 0.a)), and correspondingly the emphatic I-subject is *soi -même* (see 0.b)):

- (56) a. *On a parlé de soi(-même)*  
*On* has spoken about *soi(-même)*  
 'One has spoken about oneself'  
 b. *On l'a fait soi(-même)*  
*On* it-has done *soi(-même)*  
 'One has done it oneself'

---

<sup>40</sup> See Zribi-Hertz (1990-a/-b) for a more accurate characterization of **lui(-même)** elements. Zribi-Hertz (p.c.) points out to me that simply saying these elements are anaphoric (and pronominal at the same time) is too simplistic. I agree, but my essential point is that whatever is used as a reflexive anaphor in object position will be also used as an anaphoric I-subject. I can remain neutral w.r.t. what apparently pronominal/anaphor-neutral elements actually are. I will nonetheless suggest an account for the existence neutral forms in Chapter 4, section 2.1.2.

So French supports our hypothesis in an interesting way: it has two different elements used as non-clitic reflexives (*lui(-même)/soi(-même)*) depending on the nature of the subject; correspondingly, it has two different emphatic I-subject forms, as expected if I-subjects are anaphoric in this language.

Now, let us consider the Germanic cases. The emphatic I-subject (*selbst/selv/etc.*) is not actually the anaphoric (reflexive) element in these languages, but rather the second element of the compound anaphors these languages use: *sich selbst* (German), *sig selv* (Danish), etc., namely, a SELF element.

I think, however, that the generalization in 0 can be maintained for these cases. Let us see how.

The German and Danish emphatic I-subjects, and even the English ones, are likely to be floating elements. On the one hand, as SELF ELEMENTS, they can be adjoined to an overt DP, to which they add emphasis:

- (57) a. John himself/ he himself (English)  
 b. John selbst / er selbst (German)  
 c. John selv (Danish)  
 d. Jón sjálfur (Icelandic)

with a uniform interpretation in all cases. In 0 we exemplify the English use of emphatic *himself*, which is similar to the use of *selbst/selv/etc.* in the other Germanic languages:

- (58) a. I talked to John's wife but I didn't talk to *John himself*.  
 b. This book does not address *the problem itself*, but rather its consequences.  
 c. I didn't complain to *her herself*, but rather to her secretary.

This suggests that in these languages, the emphatic element in 0.a/b/c) is a SELF element<sup>41</sup> left floating by DP-movement to Spec of AGR. So the structure of 0.a) would be 0.b) (and the same could be claimed for the other two languages):

- (59) a. John has done it himself  
 b. [AGRP John<sub>i</sub> [AGR' has done it [DP t<sub>i</sub> himself] ]]<sup>42</sup>

We could call the emphasis added to a DP by SELF elements 'referential emphasis': in *John himself*, what is emphasized is that we are referring precisely to *John*. That is why these elements cannot be adjoined to a non-referential DP (whatever 'non-referential' means): *\*everybody herself*, *\*nobody himself*, etc. For some reason, however, when these elements are separated from the element they are construed with they can be used with non-referential DPs:

- (60) *Everybody/nobody* did the work *herself/himself/themselves*

Therefore, if, as we argued, these elements are floating elements in 0, floating is relevant

---

<sup>41</sup> The etymology could be misleading here: English **himself** in **John himself** is a SELF element, whereas **self** alone is not (**\*John self**). The same is true for French **lui-même** (**Jean lui-même**, **\*Jean même**). I think that the fact that the spelling of **self/même** indicates attachment to the pronoun is significative (as is often the case with spelling): these elements have lost their original status of independent morphemes (English **self** can in addition be a prefixed form -as in **self** (**\*balanced criticism**)).

<sup>42</sup> So (i) and (ii) differ in that in (i) **himself** has not been left floating, while in (ii) it has:

- (i) John himself has done it  
 (ii) John has done it himself

The interpretation of (i) and (ii) is not the same (as S. Vikner pointed out to me). I think this interpretative difference should be derived from the Focus interpretation that is associated with the sentence-final **himself**, which is lacking in (i).

for interpretation, in that it allows 'referential emphasis'. The *ec* accompanying the floating element is, however, an empty anaphor. Therefore, we could argue that a) these floating elements cannot add referential emphasis to a non-referential DP by S-structure, and b) empty anaphors, even if bound by a non-referential DP, count as referential.

Thus, it is likely that our initial idea that English has overt anaphoric I-subjects may be false: a null anaphor is always involved. This fact, however, does not falsify the generalization 0, which predicts that I-subjects for non-NSLs have to be anaphoric: in those languages, non-null I-subjects may be anaphoric either as overt anaphors (this is the case in French, as we will see), or as null anaphors with a floating emphatic element (Germanic languages). In West Flemish,<sup>43</sup> the claim that the emphatic element is a floating one, and not a full anaphor itself, is even more obvious than in other Germanic languages, because the SELF element (*zelve*) is not used in reflexive constructions (where the reflexive element is *zen eigen* 'his own' or a weak pronominal is used). Like in German, *zelve* is an emphasizer that can be adjoined to a DP:

- (61) a. da *Jan zelve* t-eten gisteren nie gekookt eet  
           that *J. SELF* the-food yesterday not cooked has  
           '... that Jan himself didn't cook the meal yesterday'
- b. da *Jan* t-eten gisteren nie *zelve* gekookt eet  
           that *J.* the food yesterday not *SELF* cooked has  
           '... that Jan didn't cook the meal himself yesterday'

Now let us consider French in the light of the preceding analysis. Are French I-subjects

---

<sup>43</sup> Thanks to L. Haegeman for the examples and the discussion. Like in other Germanic languages, the floating element **zelve** in West Flemish appears not only in a low (VP) position (to the right of negation and preceding the participle), but in positions more to the left (thus higher). If it is a floating element, as we contend, and subject raising is not in one step, it would be possible for the floating element to be left floating at any of the intermediate steps, as we suggested for Danish and other Germanic languages.

analyzable as floating elements? The answer seems to be that some are and some are not. Consider first *lui-même*. Like English *himself* or Germanic *selbst/selv/etc.*, *lui-même* is a SELF element that can be adjoined to a DP to add referential emphasis to it:

- (62) a. Ça concerne pas *le problème lui-même*, mais plutôt...  
 This concerns not the problem SELF, but rather...  
 'This doesn't concern the problem itself, but rather...'
- b. J'ai pas parlé avec *lui lui-même*, mais avec sa femme  
 I-have not spoken with him SELF but with his woman  
 'I haven't talked to him himself, but rather to his wife'

So it is likely that *lui-même* as I-subject is a floating element. This analysis, however, cannot be generalized to the other two I-subjects (*lui* and *soi(même)*), since they cannot be used as DP-adjuncts:<sup>44</sup>

- (63) a. \**Jean lui* l'a fait  
 J. he it-has done  
 (Cf. *Jean l'a fait lui*)
- b. \**Chacun soi(-même)* l'a fait  
 Each SE (himself) it-has done  
 (Cf. *Chacun l'a fait soi(-même)*)

---

<sup>44</sup> Even for **lui-même**, its floating status could be questioned in cases where the subject is a weak pronominal, because weak pronominals do not allow adjoined elements:

- (i) Il l'a fait lui-même  
 He it-has done he-self
- (ii) \*[Il lui même] l'a fait

But this might be simply due to cliticization requirements on weak Nominative pronominals. As for **soi(-même)**, the argument is not conclusive for the reason we mentioned above: by S-structure, an emphatic element cannot be adjoined to a non-referential DP.

The conclusion is then that French can display three types of I-subjects:

- empty anaphors [DP t ]
- empty anaphors + floating SELF: [DP t lui-même ]
- overt anaphors: [DP lui/soi(-même) ]

Germanic languages seem to allow only the first two possibilities. We will provide an account for this difference in the next chapter.

\* \* \*

0 predicts, on the other hand, that NSLs do not allow null anaphors (DP-traces) as I-subjects, so they will not allow floating emphatic elements, if floating is a result of movement.<sup>45</sup> So, while English allows both 0.a) and b), Catalan and NSLs (which have SELF elements adjoined to DP's) only allow 0.c), not 0.d):

- (64) a. *John himself* will do it  
 b. *John* will do it *himself*  
 c. *En Joan mateix* ho farà / Ho farà *en Joan mateix*  
     The J. SELF it-will-do/It-will-do the J. SELF  
 d. <sup>\*\*</sup>*En Joan* ho farà *mateix*  
     The J. it-will-do SELF

Similar facts hold for all the NSLs considered (Spanish, Italian, Portuguese):

---

<sup>45</sup> What happens, then, with floating quantifiers in NSLs? Within the present theory, I have to assume that they are not left floating by a moved preverbal subject. This is what we suggested in Chapter 1. In the next chapter we will argue that FQs are not left floating by movement in NSLs, but rather material adjoined to a **pro**.



- (65) a. *Juan mismo lo harà* /\*\**Juan lo harà mismo*  
 J. SELF it-will-do J. it-will-do SELF
- b. *Lui stesso lo farà* /\*\**Lui lo farà stesso*  
 He SELF it-will-do G. it-will-do SELF
- c. *Ele mesmo fă-lo* /\*\**Ele fă-lo mesmo*  
 He SELF did-it He did-it SELF

So, to summarize, non-NSLs allow, as I-subject, either an overt anaphor or a floating emphatic element cooccurring with an empty anaphor. NSLs allow neither, because they do not allow anaphoric I-subjects.

The conclusion is, instead, that null I-subjects in NSLs are pronominal, a controversial conclusion, to which I return later.

Romanian is potentially problematic as a NSL: although *el însuși* is apparently the same as Italian *lui stesso* (he SELF), it is not equivalent: the SELF element is actually *el însuși* in cases like 0.a):

- (66) a. [<sub>DP</sub> Ion el însuși ] (cf.: \*\*Ion însuși)  
 Ion himself
- b. Ion a scris el însuși acest proiect  
 I. has written *el însuși* this project

Thus *el însuși* in 0.b) could be analyzed as [<sub>DP</sub> t el însuși ], so that Romanian would not have reflexive I-subjects (*și însuși*) but it *would* have empty anaphoric I-subjects with a floating SELF. However, since these elements are often ambiguous (e.g., English *himself* is ambiguous between reflexive and SELF), we can perfectly contend that *el însuși* is ambiguous between a SELF element and an emphatic/logophoric element. A more detailed study should address the issue of how these ambiguities are possible and so frequent and why they arise in natural

language.

Russian and Georgian raise a similar problem: in these languages there is SELF element (resp.: *sam/tuiton*) which can be adjoined to a DP:

- Russian: Vanja/on sam : John/he SELF

- Georgian: Vanom/man tuiton : John/he SELF

This element can float when construed with the subject:

(67) a. Russian: Vanja sdelal eto sam

V. did it SELF

b. Georgian: Vanom gaalata es tuiton

V. did it SELF

Actually Russian is not a full-fledged NSL. Georgian is a strongly non-configurational language (*tuiton* can in fact float almost anywhere, the only restriction being that it cannot non-adjacently precede the DP it is construed with). So it might be that there is something to these languages which cannot be captured in our analysis. In any event, our claim is that [<sub>DP</sub> *t* SELF] (where *t* is a trace) is not a possible I-subject in a NSL is not strictly falsified in Russian or Georgian, for *sam/tuiton* can cooccur with a null subject in preverbal position:

(68) a. Sam sdelal eto / Sdelal eto sam

SELF did it Did it SELF

'He himself did it/'He did it himself'

b. Tuiton gaalata es / Gaalata es tuiton

SELF did it Did it SELF

'He himself did it/'He did it himself'

We could then claim that *sam/tuiton*, when I-subjects, occur in the configuration [<sub>DP</sub> *pro*

sam/tuiton ]. When they are preverbal (perhaps this is not significant in Georgian) they would be licenced with a resumptive strategy, as we will claim is the case for preverbal (Spec of AGR) subjects in NSLs.

Finally, there is a language that provides evidence pointing to a possibility not contemplated so far: Modern Standard Arabic (MSA).<sup>46</sup> In this language, inverted subjects appear between the verb and the complements (VSO) in the unmarked case (while preverbal subjects are likely to be left-dislocated elements). It is reasonable to assume that VSO is an instance of 'subject inversion', with V<sup>0</sup> raised to INFL<sup>0</sup> and the post-verbal subject left adjoined to VP.<sup>47</sup>

Since MSA is a NSL, our prediction that I-subjects shall be [-anaphoric]. And indeed, I-subjects in the VSO word-order are:

(69) 'Amala Yusuf/huwa/\*nafsu hu al-'amal

Did Y. /he / himself the work

'Yusuf/he did the work'

But the facts are a bit more complex. If we translate 'Joseph/he did the work himself' into MSA, we have the following sentence:

---

<sup>46</sup> Thanks to Elisabet Nebot and Isabel Herrero (Un. of Barcelona) for the data and comments. They are not native speakers (hardly anybody is, for MSA is never a colloquial language, although it is closer to Saudi dialects than to Western dialects). In any event, their judgements sounded steady and confident, giving the impression that they were asked about trivial and well-known facts.

<sup>47</sup> In addition, MSA subject inversion would be similar to subject inversion in some Northern Italian dialects, where subject inversion blocks number agreement in the third person. We will consider this fact in the next chapter.

(70) 'Amala Yusuf/huwa al-'amal nafsuhu

Did Y. /he the work himself-NOM

'Yusuf/he did de work himself'

Like English *himself*, *nafsuhu*(NOM)/*nafsahu*(ACC) is both a SELF element and an unambiguous reflexive (i.e., it has to be either bound, adjoined to a DP or floating):

(71) a. [DP [DP Yusuf/huwa ] nafsuhu ]

Y. /he SELF<sub>Nom</sub>

b. Karaha Yusuf nafsahu

Hated Y. himself<sub>Acc</sub>

'Yusuf hated himself'

So, to some extent, we could say that MSA has [+anaphoric] I-subjects. In fact, it appears that MSA has pre-VP [-anaphoric] I-subjects, and post-VP [+anaphoric] I-subjects. We can interpret this in the following way: the subject basic position is post-VP, but in MSA there is a 'second internal subject position'. Suppose this means that this is a derived position (so not really the 'internal subject position' proper), but low enough to have essentially the same properties as an I-subject w.r.t. BT. In fact, as we pointed out in Chapter 1, whether inverted subjects are really in the subject basic position or in a close higher up position is not clear, and will not really matter for the purposes of this thesis. In the next chapter, the facts in MSA will follow straightforwardly from the theory, with minimal additional theoretical cost w.r.t. the account for other languages.

The facts in Arabic might be reminiscent of some facts in Italian we have overlooked thus far. In Italian, some speakers report,<sup>48</sup> *lui* when used as an emphatic I-subject precedes the object, while *lui stesso* follows it:

---

<sup>48</sup> G. Longobardi's judgements were clear cut.

- (72) a. Gianni ha fatto (*lui*) il lavoro (*\*lui*)  
G. has done *he* the job *he*
- b. Gianni ha fatto (*\*lui stesso*) il lavoro (*lui stesso*)  
G. has done *he SELF* the job *he SELF*

These facts sound familiar if we have MSA in mind, in that they involve two different kinds of I-subjects depending on the pre- or post- VP position of the I-subject. But they are not the same for obvious reasons: *lui stesso* is neither a SELF element (rather *stesso* is) nor a reflexive anaphor in object position. And neither the pre-VP position for *lui* allows [-anaphoric] I-subjects not being emphatic (subject inversion is post-VP in the general case), nor does the post-VP position allow true anaphors (*se stesso*/ [<sub>DP</sub> t stesso ]). Thus we are probably missing some basic factor that MSA and Italian share without sharing the essential properties considered here.

Summing up, we would characterize MSA as having two 'I-subject' positions, only one of which is a  $\theta$ -position. Perhaps for head-initial languages all  $\theta$ -positions are to the right of V, and pre-VP I-subjects (as in MSA, Romanian, Spanish) are always derived by short movement:

- (73) AGR<sup>o</sup><sub>i</sub> ... (I-subject) [<sub>VP</sub> V (I-subject) ]  
-θ +θ

What determines which of these I-subject positions are 'active' (e.g., can be overt) is not clear: Catalan, Italian, French, many English dialects use the post VP position, Romanian uses the pre-VP position, Spanish can use both in essentially the same way, and MSA uses each in a different way. I will not try to account for this type of variation, which might involve low-level parameterization.

In the next chapter we will argue for a theory that derives BG and, at the same time,

accounts for the Binding Theory facts presented above.

#### 4. Summary

In this chapter we have tried to derive Burzio's Generalization within more or less standard theories. We have been crucially involved in subject inversion structures, because these are among the most problematic cases for deriving BG and, in fact, are problematic of themselves too. We have met problems with BT and their possible solutions (esp. expletive replacement) and with Case Theory (which, in trying to derive BG, cannot anymore be conceived as imposing a last resort strategy in search for Case).

Moreover, we have proposed a descriptive generalization on the binding-theoretical nature of I-subjects which, under some reasonable qualifications, seems to hold true across many languages. According to this generalization, non-NSLs have the options in 0 as possible I-subjects (some of them only have 0.a) and 0.b)), whereas NSLs have the options in 0 (examples in French and Catalan):

(74) I-subjects in non-NSLs:

a. empty anaphor:

Jean a fait le travail [ $DP\ t$ ]

J. has done the work

b. empty anaphor + floating SELF:

Jean a fait le travail [ $DP\ t\ \text{lui-même}$ ]

J. has done the work SELF

c. overt anaphor:

Jean a fait le travail [ $DP\ \text{lui}$ ]

J. has done the work himself

(75) I-subjects in NSLs:

a. empty pronominal:

(En Joan) ha fet la feina  $[_{DP\ pro}]$

(The J.) has done the work

b. overt pronominal (+plus SELF)

(En Joan) ha fet la feina  $[_{DP\ ell\ (mateix)}]$

(The J.) has done the work he (SELF)

c. empty R-expression:

Qui dius que ho ha fet  $[_{DP\ t}]$

Who say-you that it-has done

d. overt R-expression:

Ha fet la feina  $[_{DP\ en\ Joan}]$

Has done the work

Probably we could add to 0 the option: empty pronominal plus a SELF element ( $[_{DP\ pro\ SELF}]$ ), which would be available in Russian, Georgian and possibly Romanian.

Standard accounts for subject inversion have little to say about this generalization, as far as I know. The next chapter is devoted to deriving this generalization and BG.





## **Chapter 3**

### **AGR and Subjects**

#### **1. On the Nature of the AGR-Subject Dependency**

In the preceding Chapter we tried to derive Burzio's Generalization (BG) from a version of the EPP plus a theory of CHAINS. There were some problems the theory faced. Let's recall the three basic problems we pointed out:

a) It appears not to be always the case that the expletive forms a CHAIN with an Argument, so that our version of the EPP would be too restrictive.

b) The requirement that Case be assigned to the top of the CHAIN had to be stipulated and could not be derived from a last resort principle. In the account below, we will propose that the I-subject, once coindexed with AGR, has to obtain Case from AGR, either from  $AGR^0$  (as in NSLs) or from Spec of AGR (as in non-NSLs). The generalization is, roughly speaking, that the I-subject has to obtain Case from an AGR element which is richly specified in phi-features, namely  $AGR^0$  in NSLs or Spec of AGR in non-NSLs.

c) The link between expletive and Argument was problematic for Binding Theory, but nevertheless necessary for the CHAIN theory to make sense.

#### **1.1 A Reformulation of the EPP**

The proposal that I will advance has very much in common with Borer's (1986) theory of I-subjects, but, at least in its initial formulation below, it is more restrictive in the way AGR (INFL in Borer's terms) is coindexed with a I-subject.

In all this section we will be abstracting away from expletive/indefinite constructions,

which we deal with in the next section.

Let us tentatively assume the following principle:

- (1) At D-structure, AGR must be coindexed with the most prominent non-(inherently-)Case-marked DP or CP in its c-command domain.

This rule is similar to Borer's (1986) rule of AGR coindexation:

- (2) Coindex AGR with an NP in the accessible domain of AGR.

There are, however, three differences between the two rules:

a) Borer's hypothesis was not framed within the Internal subject Hypothesis. So her definition of accessible domain was devised as to include Spec of INFL. This is not the case here, because at D-structure all Arguments are in the strict c-command domain of AGR they may possibly become subjects of.

b) Our rule is obligatory: AGR must find some DP or CP to coindex with. This makes sure that an object DP will have to be coindexed with AGR if the External Argument is absent. The obligatory character of this rule is reminiscent of the EPP: only that instead of requiring that Spec of IP (or AGRP) must be filled, we require that AGR must be coindexed with some Argument. In fact, this does not imply that Spec of IP will end up being filled, as we will see. Borer emphasizes that the rule should not be obligatory, precisely because Burzio's Generalization is not, according to her, always fulfilled. She adduces dialectal Hebrew data that violate BG (Borer (1986:385):

- (3) a. Haya kativ 'et ha-yedi'a ha-zot ba-'iton  
 Was-written-*m Acc* the-message the-this in-the-paper  
 'This message was written in the paper'
- b. Meforat 'et ha-dvarim ha-'ele ba-karoz  
 Specified-*sg Acc* the-things the-these in-the-leaflet  
 'These things are specified in the leaflet'
- c. Kara li kvar 'et ha-te'una ha-zot kodem  
 Happened-*m* me already *Acc* the-accident the-this before  
 'I already had this accident before'

Since BG seems to hold steadily in many other languages, we are not going to simply ignore it. Then Hebrew exceptions should be dealt with in a specific way.

c) In Borer's rule CPs are not mentioned as candidates to be coindexed with AGR. Whether this is essential to Borer's theory or not, it is crucial in our rule: while sentences such as 0.a) are excluded because AGR is not coindexed with any DP or CP, sentences such as 0.b) are acceptable crucially because the CP can be coindexed with AGR:

- (4) a.  $*(It_{Expl})$  surprises me about this question
- b. It surprises me [<sub>CP</sub> that... ]

Let's call the DP/CP coindexed with AGR the I-subject. Let us review the principle in 0 to see how it works. First of all, 0 is intended to be a principle which applies on the basis of solely structural and Case-marking information: AGR looks down for a non (inherently) Case marked DP or CP from top to bottom of its c-command domain and coindexes with the first available candidate.

Since the principle applies at D-structure, all Case-marked DPs (which are therefore not available candidates) will be inherently Case marked DPs, if we assume that structural Case is

not assigned until S-structure. Therefore the parenthetical specification in 0 is not necessary.

Now, since we make crucial use of inherent Case marking to discriminate which DPs are candidates to be coindexed with AGR, we need some characterization of inherent Case-marking. We will not discuss all the cases where inherent Case marking has been resorted to. For our purposes, the following definitions will suffice as a characterization of inherent Case:

- (5) Inherent Case is assigned at D-structure.
- (6) Inherent Case assignment is obligatorily assigned.
- (7) Oblique Case is inherent Case.

The above statements make sure that a DP in the complement of a preposition will never be coindexed with AGR. We assume that apparent counter-examples (passives like: *This was talked about*)

involve verb-preposition reanalysis of some kind (see Kayne (1981)), or at least, that in this case the preposition does not assign inherent Case. For languages where all Arguments apart from the object and the EA are prepositionally Case-marked, the above characterization suffices to exclude the undesired cases of coindexation. For languages having oblique-Case-marked Arguments not cooccurring with a preposition, we will assume that these Cases are also inherent.

With the preceding assumptions, the Argument that will be coindexed with AGR is predicted in a relatively straightforward way:

- in a transitive or unergative structure, AGR will always be coindexed with the EA, since it is generated as the highest DP (or CP) in the VP local Domain (we assume it is adjoined to VP).

- in an unaccusative structure, AGR will be coindexed with the object DP, since it is the closest non-inherently Case-marked DP in its scope.

- in a copulative structure, AGR will be coindexed with the subject of the small clause.

- in a raising structure, the closest DP or CP will be internal to the infinitival clause, since

raising verbs have no eligible DP/CP Argument themselves, assuming that the infinitival complement is not CP but IP, as has been traditionally assumed.

- In the case of weather verbs, we have to assume that they project a quasi-Argument, that would count as an Argument for the purposes of the present theory.

- in a passive structure, we assume that the EA is not projected as a DP, so the coindexed DP will be the object (*John was seen*), or an Argument internal to the IP in an ECM construction (*John was believed to...*), or the subject of a Small Clause (*John was considered intelligent*).

Roberts (1991-b) proposes that in a passive the EA is projected as PRO. If so, it would be problematic for our account, since it would be taken by AGR as a the chosen candidate (assuming this PRO is projected in the same position as the EA in active structures). I think that, even if the EA is present in a passive in a way it is not in an unaccusative structure, it is far from clear that it is projected in the same way as in an active structure. It has been noticed that the implicit Argument can control an adjunct clause -0.a)-, while no controller is available in the unaccusative structure -0.b). But the implicit Argument cannot control a complement clause -0.c):

- (8) a. The boat was sank to prove a point  
b. \*The boat sank to prove a point  
c. \*Bill was promised to go  
(intended meaning: Someone promised Bill to go).  
d. \*It was promised/wanted/hoped to go

So I think the implicit Argument should be better characterized as not projecting in the same position as the EA in an active clause, however problematic this may be for the Projection Principle. I leave the question here.<sup>49</sup>

---

<sup>49</sup> See Roberts (1987), and Baker, Johnson & Roberts (1989), for the proposal the implicit Argument is projected as the passive participial affix (i.e. a head).

A piece of support for the above hypothesis is the fact that, in English, when we have two objects, it is the first object which is coindexed with AGR in a passive. We can assume, with Larson (1988), that the first object is higher than the second, the evidence being that there is asymmetrical c-command from the first to the second (as adduced by Larson).

The general idea is, then, that the Argument becoming the I-subject is always the most prominent available DP or CP in the c-command domain of AGR. In other words, Burzio's Generalization is definable on purely structural terms, without making reference to Theta Theory: the fact that the presence of a projected external  $\theta$ -role is relevant is due to the independently assumed fact that the EA is projected in a more prominent position.

0 expresses the idea that it is not DP's that are forced to move to Spec of AGR to get Case, but rather AGR that is forced to pick up a DP to coindex with it. An implicit assumption, up to now, is that I-subjects have to end up being assigned Case as subjects (i.e., being for instance Nominative in finite sentences). We will elaborate on this issue later. If we assume, for the moment, that Accusative assignment is optional, then BG is derived from Case theory: if the I-subject is forced to be assigned Case as subject, an object that becomes an I-subject will not be able to receive Accusative, independently of whether Accusative is available or not.

0 is, on the other hand, an abstract alternative characterization of the EPP, in that it states that AGR is always coindexed with some I-subject. In the next paragraph we will qualify this generalization by considering languages where AGR is not always coindexed with an Argument.

## **1.2. Two Parameters**

In this section we will deal with two types of languages where the above formulation of the EPP does not work. In one case (German impersonal verbs and impersonal passives) we need a minor parameterization of our formulation. In the other case (Ergative languages) the parameterization affects the basic shape of the principle.

### 1.2.1 Impersonal Constructions

Our formulation of the AGR-coindexation principle in 0, as it stands, requires that AGR must be coindexed with some DP/CP Argument. This means that there must be one available. Otherwise, the requirement in 0 would not be fulfilled, and the output sentence would be ungrammatical. Suppose, though, that 0 is parameterized as in 0, where the two parameter values are absence vs. presence of the parenthesized part:

- (9) At D-structure, AGR must be coindexed with the most prominent non-(inherently-)Case-marked DP or CP in its c-command domain (*if there is one*).

Suppose that including the parenthesized part is the option for languages like German (as opposed to English and Romance languages). The prediction is that, in German, AGR will be coindexed with some DP/CP Argument only if the lexical specification of the verb provides one. As we saw in the preceding chapter, there are cases where AGR is not apparently coindexed with any Argument:

- (10) dass gestern getanzt wurde  
that yesterday danced was

- (11) a. dass (es) mir schwindelt  
that (it) me-DAT is-dizzy  
'... that I feel dizzy'  
b. dass (es) mir davor graut  
that (it) me-DAT of-it fears  
'... that I am afraid of it'

Suppose that in both cases no DP is available for AGR to coindex with:

- in the impersonal passive case, because the Agent Argument is not projected as a DP;
- in the other case, because these verbs do not subcategorize for any DP not being inherently Case marked.<sup>50</sup>

The present account, based on parameter 0, gives a unitary explanation for the existence of impersonal passives and the class of verbs in 0 in the same language: languages lacking impersonal passives do not have verbs of this kind. Icelandic would be another instance of such a correlation.<sup>51</sup> Actually, it is not clear that this correlation is a genuine one. To my knowledge, Russian is a language having verbs similar to the ones in 0 without having impersonal passives. Perhaps we could simply assume that the parametric option set for German is a necessary but not sufficient condition for the existence of impersonal passives, whose licensing would depend on other factors or, perhaps, would simply be a matter of idiosyncrasy.

As far as I know, languages allowing impersonal constructions like 0 (German, Russian, Icelandic) are all languages overt Case marking on DPs. If this generalization is genuine, it should be captured by the theory. I do not have any interesting proposal in this connection.

---

<sup>50</sup> Some of these verbs subcategorize for an Accusative:

(i) dass (es) mich     dürstet  
      that (it) me-ACC is-thirsty

We have to assume that this is an inherent Accusative, and is already Case-marked by D-structure.

<sup>51</sup> Italian has two verbs of that germanic type: **importare** ('concern') and **dispiacere** ('dislike') (thanks to L. Rizzi for this remark):

(i) A me     dispiace di questo  
      To me dislikes of this  
(ii) A me importa di questo  
      To me concerns of this

Perhaps they are a residue of an earlier period where Italian had impersonal passives (and so the German value of the parameter).

French shows another case of a verb that allows AGR not to coindex with any Argument:

(iii) Il **faut** ces livres  
      It need these books ('These books are needed')

These cases are, I think, not representative of the general patterns of the languages in question and could be assumed to be residual and not belonging to the core grammar.



### 1.2.2. Ergative Languages

There should be another more important and obvious parameter, to account for the contrast between Nominative-Accusative languages (which this thesis deals with for the most part and 0 is conceived for) and Ergative Languages (ELs). I will only consider one such case of ergativity, which is perhaps not most typical: Basque.

In Basque, the situation seems to be that there is a Agreement for the EA (which appears only when there is an EA) and an Agreement for the internal Argument, which is obligatory. The latter (the so-called Absolutive agreement) has some distinctive properties w.r.t. the former (Ergative agreement) and the also present Dative agreement. Apart from the obligatory presence of ABS-AGR (as opposed to the optionality of ERG-AGR and DAT-AGR, which are present only when there is an external/dative Argument), ABS-AGR is prefixal while the other two are suffixal (with some exceptions in the past tense<sup>52</sup>). In addition, the prefixal AGR-ABS is apparently an older and more tightly attached affix than the other two, for which its clitic origin is more obvious.<sup>53</sup>

These facts suggest that AGR-ABS is the closest correlate of Nominative AGR in non-ELs (in section 5. we argue that the behavior of AGR in non-ELs is due to its obligatory presence in finite sentences). Therefore, we could argue, the parameter which allows the Basque option is of the following nature:

---

<sup>52</sup> Ortiz de Urbina (1989) characterizes this past tense particularity as a case of morphological split ergativity: the prefixal ABS-AGR takes on the role of ERG-AGR. In our terms, the obligatory AGR-marker shifts from internal to EA, thus behaving like a non-EL in these past tense forms. But these facts are mere morphological and have no syntactic import.

<sup>53</sup> Some of the Ergative/Dative suffixes (1st and 2nd plural) are homophonous (abstracting away from stress) with the non-emphatic independent pronouns.

(12) At D-structure, the obligatory AGR (in finite sentences) must be coindexed with the most prominent non-(inherently)-Case marked DP/CP:

a) in a designated position (such as object).

b) (no restriction)

The b) option would be the unmarked one and is equivalent to 0 above. The more restrictive a) option would be the option for Basque. It is perhaps no matter of accident that ELs are less represented, as would be predicted by the markedness characterization in 0.

The obligatory character of AGR-coindexation in 0, which accounts for Burzio's Generalization when b) is taken, would account, when option a) is taken, for an interesting fact in Basque: verbs that are usually intransitive-unergative in other languages, are for the most part 'formally transitive' in Basque. So, 'to work' is translated as 'to do work', 'to sleep' as 'to do sleep', and so on. Since AGR-ABS is the obligatory AGR, it has to be coindexed, according to 0-option a), to a designated position (object), which therefore has to be always present. Although these verbs form a rather fixed expression with their objects, they are clearly independent syntactic constituents, i.e., we are not dealing with lexical incorporation, inasmuch as they are not strictly adjacent and the object admits some partitive case in polarity contexts:

(13) a. lo egin

sleep to-do 'to sleep'

b. Gauez egingo dut lo

by-night to-be-done I-have sleep

'It's by night I will sleep'

c. Bart ez nuen lorik egin

Last-night not had-I sleep-PART done

'Last night I didn't sleep at all'

Actually, there are some exceptions to this general pattern: some unergative verbs do not show any object at all (except that ABS-AGR appears in the third person singular). These verbs have increased in number as borrowings from Romance languages have entered Basque. So they could be considered marked options having a kind of null object. Or we could assume that modern Basque has set a parametric option similar to the one taken by German: ABS-AGR has to coindex with an object *if there is one*.

Basque is actually an atypical EL (see Ortiz de Urbina (1989)). The Case array is different in other ELs and the parametric formulation above could turn out not to fit other ELs. I cannot pursue the question here.

### 1.3. A Problematic Case

Since languages like English or the Romance languages have neither impersonal passives nor verbs like those in 0 in general, it seems that these languages set the other parametric option in 0. If so, in all sentences AGR is coindexed with some Argument (DP or CP). This seems to be the case at first glance, but there are some apparently problematic cases. Consider the Catalan verb *semblar* 'to seem'.<sup>54</sup> We have assumed, in accordance to the GB tradition, that in its raising version *semblar* subcategorizes for an IP. So AGR is coindexed with some DP inside the IP, since there is no other DP available in the main sentence. Now, when *semblar* subcategorizes for a CP finite complement, it is this CP itself that will be coindexed with AGR: this is at least what 0 predicts.

That CPs are possible candidates to be coindexed with AGR is a plausible assumption: this is likely the case for examples like the following:

---

<sup>54</sup> The facts we will be considering are from Catalan. To my knowledge, they are essentially the same as in Spanish. We will consider Italian below.

- (14) a. It strikes me *that*...
- b. It is evident *that*...
- c. It was known *that*...

In the corresponding cases in Catalan, the CPs are likely to be the I-subjects too:

- (15) a. Em sorprèn *que*...
- Me strikes *that*...
- b. És evident *que*...
- Is evident *that*...
- c. Era sabut *que*...
- Was known *that*...

Now let's consider the CP Argument of *semblar*. There is evidence that this CP does not behave like a I-subject. In contrast to the CPs in 0, it can be pronominalized as an object clitic *ho* 'it', just as object CPs can:

- (16) a. Ho sembla, que vindrà
- It-ACC seems, that he-will-come
- b. Ho sé, que vindrà
- It-ACC I-know, that he-will-come

None of the CPs in 0 can be pronominalized as *ho*. So, it appears that the non-raising version of *semblar* is a counterexample to our claim that AGR is always coindexed with an Argument, unless we were to admit that AGR is coindexed with an Accusative clitic, which would strongly weaken our hypothesis.

There seems to be no wide range of problematic cases: in Catalan, *semblar* (and some

other verbs we will consider below) are the only verbs with no apparent I-subject. So it seems reasonable not to give up our hypothesis and try to find a specific explanation for the behavior of *semblar*.

Suppose we assume that *semblar*, like *weather* and similar verbs, has a quasi-Argument, which is the subject in the problematic examples.<sup>55</sup> At first glance, this position does not seem tenable: if the quasi-Argument is present with *semblar* + CP, it would block raising in the *semblar* + IP construction, since two Arguments (the quasi-Argument and the raising Argument) would be competing to coindex with AGR and become subjects of the main clause. On the other hand, if a theory of quasi-Arguments is to be taken seriously, we cannot assume that the quasi-Argument is optional and does not appear in the raising construction: quasi-Arguments should be subject to the Projection Principle and the Theta-Criterion: otherwise they would be nothing but a theoretical artifact used for our convenience.

The solution to this puzzle that we propose is based on the following assumptions: there are two verbs *semblar*, one projecting a quasi-Argument and one not projecting it, the second one being the one allowing raising.<sup>56</sup> We will see that this distinction is plausible and makes the correct predictions.

I think that the existence of two verbs *semblar* can be independently motivated on both semantic and syntactic grounds. Consider the following pair in Catalan:

---

<sup>55</sup> Recall that the existence of quasi-Arguments is crucial for our account: without quasi-Arguments, all weather verbs would be problematic.

<sup>56</sup> Hernanz (1982) arrives at a closely similar conclusion for the Spanish equivalent of **semblar** (**parecer**), which behaves exactly like **semblar**, as far as I know. Her arguments are based on essentially the same motivation, except that she does not face the problem of I-subjects and Burzio's Generalization. So, she does not consider the possibility of a quasi-Argument. Other differences will be pointed out later.

- (17) a. Sembla que està cansat  
 Seems that he's tired  
 'It seems that he's tired'
- b. Sembla que estigui cansat  
 Seems that he-be tired  
 'It looks as if he's tired'/'He looks tired'

From the syntactic point of view, in the a. example *semblar* subcategorizes for a CP in indicative mood, while in the b. example it subcategorizes for a subjunctive CP. From the interpretative point of view, the a. example means something like 'there are convincing indications that he's tired', while the b. example means 'there are (mere) appearances that he's tired'. Let's call *semblar*-1 the verb in a. and *semblar*-2 the verb in b. Now let us consider *semblar* as a raising verb:<sup>57</sup>

- (18) (Ell) sembla estar cansat  
 (He) seems to-be tired

It is clear that 0 is synonymous with 0.a) and not with 0.b): this seems to indicate that it is *semblar*-1, not *semblar*-2, that is a raising verb. As a raising verb, *semblar*-1 cannot have any quasi-Argument, for the reasons we argued above. On the other hand, since *semblar*-2 is not a raising verb, no theoretic problem arises if we postulate that it projects a quasi-Argument. Suppose that it does, and that this is what makes it possible for the CP not to become subject. If we are on the right track, the prediction is that the verb *semblar* which allows its CP to

---

<sup>57</sup> Infinitival raising constructions are not accepted in Normative Catalan and, indeed, they are not a genuine construction in spoken Catalan. However, I do not find them more awkward than many other constructions that have entered standard Catalan without originally belonging to the spoken language (passives, Wh-relatives, etc.).

pronominalize as the clitic *ho* is *semblar-2*. This in turn predicts that, in sentences with such a pronominalization, the interpretation is that of *semblar-2*, namely 'there are (mere) appearances that...'. I think this prediction is fulfilled:

(19) Ho sembla (que estigui cansat).

It-ACC seems (that he's tired)

According to my intuition, whenever the clitic *ho* replaces the (dislocated or contextually recoverable) CP, the 'mere appearance' interpretation is emphasized. Judgements about this matter are not clear-cut, probably because the interpretation of *semblar-2* stands in a subset relation w.r.t. that of *semblar-1*: if 'there are convincing indications of a situation' then 'there are appearances pointing to that situation ' (not the other way around). Nevertheless, it seems clear that in a dialogue like:

(20) A- Sembla que està cansat.

Seems that he's tired

B- Ho sembla.

It-ACC seems

the reply is less than confirming the assertion: it rather suggests cautiousness about the certainty of 'his being tired' in emphasizing that only appearances are certain. In a dialogue like:

(21) A- Estàs cansat?

Are you tired?

B- No

No

A- Doncs ho sembla

Anyway It-ACC seems

'You look as if you were anyway'

the last remark of speaker A clearly means 'mere appearance' situation. Alternatively, *semblar*-2 (= *semblar* with Accusative clitic) cannot be used to make a confirming reply:<sup>58</sup>

(22) A- Estàs cansat?

Are you tired?

B- Sí.

Yes

A- #Ja ho sembla

Indeed It-ACC seems

'#You indeed look as if you were'

Another fact that makes the distinction between the two verbs plausible is the fact that they correspond to different verbs in other languages (*seem* and *look/sound/etc.* in English). The English verb *look* has the properties we postulate for *semblar*-2: it has the '(mere) appearance' interpretation, it is not a raising verb (\*John looks to have come) and it also requires the postulation of a quasi-Argument in order to explain its apparent lack of I-subject in cases like:

---

<sup>58</sup> Gemma Rigau pointed out to me the relevance of these examples.



(23) It looks as if ...

In 0 it is unlikely that the I-subject is the *as if* clause, because it does not look like a bare CP. So the conclusion would be that verbs of the *sembler-2* type involve a quasi-Argument roughly denoting 'appearances'.

Since *sembler-2* is not a raising verb, we could expect it to be a control verb when it subcategorizes infinitive. This is the case for Catalan and other Romance languages, where *sembler-2* can have a dative controller:

- (24) a. Em<sub>i</sub> sembla PRO<sub>i</sub> estar somiant (Catalan)  
Me-seems to-be dreaming  
'I have the impression of being dreaming'
- b. Me<sub>i</sub> parece PRO<sub>i</sub> estar soñando (Spanish)  
Me-seems to-be dreaming
- c. Mi<sub>i</sub> sembra di PRO<sub>i</sub> éssere in un sogno (Italian)  
Me-seems of to-be in a dream

The verb in 0 is clearly *sembler-2* as far as it conveys the 'mere appearance' meaning.<sup>59</sup>

If we are right in postulating a quasi-Argument, we should say something about its semantic plausibility. Quasi-Arguments can be conceived as Arguments referring to an entity which is vague enough not to be possibly instantiated as a full DP: there is something that 'rains' or 'is late' (roughly the weather, the time, resp.) and cannot (or simply is not) conceptualized precisely. In the case of *sembler-2*, what would constitute the quasi-Argument is, likely, '(a set of) appearances'. In this case, we cannot say it is not conceptualizable: one can precisely know which fact or thing constitutes the appearance. Nevertheless, *sembler-2* does not allow a full DP

---

<sup>59</sup> I cannot explain why the English equivalent of **sembler-2** ('to look/sound/etc.') cannot involve control:  
\*It looks to me to be dreaming

(or CP) as an alternative to the (purported) quasi-Argument:

- (25) \*Aquest soroll sembla que plougui  
This noise seems that it-rains-SUBJ

So the reason the quasi-Argument is the only option for *semblar*-2 has to be another one. My suggestion is the following. Consider *semblar* when used without a clausal complement, as in 0:

- (26) En Joan sembla cansat  
The J. seems tired  
'Joan looks tired'

As is made clear from the translation (where we use *to look*), here we are dealing with *semblar*-2. A plausible analysis of 0 is that the subject (*en Joan*) is, at D-structure, the subject of a Small Clause headed by *cansat*. Suppose we assume that *semblar*-2 always has a Small Clause complement. Then, in:

- (27) Sembla que plougui  
Seems that rains-SUBJ  
'It looks as if it was raining'

the quasi-Argument would be the Subject of a Small Clause *whose nucleus would be the CP (que plougui)*:

- (28) Sembla [<sub>SC</sub> *quasi-A* [<sub>CP</sub> que plougui ]]

Now, suppose we make the following assumption: A CP is *not* a possible predicate of a fully referential DP. So, when the predicate in a Small Clause is a full CP, its subject has no option but being a quasi-Argument. However *ad hoc* this idea may be, it seems unproblematic.<sup>60</sup>

I think there is a possible empirical argument for the existence of the quasi-Argument. It is the same kind of argument that Chomsky (1981) uses to motivate the existence of quasi-Arguments. The idea is that they can control and thus licence a PRO which otherwise would be illicit:

- (29) a. *It* is cloudy without *PRO* actually raining  
 b. \*I took the umbrella without *PRO* actually raining

The following example shows a similar pair where the controllee would be the quasi-Argument of *semblar*-2, and the controller the quasi-Argument of *plou* 'it rains':<sup>61</sup>

---

<sup>60</sup> Basque seems to challenge this assumption as it allows sentences as:

Jonek dirudi bere anaia berriro gaixo dagoela  
 J. seems his brother again ill is-that  
 'Jon looks as if his brother is ill again'

Perhaps this is due to the adverbial nature of embedded CPs in Basque (the **-ela** 'that' complementizer is plausibly an adverbial suffix). Taking the English translation as an example, the 'as if IP' complement allows a full DP as well as a quasi-Argument ('John/it looks as if his brother is ill again'). This, too, would be due to the adverbial nature of the 'as if' complement. So it would be non-adverbial CPs which exclusively allow quasi-Arguments as subjects of predication.

<sup>61</sup> As far as in the English translation of (0.a) 'to seem' is used, our claim that **semblar**-2 corresponds to 'to look/sound, etc.' is weakened. Perhaps English 'to seem' is ambiguous as well.

In (0.a) the facts are a bit obscured by the simultaneous presence of a infinitival PRO in the **semblar** clause (which we claim is a controlled quasi-Argument) and another quasi-Argument of 'to rain' in the finite clause embedded under **semblar**. The following example avoids this situation, and is certainly a bit marginal:

(i) ?Plou sense semblar que hi hagi núvols.

This could be due to a certain lack of identity between the two quasi-Arguments. Consider:

- (30) a. Plou sense *PRO* semblar que plogui  
 Rains without to-seem that rains-SUBJ  
 'It is raining without seeming to'
- b. \*Agafo el paraigua sense semblar que plogui  
 I-take the umbrella without to-seem that rains-SUBJ  
 '\*I take the umbrella without seeming to rain'

One qualification should be made to above hypothesis: the way we presented the contrast in 0, we seem to suggest that indicative/subjunctive mood is a contrastive property of the two verbs *semblar*. Now, while it is clear that subjunctive is an exclusive property of *semblar*-2 (i.e., it has the 'mere appearance' interpretation only), indicative is not clearly an exclusive property of *semblar*-1: 0.a) can be used in the 'mere appearance' interpretation. So the contrast is half-way:

- *semblar*-1: indicative
- *semblar*-2: subjunctive (or indicative)<sup>62</sup>

In this connection, it is not a problem that the clitic *ho* can pronominalize both an indicative and a subjunctive CP, which appear as dislocated: both 0.a) and (.b) would be cases of *semblar*-2:

- 
- (ii) It's raining without it/\**PRO* being late in the afternoon (in a country where it usually rains late in the afternoon)

In (ii) a weather quasi-Argument cannot control a time quasi-Argument, because they do not refer to the same entity (however cloudy to conceptualize these entities are). Here the lack of identity is too strong, while in (0.a) and (i) it would be milder.

<sup>62</sup> Hernanz (1982) does not consider the indicative/subjunctive distinction (which holds in Spanish the same as in Catalan) and simply asserts that the constructions with indicative (as in (0.a)) are ambiguous, which we agree upon.

- (31) a. Ho sembla, que plou  
It-ACC seems, that rains-IND
- b. Ho sembla, que plogui  
It-ACC seems, that rains-SUBJ

If anything, 0.b) appears to more strongly suggest that 'raining is a mere appearance' than 0.a), but they are essentially synonymous, as is in accordance with our claim that both must be cases of *semblar*-2.

Another factor can disambiguate the *semblar*+indicative construction: in Catalan or Spanish an Experiencer dative accompanying *semblar* cannot cooccur with subjunctive (i.e., *semblar*-2), as shown in 0.a) This suggests that the Experiencer dative is not compatible with *semblar*-2. And in fact when the indicative co-occurs with a clitic, only the 'convincing evidence' interpretation (i.e., *semblar*-1) is available (as in 0.b)).<sup>63</sup>

---

<sup>63</sup> Torrego (1989) distinguishes between two verbs **parecer** (the Spanish equivalent of **semblar**) and argues that the presence of the Experiencer is coextensive with one of them. So there are two verbs **parecer**: **parecer**+ experiencer and **parecer**-epistemic. They are distinguished by a set of criteria:

- a) only **parecer**-ep. can take subjunctive:
  - (i) Me parece que llueve/\*llueva  
Me seems that rains-IND/SUBJ
- b) only **parecer**-exp. can occur in a perfective tense:
  - (ii) Hoy \*(me) ha parecido que llovía  
Today (me) has seemed that was-raining
- c) only **parecer**-ep. is a raising verb:
  - (iii) \*Juan ha parecido haberlos encontrado  
J. has seemed to-have-them found

In her analysis, **resultar** is rather like **parecer**-exp. So her analysis is clearly at odds with the one we propose here, for she assumes that raising **parecer** is precisely the one allowing subjunctive (we assume the opposite) and **resultar** does not for a class with raising **parecer**.

In fact, I do not agree with some of the data either: **parecer** without Experiencer can appear in perfective tense:

- (iv) Por un momento, ha parecido que llovía  
For a moment has seemed that was-raining
- (v) ?Juan pareció querer decirnos algo  
J. seemed to-want tell-us something

- (32) a. \*Em sembla que plougui  
           Me seems that rains-SUBJ  
           '\*It looks to me as if it's raining'
- b. Em sembla que plou  
           Me seems that rains-IND  
           'It seems to me that it's raining'

There is a problem left to be addressed concerning *semblar*-1. We have claimed that *semblar*-1 has no quasi-Argument and that the CP itself is the I-subject. CP I-subjects in NSLs can be dropped as null subjects or can be dislocated with a null resumptive pronoun:

- (33) Em molesta (, que diguis això)  
       Me-bothers (, that you-say that)  
       'It bothers me, that you say that!'

This is not the case with *semblar*-1:

- (34) \*Sembla (, que plou)  
       Seems (, that it rains)

The same happens with another verb (*resultar* 'to turn out to be the case') which is likely of the same nature as *semblar*-1 (it is also a raising verb, subcategorizes indicative and admits neither *ho*-cliticization of the CP nor Experiencer datives in Catalan or Spanish):

- (35) a. \*Resulta (, que vindrà)  
 Turns-out (, that will-come)  
 'It turns out (that he will come)'
- b. \*Ho resulta (, que vindrà)  
 It-ACC turns-out (, that will-come)  
 'It turns out (that he will come)'
- c. (\*Em) resulta que vindrà  
 (Me) turns-out that will-come  
 'It turns out (\*to me) that s/he'll come'

Spanish *resultar* and Italian *risultare* have essentially the same meaning and behavior.

So we have built a theory in order to make sure that *semblar* is not a problematic case (since whenever the CP is pronominalized as Accusative, there is a quasi-Argument I-subject), but now the CP being claimed to be the actual I-subject of (non-raising) *semblar*-1 does not behave like other CP I-subjects as far as (null) pronominalization is concerned. My suggestion is that this is due to the epistemic nature of the *semblar*-1/*resultar* verbs: they are not true predicates, but a kind of aspectual, semantically adjunct, predicate.<sup>64</sup> And in fact, predicates of this kind can be paraphrased as adverbial adjuncts ('apparently' for *semblar*-1, 'in fact' for *resultar* are close paraphrases).<sup>65</sup> There are other cases of CPs that cannot be pronominalized or dislocated. One is the following, in Catalan:

---

<sup>64</sup> The idea that raising verbs are adjunct predicates is proposed by Hernanz (1982), Rothstein (1983) and Torrego (1989).

<sup>65</sup> This adjunct nature could also be an explanation for the fact that these predicates are raising predicates: since raising is a quite restricted phenomenon, we should be able to predict why it is. We could claim that only adjunct predicates (modals and aspectuals) can be raising predicates, possibly because they are not truly bi-clausal (at least at LF).

(36) Diu que es casen

Says that they get married

'They say that they are getting married' or rather 'I heard that they are getting married'

In 0, *diu que* is a fixed expression (it can only be present tense, and it differs from standard 3rd plural arbitrary constructions in that it is singular). In fact, this fixed expression does not mean 'someone/people say(s) that', but it rather has the meaning of a speaker-oriented adverb expressing the novelty or surprise the speaker feels about the fact expressed by the embedded sentence. So, it is another case or merely adjunct predicate and, probably for this reason, it cannot be pronominalized or dislocated:

(37) \*Ho diu (, que es casen)

It-ACC says (, that they get married)

Cf. with:

Ho diuen (, que es casen)

It-ACC they-say (, that they get married)

It is not clear why the purely adjunct, epistemic/ speaker-oriented interpretation for predicate should give the result that their complement CP, when I-subject, cannot be dropped in NSLs or pronominalized and dislocated in general: the idea would roughly be that the CP is 'too interpretatively central' a part of the sentence for it to be dropped or dislocated.

Summing up, here are the main characteristics of the two verbs *semblar*:



(38)	<u>Semblar-1</u>	<u>Semblar-2</u>
	'convincing evidence'	'mere appearance'
	indicative	subjunctive (or indicative)
	no quasi-Arg.	quasi-Arg.
	CP complement	(possibly CP-headed) SC complement
	adjunct predicate	main predicate
	raising (possible)	no raising
	no control	control (possible)
	epistemic reading	no epistemic reading
	= <i>seem</i> (or <i>seem-1</i> )	= <i>look/sound/etc.</i> (or <i>seem-2</i> )
	similar verbs: <i>resultar</i> , <i>diu que</i> .	

In conclusion, it seems to me that so narrow a range of potential problematic examples as *semblar* and similar verbs should not lead us to abandon our crucial claim that any clause has an I-subject (which allows us to derive Burzio's Generalization), and that a reasonably plausible and motivated way-out of the problem can be conceived of.

Italian *sembrare* does not easily fit into the preceding picture. On the one hand, the indicative/subjunctive contrast is not present. In standard Italian subjunctive is preceptive in any case. In colloquial Italian, indicative is possible, but not in any interpretative contrast with subjunctive, because subjunctive is simply disappearing from colloquial/dialectal Italian in many regions.<sup>66</sup>

On the other hand, the finite CP Argument of *sembrare* can be dropped or dislocated, in

---

<sup>66</sup> In those dialects, it is disappearing even in the complement CP of optative verbs:  
 Voglio che vieni  
 I-want that you-come-IND

contrast with Catalan (and Spanish) 0:

- (39) Sembra (, che piova/piove)  
Seems that rain-SUBJ/rains-IND  
'It does look as if it's raining'

Since 0 can be *semblar*-2 ('mere appearance') it should involve a quasi-Argument. The CP, however, can be dropped or dislocated, as other CP I-subjects. The fact that it is the CP (and not the quasi-Argument) which becomes the I-subject therefore contrasts with Catalan and Spanish. We could assume that in Italian the quasi-Argument is inherently Case-marked and that AGR can thus be coindexed with the CP. This is obviously nothing but an *ad hoc* solution. I leave the question open.

#### 1.4. The EPP and the ECP

Our reformulation of the EPP in 0/0 is a device that determines which DP or CP will become the I-subject (i.e. will be coindexed with AGR) in a clause. For convenience, we will be using the non-parameterized version in 0:

- 0 At D-structure, AGR must be coindexed with the most prominent non-(inherently-)Case-marked DP or CP in its c-command domain.

Even if we have not fully developed what this device amounts to, we can advance that our theory is intended to cover all the cases of what is standardly conceived as A-movement. Specifically, in the following D-structures, 0 determines that the underlined DPs or CPs will become I-subjects and, in English, these DPs/CPs will ultimately become the main clause subjects (in a subtler sense in the case of CPs):

- (40) a. AGR ... [VP John [VP open the door ]]  
 b. AGR ... [VP open the door ]  
 c. AGR ... [VP seem that ]  
 d. AGR ... [VP seem [IP to [VP come John ]]]

Therefore 0 predicts which A-Chains will be formed and what length they will have.

The standard theory of A-movement can be conceived as consisting of three essential modules:

- Move-alpha as a general rule allowing A-movement;
- Case theory as a trigger for A-movement;
- constraints on the output A-movement: ECP, proper movement, the Projection Principle, Chain formation, etc.

In the above examples no unique principle in the standard theory compels the underlined DPs or CPs to become the subject of the main sentence: Move-alpha is applied freely; Case theory constrains the DPs to move (or form expletive CHAINS) and the other constraints filter the undesired cases (super-raising, movement to a  $\theta$ -position, etc.).

Our theory explicitly denies that A-movement is triggered by Case-requirements: an object may be forced to become I-subject even if it can be potentially assigned Accusative. So Case theory cannot be used as a last-resort trigger for A-movement. On the other hand, 0 clearly overlaps with the other constraints on A-movement: the way it is formulated 0 makes sure that whatever DP is chosen as I-subject will abide by the standard constraints. Consider, for instance, a standard example of super-raising:

- (41) \*John seems [ that Peter hoped t to come ]

The standard account for the ill-formedness of 0 would be that this representation

violates some locality condition on A-movement: either the ECP (if A-movement takes place in one step) or constraints on proper movement (if intermediate traces in A'-positions are used). According to 0, this structure is simply not possible because the AGR in the upper clause will never coindex with a DP internal to the embedded CP, since the CP itself is the option chosen.

So it is obvious that our principle 0 is powerful enough to make it unnecessary to resort to the ECP (or other principles) to exclude super-raising. Therefore, it is highly suspicious, as far as the ECP is a well established and independently motivated principle. Of course, we can always assume that the ECP and 0 redundantly constrain A-movement, but it is always advisable to eliminate unnecessary redundancies.

What we are going to argue is that the ECP is not sufficient to constrain A-movement, and therefore 0 is a possible way of covering the gap. One problem the ECP faces in connection with A-movement is the asymmetry between A'-long-movement and A-long movement (super-raising):

- (42) a. ?*Which book* do you know who read *e*  
b. \*\**This book* seems that (it) was read *e*

The traditional account for the mild ill-formedness of 0.a) is that, in spite of the fact that the A'-Chain violates locality constraints (subjacency) it does not violate the ECP, because the empty category is properly governed by the verb *read*, and the antecedent-government option of the ECP is not required. If this is the case, it remains a mystery why 0.b) has the status of a strong (presumably ECP) violation: the empty category should be similarly properly governed by the verb and therefore only subjacency would be violated.

All solutions to this problem are based on the assumption that, for some reason, A-movement always requires antecedent government, and therefore proper government by the verb is not sufficient.

Chomsky (1986-a), in some final developments he explores, reaches the conclusion that

the ECP could be simplified to always (and only) require antecedent government. The apparent insensitivity of objects (and internal Arguments) to the requirement of antecedent government with A'-movement can be derived from the following assumptions:

- A'-movement allows adjunction to VP as an escape hatch.
- intermediate A'-traces can be deleted at LF.

As a consequence of the preceding assumptions, the object of *read* in 0.a) can adjoin to the embedded VP as a first step, and then move on:

(43) [<sub>VP</sub> t' [<sub>VP</sub> read t ] ]

In the relevant structure 0, *t'* will be able to antecedent-govern the trace in object position (*t*); *t'* cannot in turn be antecedent-governed, because its antecedent (or the next intermediate trace) is too far, but since it can be deleted, no ECP violation takes place.

In the case of A-movement, adjunction to VP is not allowed because it would be a case of improper movement: therefore 0.b) is excluded as an ECP violation, hence its strong ungrammaticality. Thus far, Chomsky's solution is highly appealing, since a very simplified version of the ECP (which only and always requires antecedent government) is resorted to in order to cover both A- and A'-movement. A problem arises in connection with A-movement, though. If nothing else is said, even the simplest cases of licit A-movement would be excluded: since VP is, crucially, a barrier, a simple passive or unaccusative construction is predicted to violate the ECP:

(44) John was [<sub>VP</sub> elected t ]

Since the object cannot adjoin to VP in its way to the subject position, VP will be an unescapable barrier. Chomsky's solution to the problem consists in assuming that V and INFL are coindexed and an *Extended Chain* can be formed of the form (John, INFL, V, t), in which

every member of the extended Chain will govern the next one.

Even if it is quite plausible that V and INFL can be coindexed (as a consequence of V-movement to INFL), I think Chomsky's solution can be objected as tricky. On the one hand, it is not the case that the verb always moves to INFL in English (and other languages), so that the assumption that there is coindexation between V and INFL need simply be stipulated (or has to apply at LF, where V movement would possibly take place). But the main conceptual problem is the notion of Extended Chain itself: it is plain that if V and INFL are coindexed, the kind of index they share should not have anything to do with A-Chain indexing, which is plausibly reference-indexing. And, in addition, an extended Chain would be a Chain consisting of both maximal projections and heads, which is certainly an awkward proposal.

Rizzi (1989-a) faces the problem in a different way. His account for long A'-movement is based on the idea that referential expressions bear referential index which is absent in non-referential expressions, such as adjuncts. All empty categories have to fulfil the ECP, which is reduced to the requirement of head-government by a head (one which is not inert for head-government). However, there is an asymmetry between empty categories bound by a referential expression and the ones bound by a non-referential expression: the former are not required to be antecedent-governed, because their indexing makes binding by their antecedent a sufficient condition (*pace* subjacency) for their licensing; the latter, instead, do need antecedent government, because their lack of index makes it necessary that there is a local connection with their antecedent.

Once these premises established, Rizzi, like Chomsky, faces a problem concerning A-movement: it looks more local than the theory predicts, given the fact that it usually involves referential expressions, hence binding (without antecedent government) should be a sufficient licensing condition. The account Rizzi proposes is based on the idea that A-Chains require strictly local linking because  $\theta$ -role transmission takes place between their members: for the  $\theta$ -role to be transmitted between the members of an A-Chain, antecedent government is required. Cinque (1990), which refines the notion of referentiality relevant for long Wh-movement

through binding, reinterprets this idea in a perhaps more perspicuous way: members of an A-Chain are not referential because it is only the Chain as a whole that is an Argument and therefore referential.<sup>67</sup>

We have seen that both in Chomsky's and in Rizzi's or Cinque's accounts something special has to be said about A-Chains: Chomsky's theory is basically too restrictive and a special device of extended Chain formation has to be adopted in order not to exclude licit A-movement. Rizzi's theory is too permissive and a constraint on  $\theta$ -role transmission has to be adopted in order to force antecedent government on A-movement. Cinque's account, although close to Rizzi's, is perhaps more appealing. In any case, it seems that the constraints on A-movement do not trivially follow from a general theory of the ECP: something more or less *ad hoc* has to be added to the theory, however plausible it may look.

Therefore, a principle like 0, which restricts the way AGR and the I-subject are coindexed (hence the possibilities for A-Chain formation when it is required) can be welcome, as far as it is independently motivated as a means of deriving Burzio's Generalization. I do not contend that, given that 0 covers the ECP account for A-movement restrictions, the ECP is necessarily not relevant for A-movement. Suppose it is. Then the prediction is that, since super-raising violates both the ECP and 0, there are cumulative violations leading to ungrammaticality, and thus super-raising will be worse in acceptability than A'-ECP-violations. It seems to me that super-raising violations are more radically unacceptable (I would say they are inconceivable sentences) than A'-movement ECP-violations, even though the judgements are not clear because both are severely ill-formed. There is, however, as far as I know, a contrast in ill-formedness between extraction of adjuncts out of islands and that-trace effects: the latter are less severely bad.<sup>68</sup> So there is a clear contrast between A- and A'-ECP violations:

---

<sup>67</sup> Rizzi (1989-a) does not assume that extended Chains are necessary, because he basically does not address the question of non-minimality barriers. Cinque (1990) specifically rejects that extended Chains are necessary, after having proposed that VP (and IP) are not inherent Barriers.

<sup>68</sup> In fact some speakers accept them: so it could be that for

- (45) a. *?\*Which book* does it seem that *e* was leafed *e* through  
 b. *\*\*This book* seems that it/*e* was leafed *e* through

Summarizing so far, even if principle 0 is redundant with an ECP account for A-movement restrictions, the fact that the latter is not sufficient without some further assumption makes 0 a possible candidate to cover the gap. Still, if the ECP and other principles proved sufficient to constrain A-movement, we could try to reduce the power of 0 to avoid redundancy.

One way of doing so could be not to stipulate the prominence requirement on the DP/CP, so that 0 would be reduced to 0

- (46) At D-structure, AGR must be coindexed with some non-(inherently-)Case-marked DP or CP in its c-command domain.

0 poses no restrictions on the position of the DP/CP that becomes the I-subject: the ECP (or other principles) will filter out the undesired cases. Since 0 is a device whose outputs are to be filtered by independent principles, we could also assume that no Case requirement is necessary:

- (47) At D-structure, AGR must be coindexed with some DP or CP in its c-command domain.

If we assume that inherent Case is not compatible with Nominative Case (or whatever structural Case is ultimately assigned to the I-subject), then coindexation with an inherently Case marked DP will be filtered out independently of 0. Even the requirement that the I-subject must

---

other speakers they are not so bad not because they are any better than adjunct-island violations, but rather because these speakers vacillate in a low level parameter setting (which could consist in treating (or not) **that** as a possible agreeing form (see Rizzi (1989-a)) or simply as a proper head governor.



be in the c-command domain could be given up, since A-movement (which will be a consequence of this coindexing), could follow from general constraints forbidding downward movement:

(48) At D-structure, AGR must be coindexed with some DP or CP.

Therefore, our principle of AGR coindexation can be simplified in a radical way.

The simplified versions 0, 0 or 0 may be problematic for our account for German impersonal verbs, which was based on the parameter 0:

0 At D-structure, AGR must be coindexed with the most prominent non-(inherently-)Case-marked DP or CP in its c-command domain (*if there is one*).

If we parameterize 0 or 0 in a similar way, we obtain undesirable results. Consider:

- (49) a. Es graut mir vor geistern.  
It fears me-DAT of ghosts  
'Ghosts frighten me'
- b. Es graut mir davor dass der krieg anfangen könnte  
It fears me-DAT it-of that the war start could  
'It frightens me that the war could start'

If no prominence requirement is used, AGR would be coindexed with *der krieg* in 0.b) (because *der krieg* is the most prominent DP not being Case-marked at D-structure) and then the structure would be filtered as a case of super-raising. If no case requirement is used, AGR would have to be coindexed with *geistern* (or *mir*) in 0.a) and then it would be incorrectly ruled out as

containing a Case conflict. In other words, since filtering the undesired results comes later in the derivation, we have no way of stating that the condition applies only if *there is one candidate available*.

Since these are very specific predictions tied to very specific formulations, it could be that the question can be a false problem. For the sake of convenience, we will continue to refer to 0 and 0, leaving as an open question whether they could be simplified as 0, 0 or 0.

## **2. AGR-identifiers and Binding Theory**

The theory of AGR coindexation in the preceding section makes sure that AGR has an I-subject. Concerning the BG problem, the rule of coindexation makes sure that an object will be coindexed with AGR in the absence of an external Argument. Nothing we have said, however, ensures that in such a situation the object is not able to receive Accusative. Recall that we cannot stipulate that Nominative is preferred to Accusative because Nominative is not the only Case option for a subject. The result we want to obtain is that whatever Case option is available for the subject (Nominative, Accusative in ECM, PRO) is to be preferred to Accusative. To express this idea in a simple way, we will introduce the notion of AGR-identifier, and we will assume that the I-subject has to obtain Case from its AGR identifier. We will present the technical notion of AGR-identifier in the following sub-section. A discussion on what is the theoretical status of the rules we will propose is deferred to section 2.5.

### **2.1. The notion of AGR-identifier**

For languages like English, the apparent situation is that I-subjects actually move to the Spec of AGR. I will propose that this is not necessarily the case for all languages. In what follows, a theory will be presented predicting why filling Spec of AGR is sometimes obligatory and sometimes not.

The idea I want to exploit is that AGR has to be 'rich' in all languages,<sup>69</sup> in the sense that it has to be able to display a complete range of phi-features: if AGR<sup>0</sup> is not rich itself, *then it is Spec of AGR that has to provide richness in features*. I think this idea is a good basis for accounting for the fact that, diachronically, subjects in non-NSLs tend to end up being AGR<sup>0</sup>-clitics and, eventually, become part of the AGR<sup>0</sup> morphology: this is the standard explanation for the evolution of Northern Italian dialects. If Spec of AGR is the element providing phi-features in non-NSLs, it is natural enough that Spec-of-AGR ends up being reanalyzed as an AGR<sup>0</sup> affix.

To implement this idea, let's assume the principles in 0.a) and 0.b) and the parameter in 0.c):

- (50) a. AGR must have an AGR-identifier.  
 b. X can be an AGR-identifier iff X is rich in phi-features (number and person).<sup>70</sup>  
 c. AGR<sup>0</sup>/Spec of AGR is the AGR-identifier of AGR.

Suppose that when a language has a rich AGR<sup>0</sup> morphology, the first option in 0.c), which would be the unmarked one, is taken. This would be the case in NSLs. When AGR<sup>0</sup> is morphologically poor, the second option of parameter 0.c) has to be taken. Suppose we assume that:

---

<sup>69</sup> Or in all languages having agreement processes: perhaps languages like Chinese and Japanese could be characterized as completely agreementless, so that AGR is absent as a FC. This would not be the case for Scandinavian languages, which do not show any AGR morphology, but are languages with some agreement processes, such as agreement between antecedents and anaphors (unlike Chinese, where phi-features do not seem to have any grammatical relevance).

<sup>70</sup> I will assume that a paradigm is rich if it can display 6 distinctive forms. It seems that the neutralization of some of the distinctions (1st-sing and 3rd singular) is not fatal for richness. See Roberts (1991-a) for some generalizations about richness in verbal paradigms.

(51) Spec of AGR is rich iff it is filled by elements bearing phi-features (=DPs).<sup>71</sup>

This implies that, in English, some DP must appear in Spec of AGR. This is indeed the apparent situation in non-NSLs: in finite sentences they always show a DP (or CP) in Spec of AGR. In NSLs this is not the apparent situation and, we will argue, not the actual situation in some cases.

In some sense, then, English AGR is equally rich as Italian AGR, the difference being that phi-features are placed in the specifier and not in the head. This fact, however, will trigger an important array of differences concerning the distribution of subjects. Specifically, from the above assumptions, we want to derive the generalization proposed in the preceding chapter, repeated here as 0:

(52) I-subjects are [-anaphoric] in NSLs and [+anaphoric] in non-NSLs.

## 2.2. Case Theory

In this section we present a Case-theoretical account for how a DP having become I-subject receives Case. We postpone to section 3. the account for indefinite I-subjects in existential/presentational constructions. We also postpone to section 4. a specific treatment of CP I-subjects.

Suppose the result we want to achieve is that in NSLs  $AGR^0$ , as a rich AGR-identifier, can assign Nominative directly to its I-subject. In non-NSLs, on the other hand, Nominative can only be assigned to the Spec of AGR. To derive this result, let us assume the following principles:

---

<sup>71</sup> If we were to assume Fukui & Speas' (1986) hypothesis, in which Specifiers are projected only when they are filled, we could dispense with this statement: if the AGR-identifier has to be present, then it has to be filled. In any case, 0 is an almost self-evident statement.

- (53) AGR<sup>0</sup> can optionally assign Nominative Case to Spec of AGR by agreement or to some other position under government.
- (54) The I-subject must receive Case from its AGR-identifier.

The notion of 'receiving Case from' in 0 is intended to cover two notions:

- Case assignment by government.
- Case transmission along an A-Chain (assuming that Spec of AGR, when filled, can form a Chain with the I-subject).

For languages such as Italian, the I-subject will receive Case from AGR<sup>0</sup>, which will directly assign Case to it by government. In English, the I-subject will receive Case from Spec of AGR, i.e., the DP in Spec of AGR will transmit its Case to the I-subject via A-Chain transmission. So, the DP in Spec of AGR has to receive Case itself. In a finite sentence, it receives Nominative Case from AGR. In an infinitival sentence, it receives Accusative (in ECM constructions) or it is PRO (we will assume that also PRO has Case).

In Italian, AGR<sup>0</sup> has to assign Nominative Case to its I-subject under government. We postpone the discussion of infinitival constructions to the next chapter. Let us assume that finite AGR directly assigns Nominative to the I-subject.

If AGR is the highest functional category, as will be crucial in section 2.3., this means that AGR does not govern the I-subject. Suppose, however, we adopt the following convention, which we will revise in Chapter 4 (for the moment let us take it as a provisory stipulation):<sup>72</sup>

- (55) If AGR<sup>0</sup> combines with a head X by incorporation, it has the same governing capacities as X as far as Case assignment is concerned.

---

<sup>72</sup> 0 is clearly at variance with Baker's (1985) Government Transparency Corollary.

0 makes the prediction that T to AGR raising is obligatory in languages such as Italian, since this is the only way AGR can combine with a head which actually governs the I-subject. If the I-subject is in object position, V (to T) to AGR raising is also necessary to make sure AGR governs the I-subject in the relevant sense expressed in 0. In sum, V to AGR raising is a necessary condition for  $AGR^0$  to be able to assign Case to its I-subject.<sup>73</sup> So we predict that NSLs will always have long V movement (at least in finite sentences). In the next chapter, we will take advantage of the same idea to account for V-raising in infinitives in NSLs.

For convenience, we will call Nominative assignment in accordance with 0 *Chain-Government*. Notice it is a similar device to Chain extension in Chomsky (1986-b): both are a means of covering an otherwise too long gap between an I-subject position and AGR. But I think it has two advantages over it:

- in Chomsky's proposal, Chain-extension involves coindexation of several heads (V and INFL), and this coindexation has to be used as referential coindexation (after some sort of indexation merging takes place), since antecedent government is, at least intuitively, government of a referentially dependent element. It does not seem natural that indexes of V- and INFL-heads should be involved in referential indexing. Since in our case Nominative Assignment through Chain government is a purely formal (not referential) process, similar considerations cannot be so compelling.

- Chain extension yields Chains consisting of both maximal projections and heads, which is again an awkward result, while our Chain-government strictly resorts to head Chains.

In our case, we in fact may need some *extended Chain-government* device. Consider the case of a an auxiliary verb plus participle construction, where the auxiliary raises (this is the case for Italian). If we have an object position I-subject,  $AGR^0$  will Chain govern the I-subject only if the auxiliary and the participle head Chains are united into an extended Chain. The foot of the

---

<sup>73</sup> Clearly, it is not a sufficient condition, in view of the fact that French has V to AGR-raising, if we reinterpret Pollock's (1989) conclusion that V in French raises to the topmost FC.

auxiliary Chain will plausibly govern the head of the participle Chain, since it is reasonable that auxiliaries subcategorize for the participial form they are constructed with. So the link between the two sub-Chains is government and this might be a necessary condition for Government-Chain extension. What are sufficient conditions is not clear, but intuitively auxiliaries and participles are closely related entities and it looks reasonable that they share some index of some sort.

In Catalan and Spanish, participles are most likely incorporated to auxiliaries (see Llinàs (1988)), so that, on the one hand, Government-Chain extension is not needed for these languages (as far as auxiliary-participle constructions with incorporation are concerned), and, on the other, this incorporation is suggestive of the close relation we claim is between auxiliaries and participles.

It is perhaps significant that the cases where Chain extension is required in Italian involve agreement with the I-subject in both the auxiliary and the participle:

- (56) Erano ... venuti<sub>j</sub> [<sub>VP</sub> *t<sub>j</sub>* *i bambini* ]  
           Be-3rd-pl come-msc-pl                      the children(3-pl-msc)

Perhaps this situation makes Chain extension easier, in that the heads of the two sub-Chains agree and so share some index. Or the other way around, perhaps this agreement is a manifestation of Chain extension. Remember that in the present account, there is no A-movement from the basic position occupied by *i bambini* to the Spec of the upper AGR (where the auxiliary stands), so that the Agreement morphology in the participle cannot be triggered by some intermediate subject trace in the specifier of the participial FC. On the other hand, in Catalan and Spanish this agreement pattern has entirely disappeared: the reason would be that since in these languages the participle incorporates into the auxiliary, Chain extension is not required.

In Chapter 4 we will use Chain-extension in another construction: raising in NSLs.

In the preceding chapter we saw that Modern Standard Arabic (MSA) had the peculiarity of having two 'inverted subject' positions (which can in fact cooccur):

- one which is post-VP (it follows the complements) and is [+anaphoric].
- one which is pre-VP (it precedes the complements, *pace* V-movement) and is [-anaphoric].

(57) 'Amala Yusuf/huwa al 'amal nafsuhu

Did Y. /he the work he-SELF

'Yusuf/he did the work himself'

For theoretical reasons that will be clear in the next section, we will assume that the latter position (*Yusuf/huwa*) is the one receiving Case by government, while the former (*nafsuhu*) receives case via Chain-transmission from the other. For some reason, then, AGR<sup>o</sup> in MSA is able to assign Case to only one of the positions, namely the higher one.

Let us address another question. We claimed that the I-subject obtains Case from its AGR-identifier, so that, at least in finite clauses, the I-subject is Nominative. Perhaps nobody would challenge the claim that I-subjects in NSLs (i.e., inverted subjects) are Nominative. This is not so clear for what we claim are I-subjects in non-NSLs (*himself*, *soi-même* in French, etc.), for they have the same morphological shape as when these elements are used as object or oblique reflexive elements.

We will assume, however implausible it may look at first glance, that these elements are not Accusative or Oblique. They are rather morphologically Case-neutral, (as all full DPs are in these languages). The only morphological Case distinction in these languages lies in Nominative pronominals, all other forms (including anaphors) being Case neutral. In Chapter 4 we will go further to suggest that Nominative pronouns in languages like French or English are forms specific to Spec of AGR.



MSA is revealing in this connection: the post-VP [+anaphoric] I-subject is unambiguously an anaphoric element: besides being an emphatic I-subject, it can only be used as a reflexive (it does not have the logophoric usage of, say, English *himself*). However, when used as an I-subject, it is unambiguously Nominative. The same can be said of Icelandic SELF element.

(58) Jón gerdhi thetta *sjálfur*

J. did that SELF<sub>Nom</sub>

Thus our contention that [+anaphoric] I-subjects (corresponding to what traditionally has been called emphatic subject anaphors) are Nominative is confirmed by languages having rich Case morphology, and is tenable for less revealing languages, where we have to challenge the traditional idea that forms like English *himself*, French *lui(-même)* are not Nominative: this idea would only be correct from the historical point of view.

## 2.3. Binding Theory

### 2.3.1. A Definition of Binding Domain

Suppose the Specifier of AGR counts as an A-position for BT, and that in English AGRP is the Binding Domain for the I-subject. In this language, Spec of AGR always binds the I-subject as a consequence of choosing the Spec-of-AGR option in 0.c): the I-subject is coindexed with AGR and, since AGR<sup>0</sup> agrees, in the unmarked case, with its Specifier, the Specifier binds the I-subject. Under the BT principles, this predicts that the I-subject can only be anaphoric, as we have assumed above.

On the other hand, we want to derive the fact that Italian I-subjects are [-anaphoric]. This result can be achieved if, in NSLs:

- (59) a. Spec of AGR need not be obligatorily filled, in order to allow for R-expressions to occur as I-subjects.
- b. If it is filled, it is outside the Binding Domain for the I-subject, so that the I-subject can be pronominal but not anaphoric.

0.a) can be seen as a consequence of the fact that Spec of AGR (and Specifiers in general) is, in principle, optionally filled. It will be obligatorily filled only if it is an AGR-identifier, which is not the case for NSLs.

To derive 0.b), we will assume the following definitions of Binding Domain (BD):

- (60) A is Binding Domain for B iff A is the minimal FC containing B, a governor of B and the Case-marked position from which B obtains Case.

In the preceding section, we argued that in NSLs, the I-subject itself is a Case-marked position, while in non-NSLs, it is Spec of AGR that is a Case-position (it is Case-marked or it contains PRO, which is intrinsically Case marked) and transmits its Case to the I-subject. So, in non-NSLs, AGRP is the Binding Domain (BD) for the I-subject, while in NSLs, it is the first FC maximal projection containing the I-subject, namely TP.<sup>74</sup>

So, if Spec of AGR is outside the BD of the I-subject, the I-subject has no antecedent in its BD and has to be [-anaphoric], i.e.:

---

<sup>74</sup> Notice that for an object anaphor as in (i):

(i) John hates himself  
the BD will be TP, and its binder will be the I-subject. In any internal subject theory, the local binder of an object anaphor is the internal subject, so this result is not problematic.

In the definition above, we stipulate that a BD has to be a Functional Category: otherwise, if VP is a maximal projection inside VP<sup>max</sup>, as in Koopman and Sportiche's hypothesis, the BD for the object would be VP, and the anaphor in (i) would be free in its BD.

- a full DP or pronominal (subject inversion).
- a variable, which accounts for the absence of that-t effects.<sup>75</sup>
- a *pro*, which is licenced by the AGR identifier, which recovers its content.

The last point is one where the present theory differs from standard assumptions. In section 2.4. we will address the question. In the next section we will provide some independent evidence for the definition of Binding Domain we have postulated.

One obvious alternative to the rather intricate definition of Binding Domain we have proposed to account for the facts would be the following: preverbal subjects in NSLs are not A-positions, and therefore do not count for Binding Theory. Then the I-subject cannot be [+anaphoric] because as such it would have no possible A-binder. In Chapter 5 we are going to discuss the status of preverbal subjects in NSLs. Although the conclusion will be that they do not have the same status as in non-NSLs, I prefer not to commit myself to the claim that they are not A-positions: nowadays, within the internal subject hypothesis, the concept of A-position is a delicate question, and I have tried to make my theory orthogonal to the issue. If the notion of A-position is to be kept, however, I adhere to Rizzi's (1991) proposal, which likely would give the result that preverbal subjects are A-positions in NSLs. For a different matter, Rizzi's definition of A-position (which makes Spec of AGR-object in French an A-position) will be crucially used in Chapter 4, section 6, to account for French exceptional behaviour as a non-NSL.

---

<sup>75</sup> We adhere to Rizzi's (1982-b) contention that in NSLs it is the postverbal subject position that (exclusively) allows Wh-movement without that-t effects. We account for the availability of the postverbal source for Wh-movement, and we are also in a position to account for the exclusiveness of this source, i.e., why is it that preverbal subjects are not extracted by some that-deletion or equivalent means in NSLs, if Rizzi is right in claiming that this never happens: as we will see in Chapter 5, preverbal subjects are not candidates for Wh-extraction because they are dislocated-like elements resumed by a pronominal in the I-subject position.

### 2.3.2. Anaphoric Copulative Constructions

We have adopted a specific definition of Binding Domain in order to capture the correlation between (non) NSLs and the [ $\pm$ anaphoric] character of I-subjects. The definition adopted is intended to be neutral w.r.t. the classical facts in BT, since what we add to a classical definition is reference to a Case position for the bindee, which is supposed to be a trivial matter for any Argument subject to BT in the standard cases (for instance, the Case position for an object anaphor will be the position where this object anaphor stands). Therefore, our redefinition of Binding Domain is an *ad hoc* device to capture the I-subject facts. In this section, we are going to see what can be taken to be independent evidence for our definition of Binding Domain. Consider the following paradigms:

- (61) After this emotional shock...
- a. John is not *himself* anymore
  - b. Jan is *zichzelf* niet meer (Dutch)  
J. is SE-SELF not more
  - c. Hans ist *sich selbst* nicht mehr (German)  
H. is SE SELF not more
  - d. John är inte längre *sig själv* (Swedish)  
J. is not longer SE SELF
  - e. On doit être *soi-même* (French)  
One must be oneself

(62) After this emotional shock...

- a. En Joan ja no és *ell* (*mateix*) (Catalan)  
The J. anymore not is he (SELF)
- b. Joan es pas mai *el* (*mateis*) (Occitan)  
J. is not more he (SELF)
- c. Juana ya no es *ella* (*misma*) (Spanish)  
J. anymore not is she (SELF)
- d. Gianni non è più *lui* (*stesso*) (Italian)  
G. not is anymore he (SELF)
- e. O Jo~,ao já n~,ao é *ele* (*mesmo*) (Portuguese)  
The J. anymore not is he (SELF)
- f. Ion nu mai este *el* *insusi* (Romanian)  
I. not anymore is he SELF
- g. Jon ez da *bera* /\**bere burua* (Basque)  
J. not is he-himself his self

Sentences in the reverse pattern (i.e., non-NSLs using *he-SELF* or NSLs using *SE-SELF*) are completely unacceptable. Let us call the above examples Anaphoric Copulative Constructions. The interesting fact is that non-NSLs use an anaphor-like element in the post-copular position, while NSLs use a pronominal-like element in the same position. Since this contrast is reminiscent of the contrast between emphatic I-subjects in non-NSLs vs NSLs, it is tempting to derive the two facts from the same premise. Before providing an explanation, let us advance some more evidence from other languages.

As expected, MSA shows a [-anaphoric] element in one position and a [+anaphoric] one in the other:

(63) Yusuf mà huwa nafsuhu

J. not he he-self

'J. is not himself'

Cf.:

Yusuf mà huwa tawil

J. not he tall

'J. is not tall'

Hebrew, which we saw was not very telling w.r.t. the anaphoric/pronominal status of I-subjects (it apparently allows neither option), provides very interesting data in this connection. As is well-known, Hebrew is a mixed language w.r.t. null subjects: it allows them only with certain verbal forms (past tense 1st/2nd person) and contexts (embedded sentences having a subject bound by the superordinate subject). Even if it is not clear what this pattern amounts to from a theoretical point of view, one possible prediction could be that in the kind of constructions we are considering are sensitive to the (non) null-subject context. This prediction is roughly borne out: in null-subject contexts, a pronominal is preferred, as shown by the contrast in 0, whereas in non-null-subject contexts an anaphor is preferred, as shown in 0:<sup>76</sup>

---

<sup>76</sup> Thanks to Uri Shlonsky, who first suggested the possibility of having such a contrast, Tali Siloni, Hagit Borer and Erez Bronstein. Not all of them agree on the judgements, but their disagreements consist in simply not seeing some of the contrasts rather than having opposite values. Only Tali Siloni points out that for past-3rd person (which is a non-null-subject context) she prefers a pronominal -(i)- or even better a pronominal-SELF expression -(ii)-:

- (i) Hu lo haya hu/??acmo ba-mesiba  
He not was he himself in-the-party
- (ii) Hu lo haya hu-acmo ba-mesiba  
He not was he-SELF in-the-party

- (64) 'I think that you were not yourself at the party' =
- a. 'An xoshev she-lo hayyta 'ata ba-mesiba  
I think that-not you-were you at-the-party
- b. ??'An xoshev she-lo hayyta 'acmexa ba-mesiba  
I think that-not you-were yourself at-the-party
- (65) 'I think that you are not yourself these days' =
- a. 'An xoshev she-'ata lo 'acmexa be-yamin 'elu  
I think that-you you-are yourself in-days these
- b. ??'An xoshev she-'ata lo 'ata be-yamin 'elu  
I think that-you you-are you in-days these

In fact, judgements are subtle and slightly varying. But even if some speakers do not see any contrast, the mere fact that both a pronominal and an anaphor are possible is not at odds with the mixed status of Hebrew w.r.t. the null subject phenomenon.

In order to account for the above facts, we will resort to our definition of Binding Domain above, although not in a trivial way, as we will see.

Let us consider what is the structure of the above constructions. Let us assume, with many linguists, that copulative constructions always involve predication: even in apparently equative copulative constructions one of the two elements is the predicate. Let us also assume that copulative constructions involve a Small Clause, so that the S-structure of *John is intelligent* is 0:

- (66) a.  $\text{John}_i$  is [<sub>SC</sub>  $e_i$  intelligent ]

where  $e_i$  is the I-subject (the order I-subject/predicate is irrelevant).

If we apply the same analysis to the above examples, their essential structure would be,

taking English and Catalan as representative examples:

- (67) a. John<sub>i</sub> ... is [<sub>SC</sub> e<sub>i</sub> himself ] ...  
b. En Joan<sub>i</sub> ... és [<sub>SC</sub> e<sub>i</sub> ell (mateix) ]

Suppose that, in this particular kind of structure, the predicative element has the *exceptional property of being coindexed with its subject in the kind of coindexing relevant for Binding Theory*.

What kind of indexation is it? One possibility is reference indexation. In the present case, however, we cannot simply say that the subject (*John/en Joan*) and the predicate (*himself/ell*) actually co-refer, for one is an Argument and the other is a predicate. Suppose, however, that the type of indexation relevant for BT is that of denotation. In fact the subject and the predicate denote the same ('John'), even if, respectively, in one case it is a token ('the actual instantiation of the entity John') and in the other a prototype ('prototypical John').

With such a structure in 0, however, if we want to treat *himself/ell (mateix)* as an element subject to BT, we face an obvious problem: if *himself* is an anaphor bound by the subject in 0.a), it will be locally bound by the I-subject. Since this local binding holds the same for the Catalan example, the prediction would be that there should be no contrast between the two (sets of) languages: both should have an anaphor.

On the other hand, though, since we have assumed that *himself/ell (mateix)* are predicates in these constructions, they should not be subject to Binding Theory, if BT only applies to A-positions. So, suppose that their insensitivity to binding by the I-subject is due to the simple fact they are not Arguments, BT being a theory only relevant for Arguments.

Suppose, however, that the SC which appears in these constructions, as a propositional constituent, can be assumed to be an Argument: from a strictly formal point of view, it would be the Argument subcategorized by the copular verb. Although no predication relation holds between a copular verb and its Small Clause complement, because the copular verb does not



convey any meaning, we can assume that from a formal point of view, the Small Clause counts as an Argument, and therefore as an element subject to BT. Then, the anaphoric/pronominal contrast observed above will not be a manifestation of the BT-status of the predicate itself (which is simply not definable) but rather of the status of the Small Clause, although *it will be the nucleus of the Small Clause, namely the predicate, that will show the morphological contrast*. On the other hand, treating a nominal Small Clause as a formal Argument of 'to be' explains why nominal predicates (specifically [+N] predicates) often manifest Case: it would be a manifestation of the Case that the Small Clause they head requires as an Argument.

Coming back to our line of argumentation, the I-subject being internal to the Small Clause, it cannot bind it. The subject in Spec of AGR, instead, *does* bind the Small clause if we assume that:

- it is coindexed with the predicate with co-denotation coindexing, as we have proposed;
- the index of the predicate percolates to the whole Small Clause it is the nucleus of.

If so, let us try to apply our definition of Binding Domain, repeated here as 0, to the Small Clause:

(68) A is Binding Domain for B iff A is the minimal FC containing B, a governor of B and the Case-marked position from which B obtains Case.

Assuming that the copular verb is a governor of the Small Clause, the crucial step is to determine which is the Case position for it, if there is one at all. Let us assume that the Case of the Small Clause will be the same as the Case manifested by its head, the predicate. The question is: do Small Clauses (or their predicates) have Case? As far as languages with overt Case can tell, Small Clauses (or their predicates) do seem to have Case. The general pattern is that the Case manifested by the predicate of a Small Clause is usually the same as the Case of its subject (examples from Greek and German):

(69) Greek:

a. O Kóstas<sub>Nom</sub> íne [<sub>SC</sub> t<sub>Nom</sub> o kalíteros ipopsífi<sub>Nom</sub> ]

Kostas is the best candidate

b. Theorí [<sub>SC</sub> ton Kósta<sub>Acc</sub> ton kalítero ipopsífi<sub>Acc</sub> ]

He-considers Kostas the best candidate

German:

c. Dein Bruder<sub>toNom</sub> ist [<sub>SC</sub> t<sub>Nom</sub> ein guter Mensch<sub>Nom</sub> ]

Your brother is a good person

d. Ich finde [<sub>SC</sub> deinen Bruder<sub>Acc</sub> einen guten Mensch<sub>Acc</sub> ]

I find your brother a good person

Suppose that, somehow, a Small Clause inherits its Case from its subject: then the source for Case for the Small Clause in the copulative constructions under consideration is the same as that of its subject, namely Spec of AGR in non-NSLs and the I-subject itself in NSLs. Therefore, the Binding Domains will be, respectively, AGRP and TP. If it is AGRP (in non-NSLs), the Small Clause is bound by the preverbal subject inside the Binding Domain and is, therefore, anaphoric. In NSLs, on the other hand, the Small Clause will be pronominal, because there is no binder inside TP.

Although the preceding analysis is far from crystal-clear (we need some specific stipulations about the nature of these constructions) I think something of what we assumed must be on the right track: the contrast 0/0 looks significant: the judgements about the examples are steady, in spite of the fact that they are unlikely candidates to having been learned as idiosyncratic facts, because of the unusual character of the sentences. In fact, in Romance these constructions are not as usual as in English. So some speakers react with a certain reluctance to them. But when they are asked to confront them with the ones having the wrong element (i.e., the reflexive in a NSL), their judgements are sharp. This is a further argument for the poverty of

stimulus: probably the reluctant speakers had never used or even heard the sentences predicted to be good, but they 'know' that they are at least much better than the ones predicted to be wrong.

Summing up, we have established and accounted for the following generalization:

- (70) In Anaphoric Copulative Constructions, the post-copular element is [-anaphoric] in NSLs and [+anaphoric] in non-NSLs.

There are some cases that seem to be problematic or constitute counter-examples to this generalization. First of all, some languages simply do not have the construction: Brazilian Portuguese, Russian and Georgian cannot, as far as I know, express the idea of 'being oneself' with a similar construction. They instead use alternative paraphrase (such as 'to be the same person he was', 'to have changed radically', etc.). I think this is to be expected: we are dealing with a rather idiosyncratic construction.<sup>77</sup>

There are, in addition, two cases that are more problematic. One is Greek. In Greek, it is the non-clitic reflexive which is used in this case, rather than a pronominal:

- (71) O jánis dhen íne pja o eafτός tu  
The J. not is anymore the self his

where *o eafτός tu* is the Nominative form of the reflexive in Greek. I think, however, that this is not necessarily a counterexample. Notice that this reflexive form has an internal structure where the actually bound element is a pronominal in the genitive position:

---

<sup>77</sup> One could even think that 'John is (not) himself' is a tautology/contradiction, and that only pragmatic efforts to make the best out of any expression (in the spirit of Sperber & Wilson (1986)) rescue it.

(72) [DP o [NP eaftós [DP<sub>Gen</sub> tu ] ] ]  
           the self his

Iatridou (1988) argues that the genitive pronominal inside the reflexive phrase is not an anaphor itself. Rather it is forced to have a proximate antecedent by being inside a reflexive DP (which is not bound as such). Suppose this particular device for reflexivity has the following properties:

- a reflexive DP of this type is not an anaphor from a formal point of view. However, to be licenced at LF (under Full Interpretation -Chomsky (1988)), it has to inherit the index of its genitive: then it will be interpreted as reflexive.

- when this reflexive DP is the predicate of a Small Clause, however, this process of inheritance does not extend to the Small Clause, perhaps because the denotational indexing relevant for predication percolates earlier than LF.

I think that this could be the basis for explaining the otherwise problematic behavior of Greek for Anaphoric Copulative constructions. This explanation may be tricky, and obviously a more accurate study of these kind of reflexives (where the bound element is a genitive) would be required. On the other hand Basque has a similar type of reflexive:

(73) [DP [NP [DP<sub>Gen</sub> bere ] buru ] -a ]  
                                   his head the  
                                   'Himself'

and does not use this reflexive form in Anaphoric Copulative Constructions (it rather uses the logophoric/emphatic form). I cannot say anything about the issue.

Another potential counterexample is found in Icelandic: unlike in Mainland Scandinavian, the element used in this construction is not the one used as subject oriented

reflexive, but rather the one used as anti-subject oriented reflexive (which is in fact a logophoric pronoun in shape):

- (74) Jón er ekki lengur *hann sjálfur*/\**sig sjálfur*  
J. is not longer him SELF SE SELF

Since this is obviously a case of subject orientation, the facts are not expected. H. Sigurdhsson (p.c.) pointed out to me that perhaps the problem is that *sig* cannot be used because it lacks a nominative form. For all the other germanic languages we have been implicitly assuming that *sig selv/sich selbst/etc.* are Nominative forms in the Anaphoric Copulative Constructions. Then, why is Icelandic reluctant to use *sig* as Nominative?

I think there is a notable difference between Icelandic and Mainland Scandinavian, English or French: in the latter, there are some Nominative pronominals that can only be used in Spec of AGR position:<sup>78</sup>

- (75) a. *He* is sick  
b. It is *him*/\**he*  
c. *Him*/\**he*, he's a liar

In these languages, the so-called Nominative pronouns are special forms confined to Spec of AGR. The forms usually taken to be non-Nominative are, I assume, unmarked. So let's assume in these languages, by extension, even reflexive forms would be unmarked, and therefore can be used as Nominative forms unproblematically. We could say that in some languages having unmarked forms is the unmarked case.

The situation is quite different in Icelandic: Case morphology is not confined to

---

<sup>78</sup> I assume that Normative pressure concerning (0.b) is forcing an ungrammatical construction.

pronominals and it is quite rich. Therefore, taking *sig* as Nominative would contradict the general pattern, namely that DPs have no neutral Case-forms. As for German, which also has Case morphology across the board, we have to assume that this language does have neutral Case-forms (at least for *sich selbst*). In fact, German Case-morphology is much poorer than in Icelandic. So, unlike in Icelandic, the German *sich selbst* element does not distinguish between Nominative and Accusative:

(76) German:

- a. Hans ist *sich selbst* nicht mehr  
H. is SE SELF<sub>Nom</sub> not more
- b. Hans sag *sich selbst*  
H. saw SE SELF<sub>Acc</sub>

Icelandic:

- c. Jón er ekki lengur *hann sjálfur*  
J. is not longer he SELF<sub>Nom</sub>
- d. Jón meiddi *sig sjálfan*  
J. hurt SE SELF<sub>Acc</sub>

Therefore, the problematic case in Icelandic could be assumed to be due to morphology: no Nominative or neutral form being available for *sig*, the Nominative form *hann* has to be used as a suppletion for *sig*.

Even if there are some problematic cases, I think that the contrast shown in 0/0 is significant enough not to disregard the issue.

### 2.3.3. On the Status of Binding Theory

We have proposed a definition of Binding Domain which is intended to cover two empirical phenomena (the [ $\pm$ anaphoric] status of I-subjects and of copulative predicates) that, as far as I know, had not been addressed in the literature thus far and were not even considered relevant for Binding Theory.

Binding Theory has often been conceived of as a means for accounting for co-reference restrictions between DPs having independent  $\theta$ -roles:<sup>79</sup> in the above proposal, however, I-subjects do not have independent  $\theta$ -roles from their (possible) preverbal antecedents or, in the case of copulative clauses, the post-copular element is not a referring entity (it is a predicate). So they are not cases of co-reference in a reasonable sense of the term.

Then, if the above approach is on the right track, it strongly suggests that Binding Theory is a purely formal device which blindly extends beyond the scope of co-reference.

A second issue to be addressed concerns the empirical complexity of Binding Theoretical facts: recent research on a variety of languages has shown that Binding Theory is much more complex than early studies about English and similar languages suggested (one has to face complexities as long distance binding, subject (anti-)orientation, logophoricity, etc.).<sup>80</sup> Our reformulation of Binding Theory does not say anything about these issues, and one might suspect it is too naively tied to a simplistic view of classical Binding Theory.

Although I admit that a more comprehensive approach to Binding Theory is necessary, I think the present proposal has several advantages:

- as we pointed out, it is neutral w.r.t. the standard cases of Binding Theory: the

---

<sup>79</sup> See for instance Reinhart & Reuland (1991) for such a view.

<sup>80</sup> See for instance Hestvik (1990) and Vikner (1985) for Scandinavian, Everaert (1986) for Dutch, Koster & Reuland (1991) for long distance anaphora.

additional requirement of a Case-position for A in the determination of the Binding Domain for A does not affect objects or oblique Arguments, for it is implicit in standard accounts that for them the Case position is the position where they stand at S-structure.<sup>81</sup> Clitics (especially in clitic doubling structures) could be a case where this is not true, but I assume that clitics involve head Chains and head Chains as such are not subject to Binding Theory.

- it seems to me that, in the field of anaphora, there are two privileged, 'more central', unmarked cases: these are *subject-oriented* anaphors and *clause-bound* anaphors. The binding-theoretical account advanced above for I-subjects and anaphoric copulative constructions clearly belongs to the field of subject-oriented clause-bound anaphors.

- The definition of Binding Domain we proposed does not make reference to accessible SUBJECTS or to Complete Functional Complexes (see Chomsky (1981)/(1986-b)). Since any FC (including CP) can be a Binding Domain, the traditional problem of excluding examples like:

(77) \*I think that myself am sick

is solved provided we can assume that C<sup>0</sup> is an appropriate governor for this definition.<sup>82</sup> I think that Accessible SUBJECT is a tricky notion, and that the notion of CFC cannot be used once we assume the Internal Subject Hypothesis. We will discuss this issue in the next chapter in connection with PRO.

So, the present proposal, although not intended to provide any new insight into the standard cases of BT, seems to fit into it without problems.

---

<sup>81</sup> Chomsky's (1992) proposal of movement of the object to an AGR-Obj specifier does not challenge this idea: an object will be anaphoric w.r.t. the external Argument in a domain where the object has a case position, namely Spec of AGR-Obj.

<sup>82</sup> A potential problem could be:  
 ??They want very much for each other to be happy  
 If acceptable, CP does not count as a Binding Domain. Perhaps the rather acceptable status of this sentence would be due to the possibility for **for each other** to be interpreted as a benefactive controlling the infinitive.



## 2.4. The Position of *pro* and Floating Quantifiers

In the above account, it is crucial that *pro* is able to occur as an I-subject, contrary to the standard assumption that subject *pro* is only licenced in the Spec of INFL/AGR. In the present account, pro-drop is in some sense an epiphenomenon: the I-subject can be null in all languages: it will be *pro* in NSLs and a null anaphor in non-NSLs. So Taraldsen's (1980) original idea that AGR morphological richness is the key to pro-drop is here expressed in some indirect way: richness makes AGR<sup>0</sup> the AGR-identifier and this leads, through Binding Theory, to the existence of a null pronominal as I-subject. Therefore, we cannot assume that *pro* is licenced in Spec of AGR.

Rizzi has adduced two pieces of evidence in favor of the preverbal position for *pro* (the argumentation is reproduced in Roberts (1989) and Roberts (1991-a)). One of them is expressed by the following paradigm (from Italian):

- (78) a. Essendo stanco, Gianni è andato via.  
Being tired G. is gone away  
'Being tired, Gianni went away'
- b. Essendo stanco, è andato via.  
Being tired is gone away  
'Being tired, he went away'
- c. \*Essendo stanco, è andato via Gianni.  
Being tired is gone away G.  
'Being tired, JOHN went away'

Rizzi's argument, adapted in terminology, is as follows: the gerund subject is a PRO which has to be controlled by a preverbal subject, which can m-command it. This is why the

postverbal subject in 0.c), which cannot m-command the gerund, does not licence the controlled PRO and the sentence is ungrammatical. Since both the pre-verbal subject in 0.a) and the null subject in 0.b) licence the gerund, the conclusion is that *pro* is preverbal.

Catalan (and Spanish) behave in a similar way:<sup>83</sup>

- (79) a. Estant tan cansat, en Joan se n'ha anat al llit  
 Being so tired the J. SI-off-has gone to-the bed  
 'Being so tired, Joan went off to bed'
- b. Estant tan cansat, se n'ha anat al llit  
 Being so tired SI-off-has gone to-the bed  
 'Being so tired, he went off to bed'
- c. \*Estant tan cansat, se n'ha anat al llit en Joan  
 Being so tired SI-off-has gone to-the bed the J.  
 'Being so tired, JOAN went off to bed'

There are two main objections to Rizzi's account for the facts. Firstly, if we adjoin the gerund to a superordinate sentence, we obtain the following paradigm:

---

<sup>83</sup> Other non-finite sentences, such as temporal infinitives in Catalan (and Spanish) (see Rigau (1992) for an analysis of this construction), manifest the same paradigm as 0:

- (i) A l'arribar, (en Joan) se'n va anar al llit  
 At the-to-come J./**pro** went off to bed  
 'When Joan/he arrived, he went to bed'
- (ii) \*A l'arribar, se'n va anar al llit en Joan  
 At the-to-come went off to bed J.

Rigau convincingly argues that these infinitives do not involve PRO. If so, the facts cannot be explained as a problem of control, as we will see.

- (80) a. Estant cansat, és millor que en Joan se'n vagi  
 Being tired is better that the J. goes away
- b. Estant cansat, és millor que se'n vagi  
 Being tired is better that goes away
- c. \*Estant cansat, és millor que se'n vagi en Joan  
 Being tired is better that goes away the J.

The examples in 0 are parallel in acceptability to the ones in 0 except that the gerund is in a position where c-command (or m-command) by the preverbal subject of the embedded clause is not possible. So the claim that c-/m-command is a necessary condition to licence the gerund PRO cannot be maintained. There must be some other account for the fact that the postverbal subject does not licence the gerund.

In order to preserve Rizzi's analysis, one could assume that control takes place at some level where the gerund is adjoined to the embedded sentence, the surface structure in 0 being a result of gerund raising. This level, however, cannot be D-structure, since at that level the subject of *anar-se'n* 'go away' does not m-command the gerund anymore than the postverbal subject, on the assumption that *go away* is an unaccusative verb and its subject is an object at D-structure.

But independently of this problem, there are more basic empirical problems for the control account. If we replace the gerund by a finite adjunct clause, the distribution of acceptability does not vary.<sup>84</sup>

---

<sup>84</sup> The paradigm in 0 has a correlate in Italian:

- (i) a. Come era stanco (Gianni) è andato via  
 As he-was tired (G.) is gone away  
 b. \*Come era stanco è andato via Gianni  
 As he-was tired is gone away G.

Paradigm 0 also has an Italian correlate. However, since gerunds are a little literary, dislocation of the gerund to an upper clause gives slightly awkward results:

- (ii) a. Essendo stanco è meglio che (Gianni) vada via  
 Being tired is better that (G.) goes away  
 b. \*Essendo stanco è meglio che vada via Gianni  
 Being tired is better that goes away G.

- (81) a. Com que estava cansat, en Joan se n'ha anat  
           As    he-was tired   the J.   went away
- b. Com que estava cansat, se n'ha anat  
           As    he-was tired   went away
- c. \*Com que estava cansat, se n'ha anat en Joan  
           As    he-was tired   went away   the J.

In 0 no controlled PRO is involved, so the paradigm has to be explained in an alternative way. I think the facts in all the preceding c. examples, where the postverbal subject cannot co-refer with the subject of the adjunct clause, can be accounted for as cases of WCO. Since the post-verbal subject is interpreted as Focus, and Focus triggers WCO effects, as can be seen in 0 (see Chomsky (1981)), the post-verbal subject cannot be co-referent with a pronoun it does not c-(or m-)command:

- (82) \*His<sub>i</sub> mother loves JOHN<sub>i</sub>

In fact, in Catalan Rizzi's paradigm has a correlate in cases where we are not dealing with a subject, but a dislocated/clitic object.<sup>85</sup>

---

For paradigm 0 if we dislocate the adjunct clause to an upper clause the acceptability results do not change in either Catalan or Italian.

<sup>85</sup> Italian gerunds seem to be subject oriented, and so this paradigm cannot probably be instantiated in Italian.

- (83) a. *PRO* estant malalt, a *en Joan*<sub>i</sub>, *l'anirem* a veure  
           being ill to the J. him-we'll-go to see  
           'Since he's ill, we'll visit Joan'
- b. *PRO* estant malalt, *l'anirem* a veure  
           being ill him-we'll-go to see  
           'Since he's ill, we'll visit him'
- c. ?\**PRO* estant malalt, *anirem* a veure *EN JOAN*  
           being ill we'll-go to see the J.  
           '\*Since *he's* ill, we'll visit *JOAN*'

Both 0.a) and (b.) cluster together as opposed to the non-dislocated Focus object, which triggers WCO. The only difference between this paradigm and Rizzi's paradigm is that the object, as opposed to the postverbal subject, need not be Focus (so that 0.c) improves if *en Joan* is not Focus.

So the paradigm adduced by Rizzi proves irrelevant for the position of null subjects: in 0.b) we could not claim there is a preverbal empty object c-commanding the gerund.

In fact there are cases where pre-verbal subjects have co-reference restrictions which do not cluster with null subjects. Consider:

- (84) a. El cotxe d'en Joan<sub>i</sub>, *ell (mateix)<sub>i</sub>*, no el condueix  
 The car of-the J. he (SELF) not it-drives  
 'Joan's car, *he himself* never drives'
- b. ?\*El cotxe d'en Joan<sub>i</sub>, no el condueix (-> *pro<sub>i</sub>*)  
 The car of-the J. not it-drives  
 'Joan's car, *he* does not drive'
- c. El cotxe d'en Joan<sub>i</sub>, no el condueix *ell<sub>i</sub>*  
 The car of-the J. not it-drives he  
 'Joan's car is not driven by *him himself*'

In 0 *pro* contrasts with overt subjects, which behave alike. I think the explanation for that paradigm could be based on the idea that (in some languages) a null subject always involves a topic, whether overt or not. If it is overt it can appear in several places (it can adjoin in a higher or lower position, even in superordinate clauses). A topic cannot bind an R-expression. In a strict sense, the dislocated element *ell mateix* in 0.a) does not c-command the other dislocated element *el cotxe d'en Joan* containing the co-referential R-expression (*en Joan*), because *ell mateix* appears to the right of the other left adjoined element: the first branching node including the inner adjunct will not contain the outer adjunct.

Now suppose null topics always have maximal scope. Then 0.b) would be excluded for the same reason as:

- (85) \**Ell mateix<sub>i</sub>*, el cotxe d'en Joan<sub>i</sub> no el condueix  
 He SELF the car of-the J. not it-drives  
 '\**He* does not drive Joan's car'

where *ell mateix* illicitly binds *en Joan*. I leave the issue here.

The other empirical argument presented by Rizzi is based on the following paradigm:

- (86) a. I soldati sono andati tutti via  
The soldiers are gone all away  
'The soldiers have all gone away'
- b. Sono andati tutti via  
Are gone all away
- c. \*Sono andati tutti via i soldati  
Are gone all away the soldiers

The Floating Quantifier (*tutti* 'all') can co-occur with both preverbal subjects and null subjects, but not with post-verbal subjects. This would suggest, according to Rizzi, that null subjects pattern with preverbal subjects. In more theoretical terms, we could say that a FQ is licenced by being in the c-command domain of a referential subject. This would be the case in 0.a) and even in 0.b) if we assume there is a preverbal *pro*, but not in 0.c), where the referential subject is in inverted position and only an expletive *pro* appears in preverbal position.

In fact, we can redefine the pattern in 0 by saying that FQs can appear in contexts where there is null I-subject (*pro* in this case): both 0.a) and 0.b) would have, in our theory, a *pro* as I-subject; 0.c), instead, would have *i soldati* as I-subject.

In Chapter 1 we assumed that FQs in languages like Catalan cannot be the result of movement of the element they are adjoined to at D-structure, for some of them do not form a possible constituent with this element. Now we have a theoretical reason for this fact: there is no A-movement of the I-subject in NSLs.

Putting these ideas together, we can assume that FQs in NSLs are elements adjoined to *pro*. Then 0.a/.b) would have the structure:

- (87) (I soldati<sub>i</sub>) sono andati [<sub>DP</sub> tutti *pro<sub>i</sub>*] via

0.c) is ungrammatical because there is no *pro* I-subject for the FQ to adjoin to. In fact, if the quantifier adjoins to the I-subject, the result is:

- (88) Sono andati via tutti i soldati  
Are gone away all the soldiers

which, of course, is not a case of Quantifier Floating.

This analysis, however, faces important problems for the present theory. For a start, FQs do have a freer distribution than I-subjects: they can precede the complements of the verb (whereas I-subjects in general follow the complements in Italian or Catalan). In English overt I-subjects (*himself*) and FQs (*all*) even have complementary distribution (the former are VP-final and the latter are VP-initial).

Another problem is related to emphatic I-subjects: while inverted subjects do not allow FQs, emphatic I-subjects do allow FQs:<sup>86</sup>

- (89) Els nois s'han fet tots el llit *ells* (*mateixos*)  
The boys SI-have made all the bed they (SELVES)  
'The boys have *all* made their bed *themselves*'

If we want to maintain that emphatic elements like *ells* (*mateixos*) are I-subjects, then the above claim that FQ's in NSLs are adjoined to a *pro* I-subject cannot be true, if *pro* has to occupy the I-subject position too. In fact FQs seem to have a freer distribution than I-subjects (they can precede the verbal complements). On the other hand, as we pointed out in Chapter 1, FQs are likely to occupy derived positions, rather than the I-subject position.

---

<sup>86</sup> In fact, some speakers do not easily accept these constructions. See above for parallel facts with clitic doubling.



I cannot provide any clear solution for the problem raised by 0. I will only suggest two possible approaches. We have seen above that a sentence having a null subject can be claimed to involve a null topic, detectable in cases like 0.b). We have also seen that the null topic is only present when there is no overt topic: in 0.a), the overt topic *ell (mateix)* prevents a null topic from appearing. We said that null topics are licenced when there is a null subject. Suppose we assume that null topics are licenced whenever a preverbal (non-Focus-fronted) subject is possible: if the preverbal subject appears (wherever it is dislocated) it acts like a topic. If there is no preverbal subject, then a null topic appears. Now the contexts allowing preverbal subjects are, in our theory, either a null I-subject or an emphatic I-subject. Since these are the contexts allowing preverbal subjects to appear, we can assume that these contexts also trigger a null topic to be present if there is no preverbal (overt) subject.

Now, we could assume that a FQ has to have an antecedent c-commanding it, and that the antecedent can be a null topic. This is only a vague approach. We leave the question open of how FQs are generated and interpreted unexplained. Something in the theory has to guarantee that FQs are clause-bound, which does not follow from the licensing condition of being bound by a (possibly maximal scope) topic.

Notice that this approach is close to Rizzi's contention that there is a preverbal *pro* in his example 0.b): we also propose there is an empty category. The difference is that the empty category is not in Spec of INFL (AGR), but rather has maximal scope.

Another solution could consist in adopting Rigau's (1988) proposal about *strong pronouns* in Romance languages. According to her, emphatic I-subjects would be strong pronouns. Strong pronouns, in her theory, are not in A-positions. Then, a possible account for 0 would be:

- FQs are generated in A-position and, in coherence with our approach, in NSLs they are adjoined to a *pro*.
- since emphatic strong pronouns are not in A-position, FQs and emphatic strong pronouns can cooccur.

- FQs, even if generated in A-position can/must move to some higher position (which accounts for their freer distribution).

- emphatic strong pronouns, even if not in A-position, are subject to Binding Theory, and therefore abide by the generalization about the [ $\pm$ anaphoric] status of I-subjects, even if they are not I-subjects themselves.

The problem with emphatic (pseudo-)I-subjects is similar to the problem of pronominal clitic-doubling in Catalan or Spanish:

(90) El veig a ell  
Him-I-see to him  
'I see HIM'

Clitic doubling is also at the basis of Rigau's claim that strong pronouns are not in A-positions. Clitic doubling pronouns, like emphatic (pseudo-)I-subjects, can cooccur with a FQ:<sup>87</sup>

Context: I hand the new school books to *the parents* of the small children, but...

(91) els grans, *els* dono a tots el llibre a *ells* (*mateixos*)  
the older, them-hand-I to all the book to them (SELFs)  
'the older ones, I hand the book to them all personally'.

In both 0 and 0 the (emphatic) strong pronouns would be occupying a - $\theta$ -position, thus leaving the object/I-subject position free for the FQ (independently of whether the FQ then moves to some higher position).

So we seem to draw back to the traditional assumption, which we crucially challenged before, that emphatic anaphors and pronominals are not in an A-position, without giving up the crucial idea that these elements are subject to Binding Theory. In fact, both clitic doubling strong

---

<sup>87</sup> Speakers not readily accepting 0 do not accept 0 either.

pronouns and emphatic (pseudo-)I-subjects, even if not in a  $\theta$ -position, are not far removed from it: their word order distribution is nearly the same as, respectively, standard objects and inverted subjects, and only more subtle evidence, such as their co-occurrence with FQs, leads to reconsideration of the theory. One possibility is that these elements are adjoined to their respective  $\theta$ -positions, in a structure like 0:

(92) [DP [DP *pro* (FQ) ] *strong pronoun* ]

In 0:

- the *pro* would be licenced by the clitic (if clitics licence a *pro*) or by  $\text{AGR}^0$  (in NSLs), depending on whether the DP is in object position or in I-subject position,

- the strong pronoun is an adjunct carrying emphasis. It is subject to the same BT constraints as the *pro* it is adjoined to.

- the FQ can be moved to some other position, provided it is in the c-command scope of, respectively,  $\text{AGR}^0$  or the clitic.<sup>88</sup>

This is a rather speculative and non-explanatory approach. I will abstract away from this digression in the remainder of the thesis.

## 2.5. On the Nature of AGR Requirements

In section 1., we advanced several rules requiring that AGR must be coindexed with some DP or CP Argument. On the one hand AGR must have an AGR-identifier which must provide the I-subject with Case and on the other hand:

---

<sup>88</sup> French cases like:  
 (i) Il faut tout que je face  
     It needs everything(FQ) that I do  
     'I have to do everything'  
 are atypical cases of long distance FQ. I cannot say anything about the issue.

- (93) At D-structure, AGR must be coindexed with the most prominent non-(inherently-)Case-marked DP or CP in its c-command domain.
- (94) a. AGR must have an AGR-identifier.
- b. X can be an AGR-identifier iff X is rich in phi-features (number and person).
- c.  $AGR^0$ /Spec of AGR is the AGR-identifier of AGR.
- (95)  $AGR^0$  can optionally assign Nominative Case to Spec of AGR by agreement or to some other position under government.
- (96) The I-subject must receive Case from its AGR-identifier.

Within the framework of Principles and Parameters, the obvious question is whether we should postulate any such specific rules at all. This theoretical framework should optimally consist of principles and parameters of a very general nature, and rules affecting one single category as AGR looks at odds with such a desideratum.

Two considerations, however, can be put forward in defense of these rules. One is that, even if category specific, they are not by any means language specific: they seem to hold in a pervasive way across a good deal of languages. So they cannot be considered mere ad hoc theoretical devices to account for highly idiosyncratic facts. If not genuine principles of UG, they are at least good candidates to be theorems of the grammar.

I think the above set of rules can be interpreted as a specific case of a set of the general constraints that Chomsky (1986-b) dubs Licensing and Full Interpretation. In Chomsky's view, Licensing is a condition usually holding of two items that somehow need each other in order to be fully interpreted: Operator and variable, predicate and Argument, etc.; not only do the two elements have to cooccur, they also must stand in a proper relation.

In the same spirit, we could conceive that AGR and I-subject are two elements that need

each other: AGR has to be licenced by having an I-subject, and the I-subject needs Case.<sup>89</sup>

The idea I want to pursue is the following: once AGR is present in a language, it has to be licenced by having an I-subject, providing Case to it, and being rich in features. The question is: why does AGR have to be present?

We could assume that AGR is not present in all languages: Japanese and Chinese would possibly be languages lacking AGR (see Fukui & Speas (1986)). These languages challenge Taraldsen's original idea that null subjects are dependent on rich AGR-morphology. There are, on the other hand, languages having no overt AGR-morphology and nevertheless being closely similar to the Indo-European languages we have considered, namely Scandinavian languages. If the idea that Chinese/Japanese lack AGR is to account for their exceptional behaviour, then we should ask why Scandinavian languages do not take the negative setting for the [ $\pm$ AGR] parameter. This parameter should have some trigger for one or the other value.

We could then argue that the trigger for the positive value is the presence of phi-features in the language. As argued by Fukui & Speas (1986), Japanese (and Chinese) seem to lack phi-features altogether.

This approach is, however, problematic in one sense: Japanese, which has overt Case morphology and is a Nominative-Accusative language, seems to abide by Burzio's generalization as far as its Case array is concerned. Since our account for BG is based on AGR, we cannot adopt the view that Japanese (and probably Chinese) lacks AGR.

We will suggest another possibility. *All languages have AGR.*<sup>90</sup> Not all languages have phi-features. Then we could reformulate our rule 0 as:

---

<sup>89</sup> The latter idea could be challenged: we could claim that the Case filter (or visibility requirements) are an epiphenomenon, due to the fact that AGR (and other FCs) have to discharge Case (thanks to Jeff Runner for this suggestion). I will not pursue the issue.

<sup>90</sup> In Chapter 4 we will propose that some infinitival constructions do not have AGR: if so, the correct claim would be that **all languages have AGR in finite sentences**.

(97) a. AGR must have an AGR-identifier.

- b. X can be an AGR-identifier iff X is rich in the phi-features which are present in the language.
- c.  $AGR^0$ /Spec of AGR is the AGR-identifier of AGR.

Since Japanese (and Chinese) have a null set of phi-features, condition 0.b) can be fulfilled in a trivial way.  $AGR^0$  is vacuously rich in Japanese, and it can licence *pro* as an I-subject.

Summing up, the rules we postulated for AGR can be conceived as belonging to a set universal licensing principles:

- a) AGR must be licenced by having an I-subject (as expressed by 0) and providing this I-subject by Case as expressed by 0 and 0);
- b) since AGR consists in features, it must also be licenced by being able to display a sufficiently rich set of features (as expressed by 0). This will allow the I-subject to be an empty element whose content is recoverable from by the AGR-identifier ( in a trivial way if the language has no phi-features).

All these proposals are highly speculative, and there is a good deal of vagueness in them. The main difference with other instances of licensing principles is the strictly formal nature of the licensing principles for AGR. A licensing principle for, say, Operator-variable structures is more deeply rooted in interpretation and, therefore its existence seems to be a matter of necessity if language has to be an instrument for carrying meaning. If, however, the above proposal proves to be on the right track, it constitutes an argument for the autonomy of syntax, whose formal principles would be independent from the constraints on well-formed semantic structures.

## **2.6. Inverted Copulative Constructions**

There seems to be an interesting correlation between the NSL status and the possibility of having an inverted agreement pattern of some copulative constructions of the type (examples from Catalan):

- (98) a. El president sóc jo  
The president am I  
'I am the president'  
b. \*The president am I/me

We cannot simply say that 0.a) is a case of subject inversion, because then the preverbal subject would have to be a dislocated element without a resumptive clitic. Actually, nominal predicates can dislocate, and then a clitic is used as a resumptive element:

- (99) El president, no \*(ho/el) sóc pas  
The president, not it/him-am at-all  
'The president, I am certainly not'

So the clitic is obligatory in 0 and impossible in 0, which suggests *el president* is not dislocated in 0.a).

On the other hand, the possibility of 0.a) seems tightly correlated with the NSL status: Old French had the equivalent of 0.a) and lost it as it lost its NSL status.

In order to derive these facts, we need some previous theoretical assumptions. First, I adhere to the view, defended by several authors (Andrea Moro, Giuseppe Longobardi) that all copulative constructions contain a predicate, and there are no 'equative' constructions. Therefore, *el president* in the above examples has to be the predicate element, for pronoun (*jo* 'I') cannot plausibly be a predicate. The underlying structure for the above examples would be:

(100) ser... [<sub>SC</sub>jo el president ]

To-be I the president

Both Moro's and Longobardi's proposals essentially say that, when the predicate is a DP, either the Argument or the predicate are allowed to raise to Spec of AGR. Plausible and simple though this idea is, it does not immediately explain why:

- agreement takes place with the postverbal DP.
- these examples are only allowed in NSLs.

I think our theory on AGR-identifiers is well equipped to say something on the way of an explanation for these facts. We proposed that Spec of AGR is not the AGR-identifier in NSLs. So Spec of AGR is not required to be filled. In fact, as we will argue in Chapter 5, it appears to be a position with much looser requirements for the DP filling it than in non-NSLs. In any event, the fact that the predicative DP raises to it does not interfere with AGR-identification of Nominative assignment.

In non-NSLs, raising of the predicate to Spec of AGR would not abide by the requirements that there must be an AGR identifier and that this AGR-identifier has to transmit its Case to the I-subject, which implies it has to form an A-Chain with it. Therefore, non-NSLs cannot have DP-predicate raising: French lost it as soon as it became a non-NSL.

In conclusion, even if we cannot provide a precise explanation for why the predicative DP can raise to Spec of AGR, we can explain why it cannot in some non-NSLs.

We cannot explain why, in this type of construction, the pre-verbal DP can be dropped and the postverbal DP (which would be the I-subject) cannot:



- (101) a. Sóc jo (el president)  
           Am I (the president)  
           'It's me (the president)'
- b. \*El president sóc  
       The president am

Probably 0.b) is excluded because this kind of construction is precisely used to focalize the subject, and hence it cannot be dropped. It is not clear why the preverbal DP can be dropped if it is not recovered in content by AGR.

### 3. Indefinite I-subjects

So far, the predictions are that postverbal subjects can be:

- [-anaphoric] in NSLs.
- [+anaphoric] in non-NSLs.

Both kinds of languages, however, freely admit indefinite in post-verbal object position:

- (102) a. There came a man
- b. Viene un uomo  
       Comes a man

Within the theory sketched above, the question is: why are indefinites able to occur post-verbally without violating BT in English? Recall that for NSLs the existence of indefinite subjects is not a problem anymore than definite inverted subjects are: they both would be licenced as far as they are both [-anaphoric] I-subjects. Indefinite I-subjects, however, are uniformly acceptable independently of the null/non-null-subject contrast. So they should be licenced independently of how inverted subjects are licenced in NSLs.

Since Binding Theory plays a crucial role in our account of subject inversion, so that only NSLs allow [-anaphoric] I-subjects, we will exploit the idea that something allows indefinite DPs to escape from BT effects. We will address the question in section 3.2.

In section 3.1., we will address another important question: how are indefinite DPs Case-marked? If we adopt Belletti's (1988) hypothesis that indefinite DPs receive Partitive Case, then an obvious problem arises in connection with our rule of AGR coindexation, since we crucially assumed that the DP AGR is coindexed with could not be inherently Case-marked. We will refine the notion of inherent Case in a way to allow Partitive Case to be simultaneously structurally Case marked.

### **3.1. Partitive Case**

We assumed that, for languages such as English or Catalan, at D-structure AGR has to be coindexed with a non-(inherently)-Case-marked DP (or CP). Reference to inherent Case-marking was crucial in order to prevent a DP other than an EA or an object to become the I-subject.

In Chapter 1, we noticed that Belletti's hypothesis of Partitive Case provides an interesting means of excluding sentences like:

- (103)    a.    \*I consider girls interesting            (existential reading)  
          b.    \*There seem girls to have come

which would be excluded because Partitive Case can only be assigned to positions which are  $\theta$ -marked by the head assigning Partitive.

If we assume that indefinite DPs in object position are assigned Partitive Case, then they should not be candidates for AGR to coindex with, according to our rule of AGR coindexation, repeated here:

- (104) At D-structure, AGR must be coindexed with the most prominent non-(inherently-)  
)Case-marked DP or CP in its c-command domain.

However, it is crucial for our theory that Partitive DPs may be possible I-subjects. Otherwise, sentences like:

- (105) There are many children

would have no I-subject, and should be excluded, as 0 is obligatory.<sup>91</sup> On the other hand, it is plausible that the indefinite DP in 0 is the subject in some sense, since it agrees (in number) with the verb in English and many other languages.

To solve the problem, we will assume that partitive Case differs from other inherent Cases in that Partitive alone is not sufficient to Case-licence a DP. Suppose we postulate that:

- (106) Partitive has to be realized as Structural Case at S-structure.

If we want to maintain that Partitive is an inherent Case, then it must not be one relevant for 0. It would anyway be a defective Case, which has to be supplemented by a structural Case. It is not clear why 0 should hold. There is, though, some evidence for its empirical validity. Consider the following pair of examples:

---

<sup>91</sup> For this reasoning to hold, we must exclude the possibility that the expletive itself becomes the I-subject (thanks to Jeff Runner for pointing this out to me). We can obtain this result by assuming that:

a) The expletive is directly generated in Spec of AGR.  
b) The notion of c-command relevant for 0 is strict c-command: AGR<sup>o</sup> does not c-command its specifier.

Alternatively, we could stipulate that the I-subject has to have 'semantic content'.

- (107) a. There have developed typhoons here  
 b. John developed theories on that issue

Suppose both indefinite DPs in 0 are assigned Partitive at D-structure, and some structural Case at S-structure. As we remarked, there is some evidence that Nominative is assigned to the indefinite DP in 0.a), which would naturally account for the fact that there is agreement between the verb and the indefinite DP. As for 0.b), it can be naturally assumed that the indefinite DP receives Accusative Case at S-structure. There is some evidence from Romance languages pointing to that conclusion. Consider the following paradigm of Causative constructions (examples from Catalan):

- (108) a. Li fa menjar les patates  
           Him-DAT makes eat the potatoes  
           'S/he makes him eat the potatoes'  
 b. El fa menjar  
           Him-ACC makes eat  
           'S/he makes him eat'

The generalization is that in Romance (and many other languages') causative constructions the subject of the infinitive clause is assigned Dative if the infinitive has an Accusative Case-marked complement; when the infinitive does not have any Accusative complement, then its subject is assigned Accusative.

According to Baker (1983), the generalization could be expressed as follows: the complex formed by the causative verb and the infinitive can at most assign one Accusative Case; if that Case is required by the object of the infinitive, then the subject of the infinitive has to receive Dative; if not, it is the subject of the infinitive that is assigned Accusative.

Now consider the following example:

- (109) Li fa menjar patates  
 Cl-DAT makes eat potatoes  
 'S/he makes her/him eat potatoes'

which minimally differs from 0.a) in that the object of the infinitive is indefinite. If it is indefinite, it has been assigned Partitive (it is enough for our argument that this is at least a possible option). If Partitive did not additionally require a structural Case, as we are claiming, the CP *patates* would not spend the Accusative Case which the complex of verbs can afford, so the subject of the infinitive could be assigned Accusative, contrary to fact:

- (110) \*El fa menjar patates  
 Him-ACC makes eat potatoes  
 'S/he makes him eat potatoes'

So, it is reasonable that Partitive Case is not sufficient by itself to Case-licence a DP. We could then redefine 0 as 0:

- (111) At D-structure, AGR must be coindexed with the most prominent non-Case-saturated DP or CP in its c-command domain.

where a DP is Case-saturated if the Case it bears is sufficient to licence it, Partitive Case-marked DPs at D-structure not being so.

Chomsky's (1986-b) characterization of inherent Case requires that inherent Case has to be realized at S-structure, so it could appear that our special characterization of Partitive Case as a special inherent Case additionally requiring a structural Case can be accommodated within this

general view.<sup>92</sup>

I will not develop the question further. It suffices for the present purposes that Partitive Case, if at all an inherent Case, is different from other inherent Cases in that it additionally requires structural Case, so that 0 is an accurate rule for AGR coindexation. It could turn out, as we pointed out in Chapter 1, that Partitive is not a Case, but rather a special interpretative option for DPs. For convenience, though, I will keep using the term *Partitive Case*.

### 3.2. Indefinite DPs and phi-features

Now let us address the other problem raised by indefinite subjects, namely that they are possible I-subjects in non-NSLs, which otherwise do not allow [-anaphoric] I-subjects.

A solution to this problem can be formulated in the following terms. Let us assume that:

---

<sup>92</sup> I think, however, that Chomsky's proposal cannot be trivially adopted nowadays. First of all, Chomsky's assumption that inherent Case has to be 'realized' does not involve structural Case-assignment. Secondly, the facts Chomsky's theory is intended to cover are basically related to Genitive case. Since it is crucial in his account that Genitive is both assigned and realized in the government domain of N, the theory should be carefully revised in the light of the Determiner Phrase hypothesis, if we are to accommodate it to present day common assumptions.

- (112) a. Partitive DPs do not have person features.<sup>93</sup>  
 b. Partitive DPs may/may not have (grammatical) number features.  
 c. AGR and Spec of AGR need not agree in number features when person features are not present.  
 d. A binds B if A c-/m-commands B and A and B share some phi features.

The parameter in 0.b) is intended to account for the variation languages seem to exhibit with respect to agreement with an indefinite:

- (113) Agreement in number:  
 a. There *are* children in the garden (Standard English)  
 b. Es *sind* Kinder in den Garten (German)  
     It are children in the garden  
 c. Ci *sono* bambini nel giardino (Italian)  
     there-are children in-the garden  
 d. *Arriben* turistes (Standard Catalan)  
     Arrive tourists

---

<sup>93</sup> This is also assumed in Rigau (1991). Rigau also assumes that Person Agreement and Number Agreement are different functional categories. I think this idea could be adopted within the present theory, but it would require some careful elaboration, for it is crucial for the present account to work that AGR<sup>o</sup> is coindexed with the I-subject, contrary to what Rigau assumes for partitive constructions: at least Person-Agreement is not coindexed with the partitive in her view. Here I contend that there is always coindexation, which may be devoid of content if the binding features are absent.

(114) No agreement:

a. Il y *a* des enfants au jardin (French)

It there-has of-the children in-the garden

'There are children in the garden'

b. *Arriba* turistes (North-Western Catalan)<sup>94</sup>

Arrives tourists

c. There's children in the garden (Coll. English)

0.b) could be characterized as follows. Suppose indefinite DPs involve a (possibly empty)  $D^0$  which may or may not inherit the number features of its complement NP.<sup>95</sup> If it does, then we have Partitive DPs cum number. Otherwise, Partitive DPs do not have number features.<sup>96</sup>

Let us consider the following simplified D-structure:

---

<sup>94</sup> See Rigau (1991), where an explanation is provided for the contrast between languages showing verb number-agreement with the indefinite and languages with no such agreement.

<sup>95</sup> Actually, it is not crucial for the present purposes that we have an empty  $D^0$ : it could as well be an empty quantifier, if we assume indefinite quantifiers and determiners are different FCs, as proposed by many authors with several implementations (Cardinaletti & Giusti (1991), Rigau (1991)).

<sup>96</sup> French would be a language where Partitive DPs do not have number features. We could argue that this is related to the fact that a preposition usually precedes the NP in French Partitive DPs:

(i) Je mange beaucoup **de** pommes  
I eat many of apples

(ii) Je ne mange pas **de** pommes  
I Neg eat not of apples

and this preposition blocks number inheritance by the determiner. However, other languages having similar prepositions allow number in Partitive, so this account cannot be trivially correct:

(iii) Catalan: moltes de pomes  
many-**fem-pl** of apples (-**fem**) -**pl**



(115) AGR developed [<sub>DP</sub> typhoons ]<sub>Part</sub>

In 0 AGR is coindexed with the Partitive DP. In a non-NSL where the Partitive DP does not have number features, AGR and the Partitive do not share any phi-features. So, when the AGR-identifier (namely Spec of AGR) is filled by an (expletive) DP, this DP will not share any features with the Partitive DP either. According to 0.d), there is no binding relation between the AGR-identifier and the Partitive, so BT does not force the I-subject to be [+anaphoric] (and in fact prevents it from being so), as is otherwise the case with non-NSLs.

Suppose, however, that the partitive DP has number features. Then the prediction is that in 0 AGR<sup>0</sup> and the indefinite DP will share number features. If the AGR identifier (= Spec of AGR) shared these features, then a Binding relation would stand between Spec of AGR and the I-subject, and the indefinite DP would be excluded as a BT 3rd principle violation. However, non-NSLs of this kind can have recourse to the (possibly marked) option admitted in 0.c), namely that AGR<sup>0</sup> and Spec of AGR do not agree in number features, so, again, no BT violation ensues if the Partitive remains in place.

In the preceding account, it is not clear why some non-NSLs allow null expletives (or even null quasi-Arguments). Perhaps the residual character of AGR-identifiers in indefinite I-subject constructions, where the AGR-identifier does not display any features, allows for it to be dropped.

Another possibility could be developed. Suppose non-NSLs can take the marked option of having AGR<sup>0</sup> as the AGR-identifier in constructions where the I-subject does not have person features (i.e., it is Partitive). Suppose this option is subject to the constraint that AGR<sup>0</sup> is rich enough to recover the content of the I-subject: if it is the case that the I-subject has number features, then AGR<sup>0</sup> has to be rich in number features, i.e., it has to at least distinguish between 3rd-singular and 3rd plural. From this we could derive the following descriptive generalization:

- (116) If a non-NSL does not show overt *there*-expletives, the verbal morphology distinguishes between 3rd singular and 3rd plural.

Recall that the implication does not hold in the other direction: English distinguishes between 3rd-sng and 3rd-pl (only in the present tense, though). German, Icelandic or Brazilian Portuguese are all languages abiding by this generalization.<sup>97</sup>

Many questions are left open here:

- a) Why is Partitive restricted to certain verbs in English (There came a man/\*There broke a glass)?
- b) Why is Partitive allowed for non-objects in some languages (German)?
- b) Why do some of the languages allowing null indefinite-expletives also allow null quasi-Arguments?<sup>98</sup> Are quasi Arguments also lacking person features?

I cannot properly address these issues.

#### 4. CP I-subjects and Small Clauses

Thus far, the theory we have presented essentially consists of the following components:

- a rule of AGR indexation that coindexes AGR with some DP or CP, which becomes

---

<sup>97</sup> Some Brazilian dialects do not have the 3rd-sng/pl morphological distinction anymore and, as far as I know, they still have null expletives (thanks to Cristina Figueiredo for pointing this out). Maybe we could simply say that Brazilian Portuguese (or at least these dialects) has Partitive DPs without number features, which implies that AGR<sup>o</sup> is vacuously rich to licence a Partitive DP.

<sup>98</sup> Rizzi (1986) proposes there is a gradation expletive -> quasi-Argument -> full-Argument, and that if **pro** is licenced in one of the options, it is also licenced in the preceding options. See next section.

the I-subject.

- some principles of Case theory to the effect that the I-subject has to receive Case from the AGR-identifier.

We have crucially assumed that CPs are candidates for becoming I-subjects, but we have said nothing about whether they also are assigned Case. Since Stowell (1981), a widely accepted hypothesis has been that CPs do not accept Case. For the present theory it is essential that I-subjects obligatorily obtain Case from their AGR-identifier, in order to explain how Accusative is never assigned to an I-subject.

So if I-subjects obligatorily obtain Case from their AGR-identifier, we seem to be compelled to the conclusion that CP I-subjects receive Case too, contrary to Stowell's hypothesis.

In any event, it is clear that CPs show a distribution which is clearly different from that of DPs. These are some of the essential facts:

- object CPs do not have to be adjacent to the verb:

(117) I said the other day that...

- CPs cannot occupy the subject position of sentences and Small clauses:

(118) a. \*Does that John comes bother you?  
b. \*I Consider that John may come impossible

- (in many languages) CPs cannot be the complement of a preposition and do not require 'of' insertion:

(119) a. \*I talked about that...  
b. I am sure (\*of) that...

I think the above evidence is rather compelling. So, we will adhere to Stowell's claim that CPs cannot be assigned Case. Before proceeding, let us consider another aspect of argumental CPs that we have not addressed thus far.

The fact is that argumental CPs can (and sometimes have to) be 'doubled' by an expletive:

- (120) a. *It* is evident *that*...  
b. I consider *it* evident *that*...  
c. I can accept (*it*) *that*...

Suppose that the expletive in the above examples is linked to the CP in some way, and that it is the expletive that occupies the  $\theta$ -position at D-structure and is Case-marked at S-structure. In other words, suppose that it is the expletive that fulfils all the requirements of being an Argument, except that it transmits its  $\theta$ -role to the extraposed CP.

Suppose that the linking device between the expletive and the CP is coindexation. In the preceding section we have provided a means of preventing BT effects in structures like:

- (121) *There* came a man

by assuming that coindexation does not imply binding if there are no phi-features shared. Similarly, we can assume that CPs are not sensitive to being coindexed and c-commanded by their *it*-expletive, because they have no phi-features. That CP I-subjects cannot have phi-features is suggested by the fact that coordinating them does not give plural agreement:

- (122) It is/\*are well-known that ... and that ...

Coordinated preverbal CPs can trigger plural agreement:

(123) That ... and that... are two well known facts.

Since, as we and many authors argue, these CPs are not in subject position (they would be dislocated), we could assume that the real subject (some empty category) is a resumptive element having plural features.

In fact expletives linked to CPs are exceptional in that they are the only ones occupying  $\theta$ -marked positions, as shown in 0.b/.c). So let's propose there are two kinds of expletives:

- expletives of the *there*-type, which merely fulfil the formal requirement of filling the AGR-identifier in (some) non-NSL and transmitting Case to the I-subject. They are not  $\theta$ -marked, since they do not play any role in interpretation.

- expletives of the *it*-type, which are linked to a CP and behave like Arguments (they are Case- and  $\theta$ -marked) except for the fact that they transmit their  $\theta$ -role to the CP.

We cannot claim that CPs are always associated to an *it*-expletive, because there are obvious counterexamples:

(124) I think (\*it) that...

The distribution of CPs and *it*-expletives can be accounted for on the following assumptions:<sup>99</sup>

---

<sup>99</sup> This proposal is inspired on ideas in Authier (1991), who deals with the contrast between French and English CP-expletives. Here we will not go into a detailed discussion of the issue.

- (125) a. CPs cannot be (directly) Case-marked. CPs can escape being Case-marked by having recourse to:
- extraposition.
  - being associated to an *it*-expletive.
- b. CPs cannot occupy specifier positions at any level of representation.

According to 0.a) two options are available for the CP complement of *accept*:<sup>100</sup>

- (126) a. I accept  $t_i$  [<sub>CP</sub> that... ]<sub>i</sub> (extraposition)
- b. I accept  $it_i$  [<sub>CP</sub> that... ]<sub>i</sub> (*it*-expletive)

According to 0.b), only the *it*-expletive option is available for subject CPs:

- (127) a. *It* is evident *that*...
- b. I consider *it* evident *that*...

Now consider NSLs. We have claimed that *there*-expletives do not exist in NSLs. We also claim that no expletive exists in subject inversion constructions. This is a natural assumption under the view that such kinds of expletives are merely formal AGR-identifiers, and are not necessary in languages where Spec of AGR is not the AGR-identifier. *It*-expletives, on the other hand, are of a different nature, as we have argued. Since they fulfil the important role of linking CPs to A-positions, they should be in principle universally available.

Rizzi (1986) argues convincingly that null expletives exist in V-governed positions in Italian. The essential of his argumentations goes as follows. Some languages (such as Italian)

---

<sup>100</sup> Only the extraposition option is available for the CP complement of other verbs like **say** or **think**. The explanation for these facts could lie on the different status of subcategorized CP complements.

allow null objects in a way Modern English does not. Rizzi claims that such null objects are *pro*:

- (128) a. Questo conduce *pro* a concludere quanto segue  
          This leads to conclude what follows  
      b. This leads \*(people) to conclude the following

See Rizzi (1986) for an explanation of how such a *pro* is licenced. One appeal of Rizzi's characterization of the facts is that it predicts an interesting correlation: languages allowing null objects as in 0.a) also allow null expletives as in 0.a); languages not allowing null objects (as English, see 0.b)) do not allow null expletives either (see 0.b):

- (129) a. Gianni ritiene *pro* probabile che Mario venga  
      b. John considers \*(it) probable that Peter comes

Early Modern English (until the XVIIIth century) allowed both null objects and null expletives, so the correlation is highly plausible, as it holds of as closely related languages as early Modern English and present Modern English.

In the light of these facts, then, we conclude that:

- *there*-expletives are an exclusive property of (some) non-NSLs.
- *it*-expletives are universally available. They are null in contexts where null pronominals are allowed.

With the above assumptions, we can address the problem of CPs and obligatory Case for I-subjects: whenever a CP is coindexed with AGR, there must be an I-subject to which AGR provides Case. Since the CP itself cannot be assigned Case, it will have the option of being

linked to an *it*-expletive, which does not transmit Case to its associated CP.<sup>101</sup>

One question we could address is where the CP linked to the expletive is attached to. Non-dislocated sentential CP I-subjects seem to occupy a VP-final position. When a CP is a subject of Small Clauses (as in 0), it also occupies a Small-Clause-final position. The underlying place of attachment of the CP should not be far removed from the underlying position occupied by the *it*-expletive it is linked to. Therefore, the fact that the CPs in 0 are steadily Small-Clause-final seems to suggest that this is also the basic position for the *it*-expletive.

In other words, we suggest that:

- Small clauses have some FC structure which provides the specifier for Small-Clause subjects preceding the predicate.

(130) I consider [<sub>XP</sub> John [<sub>X'</sub> [<sub>AP</sub> intelligent ] ] ]

- The underlying position for Arguments of a predicate is always to the right of the predicate (this is true for any lexical category being a predicate, in both clauses and Small Clauses).

- Therefore, the underlying position for *it*-expletives (and their CP, which stands nearby) is to the right of the predicate even in Small Clauses: this is why, even if the *it* expletive moves to a FC specifier, the CP remains in SC-final position.

(131) I consider [<sub>XP</sub> *it*<sub>i</sub> [<sub>X'</sub> [<sub>AP</sub> strange *t*<sub>i</sub> that... ] ] ]

- If this analysis for Small clauses is correct, we predict other types of I-subjects,

---

<sup>101</sup> Our proposal, then, is that **it**-CP CHAINS are the only type of CHAINS which are available in all languages: there-Partitive CHAINS are only used in non-NSLs, and null-expletive/inverted-subject CHAINS do not exist.



beyond CPs, to appear to the end of the Small Clause. The following example suggests this is on the right track:

- (132) - John is too stupid to do the job  
- Why don't they resort to Bill?  
- Because they consider Bill stupid himself
- (133) They consider [<sub>XP</sub> Bill<sub>i</sub> [<sub>X'</sub> [<sub>AP</sub> stupid [<sub>DP</sub> t<sub>i</sub> himself ] ]]]

The Catalan equivalent of 0 is not well-formed:

- (134) ?\*Considero en Joan estúpid ell mateix  
I-consider the J. stupid he SELF

In fact, Small Clauses are not Null-Subject structures even in NSLs like Catalan, so the prediction would be that they should not have [-anaphoric] I-subjects. They do not have [+anaphoric] I-subjects either:

- (135) \*\*Considero en Joan estúpid (si) mateix  
I-consider the J. stupid (SE) SELF

I leave the issue open.

## 5. French Stylistic Inversion and Northern Italian Dialects

There is one case of subject inversion that does not easily fit into our theory: French Stylistic Inversion:

- (136) a. Où (dis-tu qu') est allé Jean?  
 Where (say-you that) is gone J.  
 'Where do you say John went'
- b. La personne avec qui (je crois que) viendra Jean  
 The person with who (I believe that) will-come J.  
 'The person with whom I think that Jean will come'
- c. J'espère que vienne Jean  
 I-hope that come-SUBJ J.  
 'I hope Jean will come'

There is a general agreement that it is essentially of the same nature as subject inversion in Romance NSLs (see Kayne & Pollock (1978) for an initial characterization). If French is a non-NSL, I-subjects should be, according to the present theory, [+anaphoric]. There is a fact, however, that looks consistent with the present approach: when there is stylistic inversion, no preverbal subject is present. Standard accounts would say that only null expletives licence subject inversion. In our alternative theory not involving null expletives in subject inversion we should say that in French Stylistic inversion, since there is no preverbal subject, there is no binder for the I-subject and therefore it can be and has to be [-anaphoric].

One aspect of stylistic inversion which is not easy to deal with is the characterization of the contexts allowing it. It is triggered by Wh-movement and subjunctive mood. Concerning the former, one cannot say it is licenced by a [+Wh] COMP, because it can be licenced in the clause where the Wh- has been extracted from even if this clause is not [+Wh] itself (see 0.a/.b) with the parenthetical part).

One way of characterizing the licensing contexts in a unitary way is the following. Suppose we assume that stylistic inversion is licenced *whenever the clause involving it is non-assertive*. Clauses containing a Wh-gap are non-assertive, and so are clauses in subjunctive mood. How could we relate this to (stylistic) subject inversion?

We have crucially assumed that I-subjects in NSLs are directly assigned Case by  $AGR^0$  through Chain-government, which is made possible if V-raising to  $AGR^0$  takes place. French, in fact, has V-movement to  $AGR^0$  (V-movement to the top INFL category, in Pollock's (1989) terms). Therefore, if French does not have subject inversion in the general case, it is only because it is not a NSL, not because the requirement of V-movement to AGR is not met. Let us tentatively make a rather speculative proposal in this connection.

Suppose that what in fact makes French a non-NSL is not that  $AGR^0$  is not intrinsically rich enough to make French a NSL, but rather that  $AGR^0$  cannot manage to govern its I-subject through Chain-Government. Suppose the reason is that there is something between  $AGR^0$  and the I-subject blocking government. Suppose this blocking element is absent in non-assertive sentences. Let's implement the idea.

French (like English and many non-NSLs) has the negative particle *pas* (Cf. English *not*, German *nicht*, etc.) below the inflected verb. This is what led Pollock to assume NegP is a FC placed between T and AGR. Suppose this is right. Suppose, however, that this category is not NegP, but  $\Sigma$ P, as Laka (1990) suggests,  $\Sigma$  being a FC which includes Negation and Affirmation. Suppose we assume that  $\Sigma$  also contains the feature  $[\pm\text{assertion}]$ , and that  $[+\text{assertive}]$   $\Sigma$  has some blocking effect. If  $[-\text{assertive}]$   $\Sigma$  did not have this blocking effect, then  $AGR^0$  would be able to Chain-govern its I-subject and French will possibly behave like a NSL in  $[-\text{assertive}]$  sentences.

The above suggestion is only tentative and faces several problems:

- it is not clear at all what the blocking effect of  $[+\text{Assertive}]$   $\Sigma$  could be: it does not block V-head movement, for there is long V-movement in French assertive clauses. It would block only Chain-government.

- in  $[-\text{assertive}]$  clauses, French is not actually a NSL: it allows subject inversion but not null subjects.

In any event, it is significant that non-NSLs often have post-verbal negative particles, which suggests that  $\Sigma$  intervenes between AGR and T, while NSLs often have preverbal negative

particles.<sup>102</sup> The correlation seems to have some significance: languages losing the NSL-status often shift from preverbal negation to postverbal negation: this happened in the transition from Old English to middle English and in the transition from Middle French to Modern French (see Pollock (1989)).

The fact that French [-assertive] clauses can have subject inversion, but not null subjects, suggests that the two facts need not correlate. Let us rephrase the facts in our theoretical terms.

We proposed that when  $AGR^0$  is rich enough, it is the AGR-identifier and then:

- a) it can directly Case mark the I-subject.
- b) the I-subject can be a null pronominal (in non-NSLs it can only be a null anaphor).

Suppose that a) is a necessary but not sufficient condition for b) to hold. French [-assertive] clauses would be a case where a) but not b) holds.

There is another well-known case where something similar happens: some Northern Italian dialects (see, e.g., Rizzi (1982), Brandi & Cordin (1989)). In these dialects, the agreement morphology in a strict sense is not rich enough to allow null subjects. But, as initially proposed by Rizzi (1982), subject clitics are part of AGR and make AGR rich enough to allow null subjects. Now, in some of these dialects (Paduan -Rizzi (1982), Trentino and Fiorentino -Brandi & Cordin (1989)-) subject clitics are required to allow null I-subjects, but they are not necessary (and in fact not possible) to allow an overt I-subject:

---

<sup>102</sup> Zanuttini's (1991) typology of languages concerning negation is based on the same observation. Actually, the correlation with the (non-)NSL status is not strict: there are Italian dialects with post-verbal negation which are NSLs; and Brazilian Portuguese is a non-NSL having preverbal negation. We will speculate on the parameterization of the relative hierarchical position of  $\Sigma P$  in Chapter 5.

- (137) Paduan (Rizzi (1982)):
- a. (Giorgio) \*(el) vien  
G. he-CL comes
  - b. (\*El) vien Giorgio  
He-CL comes Giorgio

Like in the case of French Stylistic inversion, this suggests that the requirements for null subjects are stronger than the requirements for subject inversion. In our terms, the requirements for *pro* I-subjects are stronger than the requirements for overt [-anaphoric] subjects. The former require some minimal richness (as the one displayed in French AGR-morphology or Northern Italian strict AGR-morphology). The latter requires a full range of AGR-distinctions (which is simply not possible in French, and possible by resorting to clitics in Northern Italian dialects).

In order to capture these facts, we should refine our parameter for the AGR-identifier, repeated here:

- (138) a. AGR must have an AGR-identifier.  
b. X can be an AGR-identifier iff X is rich in phi-features (number and person).  
c.  $AGR^0$ /Spec of AGR is the AGR-identifier of AGR.

0.b) mentions richness in number and person as necessary for  $AGR^0$  to be the AGR-identifier. In fact, if we look at the Italian dialects mentioned above, subject inversion involves person agreement, but not number-agreement (see Brandi & Cordin (1989:fn 10)).<sup>103</sup>

---

<sup>103</sup> Modern Standard Arabic behaves the same as these dialects.

(139) *6-person paradigm with a pronominal inverted subject:*

<u>Trentino</u>		<u>Fiorentino</u>	
vegno	mi	e vengo	io
te vigni	te	tu vieni	te
ven	elo/ela	e viene	lui/lei
vegnim	noi	si vien	noi
vegní	voi	vu'venite	voi
ven	lori/lore	e vien	loro

In 0 the AGR-affixes+clitics makes all person/number distinctions except for 3rd-sng/3rd-pl, which are not distinguished. On the reasonable assumption that 1st-sng is a different person from 1st-pl ('we' is not simply the plural of 'I'), and similarly 2nd-sng is a different person from 2nd-pl, so that only 3rd-sng and 3rd-pl are really the same person (or non-person), we can interpret 0 as indicating that subject inversion in those dialects involves person agreement but not number agreement (the number agreement for 1st and 2nd persons being parasitic on the person: e.g. the 4th person 'we' is inherently plural, as opposed to the 1st person 'I'). Null I-subjects, instead, require both person and number agreement, as we saw. Thus, we could reformulate 0 as 0, and then add 0:

- (140) a. AGR must have an AGR-identifier.  
 b. X can be an AGR-identifier iff X is rich in *person* phi-features (five distinctions: 1st, 2nd, 3rd, 4th and 5th).  
 c. AGR<sup>0</sup>/Spec of AGR is the AGR-identifier of AGR.
- (141) A null I-subject (be it [ $\pm$ anaphoric]) requires an AGR-identifier being rich in person and number features (six distinctions).

Let us assume that the Northern-Italian dialects under consideration choose the AGR<sup>0</sup>-

option in 0.c), because the AGR-morphology, with the help of AGR-clitics, makes AGR<sup>o</sup> rich enough (with at least 5 distinctions). But then these dialects use AGR-morphology+clitics in a parsimonious way: they use all 6 distinctions to abide by 0 when the I-subject is null (a null pronominal), but they use only 5 distinctions when the I-subject is overt.

The fact that inverted subjects seem to require Person features (except for the special Case of French Stylistic Inversion), does not support Rizzi's (1986) view that expletives only require formal licensing (not feature identification), if, in consonance with Rizzi's view, subject inversion involved an expletive. In fact, Rizzi considers three levels of feature requirements for *pro*: person and number (referential *pro*), only number (quasi-Argument *pro*) and no requirement (expletive *pro*). We see, however, that:

- subject inversion requires person features (and in some languages even number). We propose that there is no expletive involved.

- null 'there'-expletives require number (e.g. Brazilian, German) or no feature, depending on the language (see section 3.2.).

- null 'it-CP' expletives do not apparently require features (at least in V-governed position, see Rizzi (1986)), but

- both null 'it-CP' expletives and null 'there'-expletives (and quasi-Arguments) seem to be favored in in non-NSLs at least distinguishing 3rd singular from 3rd plural (Brazilian Portuguese, Icelandic, German). So it seems that number features are important to licence null expletives.

So Rizzi's three-level distinction appears not to be straightforwardly confirmed by cross-linguistic evidence.

One tentative alternative proposal could be the following:

- a) Null *there*-expletives do not exist in NSLs.
- b) Null *there*-expletives require number features (i.e., a morphological distinction between 3rd-sng and 3rd-pl) (German, Brazilian Portuguese, Icelandic).
- c) Null Quasi-Arguments require number features and some additional condition

(which is met by Brazilian Portuguese and Icelandic but not by German).

d) Subject inversion requires person features (in our technical terms,  $AGR^0$  can be the AGR-identifier only if it is rich in person features).

e) Referential *pro* requires number and person.

This is only a tentative approach, conceived solely on the basis of finite clauses. For infinitival clauses, see next chapter. What the present approach shares with Rizzi's (1986) proposal about null pronominals is the idea that, beyond requirements on formal licencing, empty pronominals are subject to requirements on *recovery of content*, which can be more or less stringent (even possibly vacuous) depending on the nature of the pronominal. We propose, however, a reduction of the empty pronominals available.

French does not fit into the theory yet: in [-assertive] contexts, where overt [-anaphoric] I-subjects are allowed, it is not the case that 5 distinctions are provided by the verbal morphology (and subject clitics cannot be used, because they are not AGR-clitics, but merely phonological clitics). I cannot provide an account for this fact. Two possibilities could be pursued:

- French Stylistic Inversion is some marked option in which 0.b) is relaxed.

- 0.b) should be relaxed for all languages, so that  $AGR^0$  can in principle be the AGR-identifier in French (and *a fortiori* in Northern Italian dialects and other Romance NSLs) but not in English or Germanic Languages, where not even [-assertive] sentences allow subject inversion.

All the preceding proposals in this section are highly speculative and only tentative. For convenience, I will continue to use 0 in the remainder of the discussion.

There are, in addition, some remaining problems which appear to be even harder to account for within the hypotheses advanced here. One is Old French, where Null Subjects and Subject Inversion are restricted to V-second contexts, i.e., root contexts for the most part (see Adams (1987)). The other is Corsican: in spite of the fact that this language has rich AGR morphology, it only behaves like Italian in root contexts. In embedded contexts a subject clitic is



required for both Null Subjects and subject inversion (see Agostini (1956)):

- (142) a. Quand'\*(ellu) canta...  
When he-CL sings  
'When he's singing...'
- b. Quand'\*(ellu) canta Petru...  
When he-CL sings P.  
'When Petru is singing...'
- c. Quand Petru canta...  
When P. sings

Our theory the null/non-null subject contrast is based on morphological richness and does not say anything about the root/embedded distinction. It is therefore unable to deal with the issue as it stands. I think, however, that the key to account for these facts lies precisely in morphological richness: both Old French and Corsican have more syncretisms in the agreement paradigm than other Romance NSLs (in the case of Old French, spelling often showed distinctions which had no pronounced counterpart). This seems to suggest that 'slightly' impoverished paradigms can still give birth to a restricted form of Null-subject-hood, only available in root clauses.

## 6. A-dependencies and Minimality

Our characterization of I-subjects can be summarized as follows:

- in NSLs, a dependency is created between  $AGR^0$  and an I-subject, which ultimately implies that the I-subject will directly receive Case from  $AGR^0$ .
- in non-NSLs, the same dependency is created, but since  $AGR^0$  itself is too poor, Spec of AGR has to be filled by a DP.

We have tacitly assumed that in the latter case, the DP in Spec of AGR forms a Chain with the I-subject. If Chains are formed by movement in the standard case, this suggests that the I-subject itself moves to Spec of AGR, leaving an empty anaphor as I-subject at S-structure. Is this result correct?

For the Germanic languages, we saw that, even when the I-subject is not empty, a trace can always be claimed to be present, as exemplified in 0, where the non-empty I-subject is in fact a floating element:

- (143) a. John<sub>i</sub> did it t<sub>i</sub>  
 b. John<sub>i</sub> did it [<sub>DP</sub> t<sub>i</sub> himself]

In French, however, the I-subject cannot be claimed to always involve an empty anaphor:

- (144) Jean l'a fait *lui*  
 J. it-has done he  
 'John did it himself'

As we saw, in 0, the I-subject *lui* cannot be analyzed as a floating element cooccurring with a trace as it does not form a possible constituent with the preverbal subject. Then no trace is involved in this Case. So the dependency between *Jean* and *lui*, even if similar to a Chain, is not a standard case of Chain (nor is it a standard case of CHAIN either, since there is no expletive).

We have not presented any account for why French allows such a non-standard kind of Chain formation while Germanic languages do not. We will try to derive this contrast from some independent linguistic facts.

Kayne (1987) assumes that French has Object Agreement. Although he is not very precise as regards the characterization of this Agreement, it could be naturally conceived, within

the present widespread conception of Functional Categories, as Functional Category to which the (participial) verb raises.

The present day multiplication of hypotheses about Functional Categories after Pollock (1989) initial proposal raises a fundamental question about their universality. One possible view is that all of them are universal, although possibly not morphologically realized in some languages. According to this view, English would have Object-Agreement, which, as opposed the French one, would have no morphological manifestation. An alternative, and perhaps more realistic view, is that only those FCs exist in a language which have some morphological manifestation. In this view, English would not have object Agreement.

Of course, an intermediate hypothesis is possible: some FCs are universal (whether they have overt manifestations or not) and others are not (and will be present only in languages where they can be detected by the morphology). Chomsky's recent hypotheses (See Chomsky (1992)) postulate that Object-Agreement belongs to the universal type of FC, since it is the universal means of Accusative-assignment. In fact, this hypothesis is not at odds with our theory on Accusative, as far as we claim that Accusative Case is in principle available independently of lexical idiosyncrasies.

I think, however, that French Object-Agreement and Chomsky's abstract Object-Agreement need not be assimilated as a single concept. French object Agreement is only operative in a restricted class of sentences, namely the ones involving a moved object of some kind (an object clitic, a Wh-moved object or a A-moved object). To account for this limited distribution, Kayne assumes that the Specifier of Object Agreement in French is not a Case position. Since Accusative assignment in French is quite independent of the Object Agreement restricted paradigm, we are led to the conclusion that, if a universal Object agreement is to be postulated, it has nothing to do with French overt Object Agreement.

We will ignore for the moment the possibility of non-overt Object Agreement, and assume that only French has an Object Agreement FC. I come back to the issue below.

Suppose that Specifier of Agreement FCs counts as an A-position, as would be

predicted by Rizzi's (1991-b) theory. If we adopt Relativized Minimality (RM), then the existence of an Object Agreement poses an immediate problem for the Internal Subject Hypothesis. Since the Spec of O-AGR counts as an A-position, raising of the external Argument to Spec of Subject-AGR will skip such an A-specifier, which is forbidden in RM. A French sentence like 0 should violate RM:

(145) Jean<sub>i</sub> les<sub>j</sub> a [<sub>O-AGR</sub> t<sub>j</sub> repeintes<sub>v</sub> [<sub>VP</sub> t<sub>i</sub> [<sub>VP</sub> t<sub>v</sub> t<sub>j</sub> ]]]

Since 0 is well-formed, the prediction is incorrect. To solve this problem without giving up RM, let us assume the following UG options.

Suppose that, whenever Spec of (Subject) AGR has to be filled because it is the AGR-identifier, the unmarked way of supplying the required DP is for the I-subject itself to raise to this position. However, when a language (such as French), has A-specifiers intervening between the I-subject and Spec of (Subject) AGR, then this option will not be available in sentences where Object Agreement is present, namely, sentence with compound tenses. In this case, a marked option can be adopted by which a DP is directly inserted in Spec of AGR and it is coindexed to the I-subject in a resumptive-like way. Since French is forced to take this option, then overt I-subjects in French need not involve an empty anaphor, as is the case in 0. In Germanic languages, on the other hand, no Object Agreement being present, the unmarked option is taken and Spec of AGR is always filled by movement, which accounts for the floating character of overt I-subjects.

In Germanic languages, both overt I-subjects and FQs would be elements left floating by A-movement. We argued in section 2.4. that in NSLs FQs need not form a possible constituent with the preverbal subject. We argued this is because in NSLs no A-Chain is formed between the preverbal subject and the I-subject, so there is no reason why FQs should be analyzable as elements left behind by movement. Now, for French we assume there is a Chain, but this Chain is not (always) a movement Chain, but rather a resumptive-strategy Chain. The

prediction is then that in French FQs are not necessarily left behind by movement and therefore need not form a possible constituent with their preverbal subject. Specifically, *tous les deux/trois* 'all the two/three' (= 'both'/'the three of them') are FQs that do not form a possible constituent with their antecedent:

- (146) a. *Les enfants /ils* sont allés *tous les deux* au cinéma  
 The children/they are gone all the two to-the movies  
 'The children/they have both gone to the movies'
- b. \**Tous les deux les enfants* / \**tous les deux ils*  
 All the two the children all the two they  
 'Both the children' / 'They both'

For NSLs having object agreement with a pattern similar to the French one (Italian, some Catalan dialects), the minimality problem does not arise, since we assume that those languages never involve I-subject raising.

Let us briefly speculate on two issues that the above approach raises. One is the existence of resumptive A-Chains, as we could call the Spec-AGR/I-subject dependencies in French cases like 0. How do they escape being subject to the ECP? In fact, a representational point of view (which is always worth keeping an eye on) would not easily distinguish between resumptive and movement A-Chain.

It seems, however, that the distinction between movement dependencies and resumptive dependencies is an irreducible complexity: in the case of A'-dependencies it would be hard to reduce one to the other.<sup>104</sup> It is clear as well that resumptive strategies are more permissive than movement strategies. Then it seems that it could be interesting to extend the distinction to A-Chains, provided we manage to put some constraints on the availability of the options (such as the markedness strategy we propose).

---

<sup>104</sup> See Cinque (1990) for extensive discussion on the issue.

Another issue to consider is Chomsky's Object-AGR. As we said, the existence of universal Object-AGR is in principle a welcome possibility for a theory claiming that Accusative is always available in principle. However, Chomsky's proposal raises several problems. One is that it is considerably theory internal and empirically under-determined: for many languages there is little evidence that object agreement exists. If objects receive Accusative by moving to Spec of Object-AGR, then we could possibly expect some head initial languages to have the order Object-Verb (i.e., languages having overt object movement and covert V-movement). This unattested possibility can be excluded somehow, but then the hypothesis has little predictive power for comparative syntax.

Perhaps part of Chomsky's theory could be kept without conflicting with RM. In fact, as far as I know, Chomsky's recent proposals have been built as an alternative to RM, and it is not surprising that the two theories conflict. However, sometimes it happens that alternative research strategies are not pointing at the same theoretical domain and are not, therefore, theoretically incompatible. It could turn out, for instance, that Chomsky's Spec of Object-AGR position is not an A-position as far as RM is concerned. For instance, we could try to define, in the spirit of Rizzi (1991-b), (non- $\theta$ ) A-positions as Specifiers of FC showing overt agreement.<sup>105</sup> Object-AGR would never be involve overt agreement in many languages.

Another possibility we could contemplate is that Object-AGR assigns Accusative by government, and that then its Specifier is not (necessarily) projected, so that no problem for RM arises. I leave the issue here.

---

<sup>105</sup> In that sense, English null  $C^{\circ}$ , which is a manifestation of agreement in Rizzi's theory, would be 'overt' agreement in the sense that its obligatory null form in the relevant cases is in 'overt' contrast with the optionally overt form (**that**).

## 7. Summary

In this chapter we have advanced a theory for deriving Burzio's Generalization and characterizing the status of the Null Subject Phenomenon. Concerning the former, we have proposed some parameters for accounting for impersonal constructions in languages like German on one hand, and for Ergative languages on the other hand.

Concerning the Null Subject phenomenology, this theory tries to minimize the differences between NSLs and (non)-NSLs: all languages have some form of null subjects and subject inversion, the difference lying in the [ $\pm$ anaphoric] character of both.

Our theory crucially relies on the properties of AGR, which we claim are universal (*modulo* some parameters concerning Ergative Languages and the presence of phi-features in a given language), and can be conceived of as instances of licensing principles.

This formulation led us to revise Binding Theory and Case Theory. We have provided some independent motivation for our revision of Binding Theory (anaphoric copulative constructions).

The strong requirements AGR imposes under our characterization of the facts led us to consider whether the ECP should be the appropriate means of constraining A-movement. We noticed that there is potential redundancy between the ECP account and our account, and that we could eliminate this redundancy by reducing the power of our rules for AGR.

We have also addressed a variety of empirical problems (inverted Copulative constructions, French Stylistic Inversion, Northern Italian dialects, etc.) and theoretical problems (Minimality, Chain formation, etc.) in a rather speculative and often inconclusive way: it could not be otherwise, given the wide range of implications the present thesis is involved in.





## Chapter 4

### Infinitival Constructions

In the preceding chapter we have developed a theory on how AGR determines what is the I-subject in a clause, in order to derive Burzio's Generalization and explain the contrast between NSLs and non-NSLs w.r.t. subject inversion. We claimed that AGR morphology plays a central role in determining a good deal of cross-linguistic contrasts.

If this is correct, then our account of infinitival clauses cannot be a trivial extension of our theory for finite clauses, for non-finite clauses have the central property, in many languages, of not showing any AGR morphology. So, two possibilities come to mind: either non-finite clauses have a radically different behavior w.r.t. the phenomena discussed in the previous chapter, or morphology is not so crucial as we claimed in accounting for those phenomena.

Our proposal will be that neither situation is exactly true: although non-finite clauses have more restricted possibilities, they are in many essential respects similar to finite clauses, because, on the one hand, they have alternative means of recovering AGR content apart from morphology and, on the other hand, they are subject to some *parallelism* principles w.r.t. the finite clauses in the same language. In other words, the speaker recovers the lack of information in non-finite clauses from both UG and some parametric options fixed on the basis of finite clauses, ultimately, from the richness of AGR in finite clauses.

## 1. Some Properties of Infinitival Constructions

A mainstream of generative analyses of infinitives is based on the hypothesis that infinitives are full clauses differing from finite clauses only in their impoverished inflectional content. The highly restricted possibilities for subjects in infinitives would be derived from the weak character (or perhaps absence) of inflectional content. The standard analysis since Chomsky (1981) assumes that infinitival INFL can neither govern nor Case-mark its subject, so that either this subject has to be PRO (which need not be governed nor Case-marked) or obtains Case independently of the infinitival INFL (in ECM constructions, 'for'-infinitives or raising constructions).

Concerning PRO, its restricted distribution has been made to derive from the postulation that it is a [+pronominal, +anaphoric] DP. Since BT requirements on [+pronominal] and [+anaphoric] elements are contradictory, PRO has to escape such requirements by being ungoverned, hence having no Binding Domain. Such a theory faces two main problems:

a) Since PRO escapes all binding requirements, it remains a mystery why its reference requirements are so highly restricted: either it is controlled or it receives arbitrary interpretation, the choice not being free in most cases. We will not propose any interesting solution for this fact, but we will contend that it is not possible to derive control from other modules of the grammar (such as Binding Theory).

b) Given a theory of Visibility as defined in Chomsky (1986-b), if PRO does not receive Case, it should not be licenced as an Argument.

Kayne (1991) presents a hypothesis that avoids at least the former problem: PRO is always governed (at least at LF) and therefore it is subject to binding requirements, so that control is reduced to BT. We will argue that this position is untenable. Kayne does not say anything about the possibility that PRO may receive Case, but his theory could be extended in that direction more easily than standard ones. We will discuss Kayne's proposal in detail in

### 1.1. Parallelism between Finite and Infinitival Constructions

Infinitives are like finite clauses except for their lack of explicit content in INFL features, from which their dependent character in both temporal interpretation and subject possibilities results. Essential for our concerns is the fact that I-subjects in infinitives show the same restrictions as in finite clauses, as far as the generalization we proposed in Chapter 2, repeated here as 0, is concerned:

- (1) I-subjects are [-anaphoric] in NSLs and [+anaphoric] in non-NSLs.

The following examples show that 0 holds in infinitives the same as in finite clauses:<sup>106</sup>

- (2) English, French, German:

- a. John decided [ to do it *himself*/\**him* ]  
b. C'est mieux [ de le faire *soi-même*/\**on* /\**lui(-même)* ]

It is better to it-do SE-SELF/ one/ he (-SELF)

'It is better to do it oneself'

- c. Hans beschloss [ es *selbst*/\**er* zu machen ]

H. decided it SELF / he to do

'Hans decided to do it himself'

---

<sup>106</sup> In controlled infinitives, NSLs allow pronominals as I-subjects, but not R-expressions. We will derive this fact from our theory of control.

(3) Catalan, Spanish, Italian:

- a. En Joan prefereix [ de fer-ho *ell (mateix)/\*si mateix* ]  
The J. prefers to do-it he (SELF) / SE SELF
- b. Juan prefiere [ hacerlo *él (mismo) /\*si mismo* ]  
J. prefers to-do-it he (SELF) / SE SELF
- c. Gianni preferisce [ di farlo *lui (stesso) /\*se stesso* ]  
G. prefers to do-it he (SELF) / SE SELF

Given that our account of the distribution of I-subjects crucially relies on the Case position for the I-subject and, on the other hand, PRO-infinitives are standardly assumed to contrast with finite clauses as far subject-Case is concerned, the facts in 0/0 do not trivially follow from the assumptions we made to account for 0 in finite clauses.

Suppose we assumed that PRO does not have Case. Then the A-Chain containing the I-subject would not have a Case position, and no Binding Domain would be definable for the I-subject, according to our definition of Binding Domain, repeated here:

- (4) A is Binding Domain for B iff A is the minimal FC containing B, a governor of B and the Case position from which B obtains Case.

If no Binding Domain is definable, no prediction is made concerning the [=anaphoric] character of the I-subject. This is an undesirable result, since the infinitives show exactly the same behavior as finite sentences, as we see in 0/0, and this should not be a matter of accident. In addition, what we called reflexive copulative constructions, whose behavior we derived from the same definition of Binding Domain behave exactly the same in finite and infinitival constructions:

- (5) a. He tried [ to be *himself*/\**him* again ]  
 b. On doit essayer [ d'être *soi-même*/\**on*/\**lui(même)* ]  
 One must try to be SE-SELF / one/he (SELF)  
 c. Er versuchte [ wieder *sich selbst*/\**er* zu sein ]  
 He tried again SE SELF / he to be  
 'He tried to be himself again'
- (6) a. Intentava de tornar a [ ser *ell (mateix)*/\**si mateix* ]  
 He-tried to return to be he (SELF) / SE SELF  
 'He tried to be himself again'  
 b. Intentaba volver a [ ser *él (mismo)*/\**si mismo* ]  
 He tried to-return to be he (SELF) / SE SELF  
 c. Non riusciva ad [ essere *lui (stesso)*/\**se stesso* ]  
 Not managed to be he (SELF) /\*SE SELF  
 'He didn't manage to be himself'

So we are led to the conclusion that infinitival I-subjects have a Case position, and that the Case position is, for as given language, the same as that of the finite clauses.<sup>107</sup> More specifically, infinitival I-subjects inherit Case from Spec of AGR in non-NSLs, while they are directly Case-marked in NSLs. Since this is a necessary requirement for our theory to be extendable to infinitives, we have to build some plausible theory that achieves this result. As far as the theory of visibility or, more generally, any theory pointing to uniform Case requirements for all Arguments is to be welcome on the grounds of simplicity, our proposal will have some independent plausibility.

---

<sup>107</sup> We will propose that this is not the case for some languages such as Occitan and Sardinian.

## 1.2. V-movement in Infinitives

Pollock (1989) shows that verb movement (for lexical verbs) in infinitives is shorter than in (French) finite sentences:

- (7) a. Jean ne pense *pas toujours* au futur  
J. ne thinks not always of-the future  
b. Ne *pas (toujours)* penser (*toujours*) au futur...  
ne not (always) to-think (always) of-the future.  
c. *Not to always* think of the future...

Pollock's theory derives this fact from the 'weak' or 'poor' character of Tense in non-finite sentences. Since all Romance infinitives show no Tense morphology on the infinitival verb, the prediction should be that infinitives in these languages do not allow long verb movement, contrary to fact: in Italian, Spanish and Catalan, infinitival verb movement is apparently as long as finite verb movement.<sup>108</sup>

- (8) a. Non ama *più* Maria (Italian)  
Not loves *anymore* M.  
b. Non (*\*più*) amare *più* Maria...  
Not to-love *anymore* M.

---

<sup>108</sup> If not longer: we will argue that enclisis in infinitives should be analyzed as extra verb movement (our proposal will differ from Kayne (1991), who argues for a very short kind of extra movement). See Belletti (1991) for the idea that elements like **più**, **mai**, etc. occupy a position similar to French **pas**, **plus** and other negative elements. On the other hand, there are NSLs with short V-movement in infinitives. We will address the question in section 1.3.

- c. No diu *mai* la veritat (Catalan)  
Not tells *never* the truth
- d. No (\**mai*) dir *mai* la veritat...  
Not to-tell *never* the truth
- e. No està *nunca* cansado (Spanish)  
Not is *never* tired
- f. No (\**nunca*) estar *nunca* cansado...  
Not to-be *never* tired

Thus, there seems to be a correlation between the Null-Subject status of the language and the possibility for the infinitival verb to raise to a high (the highest) functional category: AGR in Belletti (1991)'s theory and our own. Another fact, which is likely to be parasitic on the former, is the possibility of clitic climbing. Kayne (1989) argues that clitic climbing (and also long V-movement in infinitives) is due to the strong character of INFL in NSLs languages.

The problem is how to express the correlation between the strong character of AGR in finite clauses and the purported strong character of AGR in non-finite clauses. In infinitival clauses, AGR is not apparently strong in Italian or Spanish, as far as morphology can tell us. To simply stipulate that infinitival AGR is strong because finite AGR in the same language is strong appears to be a mere stipulation.

What I want to propose is the idea that what extends from finite to infinitival clauses is a parameter value. But before proceeding, let us consider the facts considered in Kayne (1991) concerning V-movement and clitic placement in infinitives, and the proposal Kayne presents to account for the facts.

### 1.3. Kayne's Proposal

Kayne (1991) shows that in Romance languages there is a correlation between enclisis in infinitives and the possibility of having interrogative infinitives with a COMP particle of the type 'if', which he convincingly argues is a C<sup>0</sup> particle:

(9) *Italian:*

a. *andarci / vederlo* (enclisis)

*to-go-there to-see-it*

b. *Non so se andarci*

*Not know-I if to-go-there*

*'I don't know whether to go or not'*

(10) *French:*

a. *y aller / le voire*(proclisis)

*there-to-go it-to-see*

b. *\*Je ne sais pas si y aller*

*I ne know not if there to-go*

*'I don't know whether to go or not'*

Kayne argues that this is a genuine correlation holding of many Romance languages, and that it does not correlate with the (non) NSL status of the language.<sup>109</sup> So Catalan, Spanish, and some Italian dialects (Piedmontese, Milanese and Paduan) are like Italian, while Occitan, Sardinian and some Italian dialects (Gardenese) are like French. As for English (and languages having short V-movement in infinitives in general), they would pattern with French in not

---

<sup>109</sup> Kayne argues that the NSL status is a necessary but not sufficient condition for a language to have enclisis and thus 'if'-infinitive constructions.



allowing 'if'-infinitives:

- (11) a. No sé *si* anar-*hi* (Catalan)  
Not know-I if to-go-there
- b. No sé *si* hacer*lo* (Spanish)  
Not know-I if to-do-it
- c. \*Sabi pas *se hi* anar (Occitan)  
Know-I not if there-to-go
- d. \*No'isco *si* andare (Sardinian)  
Not know-I if to-go
- e. \*I don't know if to go (English)

Kayne's analysis of these facts accounts for both the (im)possibility of 'if'-infinitives and the enclisis phenomenon. Concerning the latter, he assumes that clitic attachment is (at least in Romance) *left*-adjunction to a functional head, for principled reasons having to do with morphological headedness. Therefore, whenever a clitic is strictly left-adjacent to the verb, the structure would be:

- (12)  $[I^0 \text{ cl } [I^0 V^0 ] ]$

where  $I^0$  is some functional (inflectional) head. In fact, the general condition Kayne imposes on clitic attachment is that clitics must attach to a functional head, whether it contains the moved  $V^0$  or is silent. However, he follows Baker (1985) in assuming that the head the clitic adjoins to cannot contain a trace. A consequence of this assumption would be that, since verb movement is cyclic for principled reasons, the clitic cannot adjoin to a functional head lower than the one where  $V^0$  has moved, because this would imply that the clitic is adjoined to a head containing a trace:

(13)  $*[X^0 V^0] \dots [Y^0 \text{cl} [Y^0 t]]$

where  $X^0$  and  $Y^0$  are functional heads and  $t$  is a necessary intermediate trace of  $V^0$ -movement. The reverse situation, where the clitic is in a higher functional head than  $V^0$  is not excluded:

(14)  $[X^0 \text{cl} X^0] \dots [Y^0 V^0]$

This would be the structure for languages having the clitic preceding the infinitival verb but not necessarily adjacent to the verb (Occitan, earlier French).

Since both  $V^0$  and the clitic are heads, why is it that the clitic can move non-cyclically (giving 0) while  $V^0$  cannot (as far as 0 is not allowed)? Kayne's solution is based on the idea, developed in Kayne (1989), that for NSLs infinitival INFL (or some of its members) is strong enough to l-mark and void some potential barriers,<sup>110</sup> so that long head movement is allowed to some extent. However,  $V^0$ -movement is subject to a further constraint: it has to pick up the affixal functional heads. This would explain the contrast 0/0: in 0 the verb has to move through  $Y^0$  to merge with this affixal head, while the clitic in 0 meets no similar requirement.

Since cyclic movement (as expressed by the HMC) is not a matter of principle, but is rather derived (from the ECP and affixation requirements), there might be structures where  $V^0$  skips the clitic position without violating any principled requirement. Kayne argues that there is one such structure. Suppose the clitic moves to a functional head  $I^0$ , left-adjoining to it (as it has to); suppose  $I^0$  has no content, so that  $V^0$  is not required to move to it to pick up any affixes; suppose finally the verb adjoins to the  $I'$  projection, giving:

---

<sup>110</sup> Specifically, in Kayne (1989) VP would be the potential barrier that is voided by  $\text{INFL}^0$  l-marking. Although Kayne (1991) does not address the question, at least some of the 'heirs' of INFL after the INFL splitting hypothesis should be l-marking elements for the hypothesis in Kayne (1989) to be extendable to the proposal in Kayne (1991).

(15) [I' V<sup>0</sup> [I' [I<sup>0</sup> cl I<sup>0</sup> ] ] ]

0 does not possibly violate the ECP, since I' is not a minimality barrier (in the sense of Chomsky (1986-a)) because it does not exclude V<sup>0</sup>; it does not violate any affixation requirement as far as I<sup>0</sup> is content-less. Therefore it is a possible structure. Kayne claims that 0 is actually the structure for enclisis in Romance infinitives: the apparent fact that the clitic is right adjoined to the (functional head containing the) verb is due to the verb having skipped the position of the clitic. Kayne assumes that I<sup>0</sup> is T<sup>0</sup>, which is likely to be content-less in infinitives. On the other hand, the proposal that clitic attachment (in Romance) is left adjunction to a functional head can be assumed to hold without exception. And the very short span by which V<sup>0</sup> out-raises the clitic accounts for the apparent right-adjunction of the clitic to the verb. For more details on this proposal, see Kayne (1991).

The existence of a structure like 0 is only partially derived in Kayne's account: since it involves long clitic movement, it should be restricted to NSLs (as argued in Kayne (1989)), but being a NSL is not a sufficient condition for allowing 0, as far as there are NSLs not having enclisis in infinitives (Occitan, Old French, Sardinian). In any case, Kayne develops a proposal that makes the existence of such a structure highly appealing, in that it allows for an explanation of the generalization we mentioned above, namely that only languages having enclisis in infinitives allow 'if'-infinitives (recall the contrast 0/0).

Kayne's account is based on several assumptions. One is that 'if' is a C<sup>0</sup>-particle that, contrary to empty infinitival C<sup>0</sup>, is able to govern the specifier of IP, which is occupied by PRO in control infinitives.<sup>111</sup> If we simply assume that PRO cannot be governed (due to the PRO-

---

<sup>111</sup> Kayne crucially assumes that prepositions preceding (controlled) infinitives in Romance, as in French (i), are not in C<sup>0</sup>:

(i) Jean essaie **de** comprendre  
for they do not have the same effects as 'if'. I agree on this

theorem), then we explain that 'if'-infinitives are not allowed in French or English. We do not thus far explain, however, why languages having enclisis allow 'if'-infinitives, since the existence of the structure 0 does not trivially bear on the PRO-theorem problem.

In order to account for the facts, Kayne adapts a proposal by Chomsky (1986-b) on the definition of Binding Domain. Chomsky wanted to account for cases where pronominals and anaphors are not in complementary distribution, such as:

- (16) a. *They* like *their* pictures  
 b. *They* like *each other's* pictures

Chomsky's proposal is that the definition of Binding Domain is sensitive to the [ $\pm$ anaphoric] character of the element whose Binding Domain is determined. Without going into the details of Chomsky's technical definitions, the essential idea is that the Binding Domain for a pronominal/anaphoric element has to fulfil the condition of virtually allowing for the pronominal/anaphor to be respectively free/bound. In more concrete terms, this means that X is Binding Domain for an anaphor only if X contains an A-position c-commanding the anaphor; since pronominals do not have to (in fact cannot) be bound in their Binding Domain, there is no requirement of virtual binding for pronominals.

With these assumptions, the facts in 0 can be explained. Consider whether the NP containing the genitive is a possible Binding Domain for the pronominal or anaphor in this

---

point, and this will be crucial for my account too. However I think the preposition **de** (and its Romance counterparts) cannot be in Spec of CP, as Kayne assumes. On the one hand, since a specifier cannot be occupied by a head, we would be dealing with a complement-less PP. On the other hand, if **de** is in Spec of CP, **de**-infinitives would be like Wh-islands, contrary to fact:

- (ii) **Quand**<sub>i</sub> as-tu essayé [ **de** venir t<sub>i</sub> ]  
 When have-you tried **de** to-come  
 Cf. \***Quand**<sub>i</sub> m'as-tu demandé [ **où** Jean allait t<sub>i</sub> ]  
 When me-have-you asked where J. went

We could assume that **de** is outside CP. This is not in fact incompatible with Kayne's or my theory, nor is it crucial to either, provided **de** is not in C°.

genitive position. Suppose this NP fulfils that part of the definition of Binding Domain which is common to pronominals and anaphors, namely:

- (17) B is a Binding Domain for A iff B is the minimal CFC containing A and a governor of A.

For the anaphor (*each other*), however, there is the further requirement that the Binding Domain has to contain an A-position c-commanding it. Since the anaphor is in the specifier of the NP, i.e., the highest specifier in this NP, such a requirement is not met, and the Binding Domain will be the next CFC up, namely the whole sentence, where the anaphor will be correctly bound. For the determination of pronominal's Binding Domain, there is no such requirement, so the NP itself can be the Binding Domain, allowing the pronominal (*their*) to be free in it.

PRO, like genitives, can be assumed to occupy the highest specifier of the IP. Since IP is a CFC, Chomsky's proposal can easily be made to bear on the PRO distribution. Since Chomsky's proposal allows for the determination of the Binding Domain to give different results depending on the pronominal/anaphor status of the element in question, then the simultaneous pronominal/anaphoric status of PRO need not lead to the PRO-theorem. Specifically, if a CFC XP contains PRO, a governor of PRO and no A-position c-commanding PRO, then PRO as pronominal will have XP as its Binding Domain, where it will be free as required, but XP will not be the Binding Domain for PRO as anaphor: the Binding Domain will be the whole IP next up, where it will be bound by the controller in the standard case.

Now consider the structure of an 'if' infinitive in English or French, where it is excluded:

- (18)  $[_{C^0} \text{if}] [_{IP} \text{PRO} [_{I'} I^0 ]]$

Since there is no governor of PRO inside IP, IP cannot be the Binding Domain for PRO. In fact, the governor for PRO is  $C^0$  (containing 'if'), but since  $C^0$  is outside the IP CFC, the

Binding Domain has to be extended to the next CFC, namely the superordinate clause, where PRO, although being correctly bound by the controller as an anaphor, is incorrectly bound as a pronominal. Therefore the 'if'-infinitive is correctly excluded by the second BT principle.

Now consider the structure of an 'if'-infinitive construction in a language having infinitive enclisis. In these languages,  $V^0$  is adjoined to  $I'$ :<sup>112</sup>

(19)  $[C^0 \text{ if}] [IP \text{ PRO } [I' V^0 [I' I^0 ]]]$

In 0, Kayne argues,  $C^0$  cannot govern PRO because there is a closer governor, namely  $V^0$ , which creates a minimality effect. Therefore IP is a CFC containing a governor for PRO (namely  $V^0$ ). For IP to be the Binding Domain for PRO as pronominal nothing else is required, and PRO as pronominal is correctly free within IP. On the other hand IP does not fulfil the virtual binding requirements of PRO as anaphor, so the Binding Domain for PRO-anaphor is extended to the superordinate clause, where it is licitly bound by the controller.

So NSLs having infinitival enclisis allow for PRO to be governed without violating BT. At S-structure, the other set of languages do not have a governed PRO. Since, however, the hypothesis that PRO may be governed and, therefore, subject to BT is a natural way of accounting for the strict referential constraints on PRO (it is controlled or arbitrary), Kayne proposes that the governed status of PRO is universal (i.e., PRO never escapes BT). In languages where it cannot be licitly governed at S-structure, it is governed at LF, where V-raising would create a structure similar to 0.

Kayne's analysis is appealing in several respects:

- It accounts for an interesting generalization concerning enclisis and the existence of 'if'-

---

<sup>112</sup>  $V^0$  adjunction to  $I'$  would take place in all infinitives in these languages independently of whether an (en)clitic is present or not: enclisis would only be a manifestation of the phenomenon.

infinitives.<sup>113</sup> This generalization could hardly be a matter of accident and, even if it was, it would pose a problem for learnability as far as, for languages not having 'if'-infinitives, the learner would need negative evidence in order not to generalize Wh-interrogative infinitives to 'if'-infinitives on the basis of finite clauses, where both Wh- and 'if'-interrogatives are possible.

- It sets a plausible basis for reducing control to BT, a desirable result in view of the hitherto poorly understood phenomenon of control.

- it accounts for (the possibility of) enclisis in a highly principled way, on the basis of the ECP and affixation constraints in V-raising, as well as morphological-headedness constraints on clitic attachment.

However, Kayne's proposal is far from crystal-clear in many respects, which we will consider in turn.

One question Kayne's proposal raises concerns his crucial claim that  $V^0$  adjoins to  $I'$ . There are two possible problems with this proposal. One is whether adjunction of  $X^0$  to  $Y'$  should be allowed at all. The other problem is that Kayne's analysis of proclisis vs. enclisis analyses the former as having the clitic more closely attached to the (functional head containing the) verb than the latter. There is some evidence pointing to the opposite way (see Benincà & Cinque (1990)). Since, however, these two possible objections are extrinsic or peripheral to Kayne's discussion, we will not pursue them here. We will concentrate on intrinsic problems Kayne's theory cannot escape facing.

There is a problem that is essentially connected to Kayne's proposal, although he does not explicitly address it. Kayne seems to tacitly assume that the only means of having PRO governed inside IP in structures like 0/0 is having the governing head adjoined to  $I'$ . Let us see why this tacit assumption is necessary. Kayne analyses Sardinian as a language having infinitival V-raising to  $T^0$  ( $= I^0$  in 0), i.e., to the functional head the clitic is attached to. Since this language does not allow 'if'-infinitives, we must imply that V in  $I^0$  cannot govern PRO while V adjoined to

---

<sup>113</sup> Portuguese is a potential problem for the empirical generalization, as we will see above.

I' can. Otherwise, Sardinian would allow 'if'-infinitives. Or would it not? Let us consider the possibility that a head governs its specifier.

Kayne seems to emphasize the idea that V-adjoined-to-I' plays the role of blocking government by C<sup>0</sup> by minimality; so, Kayne could argue, I<sup>0</sup>, even if able to govern its specifier PRO when filled by V<sup>0</sup>, would not block government by C<sup>0</sup>. But notice that the government requirement in the definition of Binding Domain (as in 0) is not a requirement of exclusive government: if I<sup>0</sup> governs PRO in its specifier, that is sufficient for IP to become a potential Binding Domain (and an actual one for PRO as pronominal). So, to the extent we allow a head to govern its specifier, and we assume that functional head containing V<sup>0</sup> is a governor, then it is immaterial whether C<sup>0</sup> governs PRO or not whenever PRO is governed by I<sup>0</sup> containing V. So Sardinian, for which Kayne assumes V raises to the functional head whose specifier contains PRO, would be predicted to allow 'if'-infinitives, contrary to fact.

An option for solving this problem could be that heads do not govern specifiers, so that adjunction to I' is the only means for V to govern PRO from inside IP.

But this is not easily tenable either. Consider again Chomsky's original proposal. If heads do not govern specifiers, we would have to assume that in a case like:

(20) *They* like [<sub>NP</sub> *their* books ]

the genitive pronoun (*their*) has to be governed by a head adjoined to N' (or D' if we adopt the DP hypothesis) in order to prevent the verb (*like*) from governing the genitive and enlarging the Binding Domain to the whole clause. Put in general terms, Kayne's proposal, although designed to account only for the PRO distribution, leads to the conclusion that any case of non-complementary distribution of pronominals and anaphors involves X<sup>0</sup> adjunction to Y'.

If we go further into the consequences of this proposal, even more basic problems appear. Consider a structure like:



(21) \**John* doesn't remember if [<sub>IP</sub> *himself* actually won ]

If the 'if'  $C^0$  is a governor of the Spec of IP in infinitives, it is likely to be so in finite contexts as well. Therefore, the IP in 0 should not count as a Binding Domain unless, again, we assumed that the some element adjoined to I' blocks this government. In this case, to impair things, the element in question would be empty, in fact radically empty if affix lowering takes place at S-structure, as in Chomsky (1988).<sup>114</sup> Kayne crucially assumes for independent reasons that an anaphor has to abide by BT requirements at S-structure, so that 0 could not be rescued at LF.

To try to rescue Kayne's hypothesis from both 0 and 0 there are some ways-out to try. We could assume that in genitives the 's particle (which would be morphologically irregular for *their*) is actually a  $D^0$  element adjoined to D', which would elegantly account for its enclitic character.<sup>115</sup> To avoid the problem in 0, we could simply assume Rizzi's (1989-b) idea of the 'Anaphor Agreement Effect', by which anaphors in subject position of tensed clauses would be excluded on independent grounds. However no similar account is possible for the grammaticality of 0, which would be incorrectly ruled out because of BT second principle, in parallel with 0 being incorrectly ruled in:

(22) *John* does not remember if [<sub>IP</sub> *he* actually won ]

In conclusion, although Kayne unproblematically accounts for the impossibility of 'if'-infinitives in short infinitival V-movement languages, the account is less clear for enclisis

---

<sup>114</sup> Notice that 0 is a problem for Chomsky's (1986-b) proposal as well: he has to assume (and he does quickly in passing) that AGR counts as a virtual antecedent without being a possible antecedent, in order for the IP to be a Binding Domain for **himself** without possibly being one in which **himself** can be bound.

<sup>115</sup> We would assume, as in Fukui & Speas (1986), that 's is the head of  $D^0$  being a genitive Case assigner.

languages, and it potentially runs into problems for Sardinian, where it is not clear why 'if'-infinitives are not possible, and, besides, for cases like 0, 0 and, more conclusively, 0.

If we assume that  $I^0$  can govern its PRO specifier then Sardinian ought to allow 'if'-infinitives as we argued. Suppose we assumed that Sardinian infinitival V-movement, although very 'long' (almost as long the Italian one) does not however reach the  $I^0$  whose specifier contains PRO, but rather stops a step short of it. Then, the Sardinian status would be essentially the same as the French one: an empty  $I^0$  is unable to govern and the 'if'  $C^0$  really matters. Although this position is tenable for Kayne, it undermines the cruciality of the V-adjoining-to-I' proposal: we could simply assume that long movement is movement to  $I^0$ , (which for some reason implies enclisis), and that suffices to prevent  $C^0$  to be the exclusive governor of PRO. In languages with short movement (and now Sardinian is not relevantly different from English, French or Occitan)  $I^0$  is empty and  $C^0$  is the exclusive governor. So the idea of adjunction to I' is, at best, less motivated than Kayne claims it is.

There is another problem of a rather speculative nature. In both Chomsky's and in Kayne's proposal (implicitly in the latter), the notion of CFC plays an essential role in the definition of Binding Domain: since only a CFC can be a Binding Domain, whenever a CFC (such as the infinitival IP) fails to be a Binding Domain (because the governor is outside it or it does not satisfy virtual binding requirements) then the Binding Domain switches to the whole next IP up. The CP or VP immediately dominating the IP are not possible candidates simply because they are not CFCs. At the time Chomsky formulated his proposal, he assumed a still fairly simple structure for the sentence (the S'/S category) and, specifically, he did not consider the internal subject hypothesis at all: (NP,S) was the  $\theta$ -position for at least external Arguments, so S (=IP) was clearly the minimal constituent containing all the  $\theta$ -positions of a clause (at least in agentive clauses), thus a CFC.

The moment the internal subject hypothesis is assumed, however, things are not so simple: it could be argued that  $VP^{\max}$  (in Koopman & Sportiche's (1988) sense) is the (minimal) CFC. Similar problems would arise in connection to the DP hypothesis if we generate nominal  $\theta$ -

positions inside the (strict) NP. In order to preserve Chomsky's results, there is an obvious solution: we define CFC as the minimal constituent containing all the  $\theta$ -positions of a predicate plus the A-specifiers of the FCs locally dominating it. If the specifier of the highest IP (or the genitive specifier of DP) are A-positions, then the definition will give CFCs which coincide in essence with Chomsky's and Kayne's proposal: the highest IP for clauses and DP for nominals. (which allows Spec of CP to be an A-position under certain conditions) could be problematic. I will not speculate further on these ramifications.

This simple

Crucial for the present thesis is, however, the fact that we postulate that, at least in the set of NSLs considered, I-subjects (which are the most prominent  $\theta$ -positions in a clause) do not form an A-Chain with Spec of IP, then there is no plausible motivation for extending the CFC from  $VP^{\max}$  to IP. In fact, our definition of Binding Domain basically restricts the candidates by requiring them to be functional categories:

- (23) A is Binding Domain for B iff A is the minimal *FC* containing B, a governor of B and the Case position from which B obtains Case.

In all hypotheses about FC structure which simultaneously assume the internal subject hypothesis, FCs (as the various members of IP or DP) form a 'shell' around the predicative lexical categories (resp.  $VP^{\max}$  and NP). Therefore requiring the Binding Domain to be a FC amounts to ensuring that a Binding Domain will always contain the whole CFC (if the latter is defined on the basis of  $\theta$ -positions only).

There is another problem Kayne himself points out: if enclisis is the manifestation of close adjacency between V and the clitic without both forming a constituent, then the prediction is that V-cl cannot move as such. Since AUX-to-COMP (as characterized in Rizzi's work) has enclisis on the AUX, this could only follow from the accidental fact that the clitic happens to (left) adjoin to  $C^0$  and the AUX adjoins to C'. Although not an impossible accident, it is rather suspicious that enclisis is preserved in AUX-to-COMP.

Kayne points out an additional problem: European Portuguese allows both enclisis and proclisis in infinitives. Nevertheless, it allows 'if'-infinitives. Kayne's proposal would be strongly confirmed if Portuguese speakers allowed 'if'-infinitives *only* in the enclisis option, which is not the case: there seems to be no contrast at all between the two options, both allowing 'if'-infinitives with total naturalness. The alternative account of Kayne's data I will propose is, I think, less in trouble with Portuguese than Kayne's, even if this is at the cost of being less restrictive.

Let me point out a final problem. Kayne's attempt to reduce control (and arbitrary PRO would also be a form of control) to Binding Theory (BT) is appealing, for control is thus far one of the modules of grammar that has remained most obscure of all. However it is well-known that the mysterious nature of control is hard to reduce to the standard locality conditions in BT. So in (24) there is an unexpected minimal contrast between PRO and genitive (the latter being akin to PRO in not showing complementary distribution between pronominals and anaphors):

- (24) a. \**Each other's* accusation triggered many problems for *them*  
b. *PRO* to accuse Nixon triggered many problems for *them*

The problematic cases are numerous and diverse. To mention another two, both Wh-movement and dislocation may bleed or feed BT possibilities while control remains immune to both. More importantly, there are cases of dislocation where control and BT seem to meet contradictory requirements:

(25) Votar      *en Joan*, em sembla que ni      *ell mateix*  
To-vote-for the J.   me-seems   that not-even he SELF

s'imagina que t'ho proposessis

guesses   that you-intended

'Vote for *Joan*, I think not even *he himself* could guess you intended to"

In 0 BT requires *en Joan* not to be c-commanded by *ell mateix*, whereas the controlled PRO in the dislocated infinitival clause would have to be c-commanded by the subject of *t'ho proposessis* 'you intended to' if control is to be reduced to binding. In whatever level of representation one requirement is met, the other is not.

In addition, I have serious doubts that control can be reduced to purely structural conditions: at least for typical control verbs (*hope*, *convince*, etc.) there seem to be irreducible semantic or lexical factors (consider, e.g., *persuade* and *promise*) on the determination of the controller.<sup>116</sup> A sad conclusion though it may be, control seems to remain a poorly understood field of the grammar. What can easily be determined on a structural basis is only the controllee, not the controller. Our proposal, as the standard ones, will capture this fact.

Summing up, Kayne's hypothesis appears problematic for at least the following reasons:

- it challenges the structure preserving hypothesis concerning adjunction: a head is adjoined to a X'.

---

<sup>116</sup> There is a recent proposal by Larson (1991) which tries to derive whether the object or the subject is the controller from purely structural (c-command) conditions. Although plausible for some basic cases, the proposal is bound to adopt rather prolix assumptions to cover only the English data. In addition, the minimal contrast he postulates between objects c-commanding and not c-commanding the infinitive should optimally be derived from some Universal Theta Assignment Hypothesis, for they are unlikely to be learned and, in addition, object/subject control status seems to be relatively uniform for synonymous verbs across languages.

- it predicts a contrast in degree of attachment in enclisis vs. proclisis which is the opposite of what independent morpho-phonological evidence suggests.

- it implicitly assumes that a head does not govern its specifier, a problematic assumption for standard accounts of Binding Domain as Governing Category, unless we generalize  $X^0$ -adjunction-to- $Y'$  to all the parallel cases.

- it crucially relies on the notion of CFC in a way that is not trivially adaptable to present-day assumptions on clause structure.

- it represents V+enclitic as not forming a constituent, thus predicting that V-cl cannot move.

- Portuguese is a potential counter-example.

- control cannot be easily reduced to BT.

Perhaps it is not advisable to simply dismantle a proposal with such a basic appeal and perspicuous insights as Kayne's, on the basis of technical problems: these problems might disappear as a deeper understanding of the facts develops. However, at least one of the problems, namely the pervasive consequences of taking the proposal of  $X^0$ -adjunction-to- $Y'$  seriously, which has dramatic effects on cases like 0 and 0, does not seem to be a side-problem and casts serious doubts on the first glance plausibility of the theory.

The alternative account we will present is, admittedly, less ambitious, for it gives up some of the promising achievements in Kayne's theory, such as reducing control to BT.

## **2. The AGR-identifier of Infinitives**

In the preceding chapter, we adopted a parameter that determines which is the AGR-identifier for a (finite) clause, repeated here as 0:

(26)  $AGR^0$ /Spec of AGR is the AGR-identifier of AGR.

We assumed that  $AGR^0$  is the unmarked value, and that languages having poor AGR-morphology were forced to adopt the marked value (Spec of AGR).

Suppose 0 is a parameter which is set once for all types of clauses in a given language. The trigger for setting the parameter would be finite clauses, but the value chosen would be relevant for infinitival ones in the unmarked case.<sup>117</sup>

That the AGR identifier must be uniform across sentence-types (finite/infinitival) is a necessary assumption for the theory above to work. The reason is that the [ $\pm$ anaphoric] character of I-subjects does not seem to vary from finite to infinitival clauses, as we have seen:

- (27) a. Gli dispiace dover farlo *lui (stesso)*  
Him-dislikes to-have-to do-it he (SELF)  
'He dislikes to have to do it himself'  
b. It bothers him to have to do it *himself*

As for R-expressions, they are not allowed as I-subjects in controlled infinitives in NSLs, but this will follow from the assumption that control involves a PRO which binds the I-subject, and R-expressions cannot be bound.

## 2.1. Control

To make clear what we want to arrive at, let us advance the following idea, which we will try to motivate as we proceed:

---

<sup>117</sup> Occitan and Sardinian will be claimed to be marked languages in this connection. See below.

- (28) In the unmarked case, controlled infinitives have the same AGR-identifier option (AGR<sup>0</sup>/Spec-AGR) as finite sentences, in a given language.

Let us see what 0 predicts for Italian and English (we assume they are 'the unmarked case'). Consider the D-structures in 0:

- (29) a. John wants [AGRP AGR<sup>0</sup> to come DP ]  
 b. Gianni vuole [AGRP AGR<sup>0</sup> venire DP ]  
 G. wants to-come

In both cases, AGR is coindexed with the DP. Consider first the English case construction. In this case it is Spec of the infinitival AGR which is to be filled to be the AGR-identifier. Let us assume that this specifier is occupied by PRO (whose precise status we will discuss below). Let us assume that, contrary to what is standardly claimed, PRO can (and has to) receive Case, in order to abide by the Visibility Condition or some equivalent requirement. Let us assume that it is AGR<sup>0</sup> which assigns Nominative to PRO by agreement.<sup>118</sup> Suppose that infinitival AGR<sup>0</sup> can assign a Nominative only to PRO (a kind of weak Nominative that is specific to PRO).

Given the above assumptions, PRO is the Case-position for the I-subject, and the whole IP (= AGRP) is the Binding Domain for the I-subject. The I-subject will then be anaphoric, as desired.

In the Italian construction, on the other hand, it would be AGR<sup>0</sup> itself which would be

---

<sup>118</sup> We will crucially assume that heads do not govern their specifiers, so that, as we assumed for finite sentences, Nominative is assigned to Spec of AGR by agreement, not by government.

I think that the assumption that heads do not govern specifiers it is a tenable position in general. In addition, it subsumes much of the canonical-government proposals in earlier literature: non-canonical government is impossible in the general case because it would be government of a Specifier.



controlled. To obtain this result we could tentatively assume that:

- AGR<sup>o</sup> obtains person features by control, and becomes rich.
- so V raising to AGR is allowed (assuming, in the spirit of Pollock (1989) that only heads having some content allow V<sup>o</sup> to move to them).
- AGR<sup>o</sup> has to assign Nominative directly to the I-subject, by government, which requires its combining with T.
- so T has to raise to AGR, and it will if cyclic V-movement to AGR takes place.
- as in finite sentences, the Case position for the I-subject will be the I-subject itself, so that the Binding Domain for the I-subject will exclude Spec of AGR, and the I-subject will be [-anaphoric].

This is a first characterization of the facts. Let's address more specific questions. One is that in NSLs like Catalan or Italian, the emphatic I-subject in controlled infinitives is pronominal, but it cannot be an R-expression. So to say that I-subjects are [-anaphoric] is not enough, since the [-anaphoric, -pronominal] option is excluded:

- (30) a. Gianni vuole [ venire lui/pro ]  
G. wants to-come he/pro
- b. \*(Lui) vuole [ venire Gianni ]  
(He) wants to come G.

We could simply assume that the R-expression *Gianni* in 0.b) is excluded because it is bound by the controller in the upper clause (*lui*). Although this account is correct, it is not sufficient. As we pointed out above, there are cases where the controller seems not to c-command the infinitive at any level of representation. This is the case with dislocated controlled infinitives (see the discussion in section 1.3., where we conclude that certain controlled infinitives cannot be c-commanded by the controller at any level of representation).

We will assume that the impossibility of R-expressions as subjects of controlled

infinitives is due to a violation of the 3rd principle of BT, but the offending binder will not be the controller itself, but PRO. Let us assume that PRO is always present in control structures (and in cases of PRO<sub>Arb</sub> as well). We will later elaborate on the nature of PRO. Let's take it for granted that it is necessarily present in both NSL and non-NSL control structures.

How can this be accommodated within the above assumption that (in the unmarked case) NSLs have a controlled AGR<sup>0</sup>, while non-null subject languages have a controlled Spec of AGR? At first glance, this looks at variance with the assumption that both types of languages involve a controlled PRO, which is in Spec of AGR. In order to solve this apparent contradiction, let us make the following assumptions:

- PRO is a necessary element of control (and PRO<sub>Arb</sub>) structures: only PRO can be controlled (perhaps PRO<sub>Arb</sub> is controlled by an empty Argument or operator).
- PRO cannot be governed for the familiar reasons (we adhere to the PRO theorem).
- principle 0 requires (or, being a markedness principle, favors) AGR<sup>0</sup> as the controllee in NSLs. If this option is to be fulfilled, the only means it can be is that PRO adjoins to AGR<sup>0</sup>. Let us assume that PRO is a maximal projection, but that only its head (call it PRO<sup>0</sup>) has real content, so that adjunction of this head to AGR<sup>0</sup> actually implies placing the controller inside AGR<sup>0</sup>.

Under this view, NSLs would have, in the unmarked Case PRO<sup>0</sup> adjoined to AGR<sup>0</sup>, thus fulfilling 0. Now consider the following infinitival structure:

(31) [AGRP [PRO PRO<sup>0</sup>] [AGR' AGR<sup>0</sup>]]

Suppose PRO is to adjoin to AGR<sup>0</sup>. We have assumed that a head does not govern its specifier. Therefore, movement of PRO<sup>0</sup> into AGR<sup>0</sup> would violate the ECP. Therefore, with such a structure, PRO<sup>0</sup> could never adjoin to AGR<sup>0</sup> in order to make AGR<sup>0</sup> to controllee element.

Suppose, however, that above AGRP, but below CP, there is an intermediate functional category (XP) (which we will try to characterize later) to which AGR<sup>0</sup> can move, giving:

(32)  $[X^0 \text{ AGR}^0_i] [_{\text{AGRP}} [\text{PRO} \text{ PRO}^0] [_{\text{AGR}} t_i]]$

In this configuration,  $\text{PRO}^0$  is allowed to move to the  $X^0$  head containing  $\text{AGR}^0$  without violating the ECP. Therefore, 0 is a structure that allows for the controllee to end up in a head position which at the same time contains  $\text{AGR}^0$ . This would be the case for NSLs as Catalan or Italian.

Can 0 be somehow motivated? I will assume that enclisis, which is typical of NSLs, is a manifestation of the structure 0: whenever  $\text{AGR}^0$  (which contains  $V^0$ ) raises to  $X^0$ , enclisis is manifested. So we are in a position to account for enclisis and the empirical facts discussed in Kayne (1991). Let us see how.

In order to motivate our account for the enclisis facts, we will concentrate on two questions: How is 'longer' V-movement related to enclisis? and, What are the effects of adjoining  $\text{PRO}^0$  to  $\text{AGR}^0$  (or a head containing  $\text{AGR}^0$ ).

### 2.1.1. Enclisis

Let us first consider a well known case of enclisis not concerning Romance infinitives: French (complex) verb/subject-clitic inversion. An interesting and recent account for this phenomenon is Rizzi & Roberts (1989). The literature on this topic agrees on the point that V-INFL movement to COMP is a crucial factor for this construction. What concerns us here is the following fact: once the V-INFL raises above INFL into  $C^0$ , something forces the clitic subject to cliticize to  $C^0$ . I think this can follow from the clitic's need for a host, but there is no obvious reason why the clitic does not *left*-adjoin to  $C^0$  instead of right adjoining. Whatever the account is, we could take it to have a rather pervasive nature. The situation, then, would be:

- there is a designated landing site where clitics *left*-adjoin:  $\text{AGR}^0$  for both object clitics in Romance and subject clitics in French and Northern Italian dialects.
- whenever the content of this head moves a head up, the clitic is left without an

appropriate host and it has to cliticize into the new host position; in this case (for whatever reason) the clitic right-adjoins to the upper host.

Now, if French (complex) inversion involves longer V-movement than in assertive clauses, and this longer movement is not a minimal further step as in Kayne's (1991) account for enclisis, but rather a one-head-up-more step, is there evidence that Romance enclisis also involves further head-to-head movement? I think there is some evidence, at least in Catalan and Spanish (I will be presenting the Catalan examples; Spanish equivalents behave identically as far as I know).

One way of measuring the length of the further-step movement in enclisis constructions is to consider word-order phenomena. If enclisis involves V-INFL moving one more head up w.r.t. proclisis, then this extra movement will skip over possible specifiers or adjuncts of the maximal projection in between:

$$(33) \quad [X^0 V-AGR_i] [AGRP \textit{adjunct} [AGRP \textit{Spec} [AGR' t_i]]]$$

^-----'

so that the order will be adjunct-V or Spec-V in proclisis constructions and V-adjunct or V-Spec in enclisis constructions.

I think there is evidence of precisely this kind. In Catalan and Spanish, there are several kinds of adverbs (Catalan *sempre* 'always', *mai* 'never', *ja* 'already'/'yet', *encara* 'still', etc.) which usually precede the verb in finite sentences.<sup>119</sup> In fact, when they precede the verb, they must be adjacent to it:

---

<sup>119</sup> They can also follow the verb with a slightly lower degree of naturalness. This does not affect the argument below.

- (34) Sempre/mai /ja /encara (\*en Joan) hi va  
 Always/never/already/still (the J.) there-goes  
 'Joan always/never/already/still goes there'

So these elements are likely to occupy the specifier of the head where the verb stands in finite sentences.<sup>120</sup> In fact, the generalization is that these adverbs can precede the verb in precisely the constructions having proclisis: finite sentences (except affirmative imperatives, but including negative imperatives), and cannot in the constructions having proclisis: infinitives, gerunds and affirmative imperatives. In the latter case, they obligatorily follow the verb:

- (35) a. infinitives: (\*ja) anar-hi (ja) avui  
 (\*already) go-there (alr.) today  
 'to already go there'
- b. gerunds: (\*ja) anant-hi (ja) avui  
 (\*already) going-there (alr.) today  
 'already going there'

---

<sup>120</sup> Remember that in NSL Spec of AGR is not necessarily filled by the subject DP, so there is no problem if we assume this specifier is Spec of AGR. **Ja** 'already/yet' and **encara** 'still/yet' can precede the other adverbs in preverbal subject, giving marginal results:

(i) ?Ja sempre ve a classe  
 Already always comes to class

Possibly these two adverbs can be adjuncts to AGRP in addition to being specifiers.

In fact, we will assume that Spec of AGR is filled by PRO in control structures, but the facts in 0 are independent of control (they show up in non-controlled infinitives and imperatives).

- c. aff. imperatives: (\*ja) ves-hi (ja) avui!  
 (\*already) go-there (alr.) today  
 'do go there right today'
- d. neg. imperatives: (encara) no hi vagis (encara)!  
 (yet) not there-go (yet)  
 'Don't go there yet'

The acceptability judgements are clear cut. Especially significant is, I think, the case of imperatives: a minimal contrast such as affirmation/negation involves a change from enclisis to proclisis, and correlatively, a sharp change in the possibility for those adverbs to precede the verb. A natural and simple explanation is that, since enclisis involves an extra step in head-to-head movement, the adverbs in the specifier of the lower head shift in word order w.r.t. the verb once the verb moves a head up. This account is not possible within Kayne's hypothesis, where the extra movement is adjunction to next I', since this does not predict any word order change w.r.t. specifiers or adjuncts.

I will not develop in depth the question of what the upper head in this enclisis account is. In the imperatives it could be  $C^0$ : as evidence for this we have the fact that affirmative imperatives strictly forbid COMP material preceding them, whereas negative imperatives do not:

- (36) a. (\*\*Que) ves-hi!  
 (That) go-there!  
 'Do go there!
- b. (Que) no hi vagis!  
 (That) not there-go!  
 Don't you go there!

In infinitives and gerunds the upper head for extra-movement could be respectively a

'nominal' and 'adverbial' FC head. At least for infinitives, it cannot be COMP, because COMP can be filled by the 'if' particle as we have seen. Let us assume it is a FC intervening between CP and AGRP. We will call it NOM.<sup>121</sup> This head is probably not available in languages without enclisis. This can be a good explanation for the strongly nominal character of infinitives in languages like Italian, Catalan and Spanish, where it is used with a preceding definite article in some constructions (a. examples), and, besides can be readily used as a nominalization (b. examples).<sup>122</sup>

(37) Catalan:

- a. A l'arribar  
At the-to-arrive  
'In arriving'
- b. El desvetllar-se de la natura  
The to-awake of the nature  
'The wakening of the nature'

---

<sup>121</sup> Kayne (1991) suggests that there is an INFN ('infinitival') functional category which is specific to infinitives. I agree on this point in proposing NOM. What is odd about Kayne's proposal is that INFN is rather at the bottom of the clausal FC-hierarchy. If this extra FC is to be responsible for the categorial specificity of infinitives (their nominal behavior, for instance), it is more natural that it is the shell containing the other, sentence specific, FCs.

<sup>122</sup> Actually, even Sardinian, which has proclisis in infinitives, hence, in our terms, no movement of AGR° to NOM°, readily allows for infinitives as nominalizations. Perhaps Sardinian really has this nominal head but does not use it. After all, the nominal (or adverbial) character of English gerunds does not imply long verb movement. So the existence of an extra nominal FC would be a necessary but not sufficient condition for verb extra-movement (amounting to enclisis) in infinitives.

Spanish:

- a. Al llegar  
At-the to-arrive  
'In arriving'
- b. El cantar de los pajaritos  
The to-sing of the birdies  
'The singing of the birdies'

Italian:

- a. L'aver (lui) affermato questo  
The-to-have (he) stated this  
'(His) having stated this'
- a. Lo svegliarsi della natura  
The to-wake of-the nature  
'The wakening of the nature'

Raposo (1987-a) argues, that Romance Infinitives' nominal properties account for the apparent fact that they have to be Case-marked.

Even if, as I said, I cannot explain why enclisis and not proclisis takes place when there is extra head movement, at least we can generalize the phenomenon to the case of French (and Northern Italian) (complex) inversion, both being cases of extra head-movement which force the proclitic of the lower head to become an enclitic of the upper head. Perhaps the reason of the change of directionality of cliticization is a matter of diachronic change, having to do with the 'easiest' way a new generation can reanalyse the parents data when reanalysis takes place: for instance Old French subject inversion in interrogatives had nothing to do with cliticization even when the subject was a pronominal; when cliticization started to be active, the easiest way for the new generation to accommodate facts was to assume right-adjunction of the subject clitic to C<sup>0</sup>. Perhaps similar accounts could be given for Romance clitics, assuming their original position



was less fixed (subject to the Tobler-Mussafia law -see Benacchio & Renzi (1987)). I will not elaborate on the matter.

Since the enclisis/proclisis facts are not predicted from a synchronic point of view, Portuguese is not necessarily a counter-example: since in this language clitics have had a development which is rather different from the one in other Romance languages (there is nothing similar, in the latter, to Portuguese clitic-order alternations in main clauses being sensitive to negation and preverbal quantification), it might simply be that extra-movement to NOM<sup>o</sup> does not necessarily correlate with enclisis (maybe because of a different reanalysis process in the history of Portuguese). This conclusion is too loose to be of great interest. I think that the co-occurrence of proclisis and enclisis could be the key to a more accurate analysis of the facts. It might also be that enclisis is the core option.<sup>123</sup> I leave the issue here.

In the next section we will see how the above assumptions interact with the PRO-theorem.

Interrogative infinitives in Romance NSLs present a problem which also appears in finite embedded interrogatives. Consider the following finite paradigm from Catalan (which is representative of all Romance NSLs as far as I know):

---

<sup>123</sup> Joana Louro (p.c.) pointed out to me that enclisis in standard Portuguese requires some morpho-phonological alternations which are not used in the colloquial version of enclisis and that are rather annoying. So it might be that:

- enclisis is the unmarked option, but proclisis is also allowed as a marked variant.
- to speak Portuguese 'correctly', you have to learn cumbersome rules affecting enclisis.
- so speakers shift to proclisis to avoid both cumbersome school grammar rules and speaking 'incorrectly'.

- (38) a. Crec (\*en Joan) que (en Joan) el veuré  
 I-think (the J.) that (the J.) him-will-I-see  
 'I think that Joan I'll see'
- b. No sé (en Joan) quan (\*en Joan) el veuré  
 Not know-I (the J.) when (the J.) him-will-I-see  
 'Joan, I don't know when I'll see (him)'
- c. No sé (en Joan) si (en Joan) el veuré  
 Not know-I (the J.) if (the J.) him-will-I-see  
 'Joan, I don't know if I'll see (him)'

0.a) shows that dislocated elements and clause adjunct adverbs cannot precede the complementizer *que* 'that', but only follow it. So they would be IP adjuncts. 0.b) shows that these elements cannot follow a Wh-word (i.e., cannot intervene between the Wh-phrase and the verb), which could suggest that there is V-INFL to COMP movement (as in English main interrogatives), but the fact that now these elements can precede the Wh-phrase seems to suggest that it is the Wh-phrase which is lower (in some IP-specifier), not the V-INFL that raises. To make things worse, 0.c) shows that 'if' interrogatives allow the adjunct to both precede or follow the 'if' particle. In Chapter 5 we will provide some way of explanation for the facts.

What is to be noticed now is that infinitive interrogatives display a rather similar paradigm:

- (39) a. No sé (en Joan) quan (\*en Joan) visitar-lo  
 Not know-I (the J.) when (the J.) to-visit-him  
 'Joan, I don't know when to visit (him)'
- b. No sé (en Joan) si (??en Joan) visitar-lo demà  
 Not know-I (the J.) if (the J.) to-visit-him tomorrow  
 'Joan, I don't know if to visit (him) tomorrow'

To the extent these data cast doubts on the idea that 'if' is a  $C^0$  particle, they are a potential problem. I think, however, that these facts do not seriously challenge our analysis: It may well be that the assumption that Wh-elements and interrogative 'if' in Romance NSLs are in the same functional category as the one hosting *que* 'that' is too simple, and there is an intermediate projection. So provided 'if' is in the head of the FC immediately dominating  $NOM^0$  (or  $AGR^0$ ), our account is tenable, whether this FC is CP or not.

### 2.1.2. The PRO Theorem

Before going into the discussion of the PRO-theorem effects, let us say something about what a theory on PRO and control has to minimally specify. We have argued that control cannot be reduced to Binding Theory in a positive way (i.e. with a theory predicting the reference of PRO as a case of BT, as is intended in Kayne (1991)). We will adhere, however, to the more traditional view that PRO's distribution is determined by BT in a negative way: PRO has to escape binding requirements by being ungoverned (or not having Case, see below).

Since control is, under this view, still a mysterious module of the grammar, nothing of great interest can be said about it. What I want to suggest, however, is that control can be made minimally an interesting phenomenon if we relate it to another phenomenon which is apparently akin in nature: subjunctive obviation.

For both control and obviation there have been authors trying to derive either from Binding Theory (Kayne (1991) and Picallo (1985), resp.). Both phenomena, however, are reluctant to such accounts. To mention a major problem, dislocation or movement of the infinitive/subjunctive does not have any effect on the control/obviation facts, contrary to what happens with genuine BT facts. Recall our discussion with examples such as 0 and 0 (repeated here as 0 and 0). Similar examples can be built for obviation (consider 0):

- (40) a. \**Each other's* accusation triggered many problems for *them*  
 b. *PRO* to accuse Nixon triggered many problems for *them*

(41) Votar *en Joan*, em sembla que ni *ell mateix*

To-vote-for the J. me-seems that not-even he SELF

s'imagina que t'ho proposessis

guesses that you-intended

'Vote for *Joan*, I think not even *he himself* could guess you intended to'

(42) \*Que votis *en Joan*, em sembla que ni

That you-vote-SUBJ the J. me-seems that not-even

*ell mateix* s'imagina que t'ho proposis

he SELF guesses that you-intend

'Vote for *Joan*, I think not even *he himself* could guess you intend to'

In 0 the dislocated subjunctive *que votis en Joan* is subject to the obviation constraint w.r.t. the subject of the embedded clause (this is the reason for the ill-formedness of the structure), in spite of the fact that, were it not for obviation, the example would be fine and it would involve crucial non-c-commanding of *en Joan* by *ell mateix* 'he himself'.

I will leave the question here, only suggesting that control and obviation are closely related phenomena which could constitute an autonomous module involving some sort of obligatory coindexation/anti-coindexation without recourse to c-command.

Since we assume that control is outside Binding Theory, the traditional assumption that PRO is an element having to escape BT seems reasonable.

Now let us try to account for Kayne's (1991) generalization that enclisis is a necessary (perhaps sufficient) condition for having 'if-infinitives. Let us begin with languages not having

enclisis (English, French, Occitan, Sardinian). Let us simply assume that in those languages a controlled infinitive has the following essential structure:

$$(43) \quad [_{CP} C^0 [_{AGRP} PRO [_{AGR'} AGR^0 ] ] ]$$

PRO is not governed by  $AGR^0$ : in English, French and Occitan, simply because  $AGR^0$  is not filled and therefore is not intrinsically able to govern. For any language, however, there is a major reason:  $AGR^0$  is not structurally able to govern PRO, because, we assumed, heads do not govern specifiers. If, however,  $C^0$  is filled with an intrinsically possible governor (such as 'if'), then PRO is governed and there is a conflict with the PRO-theorem. So we adhere to Kayne's (1991) initial idea that the impossibility of 'if'-infinitives in some languages is due to an illicitly governed PRO.

In languages having enclisis, however, the structure is the following (where NOM is the nominal head of infinitives, as we proposed above):

$$(44) \quad [_{NOM^0} AGR^0_i ] [_{AGRP} [_{PRO} PRO^0 ] [_{AGR'} t_i ] ]$$

where  $AGR^0$  has moved to the next head.

We assumed that in this structure  $PRO^0$  is able to adjoin to  $NOM^0$ . Therefore:

a) Since  $AGR^0$  is merged in the same head as PRO, it becomes rich in (control-provided) features. Thus it is the AGR-identifier, and can assign Nominative directly to the I-subject, in accordance with the principles we postulated in Chapter 3:

(45)  $AGR^0$  can optionally assign Nominative Case to Spec of AGR by agreement or to some other position under government.

(46) The I-subject must receive Case from its AGR-identifier.

b) Hence (headless) PRO in Spec of AGR will not have Case, and *its Binding domain will not be definable*, according to our definition, which makes reference to a Case position. Therefore it is immaterial whether it is governed or not as far as the PRO-theorem is concerned (in 0 it would be governed by  $[_{NOM}^o AGR^o]$ ). In the present account, then, PRO can escape BT either by not being governed or by not having a Case position.

c)  $PRO^o$ , on the other hand, once adjoined to a head, is not in an A-position and is not subject to BT.<sup>124</sup>

d) (headless) PRO in Spec of AGR is coindexed with  $AGR^o$  (by Spec-head agreement) and  $AGR^o$  is coindexed with the I-subject (by the very same reason it is in finite sentences). Then PRO A-binds the I-subject, with the result that the latter cannot be an R-expression (in the following example, the infinitive is dislocated, so that the reason for the BT-3rd Principle violation cannot be binding by the main subject):

- (47) a. Fer això ell/\*en Joan, no ho desitja pas  
 To-do this he/ the J. not it-wishes at-all  
 'To do this himself/\*John, he does not wish'
- b.  $[ \text{Fer } [_{IP} [_{PRO} t] [_{VP} [_{VP} \text{això} ] \text{ell}_i/*\text{en Joan}_i] ] ]$

The Case position for the I-subject is the I-subject itself, thus the Binding Domain for the I-subject excludes PRO, and the I-subject can be (and has to be) a pronominal.

---

<sup>124</sup> If this idea is correct, then it could have far-reaching consequences for other cases of head-movement of a pronominal (thanks to Carme Picallo for pointing this out to me): if, for instance, clitics are analyzed as DP heads adjoining to some FC, then, by the same logic of our reasoning about  $PRO^o$ , no Binding Principle would apply to clitics, contrary to fact. On the one hand, however, it is not clear that clitic attachment is the output of head movement from the A-position. On the other hand, the Binding Theoretical properties of clitics are an obscure area at least in the case of reflexives (reflexive clitics often become passivizer or unaccusativizer morphemes across languages).

An important question arises here. Since we crucially derive enclisis with PRO<sup>0</sup>-adjunction, and PRO<sup>0</sup>-adjunction ultimately forces the infinitive I-subject to be pronominal, languages not having enclisis in infinitives will not behave like NSLs in infinitival constructions, even if they are NSLs in finite sentences. Specifically, we would expect those languages to have pronominal I-subjects in finite clauses and anaphoric I-subjects in (controlled) infinitives.<sup>125</sup> This would be the case for Occitan and Sardinian. Is this prediction borne out?

Quite disappointingly, both Sardinian and Occitan are languages using pronominals *qua* (strong) reflexives in an exclusive way:<sup>126</sup>

(48) Occitan:

Jacme<sub>i</sub> parla d'el<sub>i/j</sub> (mateis)

J. talks of-he (SELF)

'Jacme talks about himself/him (himself)'

Sardinian:

Gavini<sub>i</sub> l'at comporatu pro isse<sub>i/j</sub> (matessi)

G. it-has bought for him (SELF)

'Gavini bought it for himself/him (himself)'

So we cannot obtain confirmation for our prediction that these languages have a contrasting AGR-identifier strategy for finite/infinitival sentences, although, happily enough for

---

<sup>125</sup> Recall we assume that the parallelism principle 0 holds only 'in the unmarked Case'. Occitan and Sardinian would thus be marked in this connection.

<sup>126</sup> The Sardinian example comes from Jones (1990), where he explicitly states that there is no strong pronoun/strong anaphor lexical contrast. The same is true for Occitan as far as I know: even if literary Occitan has tried to retrieve the Medieval strong reflexive form (**se** (**mateis**)), it is never used in colloquial speech.

our hypothesis, our prediction is not falsified either.

In fact, I think it is not an accident that languages not distinguishing strong pronouns/anaphors are the ones both being NSLs and having proclisis in infinitives: in the absence of evidence for the [ $\pm$ anaphoric] character of infinitival I-subjects, these languages have more unproblematically adopted proclisis, because it is more unlikely to conflict with the I-subject BT data of a possible earlier period where the parallelism finite/infinitival (i.e., the unmarked case) held. Or, the other way around, the fact that they are NSLs with proclisis has favored the adoption of neutral pronoun/anaphor elements in order to avoid pronominal/anaphor switch in finite/infinitive clauses: if the learner does not observe the switch in the parents' data, it would be easier for him/her to reinterpret pronominals as neutral pronominal/anaphoric elements than to adopt the switch practice. The idea would be that the learner undergoing diachronic change tries to build grammars that are the least conflicting possible with the parents data.

This idea extends to old French, which also was a NSL with proclisis: in fact, even modern French has neutral pronominal/anaphor forms, surely a residue of its NSL period. Assuming that the existence of neutral anaphoric/pronominal elements is a marked option, the present theory provides a possible trigger for its existence in these languages: optimal data reinterpretation in diachronic change.

Summing up, we have provided an account for infinitival constructions that consists of the following assumptions:

- in the unmarked case, infinitives have the same AGR-identifier option as finite clauses for a given language.
- given that PRO is always the key element in control, and it is a maximal projection, it has only two options: either it is ungoverned (as required by its extraordinary status), or its head incorporates to the head containing AGR.
- the latter option requires the V-AGR head to move to an upper head in order to make the landing site for PRO<sup>0</sup>-movement a position governing the trace; and it allows AGR<sup>0</sup> to



become the AGR-identifier, a favored option in NSLs.

- long V-movement to  $AGR^0$  (and possibly  $NOM^0$ ) is only allowed as far as  $AGR^0$  ends up being contentful; it is in addition required if the AGR-identifier is  $AGR^0$  and has to Chain-govern the I-subject to assign it Case.

## 2.2. Raising

In Chapter 3, we crucially assumed that raising constructions are characterized as involving a non-CP infinitive:<sup>127</sup> this is why AGR in the upper clause is coindexed with an I-subject internal to the infinitive, this I-subject being the first DP/CP it c-commands. Suppose, nevertheless, that the infinitival AGR has to abide by one of the options concerning the AGR-identifier: either it is  $AGR^0$  or Spec of AGR. We provisionally assume the following parallelism principle:<sup>128</sup>

- (49) In the unmarked case, raising infinitives have the same AGR-identifier option as finite sentences in a given language.

Consider the D-structures in 0, taking English and Italian as representative languages:

- (50) a.  $AGR^0$  seem [<sub>AGRP</sub>  $AGR^0$  to have come DP ]  
b.  $AGR^0$  sembra [<sub>AGRP</sub>  $AGR^0$  esser venuto DP ]

---

<sup>127</sup> Recall we suggested that only verbs with an epistemic meaning are likely candidates to be raising verbs, at this could be the basis for accounting for their exceptional non-CP character or their complements: they would form a (semantically) mono-clausal structure with their complement.

<sup>128</sup> This principle has an obvious parallelism with 0 (which concerns control, see page 212). We will reduce them to a single principle. For the moment we distinguish them for convenience.

In both cases, both the main clause AGR and the embedded AGR (if there is one) are coindexed with the DP in the embedded clause, since it is the first DP or CP in the c-command domain of both. This means that the two AGR's end up coindexed.

For English and non-NSLs, both Spec's of AGR have to be filled by a DP to render both specifiers AGR-identifiers. This is the unmarked option in 0 and, in fact, the only one for the infinitival AGR<sup>0</sup>, given that it is not rich to be an AGR-identifier itself.

The infinitival AGR<sup>0</sup> cannot assign Case to its specifier: we assumed that infinitival AGR<sup>0</sup> can only assign Case to PRO, and PRO cannot occur in this position because it is governed by the main verb, as standardly assumed. Since we are trying to derive Burzio's Generalization, we assume that 'to seem' verbs can in principle assign Accusative. So, in principle the Spec of the infinitival AGR could be assigned Accusative. However, since principles 0 and 0 force the main AGR-identifier (Spec of AGR) to transmit its Case to the I-subject, the result would be that the I-subject would end up obtaining both Accusative Case from the verb 'seem' and (Nominative) Case from the main clause Spec of AGR.

Therefore, the only option for the infinitival Spec of AGR is to be a non-Case-marked empty category and form a Chain with the Specifier of the main clause AGR, from which it will obtain Case.

Given the definition of Binding Domain, the main clause will become the Binding Domain for the I-subject in the infinitive, so that both the I-subject in the infinitive and the infinitival Spec of AGR can only be [+anaphoric]:

- (51) a. John seems [<sub>AGRP</sub> to have done it [<sub>DP</sub> *e* (himself) ]]

As for the Spec of AGR, its anaphoric behavior cannot be instantiated by the presence of an overt anaphor or floating *himself*:

- (52) John seems [<sub>AGRP</sub> *e*<sup>\*</sup> himself to have done it ]

This fact should be due to some restrictions on the distribution of emphatic elements. I'm not able to say anything interesting about this issue, except that emphatic subjects seem to be restricted to some positions.

Now consider Italian. In this case, it is  $AGR^0$  which has to become an AGR-identifier, if the unmarked option is taken. For the infinitival  $AGR^0$  to become the AGR-identifier, it has to be rich. As before, in raising constructions the infinitival AGR and the main AGR end up coindexed with the same I-subject and, therefore, coindexed themselves. Suppose that, since the main AGR is rich, the infinitival  $AGR^0$  may inherit phi-features from it and become rich itself.<sup>129</sup> If so, it will be the AGR-identifier and will have to assign, by the combination of principles 0 and 0, Case to its I-subject by Chain-government, provided the infinitival verb raises.

Since our Parallelism Principle 0 induces, for NSLs like Italian or Catalan, the main  $AGR^0$  to assign Case to its I-subject, we have to assume that, in raising constructions in these NSLs, the main and the embedded AGR are reanalyzed as an 'extended' Government-Chain, consisting of the union of the two head Chains formed by V-movement up to AGR in both the main and the infinitival clause. The foot of one Chain ( $V^0$ 's trace of 'seem') governs the head of the other, and in addition the two  $AGR^0$ 's are coindexed. So Chain extension is a very natural

---

<sup>129</sup> We crucially assume that the main AGR is rich and its coindexation with the embedded AGR makes the latter rich. When the richness of the main  $AGR^0$  is only obtained after  $PRO^0$  incorporation (see section 2.1.2. above), we might wonder whether, since incorporation is an S-structure phenomenon, it comes 'too late'. In fact, 'seem' verbs have a degraded status in infinitival constructions in Catalan (and possibly other NSLs) :

- (i) ??Semblar estar borratxo no t'ajudarà  
To-seem to-be drunk not will-help-you
- (ii) ??Vull semblar estar borratxo  
I-want to-seem to-be drunk

I cannot pursue this idea here.

Coindexation between main and embedded AGR takes place in subject controlled constructions as well, so that in this case we predict that both  $PRO^0$  incorporation and AGR coindexation converge to the same effect.

device in this case.

In fact, Torrego (1989) argues that in Spanish the infinitival verbal form incorporates into the main verbal form in raising constructions.<sup>130</sup> If so, no device of Chain extension would have to be stipulated. Or, reversing the argument, in NSLs incorporation would be triggered by the need for Chain-extension.

Given our BT definitions, the Binding Domain for the I-subject will be the infinitival TP, and thus the I-subject will be [-anaphoric]:

(53) a. Sembra [ averlo fatto *pro*/lui/Gianni ]

Seems to-have-it done *pro*/he /G.

'He/Gianni seems to have done it'

b. Chi<sub>i</sub> sembra [ averlo fatto *t<sub>i</sub>* ]

Who seems to-have-it done

'Who seems to have done it?'

In the above examples, one could argue, it is not clear that the inverted subject is inside the embedded clause: it could be as well right adjoined to the main VP. There is, however clear evidence that this is not necessarily so (see 0) and even not possibly so (see 0) (examples from Spanish):

---

<sup>130</sup> She claims that this incorporation takes place only at LF, on the basis of examples where some adverbs intervene between the two verbs:

(i) Parecía **ayer** haber muchos mosquitos  
Seemed yesterday there-to-be many mosquitoes  
'Yesterday there seemed to be many mosquitoes'

To my ear (for Catalan and my Spanish), the presence of an intervening adverb sharply degrades the acceptability of the sentence. so incorporation would be an S-structure phenomenon in Catalan.

- (54) a. El libro, parece [IP haberlo comprado Juan en Londres ]  
 The book seems to-have-it bought J. in London  
 'That book seems to have been bought by John in London'
- b. ??El libro, parece haberlo comprado en Londres Juan  
 The book seems to-have-it bought in London J.
- (55) a. Parece [IP haberlo escrito Juan ] por la letra<sup>131</sup>  
 Seems to-have-it written J. by the handwriting  
 'It seems to have been written by John to judge by the handwriting'
- b. \*Parece haberlo escrito por la letra Juan  
 Seems to-have-it written by the hndwrng. J.

What 0 shows, is a particular instance of a more general fact: constituents of the embedded clause cannot be extraposed to the main clause:

- (56) a. Ha dit [CP que donaria el llibre als nois ] avui  
 Has said that he'd-give the book to-the boys today  
 'She said that s/he would give the book to the boys today (S/he said it today)'
- b. \*\*Ha dit [CP que donaria el llibre ] avui als nois  
 Has said that he'd-give the book today to-the boys

0.b) and 0.b) do not improve at all if the extraposed element (subject/dative) is heavy, (provided *avui* 'today' is actually interpreted as specifying the 'saying' time interval).

It is a trivial matter that the parallelism principles in 0 (for control infinitives) and 0 (for

---

<sup>131</sup> The constituent **por la letra** should be read as non-dislocated, the interpretation being then, roughly: 'It is the handwriting that persuades (me) that it has been written by Joan'

raising) can be reduced to a single principle:

- (57) In the unmarked case, infinitives have the same AGR-identifier option as finite sentences in a given language.

We assume that, at least for control, this is only an unmarked case option for there would be languages (Occitan and Sardinian) taking the marked option. Now, if we have a single principle, the prediction would be that if Occitan and Sardinian take the marked option in control, they take the marked option in raising infinitives too, namely, the AGR-identifier would be the Spec of AGR in raising infinitives and AGR<sup>0</sup> in finite clauses. We argued that Spec of AGR in a raising infinitive cannot be PRO because it would be governed by the 'seem' verb. What other options are there left for Occitan or Sardinian?

If it were an empty anaphor, as is the case in English, it could in principle, like in English, form a Chain with the main clause Spec of AGR. But the main Spec of AGR does not receive Nominative Case, because in finite clauses AGR<sup>0</sup> assigns case only through Chain-government, as we argued in Chapter 3. Therefore an empty anaphor is excluded. (Remember preverbal subjects in NSLs enter a resumptive pronoun strategy with the I-subject, and elements resumed by an expletive are not assigned Case themselves).

Since the Infinitival AGR<sup>0</sup> is rich by being coindexed with a main AGR<sup>0</sup>, it could conceivably assign Case to its Specifier by agreement. But since the main AGR-identifier (AGR<sup>0</sup>) has to assign case by forming a Government-Chain, there would be conflicting requirements on the way Case is assigned: a Government-Chain can only assign Case by government, and its foot cannot be an agreement-Case-assigner. If this (perhaps dubious) assumption is on the right track, it would leave Occitan and Sardinian with no option for the AGR-identifier (Spec of AGR) of raising infinitives. In fact, Sardinian does not have infinitival raising.<sup>132</sup>

---

<sup>132</sup> Jones' (1990) description of Sardinian, which is rather cautious in excluding non-genuine constructions (Italianisms) is categorical in this connection.

As for Occitan, it does apparently allow raising (example from Sauzet (1989)):

- (58) Lo paire semblava la me voler donar  
The father seemed it-me-to-want to-give  
'The father seemed to want to give it to me'

However Sauzet presents evidence showing that Occitan has what he dubs 'pseudo-raising': cases where there is apparent raising out of a finite clause:

- (59) Lo paire semblava que la me voliá donar  
The father seemed that it-me-wanted-he to-give  
'The father seemed to want to give it to me'

I will not go into the details of his analysis: what is essential is that pseudo-raising is an exceptional construction (at least within the Principles and parameters framework): Sauzet assumes that in 0 what 'raises' is a topic element. Therefore, we could make the following argument:

- Occitan does not allow raising constructions for the reason we mentioned above.
- for whatever reason (perhaps strong interference with French) it has acquired a construction (namely 0) which looks like raising. But it is not subject raising, but rather 'topic raising' (pseudo-raising).
- once the speaker accepts pseudo-raising with infinitives 0), the construction generalizes to finite sentences, which could not be predicted by a theory assuming that 0 is simply a standard case of raising.

So our claim that Occitan or Sardinian cannot have (standard) raising makes some sense: it predicts that either the construction is absent (Sardinian) or that, if apparently present, it is not standard raising (and there is independent motivation for pseudo-raising with finite clauses).

Since pseudo-raising is likely to be a marked option, we could exploit the idea that the trigger is a historical one: French pressure (which is very strong in Occitan areas) has forced a raising-looking construction into the language and speakers have interpreted it as pseudo-raising and then generalized it to cases where French does not allow raising (namely 0).

In fact, 'pseudo-raising' is a widespread phenomenon (a fact that could cast serious doubts on the accuracy and cross-linguistic significance of our theories for 'standard' raising). Basque is another language which has some sort of pseudo-raising:

- (60) a. Haiek dirudite hauteskundeak galdu dituztela  
           They seem election-the lost have-that  
           Lit.: 'They seem that (they) have lost the election'
- b. Bush-ek dirudi Perot-ek hauteskundeak irabaziko ditu  
           Bush seems Perot election-the to-win has  
           Lit.: 'Bush seems that Perot will win the election'

0.b) clearly shows we are not dealing with raising, since the main clause subject does not even bind an Argument inside the embedded clause. In fact, our characterization (in Chapter 3) of Basque as having a designated position for AGR-coindexing would make 'true' raising impossible in this language, for it is essential to raising that the main AGR may be able to coindex with whatever DP is the most prominent one in its c-command domain.

In Chapter 3 we assumed that our principle of AGR-coindexation could be too powerful in that they would exclude super-raising without any need for the ECP (see the section 1.4 in Chapter 3). Then we considered the possibility of reducing the power of our principle of AGR-coindexation. One way or another, super-raising is not a problem in the present theory.

Our account for raising in NSLs having proclisis does not involve the formation of an A-



Chain, so there is no trace in the Spec of the infinitival AGR. Rizzi (1982-a) (cited in Burzio (1986:206) adduces evidence in favor of the existence of such a trace. Since a trace has to be (properly) governed, the infinitive cannot be moved, dislocated or, in general, appear outside the governing domain of the 'seem' verb, for then the 'seem' verb would not govern the trace (I adapt the Italian data to Catalan):

(61) \*És [ *t* estar cansat], que en Joan sembla.

Is to-be tired that the J. seems

'\*It's to be tired, that Joan seems'

Cf. with:

És [ *PRO* estar cansat], que en Joan tem

Is to-be tired that the J. fears

'It's being tired that Joan fears'

In our account, if there is no trace, what is the cause for the ungrammaticality of 0? If our suggestion is correct that an extended Government-Chain is formed by the union of the V-movement Chains of the main and infinitival clauses, then it is reasonable that the foot of one (sub-)Chain has to govern the head of the other. We even suggested that, in Spanish or Catalan, the lower verb incorporates into the main verb: that would be an even stronger reason for the ill-formedness of this construction. And even we could suggest that an AGR-identifier has to c-command its I-subject at all levels. Notice that in control structures the main clause AGR is not the AGR-identifier of the embedded I-subject, the control relation being of another nature.

### 2.3. ECM and 'for' Infinitives

In these constructions, the infinitival Spec of AGR receives Case from the head governing it, according to the standard analysis:

- (62) a. I ... [ $V^0$  *believe*] [ $IP$  him [ $I'$  to be there ] ]  
 b. [ $C^0$  *for*] [ $IP$  him [ $I'$  to be there ] ] ...

ECM constructions are exactly like raising constructions except that the main AGR is not coindexed with an I-subject inside the infinitive, because there is a preferred candidate in the main clause (e.g., the Experiencer Argument of *believe*). In both, the infinitival AGR is poor and cannot become rich by any means. Therefore, only Spec of AGR can possibly be the AGR-identifier. Since the infinitival  $AGR^0$  is not able to Case mark (it can in principle only Case-mark PRO, which is excluded by its being governed by the main verb or 'for'), Spec of AGR has to receive Case from that upper governor.

The present theory predicts that ECM and 'for' infinitives should not be possible for NSLs taking the unmarked option by which infinitives have the same AGR identifier as finite sentences: the infinitival  $AGR^0$  would have to be the AGR-identifier, which it could not, being irreparably poor. Case-assignment to the Spec of the infinitival AGR would be useless, this position not being the AGR-identifier.

The claim that NSLs taking the unmarked option (for infinitival AGR-identifiers) have no ECM might have a counterexample in classical Latin.<sup>133</sup> In this language, however, infinitives cum Accusative are not clear cases of ECM, as far as the presence of the Accusative appears to be quite independent of the main verb's lexical characteristics and Case properties. So we have (examples from Maraldi (1983)):

---

<sup>133</sup> If it is the case that classical Latin was a language taking the unmarked option for infinitives, of which I have no evidence.

- (63) a. Dicitur eos venisse  
Said-is them to-have-come
- b. Manifestum est eum abisse  
Evident is him to-have-gone

In standard accounts, we do not expect a passive verb or a copula to assign Accusative. In the present theory any verb can assign Accusative, but:

- either the infinitival is a CP (and therefore becomes the I-subject) and the main verb cannot govern the infinitival subject in Spec of IP,
- or the infinitival is an IP and the infinitival subject becomes the main clause I-subject, and then it will be Nominative.

In whatever theory, it seems that Latin infinitives cum Accusative are not ECM, but some other construction. I have nothing to say about the issue.

Raposo (1987-b) proposes that some verbs take an IP complement in Portuguese, but, if this is correct, the resulting structure does not allow ECM (it would be one of the types of inflected infinitive): in our terms, because, as we argued, assignment of Case to Spec of AGR is pointless in such a language. Thus our conclusion that NSLs taking the unmarked option for infinitives cannot have ECM is not challenged as far as I know.

Since, we assumed, Occitan and Sardinian take the marked option of having infinitival Spec of AGR as the AGR-identifier, we predict ECM is possible in principle for these languages, which is not apparently the Case<sup>134</sup>. In fact, nothing forces a language to have ECM (or raising):

---

<sup>134</sup> Sardinian allows infinitival complements without control, as we will see in section 2.4., but the overt subject is Nominative and has the distribution of an inverted subject

As for Occitan, it does have apparent cases of ECM (see Sauzet (1989)):

(i) Pensava las vacas manjar son sadol  
He-believed the cows to-eat their fill

Sauzet argues that these cases are to be analyzed as involving a PRO controlled by a topic adjoined to CP, since,

they are anyway marked options. I do not know what might trigger (if anything other than positive evidence during acquisition) the presence of ECM or raising in a language.

French is another language not having ECM in general, except for the fact that the Wh-extracted version of an ECM construction is allowed (see Kayne (1981), Rizzi (1982-b)):

- (64) Le garçon que je croyais être intelligent  
 The boy that I believed to-be intelligent

Even Italian has this construction:

- (65) Il ragazzo che ritenevo essere intelligente  
 The boy that I-believed to-be intelligent

Verbs allowing this construction are epistemic verbs and verbs of saying (see Rizzi (1982-b:78). Since these verbs are, unlike English *believe*, control verbs in the general case, these exceptional constructions seem to require some analysis expressing their exceptional status, which does not necessarily threaten our claim that Italian cannot basically have ECM. What is essential here is that Spec of IP is not Case-marked,<sup>135</sup> for it is not an AGR-identifier and,

---

unlike in English, the absence the DP **las vacas** gives control by the subject 'he', and, on the other hand, the DP (**las vacas**) can be adjoined to a finite CP:

- (ii) Pensava las vacas que manjavan son sadol  
 He-believed the cows that they-ate their fill  
 in which case the topic binds a **pro**. In any event, the infinitival (apparent) subject is not Accusative.

<sup>135</sup> It is apparently Case-marked in AUX-to-COMP infinitives, but this is an exceptional construction (see section 2.4.). As Rizzi (1982-b) points out, even if there is a correlation between the verbs allowing 0 and the ones allowing AUX-to-COMP, the latter construction is significantly more marked (high literary speech-level). In section 2.4.2. we will challenge the view that AUX-to-COMP involves Case-marking of the Specifier of AGR.

therefore, not a Case-position. The proposal in Kayne (1981) and Rizzi (1982-b) that in 0 and 0 the main verb Case-marks the trace of the Wh-phrase in COMP does not contradict this claim.

French, even if allowed to have ECM, does not happen to. I leave the question here.

## 2.4. Infinitives with an Overt Subject

What I will call infinitives with overt subjects (IOSs) should be clarified: we have seen that both control and raising infinitives do have overt I-subjects of a [ $\pm$ anaphoric] restricted nature (depending on the (non-)NSL status) and, for NSLs, of a [ $\pm$ pronominal] nature (depending on the whether there is control or raising). I will use the term IOS to name only those infinitives that allow an overt subject in a way not predictable from control or raising. Romance languages show a variety of them:

- AUX-to-COMP infinitives in Italian and Portuguese:<sup>136</sup>

(66) Ritengo non esser lui in grado di farti niente

I-believe not to-be he in a position to do-you nothing

'I believe him not to be able to do anything to you'

(67) Penso terem os deputados trabalhado um pouquinho

I-think to-have-3pl the deputies worked a little-bit

'I think the MPs finally worked just a little bit'

- Non-subcategorized prepositional infinitives in Catalan and Spanish (examples from Catalan):

---

<sup>136</sup> See Rizzi (1982-b)/(1986) and Raposo (1987-b) .

- (68) a. Aquesta habitació és per jugar-hi els nens  
 This room is for to-play-there the children  
 'This room is for the children to play in it'
- b. A l'arribar la Maria, vaig fer-li el dinar  
 At the to-arrive the M. I cooked-her the meal  
 'Upon Maria's arrival, I cooked a meal for her'

- Dislocated infinitives (in colloquial Catalan and Spanish):

- (69) Anar-hi en Joan, no em sembla pas la solució  
 To-go-there the J. not me-seems at-all the solution  
 'John going there, I don't think it's the solution'

Interestingly, Occitan has more or less the same possibilities, but with preverbal subjects:

- (70) a. Aquesta cambra es per los dròlles jugar  
 This room is for the children to play
- b. En Joan arribar/arribant, farem lo trabalh  
 In J. to-come/coming we'll-do the job
- (71) a. Joan far tot lo trabalh, m'estonarià.  
 J. to-do all the work me-would-surprise

- Non-controlled complement infinitives in Sardinian:<sup>137</sup>

---

<sup>137</sup> See Jones (1990) .

- (72) Non keljo a vénnera tue  
 Not I-want to to-come you  
 'I don't want you to come'

The Sardinian case is specially interesting (and puzzling): what makes the infinitive in (72) a non-controlled infinitive is the presence of the preposition *a* 'to' before it. Without this preposition, the infinitive is obligatorily controlled (see (73.a)). When the non-controlled *a*-infinitive has no overt subject, its subject is interpreted as arbitrary (73.b):

- (73) a. Non keljo vénnera  
 Not I-want to-come  
 'I don't want to come'
- b. Non keljo a vénnera  
 Not I-want to to-come  
 'I don't want someone/anyone to come'

Several proposals, which I will not discuss here, have been provided to account for IOS. Rizzi (1982-b) deals with AUX-to-COMP. He proposes that an Aux in non-finite COMP acquires some capacity to assign Nominative. Raposo (1987-b) makes a much more restrictive claim (AGR has to be Case-marked itself to be a Case assigner). Although he makes several interesting predictions (he restricts AUX-to-COMP to NSLs, and predicts which contexts allow it), some of his crucial proposals on the precise nature of c-selection by verbs are far from obvious, and, in addition, his proposals on Portuguese are not easily extendable to AUX-to-COMP in Italian and even less, I think, to IOS in Romance. For IOS in Spanish, Fernández Lagunilla (1987) simply states that Nominative is assigned by default, an attractive idea only if we can provide means of excluding the default option for the cases when it is not possible, which she does not even attempt to do. Rigau (1992) accounts for the postverbal position of subjects in

Catalan and Spanish prepositional temporal IOSs (see 0.b)) not in terms of Case theory but in terms of tense interpretation (a weak  $T^0$  has to incorporate to the temporal preposition preceding these infinitives). Independently of this, Nominative is assigned -and *pro* is licenced- by an abstract AGR. Abstracting away from *pro*, this idea is much in the spirit of Reuland's (1983) proposal for English gerunds with an overt subject. Galves (1991) revises Raposo's (1987-b) proposal in an interesting way, but it is still a theory basically conceived for European Portuguese.

Some authors assume it is AGR which assigns Nominative in IOS (Raposo (1987-b), Galves (1991), Rigau (1992)); others (Fernández-Lagunilla (1987), Hernanz (1992), Delfitto (1990), Belletti (1991)) assume AGR does not take part in Nominative assignment in some non-finite sentences. Except for the case of Portuguese inflected infinitives, in which AGR is obviously present, I will propose that AGR is not present in most types of IOS, and I will contend that this is not at variance with what happens in finite clauses.

The proposals we will introduce here are highly speculative and far from precisely established. They will however lead to important qualifications to the theory sketched so far. Whether our speculations are on the right track or not, we cannot, I think, ignore the issue or treat it in a independent way from what we have assumed so far.

We will classify IOS's in two groups, assuming the distinction is of theoretical relevance: IOS's where, we will argue, the overt subject is an I-subject (the cases in Catalan, Spanish, Sardinian) and IOS's where the overt subject is in Spec of AGRP (AUX-to-COMP, Occitan).

#### **2.4.1. IOS with an Inverted Subject**

Let us assume that the IOS's where the subject follows the verb without necessary adjacency V-subject are cases of subject inversion, in our terms, cases involving an overt [-anaphoric] I-subject. In Catalan, where inverted subjects are VP-final (with the qualifications we introduced in Chapter 1), IOS's preferentially show VP-final subjects:



(74) Fer (???en Joan) la feina (en Joan), em sembla...

To-do the J. the work the J. me-seems

'That Joan does the work, well, it seems to me...'

In Spanish, instead, where the VSO word-order is a possible form of subject inversion in finite clauses, VSO word order in IOS is quite usual.<sup>138</sup>

Let us assume that the cases we have seen in 0, 0, 0 and 0.b) are cases of IOS with an inverted subject (*IOS-INV*). 0.b) shows another phenomenon: Sardinian IOS-INV allow a null subject with arbitrary interpretation. Catalan and Spanish IOS-INV show an even more surprising fact: the subject can be null and fully referential (see Rigau (1992)):

(75) En acostar-m'hi, em va mirar provocativament

In approach-me-there me-looked-at provocatively

'As I approach him/her, s/he looked at me provocatively'

The null subject in is not controlled and is fully referential. We will assume, with Rigau (1992), that this null subject is a *pro*. We will also assume that the arbitrary null subject in Sardinian IOS-INV in 0.b) is a *pro*.

There are at least two questions our theory should address:

---

<sup>138</sup> As we will see in the next chapter infinitives with a VSO word order are predicted to be possible in principle. In fact, this example is not as unacceptable as a parallel case with a finite clause. Still, the prediction is that Spanish allows VSO in IOS with no problem, while Catalan marginally allows it for other reasons. This is in accordance with the spirit of the Parallelism Principle (in whatever version, see below) we are proposing.

- (76) a. How is Nominative assigned to the I-subject in IOS-INV?  
b. What makes referential *pro* possible in IOS-INV?

Let us start with 0.a). In the present theory, the most natural assumption is that Nominative is assigned by government directly to the I-subject, in a way similar to Nominative assignment in finite clauses. This is a desirable prospect if we want to account for the fact that IOS's with an inverted subject are, as far as I know, restricted to NSLs, which have subject inversion in finite clauses.

In Chapter 3 we contended that direct Case assignment to the I-subject is a consequence of AGR<sup>0</sup> being rich and becoming thus the AGR-identifier, plus the proposal that it is the AGR-identifier which has to provide the I-subject with Case: if the AGR-identifier is AGR<sup>0</sup> it has to Case-mark the I-subject by head-governing it. In this Chapter we have contended that infinitival AGR<sup>0</sup> can end up being the AGR-identifier if some configuration allows it to be rich: either PRO<sup>0</sup> incorporation (in control structures) or, in raising structures, coindexation with the upper AGR (possibly involving incorporation). In IOS's, however, there is no apparent non-*ad-hoc* device by which the infinitival AGR<sup>0</sup> could end up being rich.

Let us first speculate on some theoretical basis for accounting for the facts which is consistent with the present theory. Our account for finite AGR in Chapter 3 is based on the assumptions that AGR is present and has to be licenced (by having an I-subject, providing Case to it and being rich in features). In this chapter we have developed the idea that AGR is also present in infinitives and, therefore, there must be some means to make the AGR satisfy the same requirements, provided everything follows from control, raising or ECM legitimate structures.

Suppose, however, we assume that, since infinitives are morphologically silent in AGR content, AGR is only optionally present. Then no infinitive has to have an AGR in principle. What determines the presence of AGR when it is morphologically silent? Let us explore some possibilities.

In the case of lexically determined control, infinitives would be forced to have AGR

because control would be a universally available grammatical option, which 'tries to apply' whenever possible. This idea could be related to the Elsewhere Condition, which essentially states that regular processes apply automatically unless they are blocked by the existence of more specific/irregular processes (e.g., in morphology, a regular verbal form is used unless an irregular form exists).

Suppose control is a universal option, in fact the most 'regular' one possible for infinitives, and applies whenever possible, unless a more specific/idiosyncratic option exists (for instance, an idiosyncratic subcategorization specification for 'believe' ([\_\_IP]) in English). Then the existence of AGR in lexical control structures is forced by the obligatory application of control: control involves PRO; PRO can only appear in Spec of AGR;<sup>139</sup> an element in Spec of AGR has only two options to be licenced, in accordance with our previous assumptions: either it becomes an AGR-identifier or it is resumed by an independently Case-marked I-subject (see next chapter). Putting aside the second option, the conjunction of the above assumptions gives the result that AGR has to be present in control structures.<sup>140</sup>

---

<sup>139</sup> It is far from obvious that theories being as rich in structural positions as the Split-INFL hypothesis, or Larson's (1988) VP-shell theory, can manage to confine PRO to Spec of AGR. Two solutions come to mind: a) PRO (or PRO°) is inherently an AGR-identifier; b) PRO can appear in other positions, but then it is not available for control.

<sup>140</sup> It is not obvious that the Elsewhere Condition, which has been used in phonology and morphology, should be relevant for syntax. I think it is at variance with Chomsky's (1988) implementation of the idea of economy of derivation, in that he explicitly states that regular/universal options take preference over irregular/idiosyncratic options. In fact Chomsky's proposal is intended to deal with subtle, theory internal problems, while the EC could be argued to be relevant for syntax in more obvious cases (e.g., if a language has object agreement morphology, it has to use it, and give up what is likely to be the more universal option of not using it). In addition, the idea of economy need not be tied to Chomsky's universal-over-particular constraint. For instance, Roberts (1991-a) uses a notion of economy based on the length of derivations.

Since, however, Chomsky is concerned with the economy of derivations in a very subtle sense, it could turn out not to be at variance with the Elsewhere Condition if some subtle distinction could be made between the field of application of

In the case of raising and ECM, since they are marked options, no problem of necessity (but rather one of mere availability) arises: these structures are allowed, not required, by UG, and will only exist if some lexical idiosyncratic (pattern of) specification(s) is learned (I think ECM and raising pose no problem for being learned on positive evidence). Since raising or ECM always involve infinitival complements (not, for instance, adjuncts), if raising or ECM configurations are not used in a language, then lexical control takes over (e.g., 'believe' is a control verb in languages not having ECM<sup>141</sup>) or the construction simply does not exist.

So the existence of AGR in complement infinitives is always forced to be present by control or simply allowed, to give raising/ECM structures. In some adjunct infinitives, however, neither lexical control nor raising is a possible option. Therefore, there is in principle no necessity by which the infinitive should have AGR. PRO<sub>Arb</sub> or dative control is still an available option in most cases, but not always a necessary one, in that it is not subject to fixed lexico-semantic control requirements. Therefore, absence of AGR would give the result that no principle of licensing concerning AGR applies.

If this is the situation, the prospects are not highly promising yet within the present theory: we are in the middle of nowhere, for we have given up the key element we had recourse to in order to account for distribution of subjects: AGR. It would be little interesting to simply propose independent constraints that apply to subjects when there is not AGR, especially because the distribution of I-subjects in the above cases is obviously reminiscent of the one in finite, control and raising/ECM cases: I-subjects are strictly [-anaphoric].

Therefore, the optimal theory should try to characterize what is the minimal common factor between control/raising/ECM infinitives and IOS. In more technical words, if the preceding considerations are on the right track, we should determine what is common between AGR-ful and an AGR-less infinitives that bears on I-subject distribution. And in addition we

---

the EC and economy. I leave the question open.

<sup>141</sup> English verbs allowing both control and ECM (e.g. **expect**) would have [\_\_\_\_ IP ] only as an optional subcategorization idiosyncrasy.

should be able to derive the existence of IOS-INV from the NSL status.

Concerning this last desideratum, it could be the case that IOS with a [-anaphoric] I-subject is parasitic on the existence on long V-movement in infinitives, which, crucially in our theory, allows Nominative assignment through Chain-government, and in turn would be parasitic on the existence of such a long movement in controlled/raising constructions, taking control to be a core case of infinitive that learners use to set parametric options in this field of grammar.

The idea then would be that once there is long V-movement, Nominative assignment by government is possible, even if AGR does not take part in the process. Our previous proposal concerning Case is that the AGR-identifier must provide Case to its I-subject. If it is Spec of AGR, it must transmit its Case to the I-subject. If it is  $AGR^0$ , we claimed, it must Case-mark the I-subject.

Suppose we assume, *alternatively*, that the requirement is of the type: the AGR-identifier must participate in providing Case to the I-subject. If it is Spec of AGR, we have Case transmission. If it is  $AGR^0$ , and therefore Case is provided by (Chain-)government, it must combine with a head that is *the actual Nominative Case-marker by government*:  $T^0$ . Thus it is not only finite  $T^0$  which is a Nominative assigner, as has been claimed so often: the fact that infinitival  $T^0$  is not apparently able to assign Case would be due to the independent requirements on AGR, that must participate in the process, if AGR is (forced to be) present. If AGR is not present, then  $T^0$  can assign Nominative by itself. Why can it not in non-NSLs? The idea would be that in NSLs the case where  $T^0$  assigns Nominative by itself (in IOS) is not at variance with the case where  $T^0$  combines with  $AGR^0$ . Suppose we rephrase the parallelism principle 0 above as:

- (77) In the unmarked Case, the I-subject in infinitives obtains Case in the *same* way as in finite sentences (where 'same' means involving sets of processes that are 'unifiable' in the set-theoretical sense).<sup>142</sup>

---

<sup>142</sup> An alternative (and only apparently simpler) formulation

Then in NSLs taking the unmarked option, T<sup>0</sup>'s assigning Case by itself to the I-subject is the *same* option as combined assignment by AGR<sup>0</sup> and T<sup>0</sup> by Chain-government. In non-NSLs, instead, this is not true, for, in finite sentences, the AGR-identifier in Spec of AGR cannot combine with T<sup>0</sup> to provide it with case: instead, we have contended, it has to be Case-marked by AGR under agreement and then transmit its Case to the I-subject.<sup>143</sup>

There remain some problems: why does Sardinian allow complement infinitives not to be controlled when a preposition precedes them (see 0)? No similar blocking effect occurs in other languages having prepositional controlled infinitives. On the other hand Sardinian apparently contradicts our claim that lexical control is a strong and pervasive requirement that applies whenever possible. Sardinian, as we saw, is exceptional in another sense: in many cases of IOS, if we take out the overt subject, the subject reference shifts to control or the PRO<sub>Arb</sub> interpretation can appear when there is no possible controller. Sardinian prepositional infinitives, however, take the arbitrary-existential interpretation steadily (see 0).<sup>144</sup> Although it is clear what the speaker has internalized as a pattern (no preposition -> control/preposition -> no control), it is not clear how this pattern can have been developed.

---

would be:

In the unmarked case, the I-subject in infinitives is Case-marked with the same option as in finite clauses, the options being: a) (Chain-)government; b) Agreement.

The trans-derivational and trans-structural character of this formulation does not change w.r.t. 0. I leave the question open.

<sup>143</sup> What we call parallelism principles should perhaps be called parameters in that they allow for variation. However, I do not know of any non-NSLs taking the marked option (i.e., having Nominative assignment under government in infinitives) If there is none, this asymmetry should obviously be captured. In any case that is why I prefer to keep to the term 'principle': as a parameter, it should have the two options freely available.

<sup>144</sup> Thus the Sardinian cases of IOS-INV we are considering do not behave like complements of ECM verbs such as English **expect**, which shift to control when the infinitival subject is null.

In order to account for these facts in line with the above considerations, let us assume that in Sardinian some prepositions are lexically specified to block control (so that the Elsewhere Condition would take this more irregular option).

Let us try to summarize and integrate all the above considerations:

a) Control is the most regular option as far as the Elsewhere Condition is concerned: it applies if the structural conditions are met and there is no more irregular, language-particular option blocking its application.

b) Languages can have lexically determined options blocking control: verbs subcategorizing for IP instead of CP (raising/ECM); prepositions specified for non-control (as in Sardinian), etc. Of course, these language-particular options have to abide by the learnability problem: they have to be easily recognizable, and allowed by UG. Thus, for instance, the existence of ECM or raising has to be learnable (and I think it is, on the mere basis of hearing the constructions). The pattern in Sardinian (absence/presence of preposition  $\leftrightarrow$  control/non-control) is, I think, not difficult to identify, for at least the prepositional IOS complement is readily identified.

c) Another language particular fact that the learner can easily identify is the existence of AGR-inflected infinitives (see next section). In this case, the issue of whether AGR is present or not is not at stake: it is obviously present and poses no special problem of learnability. This option blocks control.

One idea contained in the above proposals is that the theory must not explain too much: whenever a construction is language specific and poses no obvious problem for being straightforwardly learned, it would be simply inadequate to try to directly derive its existence from principles and wide-scope parameter settings. Principles and parameter settings should only allow the construction, not determine it.

There is a remaining important problem concerning Sardinian: if Sardinian, as we argued, takes the marked option in the Parallelism Principle, then IOS are not expected at all. As

for Occitan, it is certainly better behaved in this connection: as we saw in 0/0, IOS are (rather) constructed with preverbal subjects. As far as I know, IOS's with inverted subjects are marginal.

Now let us address question 0.b): what licences (referential) *pro* in IOS-INV? We have assumed, in the spirit of Rizzi (1986), that *pro* requires recovery of content. I think it would be senseless to assume that an abstract AGR can recover the content of *pro*, for this would void the traditional intuition that *pro* is allowed as far as its content is overtly expressed in a head.

The solution we will propose is similar in spirit to the one we have proposed for Nominative assignment in the absence of AGR. The idea is that:

- (78) AGR, if present in a clause, must be rich enough to licence an empty I-subject (which will be [ $\pm$ anaphoric] depending on the (non) Null Subject status of the language).

0 is a concise expression of what we have assumed thus far for AGR-licensing, except that we did not consider the possibility of AGR being absent. If it is absent in IOS-INV in a NSL, then the null I-subject being [-anaphoric] (*pro*) will not be required to be identified by AGR, and some other device (such as context recovery) takes place.

The idea advanced here about 'free' *pro* is obviously a simplification: we cannot account for why *pro* can be fully referential in some IOS-INV (e.g., Catalan and Spanish prepositional infinitives) and only arbitrary in others (e.g., Sardinian non-controlled complements). A comprehensive account should even account for cases of null objects (if they are indeed *pro*, as in Rizzi's (1986) proposal about Romance).

I think the account for the referentiality of 'free' *pro* should be expected to be derived from independent modules concerning context recovery (referential *pro*) and unselective Binding by sentence operators (arbitrary *pro* -and PRO), which should interact to allow only the attested cases.<sup>145</sup>

---

<sup>145</sup> See Authier (1991) for an account of Romance null objects in terms of unselective binding.



### 2.4.2. IOS with a Subject in Specifier of AGR

Let us start with AUX-to-COMP, and assume, following Rizzi (1982-b) and subsequent work, that it involves a subject in Spec of INFL (= AGR).

AUX-to-COMP, as far as it is a NSL phenomenon, falls out of all the above proposals: here we are dealing with a subject which is not an I-subject: unlike I-subjects (inverted subjects) in Italian and Portuguese, subjects in AUX-to-COMP construction *may* appear right after the auxiliary. There is another outstanding difference: while all other cases of IOS are colloquial (often substandard), AUX-to-COMP is a rather literary construction. Therefore we can take it to be a non-core construction.

In fact, I want to suggest that, in the spirit of the present theory, AUX-to-COMP can be treated as having much in common with IOS-INV. We tentatively will propose the following account:

- a) In AUX-to-COMP movement of the verb to  $C^0$  is triggered because  $T^0$  moves to  $C^0$ , for some reason (see below).
- b) In AUX-to-COMP, Case assignment to the I-subject is carried out by Chain-government by  $T^0$  (in combination with  $AGR^0$ , see below):

(79) Ritengo...

[C<sup>o</sup> avere+T<sub>i</sub>] [AGRP [AGR t<sub>i</sub>] [T<sup>o</sup> t<sub>i</sub>] fatto questo Gianni/lui]

to-have    done this G. /he

'I believe Gianni/him to have done this'

- c) As in finite clauses, once the I-subject is licenced, it can act as a resumptive element licensing a DP in Spec of AGR:

(80) (Peter had it done by a lawyer, but...) ritengo...

[C<sup>0</sup> averlo+T<sub>i</sub>] [AGRP Gianni [AGR t<sub>i</sub>] [T t<sub>i</sub>] fatto lui/*pro*]

to-have-it Gianni done he

'I believe Gianni to have done it (himself)'

In 0, *pro* is licenced without its content being recovered by a rich AGR. If we assume that AUX-to-COMP in Italian is AGR-less (just like IOS-INV in other languages), then this *pro* would not require content recovery. In fact, however, this is not a plausible approach: AUX-to-COMP in Italian, unlike IOS-INV in other languages, is a construction not allowing arbitrary or referential null subjects, and, anyway, it would be senseless to assume that the *pro* in 0 is recovered from the context, for it acts as a resumptive pronoun.

We will alternatively assume that in this case AGR<sup>0</sup> is *exceptionally* allowed to be enriched by the DP in its specifier (*Gianni*). This is certainly an unusual possibility, which we did not consider thus far:<sup>146</sup> this would be the exceptional point in our proposal, which would account for the marked character of AUX-to-COMP.

d) We want to account for why AUX-to-COMP is restricted to auxiliaries, copulas and modals. We can assume that the reason for this movement is that T<sup>0</sup> has to move to C<sup>0</sup> because it is selected (by a factive verb, for instance) or required for the interpretation of the construction (gerunds). Suppose this selection or requirement has some further restriction: T<sup>0</sup> which is allowed to move to COMP is one having the special property of *not selecting a VP* in its base position: in compound-tense clauses, as well as in copular structures and modal-verb clauses, T<sup>0</sup> does not select a VP, but rather, respectively, a participle (which is a FC of some sort), a non-verbal predicate (a Small Clause) or an infinitival (if modal-verb clauses are mono-clausal and modals subcategorize for an infinitival). It is crucial to this account that Auxiliaries, copulas and modals do not head their own VP (which would be a complement of T<sup>0</sup>), but are rather generated

---

<sup>146</sup> In fact there is circularity in this procedure: **pro** resumes a preverbal subject which enriches AGR<sup>0</sup>, which in turn licences **pro**.

under a FC: perhaps under  $T^0$  itself. If this last suggestion was correct, we could even reformulate the theory another way: only verbal forms which are generated under  $T^0$  can raise to  $C^0$  to satisfy selectional restrictions involving  $T^0$ .

In fact, there is a more appealing possibility: let us assume that the  $T^0$  which can move to COMP in AUX-to-COMP has to be one not selecting an event:

- Modals select a proposition which would take the form of an infinitive.
- As for compound-tense auxiliaries, even if the clause contains an eventful VP, the auxiliary does not carry the Tense interpretation for the event itself, but rather the tense interpretation for the reference time (*At five o'clock<sub>i</sub>, John had-T<sub>i</sub> already finished-T<sub>j</sub>*), which is not eventful itself.<sup>147</sup>
- As for copulas, it is harder to contend that they do not select an event, if the predicate is stage-level (see Kratzer (1988)), unless we assume that some (meaningful) FC intervenes between the copula and the stage-level Small Clause.

These are rather speculative considerations. I leave the issue here.

We can assume that AUX-to-COMP in Portuguese is licenced in a similar way. The obvious difference between AUX-to-COMP in Italian and in Portuguese is that in the latter AGR is morphologically present. Our theory predicts, if nothing else is said, that inflected infinitives are essentially like finite clauses as far as AGR and subjects are concerned. In fact, as shown by Raposo (1987-b), there are many restrictions on the distribution of infinitival overt AGR. We cannot address the issue here. The optimal situation for our theory would be that these restrictions are the result of the interaction of factors other than AGR, such as:

- the AUX-to-COMP trigger (which we have assumed is ('event-less')  $T^0$  raising to COMP).

---

<sup>147</sup> See, e.g., Giorgi (1992) for an implementation of the idea that in compound tenses the Reference time and the Event time are expressed in different positions in the syntactic structure.

- incompatibility between (lexical) control and overt AGR: it must be the same restriction which prohibits PRO in finite sentences.

As for the Occitan examples in 0/0, they seem involve a subject in Spec of AGR, not an I-subject. Since we have assume that Occitan takes the marked option in the parallelism principles (i.e., Occitan infinitives have Spec of AGR as the AGR-identifier), these facts are not at odds with the present theory. These examples would be similar to English gerunds with an overt subject. I will not address the issue of what licences non-finite clauses having an overt Spec of AGR (hence an AGR not involving control, raising or ECM).

\* \* \*

The above contention that  $T^0$  is the basic Nominative Case-marker by government leads to another speculation: the way we have formulated Nominative assignment through agreement,  $T^0$  cannot take part in this process, for this manner of assignment does not require T raising, manifested as long V-movement (English is an instance of language without obligatory long-V-movement in finite sentences). Therefore, Nominative through government and Nominative through agreement could actually be different Cases. It is noteworthy, in this connection, that non-NSLs tend to develop pronominal forms which show a Case form which exclusively occurs in Spec of AGR (e.g., in colloquial English *he* is not used outside Spec of AGR, contrary to Italian *lui* 'he', which can appear in postcopular position and in dislocated position). So forms like *himself*, French *lui*, which we take as Nominative when they are I-subjects, and must be non-Nominative in other cases, would be neutral forms: English (and French) would have no Case distinction except for pronominal agreement-Nominative forms.<sup>148</sup>

---

<sup>148</sup> In Chapter 3 we commented on these facts in another sense: these Spec-of-AGR-only pronominal forms would be AGR-identifier forms, which, diachronically, tend to cliticize to  $AGR^0$  to become (unmarked)  $AGR^0$  AGR-identifiers. I think both ideas (Case singularity and AGR-identifier singularity) can converge in a natural way, perhaps one being derived from the other. The fact that these Case distinctions are exclusively pronominal suggests that the AGR-identifier singularity is more basic, for pronominals, unlike full DPs, are minimal sets of AGR-features.

Summing up, Nominative assignment works in the following way:

a)  $T^0$  is the unmarked Nominative Case-marker: it is so in NSLs, which are the unmarked option: both in finite clauses and, with the unmarked parallelism option, also in infinitives.

b)  $AGR^0$  is, to use a metaphor, 'jealous' of  $T^0$ 's Nominative Case marking: if  $T^0$  actually assigns Nominative (in the cases mentioned in a)), then, if  $AGR^0$  is around (in finite clauses, control, and Raising, and AUX-to-COMP), it wants to take part (just like a jealous younger baby wants to take part in the older baby's game whenever the latter plays). Since AGR has a lot of restrictions for its own licensing, this ultimately will reduce the possibilities for subjects to control, raising and ECM (in the metaphor, the younger baby's taking part actually reduces the possibilities for the older one's games).

c) If  $AGR^0$  is not around (in IOS minus AUX-to-COMP), then  $T^0$  can assign Nominative alone, and the possibilities for subjects seem to surprisingly increase (the little baby is not pestering around).

d) In languages where  $T^0$  never assigns Nominative (non-NSLs), AGR is the only Case assigner (by agreement).

### 3. Summary

In this chapter we have extended our theory in Chapter 3 to infinitival constructions. In fact our theory on the [ $\pm$ anaphoric] status of I-subjects forced us to make some assumptions on the existence and status of AGR in non-finite clauses. Although this extension from finite to non-finite is far from being trivial (and could possibly be implemented in other ways), I think it has some plausibility in that it accounts for a range of facts that are thus far poorly understood: long infinitival verb movement in many NSLs and the existence of neutral pronominal/anaphoric forms in NSLs not having infinitival long V-movement; and it provides a reasonable alternative account to Kayne's (1991) theory concerning clitics and PRO, avoiding the problems we noticed for this proposal.

Our treatment of Infinitives with an Overt Subject (IOS) is only tentative. The literature on the issue is fragmentary and far less developed than that devoted to other types of infinitives. I think the reason for this is that Principles and Parameters (and all the research stream leading to it since the 1960's) is, as far as infinitives are concerned, intrinsically feeble to account for IOS. Most research on infinitives has ignored IOS, perhaps because English lacks it, and this fact has possibly biased research trends. This is a typical situation in any empirical science, which is rooted in the necessarily accidental component of research.

## Chapter 5

### Preverbal Subjects in NSLs

#### 1. Finite Sentences

There is an important question that remains unanswered in the above theory: what is the status of preverbal subjects in NSLs? If they are not required as AGR-identifiers, how are they licenced?

A reasonable position is that specifiers are not always filled: it would at least be difficult to contend they always are for any category and any well-formed structure (we keep neutral w.r.t. the issue whether non-filing implies non-projection). The obligatory filling of a given specifier should rather follow from principles and parameter settings. The above theory characterizes the requirements on the AGR category in a way that makes Spec of AGR obligatorily filled only for non-NSLs. So in a sentence like the following (Catalan):

- (1)    Ho ha fet (en Joan)  
       It-has done (the J.)  
       'JOAN/he did it'

Spec of AGR would be empty, at least when the I-subject is an R-expression (*en Joan*); when it is empty (*pro*), it is not logically impossible that Spec of AGR be filled by another empty category (possibly *pro*), but the null hypothesis is that it is not.

Now we have to answer at least three questions:

- Is Spec of AGR ever filled in NSLs? If it is:
- What is its status? More specifically:
- Is it filled only by elements coindexed with the I-subject?

We will address these questions in sections 1.1. and 1.2.

### **1.1. Dislocation vs. Specifier (of AGR)**

A conceivable approach to the nature of preverbal subjects in NSLs is to assume that they are left dislocated DP's. There are various arguments in favor of this view. We will see that it is nevertheless too simplistic a view.

Before addressing the issue, let me say a word about word order. It has often been claimed that inverted subject constructions cannot be the 'basic word order' for Romance languages allowing this option, for they involve a Focus interpretation which is contextually and pragmatically marked. Therefore preverbal subject constructions (which are indeed pragmatically more neutral) have to be basic structures, and a basic structure will not involve dislocation. I think that these considerations are pointless: they involve a naive conception of transformational grammar that has been largely overcome: nowadays underlying levels of representation are abstract and complex representations which simply cannot be claimed to be more or less 'basic'. When I claim that inverted subjects are Case positions in NSLs I am not committing myself to any claim about what is more 'basic' or 'neutral': inverted subject constructions are simply an available option of the grammar which happens to be pragmatically or contextually marked (in that it usually involves Focus).

A basic argument in favor of dislocation for preverbal subjects in NSLs is a theoretical one, based on indeterminacy:<sup>149</sup> even the most radical theory contending preverbal subjects in

---

<sup>149</sup> This is one of the arguments wielded in Rosselló (1986).



NSLs are not dislocated must admit that they can be. So a simple sentence like:

- (2) En Joan no ha vingut  
The J. not has come

would be ambiguous between the dislocation and Spec of AGR status of *en Joan*. More or less explicitly, many authors have contended this ambiguity is only apparent: the dislocation version would imply both a phonological pattern (usually a rise-and-fall intonation) separating the preverbal subject from the rest of the sentence, and a special interpretation by which the preverbal subject is read as 'as for Joan' or something similar.

It is true that there may be a phonological clue for dislocated elements. What is not true, at least in Romance, is that it is obligatory: any clitic left-dislocated (CLLD) element can be pronounced without any special pause or phonological clue possibly differentiating it from what would be a 'true' non-dislocated subject. Here are some examples CLLD (and the pattern generalizes to any CLLD element and to any Romance language as far as I know), where no pause or phonological clue obligatorily indicates dislocation.<sup>150</sup>

- (3) a. A en Joan li han robat la cartera  
To the J. him-have robbed the wallet  
'They stole the wallet from Joan'
- b. Amb ell no hi parlo  
With him not there-speak-I  
'To him, I never talk'

---

<sup>150</sup> Normative grammarians do not like dislocation in formal speech, especially in written formal speech. Liberal ones, they allow it only in informal speech, and then a comma should be used. This might be at the origin of the belief on the existence of the pause that has gone unchallenged by many linguists.

So both preverbal subjects and CLLD elements involve a pause only optionally, which depends on the degree of emphasis (in fact, it is not a discrete sign: one can add more emphasis by overdoing the tone break or adding a pause, in a continuous way). It may be the case that with CLLD elements the (optional) dislocation pattern is used more often than with subjects. But if so, it would be natural: subjects are more often [+human] than internal Arguments; [+human] DPs are more often D-linked for obvious reasons; D-linked DPs need less emphasis to be introduced; therefore, by diffuse-logic *modus ponens*, subjects need the emphatic dislocation pattern less often.

As for the interpretative clue, it has often been contended that a true dislocated element cannot be used in a sentence which is a natural answer to 'What is happening?'. This test, apart from being a bit vague (it implicitly relies on pragmatic factors such as speakers' optimal cooperation) does not give clear results. All of the following examples in Catalan are a reasonable answer to 'What's happening?' (e.g., when someone arrives and feels something strange in the atmosphere):

- (4)
- a. A en Joan li han robat la cartera.  
To the J. him-have robbed the wallet  
= 0.a)
  - b. Al nen l'ha mossegat una rata  
To-the child him-has bitten a rat  
'A rat bit the child'
  - c. A en Joan li han concedit una beca  
To the J. him-have-they awarded a grant  
'Joan has been awarded a grant'
  - d. A en Joan li ha vingut un atac de cor  
To the J. him-has come an attack of heart  
'Joan had a heart attack'

In fact, the non-dislocated versions of the above examples would be slightly unnatural as an answer to 'What's happening?' (they would be certainly acceptable in literary speech, for they would be the outputs of the avoid-dislocation correction effort<sup>151</sup>).

We must point out the in 0 the CLLD element is human and definite. These factors are clearly relevant. So the following examples would not be appropriate replies to 'What's happening?':

- (5) a. Una sargantana, l'he vista al jardí  
 A lizard it-have-I seen in-the garden  
 'One (of the) lizard(s), I saw in the garden'
- b. De lingüística, en parlarem després  
 Of linguistics of-it-will-we-talk later  
 'Linguistics, we'll talk about later'

But about the same happens with preverbal subjects, as the following are not appropriate replies either (perhaps they can be felt as more acceptable than the preceding because of normative pressure for SV(O) word order):

---

<sup>151</sup> There is a long tradition in traditional grammars stating that SVO is the unmarked word-order, and it is a dearly recommended one when the grammar is prescriptive. This partly can be traced back to the XVIIIth century belief that French was the language of reason, because, among other trifles, it had the virtue of systematically expressing the 'natural order of thoughts' (i.e. 'agent-action-object') properly (i.e. SVO). If this sheer nonsense deserves any theoretical attention it is precisely a warning against being misled by it as far as it has become 'common sense' belief among literate people that SVO is the unmarked order.

- (6) a. Un roc ha caigut  
 A stone has fallen  
 'One (of the) stone(s) fell'  
 (not just: 'A stone fell')
- b. Un cotxe ha passat  
 A car has gone-by  
 'One (of the) car(s) went by'  
 (not just: 'A car went by')

There seem to be other factors favoring neutral (i.e., 'What's happening?'-appropriate) dislocation or preverbal subjecthood: there is a gradation agent-dative-object-oblique going from most to least favoring. This favoring gradation is specially apparent with 'inverted' psych-verbs (It. *piacere*, 'to like' see Belletti & Rizzi (1988)), which most often have the Dative or Accusative Experiencer Argument as CLLD and the subject Theme as an I-subject. I will not pursue the issue. Suffice it to be the case that no criterion singles out, as far as we have seen, preverbal subjects as opposed to CLLD elements in a clear-cut way: both subjects and CLLD datives or objects have the same favoring conditions for (non-Focus-fronted) preverbal position, even if one of the favoring conditions is possibly being an external Argument (then a subject).

There is a more solid criterion that has become a classical test for subjecthood as opposed to dislocation: only true subjects can be quantified. I think this criterion can be a good test. But I think too that it has many times been oversimplified. It is not enough to come up with a few examples (such as 0) and conclude they are genuine cases of true subjects. On the empirical side, it may happen (and it does happen) that not all quantifiers behave alike. On the theoretical side, we know there are various kinds of quantifiers: apart from 'true' quantifiers, researchers have proposed subclasses such as D-linked quantifiers (Pesetsky (1982)), branching quantifiers (May (1985), Hornstein (1984)), indefinite DPs with a referential reading (Fodor & Sag (1982)), etc. I will not address the issue of a proper classification (these classes are not complementary), but

many of the proposals coincide in characterizing some quantifiers as being 'less quantificational' than the 'true' ones. For convenience, let us call the former 'weak' quantifiers (D-linked, referential)<sup>152</sup> and the latter 'strong' quantifiers.

Now one possible expectation would be that only 'weak' quantifiers are possible preverbal subjects or CLLD elements. The expectation is more or less fulfilled (remember, however, what we said above about other factors, which might distort the results). So it is the case that in Romance NSLs preverbal indefinite subjects tend to be 'referential' (= 'a certain') or D-linked (= 'one of the'), while the purely existential interpretation is hard to obtain (see 0). Bare indefinite DPs are simply not allowed as preverbal subjects in these languages (while they are in English) -0.a); Some negative and proportional quantifiers are not either -see 0.b/.c); 0 shows that the same restrictions hold for CLLD elements resumed by definite clitics:

- (7) a. \*Rocs cauen de la muntanya  
           Stones fall from the mountain  
       b. \*Res ha passat  
           Nothing has happened  
       c. \*Pocs estudiants han vingut  
           Few students have come

---

<sup>152</sup> Cinque (1990) uses **referential** for what we call 'weak', assuming referentiality subsumes D-linking. In fact, it seems to me that with D-linked quantification what is really referential is the set over which the quantifiers range. But it is only a matter of terminology: 'weak'/'strong' are in fact vague terms used here to avoid being committed to specific theories of quantification.

- (8)
- a. \*Estudiants els he vist  
Students them-have-I seen
  - b. \*Res ho ha fet  
Nothing it-has done
  - c. \*Pocs estudiants els ha convidat  
Few students them-has invited

Spanish and Italian display about the same restrictions (examples from Italian; recall we are not dealing with Focus-fronted elements, which would be acceptable in 0, with a different intonation):

a) Peverbal subjects:

- (9)
- a. \*Studenti sono arrivati  
Students are arrived  
'Students have arrived'
  - b. ?\*Pocchi studenti sono arrivati  
Few students have arrived
  - c. ?\*Niente è successo  
Nothing has happened

b) CLLD elements:

- (10) a. \*Studenti li ho visti  
Students them-have-I seen  
'I have seen students'
- b. \*Pocchi studenti li ho visti  
Few students them-have-I seen  
'I've seen few students'
- c. \*Niente l'ho visto  
Nothing it-have-I seen  
'I've seen nothing'

The reason these examples are not allowed is that these quantifiers are 'strong', and cannot be forced to a D-linked or 'referential' reading.

Similar considerations can be made of proportion quantifiers ('a few', 'most', 'many') and numerals: the D-linked or 'referential' reading prevails for preverbal subjects, while postverbal subjects are ambiguous:

- (11) a. Molts estrangers treballen aquí  
Many foreigners work here  
'Many of the foreigners work here'  
Not: 'There are many foreign workers here'
- b. Aquí hi treballen molts estrangers  
Here there-work many foreigners  
a) 'There are many foreign workers here'  
b) 'Many of the foreigners work here'

As for universal quantifiers, there do not seem to be sharp occurrence restrictions.

There is, in addition, an interpretative constraint which holds true of any preverbal quantifier in Catalan (and, I assume, in other Romance languages): they never have narrow scope w.r.t. a another quantifier inside their clause. Narrow scope for the subject quantifier has been reported to be less preferred in English and other languages, but I think that in Romance NSLs is not only less preferred: it is excluded.

The facts may have often been obscured because scope interactions can be interfered by the 'weak' status of one of the quantifiers: especially in the case of indefinite DPs with D-linked or referential reading, this reading is such that it gives the same extensional interpretation as the wide scope reading (consider, e.g., the sentence 'Everybody loves a certain woman', which is extensionally equivalent to 'Everybody loves a woman' with wide scope for 'a woman'). So, in the following example:

- (12) Tot estudiant ha llegit tres llibres  
 Every student has read three books

we can conceive of three readings: two with scope interactions (wide scope for 'every student' and wide scope for 'three books') and one with no scope interaction, when the numeral is interpreted as 'some certain three books'. Since the wide-scope reading for 'three books' is extensionally equivalent to the non-scope reading, our claim that 'every student' cannot have narrow scope can only rely on the intuition that 'three books', when apparently wide scope, is 'referential'. So it cannot be conclusive w.r.t. our claim that 'every student' cannot have narrow scope.

There are however cases where the wide scope interpretation of a quantifier is not equivalent to its scopeless reading. Consider the scope interactions between two numerals in the following example:



- (13) He enviat tres exemplars a quatre editorials  
Have-I sent three copies to four publishing-houses .  
'I've sent three copies to four publishing houses'

In 0 both numerals are inside the VP, and we have the two relative scope readings ('3 copies and possibly 12 publishing houses' and '4 publishing houses and possibly 12 copies'), and in addition we have the non-scope reading, with 'just 3 copies and 4 publishing houses'. If, however, one of the numerals is a preverbal subject, as in the following example:

- (14) Tres directors han dirigit quatre films  
Three directors have directed four films

we have the scopeless reading and only one of the scope readings, namely the one with the preverbal subject having wide scope. As I said, even in English the wide scope reading has been reported to be preferred, but it is simply the only option in Catalan.<sup>153</sup>

Negative quantifiers<sup>154</sup> cannot either have narrow scope when allowed as preverbal subjects or CLLD elements:

---

<sup>153</sup> There are actually some speakers who admit it, but most people steadily exclude it. My Italian reports are also less clear-cut.

<sup>154</sup> I use the term 'negative quantifier' for convenience. Zanuttini (1991) argues these elements in some Romance languages are negative universal quantifiers involving negative concord, and argues against a negative polarity item analysis. I cannot address the question here. Only recall that in Romance NSLs these elements do not cooccur with the sentence negation particle when they are preverbal.

(15) a. *Preverbal subjects:*

Ningú coneix tots els llibres

Nobody knows all the books

a) 'There is no person knowing every book'

b) \*'Every book is such that no person knows it'

b. *CLLD elements:*

A ningú el coneix tothom

To nobody him-knows everybody

a) 'There is no person being known by everybody'

b) \*'Everybody is such that he knows nobody'

Similarly, existential and proportional quantifiers never have narrow scope in the following cases:<sup>155</sup>

(16) *Preverbal subjects:*

a. Un/algun metge es cuida de tots els malalts

A /some doctor takes-care of all the patients

b. Molts/la majoria d'estudiants llegeixen tots els llibres

Many /most students read all the books

---

<sup>155</sup> For some reason, existential quantifiers of the type 'some NP or other' and numerals like 'at least one' can have narrow scope no matter how higher up from the other quantifier they are:

(i) Some book or other, I think every student has read

(ii) At least one car, I guess everybody can afford

Since this narrow scope reading is not predictable from the general clause-boundedness constraint for scope interactions, we leave it as irrelevant. Therefore, in (0.a) and (0.a) this type of reading, which is usually obtained with a special intonation pattern, should be disregarded. Interestingly, (0.a) only allows this reading if the preposition optionally accompanying the CLLD object is absent.

(17) *CLLD elements*:

- a. (A) un/algun malalt el visiten tots els metges  
(To) a /some patient him-visit all the doctors  
'A/some patient is visited by all the doctors'
- b. Molts/la majoria de llibres els llegeix tothom  
Many /most books them-reads everybody  
'Many/most books are read by everybody'

So the quantification test for preverbal subjects gives no clear-cut result because quantification is not a trivially unitary concept in natural language. A minimal characterization of the facts could be the following:

- some indefinite DPs (bare NPs, non-D-linked non-referential existential quantifiers - such as 'few'- and some negative quantifiers not having the [+human] feature - e.g. 'nothing' ) are not possible as preverbal subjects, in the same way as they are not as CLLD elements.

- other quantifiers (negative elements, universal quantifiers) are possible preverbal subjects or CLLD elements, but they always have wide scope.

Given this state of affairs, an obvious possibility is to assume that:

- preverbal subjects in NSLs are dislocated elements resumed by a *pro* (as CLLD elements are resumed by a (definite) clitic).

- only 'weak' quantifiers can be dislocated (this is less evident for universal and negative quantifiers).

We will see that this hypothesis can not be maintained. Thus far we have not considered an additional criterion that can distinguish dislocated elements from elements in Spec of AGR: word order. If we assume that:

- a) V moves to AGR<sup>0</sup> in NSLs (and we crucially assumed that);
- b) there can be no maximal projections adjoined to X' (and specifically to AGR'): this is a reasonable constraint on X'-structure;

then it follows that no XP can intervene between a true subject in spec of AGR and the verb in AGR<sup>0</sup>. This criterion is useless as far as non-quantified subjects are concerned: there is no way to tell whether adjacency is due to the preverbal subject's being in Spec of AGR or simply to the accidental fact that no XP happens to intervene. But for quantified subjects the data are clearly revealing. At least some of the quantified preverbal subjects require adjacency (we will see examples directly).

So we seem to finally arrive at some conclusion about preverbal subjects: at least some of them are in a fixed position, which is likely to be Spec of AGR. But let us raise another question first: we have seen that the restrictions on preverbal subjects are the same as on CLLD elements. So it might be that even for CLLD elements there was an adjacency requirement when they are quantifiers of a certain type. And this is indeed the case. What follows is set of pairs of examples (each pair containing a CLLD example and a preverbal subject example), in a gradation from the most ill-formed to the best well-formed cases. For the examples with preverbal subjects we abstract away from Focus fronting, which is irrelevantly acceptable without the part in the parenthesis and with another intonation, as it would be with objects and other Arguments; for the latter, however, no clitic would appear.<sup>156</sup> The asterisk at the beginning means the sentence is ill formed even without the part in the parenthesis; the asterisk inside the latter means that the part in the parenthesis impairs the sentence or makes it bad:

---

<sup>156</sup> I suspect that there are processes which are similar to Focus-fronting and nevertheless do not have the typical intonation and contrastive interpretation of typical Focus-fronting. So, in the judgements below I tried to disregard the acceptability of the preverbal subjects when a parallel (clitic-less) object fronting is available which intuitively has the same phonological, interpretative and stylistic flavor. For what is at stake here is if preverbal subjects are any different from clitic-resumed elements.

- (18) a. 1. \*Res (\*en Joan) ho ha fet  
Nothing the J. it-has done
2. \*Res (\*a en Joan) li ha passat  
Nothing to the J. him-has happened
- b. 1. \*Cap (\*en Joan) l'ha llegit  
None the J. it-has read
2. \*Cap (\*avui) ha arribat  
None today has arrived
- c. 1. \*Pocs amics (\*aquest any) els ha convidat  
Few friends this year them-has invited
2. \*Pocs amics (\*aquest any) m'han convidat  
Few friends this year me-have invited
- d. 1. ?\*Cap paquet (\*en Joan) l'ha enviat  
No packet the J. it-has sent
2. ?\*Cap paquet (\*avui) ha arribat  
No packet today has arrived
- e. 1. A ningú (\*aquesta vegada) l'han acceptat  
To nobody this time him-have-they accepted
2. Ningú (?\*aquesta vegada) ha aprovat l'examen  
Nobody this time has passed the-exam
- f. 1. A tothom (?\*aquesta vegada) l'han acceptat  
To ev.body this time him-have-they accepted
2. Tothom (?\*aquesta vegada) ha aprovat l'examen  
Ev.body this time has passed the-exam

- 0
- g. 1. A qualsevol (???al Brasil) el poden robar  
To anybody in Brasil him-can rob
2. Qualsevol (???al Brasil) et pot robar  
Anybody in Brasil you-can rob
- h. 1. A tots (??demà) els veuré a la reunió  
To all tomorrow them-will-I-see at the meeting
2. Tots (??demà) aniran a la reunió  
All tomorrow will-go to the meeting
- i. 1. A algú (?fa una estona) l'han tancat al lavabo  
To sm.body a while ago him-have locked in-the  
bathroom
2. Algú (?fa una estona) s'ha tancat al lavabo  
Sm.body a while ago himself-has locked in the  
bathroom
- j. 1. A tots els alumnes (?demà) els veuré  
To all the pupils tomorrow them-will-I-see
2. Tots els alumnes (?demà) vindran a veure'm  
All the pupils tomorrow will-come to see-me
- k. 1. A tots aquests alumnes (demà) els veuré  
To all these pupils tomorrow them-will-I-see
2. Tots aquests alumnes (demà) vindran a veure'm  
All these pupils tomorrow will-come to see-me
- l. 1. La majoria d'aquests (en Joan) no els coneix  
The majority of-these the J. not them-knows
2. La majoria d'aquests (a mi) no em coneixen  
The majority of-these to me not me-know
- m. 1. A molts d'aquests (avui) no els he vist

- To many of-these today not them-have-I seen
2. Molts d'aquests (avui) no han vingut  
Many of-these today not have come
- 0 n. 1. A dos amics meus (avui) els han arrestat  
To two friends mine (today) them-have arrested
2. Dos amics meus (avui) han tingut un accident  
Two friends mine today have had an accident
- o. 1. Alguns amics (al casament) no els hi convido  
Some friends to-the wedding not them-there-invite-I
2. Alguns amics (al casament) no hi vindran  
Some friends to-the wedding not there-will-come
- p. 1. A un meu amic (l'altre dia) el van arrestar  
To one my friend the other day him-arrested-they
2. Un meu amic (l'altre dia) va venir a veure'm  
One my friend the other day came to see-me

Examples from a. to d. are unacceptable as preverbal (and even worse if not adjacent to the verb): they are all 'strong' quantifiers hardly interpretable as D-linked or 'referential'. They are all [-human]. From e. to j., they are acceptable only if adjacent to the verb (with various degrees of ill-formedness if they are not). The rest of the examples are fully acceptable even if there is not adjacency to the verb. The gradation can be clearly related to the 'strength'/'weakness' distinction. The more a quantifier is likely to be interpreted as D-linked or 'referential', the more it is acceptable as dislocated. I think that in fact the intermediate degrees of ill-formedness are actually not 'intermediate' in grammatical status, but rather in pragmatic acceptability: what is intermediate is the chances for the speaker to imagine a likely pragmatic context where D-linking or 'referentiality' are plausible. So for instance the 'referential' use of *algú* 'somebody' implies that the speaker is wanting to be enigmatic or unexplicit in referring to a person. It is more usual that,

when one wants to refer to somebody the other person does not know, one minimally introduces a descriptive restriction for the quantifier (e.g., 'a friend of mine', 'some guy I met the other day', etc.).

So, the conclusion seems to be that:

- 'Strong' [-human] quantifiers cannot be preverbal (unless moved by Focus fronting).

- 'Strong' [+human] quantifiers can be preverbal but they have to be adjacent to the verb, which suggests they are in Spec of AGR. They involve a resumptive clitic strategy if they are not subjects. If subjects, they would be resumed by *pro*.

- 'Weak' (D-linked or 'referential') quantifiers can be dislocated.

What is of our present concern is that there is a preverbal Specifier position (we assume for the moment it is Spec of AGR) that can be filled by [+human] strong quantifiers being resumed by a minimal (clitic or *pro*) pronoun. So our previous speculation that all preverbal subjects are all dislocated is false, but now we conclude their position is not one specific to subjects.<sup>157</sup>

Another fact which is worth considering is that the acceptable preverbal quantified elements having the V-adjacency requirement (examples from e. to g. above) are clause-bound: if they appear in the upper clause they are unacceptable approximately to the same degree as if they are not adjacent (the following examples are identical to examples from 0.e) to g., except that instead of an intervening XP we have a superordinate clause segment *crec que* 'I think that'):

---

<sup>157</sup> In fact, if that position exists, nothing prevents non-quantified preverbal subjects from being there, unless we assumed it is an exclusively quantifier position. We will address this issue in the next section.



- (19) a. 1. A ningú (\*crec que) l'han acceptat  
To nobody (I-think that) him-have-they accepted
2. Ningú (?\*crec que) ha aprovat l'examen  
Nobody (I-think that) has passed the-exam
- b. 1. A tothom (?\*crec que) l'han acceptat  
To ev.body (I-think that) him-have-they accepted
2. Tothom (?\*crec que) ha aprovat l'examen  
Ev.body (I-think that) has passed the-exam
- c. 1. A tothom (???crec que) el poden robar  
To ev.body (I-think that) him-can rob
2. Tothom (???crec que) roba  
Ev.body (I-think that) robs
- d. 1. A tots (?crec que) els veuré a la reunió  
To all (I-think that) them-I'll-see at the meeting
2. Tots (?crec que) aniran a la reunió  
All (I-think that) will-go to the meeting
- e. 1. A algú (?crec que) l'han tancat al lavabo  
To sm.body (I-think that) him-have locked in-the  
bathroom
2. Algú (?crec que) s'ha tancat al lavabo  
Sm.body (I-think that) SI-has locked in-the bathroom
- f. 1. A tots els alumnes (?crec que) els veuré  
To all the pupils I-think that them-will-I-see
2. Tots els alumnes (?crec que) vindran a veure'm  
All the pupils I-think that will-come to see-me

Finally, there is still another fact that has to be considered. 'Strong' [-human] quantifiers

are not allowed as preverbal subject or clitic resumed elements in neutral sentences. But there is a factor that can rescue them in this position: they are allowed when they are somehow construed with a postverbal contrastive Focus:<sup>158</sup>

- (20) a. 1. ???Res ho ha fet EN JOAN  
 Nothing it-has done the J.  
 'Nothing has been done by Joan' (-> everything has been done by others')
2. ??Res li ha passat A EN JOAN  
 Nothing him-has happened to the J.  
 '(Many things happened to many people, but) nothing happened precisely to Joan'
- b. 1. ?Cap l'ha llegit EN JOAN  
 None it-has read the J.  
 '(We have read many of these books but) none was read precisely by Joan'
2. Cap ha arribat AVUI  
 None has arrived today  
 '(Many have arrived, but) none arrived precisely today'

---

<sup>158</sup> To my ear, examples in each pair 1./2. (resp. clitic resumed elements and subjects) are equally acceptable. Some speakers prefer the examples with subjects. I assume this is due to the fact that these speakers are assigning some fronting analysis to these examples, which is impossible with the clitic-resumed elements, as far as fronting does not involve clitics. On the other hand, there are kinds of fronting which are not Focus fronting, and do not involve the typical phonological pattern of Focus fronting:

(i) A mi, res m'han dit.  
 To me nothing me-have told  
 'Me, I was told nothing whatsoever'

It would be this 'intonation-flat' fronting which makes the examples in 2. better for some speakers. In my dialect, this kind of fronting is not used.

- 0 c. 1. ?A pocs assistents els havia convidat EN JOAN  
 To few people-present them-had invited the J.  
 '(Many people were at the party, but) few had been invited by Joan'
2. Pocs amics m'han convidat A SOPAR  
 Few friends me-have invited to dinner  
 '(Friends have invited me to lots of things, but) few have invited me to have dinner'
- d. 1. (?)Cap paquet l'ha enviat EN JOAN  
 No packet it-has sent the J.  
 '(Many packets were sent, but) none by Joan'
2. Cap paquet ha arribat PER CORREU  
 No packet has arrived by mail  
 'Packets arrived, but none arrived by mail'

To summarize, quantified subject- or clitic-resumed- elements are allowed if:

- they are 'weak' (then they can be dislocated)
- they are 'strong' [+human] (then they are in a Spec position).
- they are construed with a postverbal Focus elements (and then they are in this Spec position too).

The facts reported from Catalan extend to Spanish with a reasonable degree of approximation. As for Italian, the judgements I have been reported are less clear and systematic. Southern dialects<sup>159</sup> seem to be close to Catalan, with the difference that adjacency to the verb does not seem to play any role:<sup>160</sup>

---

<sup>159</sup> Thanks to Giuseppe Longobardi for his judgements, comments and suggestions

<sup>160</sup> Thanks to Giuseppe longobardi and Giovanni Albertocchi for the data (which I take as roughly representative of resp.

- (21) a. 1. \*Niente (Gianni) lo ha fatto  
Nothing (G.) it-has done  
2. \*Niente (a Gianni) gli è successo  
Nothing to G. him-has happened
- b. 1. \*Nessuno (oggi) l'hanno letto  
None (today) it-have-they read  
2. \*Nessuno (oggi) è arrivato  
None (today) has arrived
- c. 1. \*Pochi amici (questo anno) li ha invitati  
Few friends (this year) them-has invited  
2. \*Pocchi amici (questo anno) mi hanno invitato  
Few friends (this year) me-have invited
- d. 1. ?\*Nessun pacco (Gianni) l'ha spedito  
No package (G.) it-has sent  
2. ?\*Nessun pacco (oggi) è arrivato  
No package (today)
- e. 1. \*Ognuno (questa volta) l'hanno accettato  
Everyone (this time) him-they-have accepted  
2. \*Ognuno (questa volta) ha passato l'esame  
Everyone (this time) has passed the-exam
- f. 1. ?Nessuno (questa volta) l'hanno accettato  
Nobody (this time) him-they-have accepted

---

Southern and Northern varieties of Italian).

Italian has less unambiguously [+human] quantifiers than Catalan or English: so **nessuno** is both 'nobody' and 'none'; **tutti** is both 'everybody' and 'every one/all', etc. Nevertheless the [+human] interpretation is equally significant in the data below.

2. Nessuno (questa volta) ha passato l'esame  
Nobody (this time) has passed the-exam
- 0 g. 1. Tutti (questa volta) li hanno accettati  
All (this time) them-they-have accepted
2. Tutti (questa volta) hanno passato l'esame  
All (this time) have passed the-exam
- h. 1. Chiunque (in Brasile) lo possono derubbare  
Anybody (in Brasil) him-they-can rob
2. Chiunque (in Brasile) ti può derubbare  
Anybody (in Brasil) you-he-can rob
- i. 1. Tutti (domani) li vedrò alla riunione  
All (tomorrow) them-I'll-see at-the meeting
2. Tutti (domani) andranno alla riunione  
All (tomorrow) will-go to-the meeting
- j. 1. ?\*Qualcuno (pocco fa) l'hanno chiuso nel bagno  
Somebody (just now) him-have locked in-the bathroom
2. Qualcuno (pocco fa) si è chiuso nel bagno  
Somebody (just now) himself-has locked in-the bathm.
- k. 1. ?Tutti gli studenti (oggi) li vedrò in classe  
All the students (today) them-I'll-see in class
2. Tutti gli studenti (oggi) verranno a vedermi  
All the students (today) will-come to see-me

In Northern dialects, adjacency seems to play a role. They differ from Catalan in that preverbal subjects appear to be more acceptable than other clitic resumed elements.

- (22)
- a.
    - 1. \*Niente (Gianni) lo ha fatto  
Nothing (G.) it-has done
    - 2. ??Niente (??a Gianni) gli è successo  
Nothing (to G.) him-has happened
  - b.
    - 1. \*Nessuno (oggi) l'hanno letto  
None (today) it-have-they read
    - 2. Nessuno (??oggi) è arrivato  
None (today) has arrived
  - c.
    - 1. \*Pochi amici (questo anno) li ha invitati  
Few friends (this year) them-has invited
    - 2. Pocchi amici (??questo anno) mi hanno invitato  
Few friends (this year) me-have invited
  - d.
    - 1. \*Nessun pacco (Gianni) l'ha spedito  
No package (G.) it-has sent
    - 2. Nessun pacco (??oggi) è arrivato  
No package (today) has arrived
  - e.
    - 1. \*Ognuno (questa volta) l'hanno accettato  
Everyone (this time) him-they-have accepted
    - 2. ?\*Ognuno (questa volta) ha passato l'esame  
Everyone (this time) has passed the-exam
  - f.
    - 1. \*Nessuno (questa volta) l'hanno accettato  
Nobody (this time) him-they-have accepted
    - 2. Nessuno (??questa volta) ha passato l'esame  
Nobody (this time) has passed the-exam
  - g.
    - 1. Tutti (??questa volta) li hanno accettati  
All (this time) them-they-have accepted
    - 2. Tutti (questa volta) hanno passato l'esame

- All (this time) have passed the-exam
- h. 1. ?Chiunque (???in Brasile) lo possono derubbare  
Anybody (in Brasil) him-they-can rob
2. Chiunque (?in Brasile) ti può derubbare  
Anybody (in Brasil) you-he-can rob
- 0 i. 1. ??Tutti (???domani) li vedrò alla riunione  
All (tomorrow) them-I'll-see at-the meeting
2. Tutti (domani) andranno alla riunione  
All (tomorrow) will-go to-the meeting
- j. 1. \*Qualcuno (pocco fa) l'hanno chiuso nel bagno  
Somebody (just now) him-have locked in-the bathroom
2. Qualcuno (pocco fa) si è chiuso nel bagno  
Somebody (just now) himself-has locked in-the tathrm.
- k. 1. ?\*Tutti gli studenti (oggi) li vedrò in classe  
All the students (today) them-I'll-see in class
2. Tutti gli studenti (oggi) verranno a vedermi  
All the students (today) will-come to see-me

\* \* \*

Apart from their quantified status, there are other facts that point to the 'different nature' of preverbal subjects in NSLs w.r.t. preverbal subjects in non-NSLs. These concern overt pronouns. Montalbetti (1984) noticed that subject strong pronouns<sup>161</sup> in NSLs are not able to be bound variables. In 0, for instance, *ells* can only be interpreted as bound by *tots els estudiants* 'all the students' in the group reading:

---

<sup>161</sup> For non-subject strong pronouns, Montalbetti's Generalization only would hold when there is no 'weaker' option (such as a definite clitic). See Rigau (1988).

(23) *Tots els estudiants<sub>i</sub> es pensen que ells<sub>i</sub> aprovaran*

All the students think that they will-pass

This fact is known as Montalbetti's Generalization. Although most linguists have not challenged the descriptive accuracy of this generalization, as noticed by Rosselló (1986), it is only true of preverbal subjects. Postverbal strong pronouns are not subject to it (at least 0 allows the bound variable reading much more easily):

(24) *Tots els jugadors<sub>i</sub> estan convençuts que guanyaran ells<sub>i</sub>*

All the players are persuaded that will-win they

'For any player x, x is persuaded that x will win'

So Montalbetti's generalization would be an epiphenomenon: it is not the nature of the strong pronoun which is to blame for the lack of bound-variable-reading: it is rather its position (and, historically, the often uncritical acceptance that preverbal subjects in NSLs are the same as in non-NSLs is also to blame, I think). What happens with strong pronominal subjects (i.e., the contrast between preverbal and postverbal w.r.t. the possibility of a bound variable reading) also happens with object or oblique strong pronominals all the same (now the contrast being between dislocated and non-dislocated):

(25) a. *Tots els candidats<sub>i</sub> pensen que a ells<sub>i</sub> els<sub>i</sub> elegiran*

All the candidates think that to them them-will-choose

'All the candidates think that they (as a group/\*bound variable) will be chosen'

b. *Tots els candidats<sub>i</sub> pensen que els elegiran a ells<sub>i</sub>*

All the candidates think that them-will-choose to them

'For any candidate x, x thinks x will be chosen'



Another well-known fact points to the idea that preverbal subjects in Romance NSLs are 'different': as convincingly argued for in Rizzi (1982-b) and subsequent work, the 'that'-trace (ECP) effects in these languages are avoided by extracting the subject from the subject inversion position *exclusively*, not by any strategy like a null C<sup>0</sup> (as in English) or *que*->*qui* alternation (as in French), which could suggest that preverbal subjects are not candidates to be Wh-extracted by these strategies.<sup>162</sup>

All the preceding observations hold in full of Catalan and, as far as I know, also Spanish (perhaps with some qualifications). The Italian facts we saw are clearly different as far as quantification is concerned: the restrictions we have seen above for quantifier preverbal subjects also hold of Italian in about the same way. But CLLD (or clitic resumed) elements are much more restricted in Italian. So the Italian versions of examples from 0.a) to 0.i). would have the same acceptability only in the preverbal subject version, not in the clitic resumed version, which would be unacceptable (perhaps with a gradation of unacceptability too).

We will try to make sense of these facts in the next section, by considering what is the status and position of preverbal subjects in NSLs.

## 1.2. On the Nature of the Preverbal Specifier in NSLs

Throughout this thesis, we have assumed that AGRP is the highest FC in the sentence below CP. We have crucially assumed that V-movement to AGR<sup>0</sup> takes place in NSLs (in order for the I-subject to receive Case by Chain-government). In the preceding section, we detected a specifier position which can be occupied by some quantified elements (and perhaps even by non-

---

<sup>162</sup> One can always argue that these strategies are last resort, or more marked than postverbal subject extraction, and this is why they are not used in Romance NSLs. Kenstowicz (1989) shows that some Arabic dialects being NSLs indeed allow a **que-qui**-like strategy for preverbal subject extraction.

quantified ones). Since this position was adjacent to the verb (so to AGR<sup>0</sup>, according to our assumptions), we concluded that this position is Spec of AGR. Since this position was not exclusively occupied by subject elements, but also by clitic-resumed elements, we should explain why Spec of AGR is so 'liberal' a position.

There is one first consideration (which we have already stressed throughout this thesis): in NSLs Spec of AGR is not the AGR-identifier and therefore is not required to be filled by an element (which will be coreferential with the I-subject) in order to provide features to AGR. This would allow at least Spec of AGR to be empty, and we argued it is, e.g. in sentences where nothing precedes the finite verb and there is a null or inverted subject. However, the non-AGR-identifier nature of Spec of AGR does not seem to guarantee it can be freely filled with elements not being coindexed with the I-subject and, by transitivity, with AGR<sup>0</sup>. Whenever Spec of AGR is filled, it seems reasonable that it agrees with its head (as any specifier is basically expected to).

For Catalan and Spanish, at least, we saw that preverbal non-subject elements are allowed in apparently the same Specifier position (i.e., Spec of AGR, we assume) as preverbal subject elements. In addition, there seem to be other elements that apparently can occupy this position in Catalan and Spanish. In these languages, as we already saw in Chapter 4, section 2.1.1., adverbs like 'always', 'never', 'already/yet' and 'still/yet' can precede the verb (and cannot be separated from it), whereas in Italian they cannot:

- (26) a. Sempre/mai /ja /encara està cansat (Catalan)  
 Always/never/already/still is tired  
 b. \*Sempre/mai /già /ancora è stanco (Italian)  
 Always/never/already/still is tired

To account for the contrast between Catalan/Spanish and Italian, two options seem workable. One option is to assume that the preverbal Specifier we detected for Catalan and Spanish is not Spec of AGR, but Spec of some other FC. In other words, in Catalan and Spanish

the highest FC below CP is (in finite sentences) not AGRP but another XP (TP, NEGP or  $\Sigma$ P,<sup>163</sup> for instance). Suppose this FC is more liberal than AGRP in restricting the status of its specifier. Or alternatively that it is a position for quantified elements, either adverbial as in 0 or Argumental. Italian, instead, would have AGRP as the highest FC, then restricting elements in its specifier to DPs resumed by the I-subject.

This approach implies parameterization of the FC hierarchy, which has been advocated for by several authors (mainly Ouhalla (1988)). It implies we have to give up the assumption maintained throughout this thesis that  $AGR^0$  is the highest FC below CP (except for infinitives, where NOMP would intervene). I think one should be cautious about parameterization of the FC hierarchy:<sup>164</sup> in the present-day profusion of under-determined alternative hypotheses on FCs, hierarchy parameterization is not an advisable working hypothesis. And in any case, any theory advocating for hierarchy parameterization should carefully address the learnability problem.

In any case, the hypothesis that the highest FC below CP is not AGRP in Catalan and Spanish while it is in Italian does not explain why, even if Italian is more restrictive in only allowing subject elements in this top Specifier, it has the same restrictions as Catalan and Spanish w.r.t. the 'weak'/'strong' and [+human] factors.

Since none of these restrictions is observable in English, we cannot simply assume Italian preverbal subjects are 'true' subjects having the same nature as in English. Rather they are like in Catalan except they are further restricted to be subject elements. So the theory should capture

---

<sup>163</sup> See Laka (1990) for the postulation of a  $\Sigma$ P category which contains both negation and affirmation.

<sup>164</sup> Ouhalla (1988) claims that it is not parameterization what triggers different FC hierarchies: it is only variation in the subcategorization framework of FCs, therefore only 'lexical' variation. I cannot see this distinction has any real significance and it seems to me a merely terminological trick. For the child to learn that  $T^0$  subcategorizes for AGRP or, conversely, that  $AGR^0$  subcategorizes for TP is as difficult as if the alternative options are phrased as a parameter. On the other hand, affixal and highly abstract FCs are not 'lexical' in a reasonable sense.

two factors restricting preverbal specifiers:

- restrictions on quantification (holding of Catalan, Spanish and Italian).
- restrictions on subjecthood (holding only of (Northern) Italian).

To account for these facts we will suggest a possible approach. Thus far we have assumed AGR is a FC which projects a full X-bar structure. Our theory relies on the existence of AGR in crucial way: essentially we have derived Burzio's Generalization and the distribution of I-subjects from requirements on the licensing of AGR.

Although many researchers working in the split INFL hypothesis have assumed AGR is one of the components of INFL, some linguists contend that AGR is not a category projecting by itself, but it is rather parasitic on other, semantically 'contentful' FCs.<sup>165</sup> Let us see how this idea could help in our present concerns.

In fact, an alternative approach to the idea that AGR is parasitic, could be to assume that it is not AGR which is parasitic on X (X a FC), but rather that X and AGR are a mixed FC. In other words, we are assuming a partially non-split INFL hypothesis: some of the components of INFL are not split. Specifically, we could restrict the 'splitting' of INFL in the following way: a FC can contain one (and only one) semantically 'contentful' component (T, Mode, Asp, etc.) and (possibly) one of a purely formal component (AGR). AGR can never be the exhaustive component of a FC.<sup>166</sup>

With this idea, we could assume the following parameter:

---

<sup>165</sup> See, for instance, Laka (1990). She contends that the three AGRs in Basque (subject, object and dative AGR) are parasitic on the other FCs (TP and Modal Phrase). In Basque AGR-morphemes are 'sandwiched' between other FC morphemes, suggesting, from a Mirror Principle point of view, they are parasitic on these other FCs. Poletto (1991) proposes a similar idea on the basis of Romance finite and participial morphology, although she does not contend that the AGR FCs are parasitic on the others.

<sup>166</sup> This idea somehow reconciles the two notions of agreement, namely as a FC (AGR) and as a process/dependency: the latter would be morphologically instantiated as the former.

- (27) In a FC, the AGR component *imposes/does not impose* selectional restrictions on its specifier.

Suppose the only selectional restriction AGR can impose on its specifier is precisely agreement with itself. In some Italian dialects, (subject) AGR (we will discuss directly which FC subject AGR could be a component of) would have chosen the first value in 0, while Catalan or Spanish would have chosen the second value.<sup>167</sup> The other component of the functional category containing AGR would, in all three languages, impose restrictions on the quantified nature of its specifier. Let us try to ascertain which category could be the host of subject AGR.

Thus far, we are trying to maintain a maximally uniform structure (specifically hierarchy) of the FCs across languages. Between Italian and Catalan/Spanish, the only difference is not structural, but only concerns parameter 0. Can we contend that AGR has the same host in non-NSLs? We know that in English and non-NSLs in general, the restrictions on the quantified nature of preverbal subjects do not hold as in Italian (recall English can have bare DPs and quantifiers of any nature as preverbal subjects). So this suggests that in non-NSLs like English AGR is not hosted by the same FC as in Italian or Catalan, for otherwise the same restrictions should hold. In the next section we shall try to find independent evidence pointing to this direction.

---

<sup>167</sup> This would be the marked value. The fact that Italian varies would be due to the low level status of this parameter, the triggering evidence being thin.

### 1.3. Interrogative Wh-movement

It has been the prevailing hypothesis for at least two decades that Wh-elements move to (Specifier of) COMP in languages where overt Wh-movement takes place. Evidence for this hypothesis in languages such as English is compelling, and so many developments rely on it that any attempt to give it up should seriously consider whether the alternative proposal is not actually a terminological variant on the term COMP, for a good deal of the theory on the nature of COMP relies on precisely Wh-movement.

However, as has often been the case, the theory could have been biased by the fact that English or English-like languages have been the first to be studied in depth. I think English-like languages have two characteristics concerning Wh-phenomena which are far from being universal:

a) Wh-interrogative clauses and relative clauses use very similar strategies: moving a Wh-word. This is not apparently the case in many languages, where Wh-phrases are only used in interrogative clauses: Basque, Irish, Chinese, etc. and even languages as little exotic as colloquial Catalan and other colloquial Romance languages, where headed relatives are uniformly introduced by the complementizer 'that'. This is even true of colloquial English, where sentences 'The man who(m) I saw'/The man about whom I talked' are not used.

b) The landing site for the moved Wh-element is (likely to be) (specifier of) COMP. There are, however, languages having overt Wh-movement where the landing site is not obviously COMP. Specifically, languages having a Focus position (Basque, Hungarian) move their Wh-elements to that position.<sup>168</sup>

---

<sup>168</sup> Ortiz de Urbina (1989) argues that the Focus position is Spec of COMP. Since Basque would have a final C° head (the embedding 'that' particle is affixed to the also head-final V+INFL) but a left-branching Spec of COMP, the word-order predictions are less clear-cut than in head-initial/spec-initial languages, and doubts can be cast on the accuracy of his analysis.

I will contend that in Romance NSLs interrogative Wh-elements (unlike relative Wh-elements, when used) do not move to Spec of CP, or at least not always. This will become relevant for the argumentation we left unfinished in the preceding section.

I think the hypothesis that interrogative Wh-phrases move to Spec of CP in Romance NSLs is built on the basis of two considerations. One is universality: it is obviously the null hypothesis. The other is that word order is not extremely problematic: since interrogative Wh-phrases are sentence initial, the facts can be accommodated in a relatively easy way. However, we are going to contend that a close look at word order facts poses serious problems for the Wh-to-COMP analysis of interrogative Wh-movement in Romance NSLs and that the universality argument should be relaxed as to allow other possibilities. We will see, on the other hand, that postulating Wh-to-COMP forces the theory to adopt *ad hoc* stipulations that seriously undermine the initial universality argument.

One basic fact that lead to the postulation that the embedding COMP particle ('that') and Wh-elements occupy essentially the same position is their mutual distributional complementarity. Nowadays the theory does not predict such a complementary distribution (complementarity being an effect of some non-universal doubly-filled-COMP filter), but it predicts that Wh-in-COMP and C<sup>0</sup> are adjacent positions.<sup>169</sup> Romance NSLs, however, provide several kinds of evidence suggesting that the complementizer 'that' and interrogative Wh-phrases are not in such close positions. One argument is based on clause adjuncts, such as CLLD elements and other adverbial adjuncts. These elements, when adjoined to an embedded clause, always occur to the right of the complementizer particle:<sup>170</sup>

---

<sup>169</sup> Except, perhaps, in exceptional cases like French complex inversion (see Rizzi & Roberts (1989)).

<sup>170</sup> Occitan seems to be an exception to this generalization:

Pensi    aquel libre que    lo trobaràn            pas jamai  
 I-think that    book    that it-will-they-find not never  
 'I think that, that book, they will never find'

Occitan seems to have some kind of topic raising of an A-movement-like nature which gives other surprising results. So I

- (28) Espero (\**avui*) (\**en Joan*) que (*avui*) (*en Joan*) el veuré  
 I-hope (today) (the J.) that (today) (the J.) him-I'll-see  
 'I hope that Joan I will see today'

If we consider relative clauses with a Wh-phrase, the facts are parallel in that clause adjuncts follow the Wh-phrase (in this case pre-Wh adjuncts are in fact worse than pre-'that' adjuncts, probably because relative clauses have to be adjacent to the NP they modify):

- (29) La noia (\**avui*) de qui (*avui*) hem parlat  
 The girl (today) of who (today) have-we talked  
 'The girl about whom we talked today'

If we shift to interrogative Wh-phrases, however, the facts are reversed: adjuncts and dislocated elements can precede (at least in colloquial speech) the Wh-element but cannot follow it at all:<sup>171</sup>

- (30) No sé (*en Joan*) quan (\**en Joan*) el veuré  
 Not know-I (the J.) when (the J.) him-will-I-see  
 'I don't know when I will see Joan'

It is reasonable to assume that clause adjuncts cannot be adjoined to CP in general: this is what 0 and 0 suggest.<sup>172</sup> Adjunction to IP, instead, seems to be allowed quite freely in many

---

think it is an independent problem.

<sup>171</sup> We already discussed some of these facts in Chapter 4, 2.1.1.1.

<sup>172</sup> This is also a necessary assumption for V-second



languages. If so, in 0 *en Joan* would not be adjoined to CP, but rather to IP.

There are some ways out of this problem if we want to maintain the Wh-to-COMP hypothesis for 0. One could be the following:

a) Interrogative CPs, unlike relative or [-WH] CPs, allow adjunction.

b) The fact that the adjuncts cannot follow the Wh-element in 0 could be due to V+INFL movement to C<sup>0</sup>. So the prediction would be that IP adjuncts should follow the verb.

As for the assumption in a), I do not know of any interesting way to derive it. In addition this assumption does not receive support from other languages: English, French or German interrogative CPs do not allow adjunction (at least when embedded). As for b), we will see below that a V-to-COMP account of the adjacency requirement between interrogative Wh-phrases and the verb is problematic in Romance languages.

Another possibility would be to exploit CP-recursion. We could assume, with Chomsky (1986-a), that adjunction to Arguments is forbidden. If the embedded CP is an Argument of the main verb, then adjunction to it is forbidden. Suppose, however, that:

a) there is CP-recursion: a CP can optionally subcategorize for another CP: [<sub>CP</sub> C<sup>0</sup> [<sub>CP</sub> IP ] ].

b) when there is CP recursion, the 'that' particle and the relative Wh-phrases have to occupy the upper CP, while interrogative Wh-phrases have to (or may) occupy the lower CP.

c) the lower CP is not, strictly speaking, an Argument of the main verb, but rather a subconstituent of it, so adjunction to the lower CP is allowed.

The assumptions in a) and b) would hold only of Romance NSLs, for the asymmetry between 0/0 and 0 is not found in English, French or German. So, in addition to the unclear nature of these assumptions, we should ask why they are only relevant for some languages. I

---

languages: free adjunction to CP in main clauses would break the V-2 appearance of the language, contrary to fact. This restriction is not likely to be universal, however (recall what we said about Occitan). It could rather follow from parametric options affecting the nature of CPs. Once a language forbids this adjunction, however, it seems to be a strong prohibition.

honestly do not see any interesting way of deriving b) and even less of restricting a) (and b)) to (Romance) NSLs. But even if these possibilities were workable, there are further problems for the interrogative Wh-to-COMP hypothesis.

One is INFL-to-COMP movement. A recurrent pattern across languages where interrogative WH-to-COMP seems to hold without problems is that there is an asymmetry between root and embedded clauses: in interrogative main clauses V+INFL-to-C<sup>0</sup> movement is obligatory while in embedded clauses it is forbidden (or at least not obligatory). In Catalan or Spanish, however, interrogative Wh-phrases are obligatorily adjacent to the verb in both main and embedded sentences. In Italian, adjacency is also required except for subjunctive embedded clauses:

Catalan (same for Spanish):

- (31) a. Què (\*avui) faràs?  
           What (today) will-you-do  
       b. No sé què (\*en Joan) farà  
           Not know-I what (the J.) will-do

Italian:

- (32) a. Cosa (\*oggi) farai?  
           What (today) will-you-do  
       b. Non so cosa (??Gianni) farà  
           Not know-I what (G.) will-do  
       c. Non so che cosa (Gianni) abbia fatto  
           Not know-I what (G.) have-SUBJ done

Putting aside, 0.c) for the moment, we have to account for the adjacency requirement. There are, to my knowledge, two alternative possibilities. One is assuming that adjacency is due

to V+INFL movement to C<sup>0</sup>. Since, except for Italian subjunctive interrogatives,<sup>173</sup> there is no main/embedded contrast, whatever explanation we have for this contrast should be qualified to cover Spanish, Catalan and Italian indicative interrogative embedded clauses.<sup>174</sup> There is an additional problem: V+INFL to COMP predicts there will be V-subject inversion. This is not the case in Italian or Catalan: the order Wh- V subject object is not allowed. It is certainly allowed in Spanish, but, as we suggested in Chapter 1, this fact is not exclusive of interrogative constructions: VSO order is amply used in a variety of contexts. We argued that this is due to the fact that Spanish allows the I-subject to left-adjoin to VP.

Rizzi (1991-a) proposes a solution that accounts for the adjacency requirement in terms of V-INFL to COMP, and for the non-existence of the (WH-) V S O word order in terms of Case theory. He assumes that AGR can assign Nominative by agreement (when the subject is in the specifier of the head containing AGR) or by government (when the subject is in its government domain). To account for the apparent absence of VSO word order, he assumes that:

- INFL to COMP movement destroys the context for Nominative assignment under agreement, since now AGR is in C<sup>0</sup> and stands in a relation of government with the subject.
- Italian or Catalan do not choose the parametric option of assigning Nominative under government, unlike English or German.

In this context, only *pro* or a postverbal subject (which is independently assigned Case by T<sup>0</sup>) is possible:

---

<sup>173</sup> Subjunctive in standard Italian is more profusely used than in any other Romance language. However, colloquial Italian tends to dispense with subjunctive altogether and so, I presume, with possibility (0.c).

<sup>174</sup> Rizzi (1991-a) assumes that in Spanish INFL-to-COMP obligatory movement in embedded interrogatives is due to the fact that, in this type of languages, the [+Wh] feature, which must appear in an embedded [+Wh] C<sup>0</sup>, is always generated in INFL, as it is, in his hypothesis, in root sentences in all languages.

- (33) a. Cosa ha *pro* fatto?  
           What has (he) done
- b. Cosa ha fatto Gianni?  
           What has done G.

Thus the order Wh- V S O does exist but only with null subjects. In English, instead, AGR can assign Nominative both by agreement and by government, so INFL to COMP creates overt verb-subject order with no problem.

I think this theory, appealing as it may be, is unclear in one respect: it is not clear at all that (head) movement should be conceived as destroying a category's capabilities: in the traditional spirit of trace theory, traces are full copies of their antecedents except for phonological content, so the trace of INFL (or AGR) should in principle be able to do the same job as its antecedent. Roberts (1991-a) (who follows Rizzi in this respect) argues that this is true only as far as Baker's Government Transparency Corollary allows it, and Nominative case assignment by agreement does not involve government. The adequacy of this position could be tested in independent constructions. Perhaps the main problem for this theory could be it is too powerful: since Case assignment by government is a parametric option (so English INFL in COMP can assign Nominative by government to the subject in Spec of INFL, in addition to Nominative by agreement), then potential problematic cases could always be neutralized by assuming that government is a legitimate alternative to agreement in precisely these cases.<sup>175</sup> I will not pursue

---

<sup>175</sup> So, for instance, in Chomsky's (1992) proposal, object AGR would have the object in its specifier and, since there is no overt AGR material to its right (in any Spec-head-complement language I know of), this material has likely been carried along by V-movement. So, in the spirit of Rizzi's theory, object AGR always would have to assign Case by government, which, in this case, would void the theory of its initial variation-predicting interest. But this is only a speculation on a not necessarily consistent Case, for Chomsky's object agreement and Rizzi's Relativized Minimality are rather incompatible.

the issue here.<sup>176</sup>

The main problem with Rizzi's proposal is, I think, that it does not say anything w.r.t. the clause-adjunct problem we presented above. The clause-adjunct word order and the lack of VSO word order in interrogatives suggest there is not verb movement to COMP, but rather that interrogative WH-phrases land in a lower position: the specifier of the FC whose maximal projection allows clause adjuncts. Now the question is: could this Specifier be the same which hosts preverbal subjects (and the other phrases we showed can appear in the same position in Catalan and Spanish)?

#### 1.4. A parameter for the FC Hierarchy

In section 1.2., we ended by suggesting that:

- AGR is the guest of some functional category.
- FCs can contain two guests: a semantically contentful category and an AGR.
- the contentful guest in the FC containing (subject) AGR in English and non-NSLs cannot be one imposing restrictions on quantified specifiers (as any quantified DP can be in the spec of that category in English).
- in (some) Romance NSLs, the contentful guest must impose restrictions, based on the

---

<sup>176</sup> Another problem with Rizzi's proposal concerns his characterization of the facts: he simply assumes that (main) Wh-clauses in Romance NSLs involve either a null subject or an inverted subject. In fact, while the former possibility is always an available option, the second is far from being the unmarked strategy: the most natural way of translating 'What has John done?' is by resorting to left- or right- dislocation of the subject:

- (i) Gianni cosa ha fatto?  
G. what has done
- (ii) Cosa ha fatto, Gianni?  
What has done G.

Having an inverted subject ('Cosa ha fatto Gianni', without an intonational dislocation pattern for **Gianni**) is possible but it is rather used to convey a special interpretation: 'What has JOHN done?'.

'strong'/'weak' and [ $\pm$ human] nature of the quantifier, of the kind we saw above. The AGR guest imposes agreement in Italian, but not in Catalan and Spanish.

What category is the contentful mate of AGR in Romance NSLs?

If we compare (most) Romance NSLs with Germanic languages and French, there is an outstanding difference concerning the negative particle: in the latter this particle is apparently in a lower position than the finite inflected verb. It is instead preverbal in (most) Romance NSLs: apparently it is not above the inflected verb, but rather attached to the clitic+V+INFL head. This suggests that, if NegP is a FC, its position with respect to other FCs is parameterized. Let us tentatively assume the following:

a) the hierarchy of FCs is minimally parameterized. The two options are:

(34) a.   ModalP    $\Sigma$ P                   TP

      b.    $\Sigma$ P           ModalP           TP

          where  $\Sigma$  stands for Negation/Affirmation (see Laka (1990))

b) subject AGR is a FC-mate of the highest XP (ModalP or  $\Sigma$ P) in 0.

c) Modal<sup>0</sup> does not impose any restriction on its specifier.  $\Sigma^0$  imposes restrictions on the quantifier properties of its specifier.

d) non-NSLs use option (34).a). The reason is the following: in non-NSLs the I-subject is restricted to be [+anaphoric]. Therefore the only means for the Argument becoming subject to be expressible as a quantifier (and possibly a 'strong' quantifier) is to choose the ModalP FC as the host of AGR (option 0.a)), so that the the Specifier of (the host of) AGR can be a quantifier of any type.

e) Option 0.b) is the unmarked option. It implies that Spec of (the FC containing) AGR is restricted in its quantification possibilities. Since NSLs can have quantified I-subjects (they are [-anaphoric]), Spec of (the FC containing) AGR can be restricted for quantification and no problem of affability arises. So the unmarked option 0 is chosen.

The suggestion in d) that non-NSLs have to choose 0.a) (at least as a preferred option) is, I think, of historical interest: Languages losing the Null Subject status quickly develop postverbal negation (this is what happened in the passage from Old French to Modern French, and from Old English to Middle English:

(35) Idealized data:

- a. Old French:                   Je *ne* mange  
  I not eat
- b. Modern French: Je (ne) mange *pas*  
  I (not) eat not
- c. Old English:       I *no* come
- d. Middle English: I (no) come *nought*

These changes are contemporary of the loss of the Null Subject status. Since they are rather spectacular changes ( they are not trivial reanalysis processes), it is plausible that they were forced by UG. Our suggestion is that a non-NSL has to take the unmarked option of lowering  $\Sigma P$  and promoting ModalP as the top FC in order to make the top FC an appropriate host for subject-AGR, as explained in d) above. The reason why subject-AGR has to be the host of the top FC is probably that this is the only way for left dislocated or *pro*-resumed DPs to be easily reanalysed as AGR-identifiers when the language loses the Null Subject ability.

## 2. Infinitives

Now let us consider infinitives. For infinitives and gerunds we assumed that there is a resp. nominal/adverbial FC between CP and AGRP. We have proposed in the preceding sections that AGRP is actually  $\Sigma$ +AGR phrase; from now on we will use  $\Sigma$ P or AGRP depending on the issue at stake. Non-finite sentences in NSLs do not allow preverbal subjects in general (except in Aux-to-Comp constructions). Not even the dislocated or clitic resumed elements we considered in section 2. are easily allowed in non-interrogative controlled infinitives (the dislocated element can adjoin to the main clause, as in 0.b)):

- (36) a. Espero (???a tots aquests) veure'ls aviat  
I-hope to all these to-see-them soon  
'All these people, I hope to see soon'
- b. A tots aquests espero veure'ls aviat  
To all these I-hope to-see-them soon  
= a.

So, for some reason, the preverbal specifier (which would be Spec of NOMP) is not available for clitic- or *pro*-resumed elements nor is dislocation available. For dislocation, we could assume NOMP does not allow adjuncts. As for Specifiers, let us consider all the possibilities.

In the case of control, we assumed that Spec of AGR is filled with PRO, (whose head raises to NOM<sup>0</sup> in NSLs like Catalan or Italian). So this Specifier is not available for other material. Suppose Spec of NOMP is not available either because it is of another nature and does not licence the occurrence of pronoun resumed elements.

If non-controlled infinitives do not have PRO in Spec of  $\Sigma$ P (which is, we assumed,



AGR-less), this position should be available. In fact, Infinitives with Overt Subjects can have a post-VP subject (as we saw in Chapter 4) or an immediately postverbal subject. The Structures would be as follows: in 0.a) the overt subject is an I-subject; in 0.b) the overt subject is in Spec of  $\Sigma$  and is resumed by the *pro* I-subject:

- (37) a. En [<sub>NOM</sub><sup>o</sup> arribar ] [ <sub>$\Sigma$ P</sub> [ <sub>$\Sigma$</sub>  a casa en Joan ] ]  
           In     to-arrive           home the J.  
           'When Joan arrives/d home'
- b. En [<sub>NOM</sub><sup>o</sup> arribar ] [ <sub>$\Sigma$ P</sub> en Joan<sub>i</sub> [ <sub>$\Sigma$</sub>  a casa *pro*<sub>i</sub> ] ]  
           In     to-arrive   the J.     to house  
           'When Joan arrives/d home'

Clitic resumed elements cannot appear in Spec of  $\Sigma$ P probably because they would not c-commanded their clitic, which is attached to  $\text{NOM}^o$ .

So in conclusion, infinitival Spec of  $\Sigma$  in NSLs is in principle available for pronoun resumed elements, as it should be other things being equal, but there are some restrictions:

- in control structures, Spec of  $\Sigma$  is occupied by PRO.
- in other non-finite structures it can be filled only by an element resumed by the I-subject *pro*: the other (clitic resumed) elements cannot be resumed by a clitic which appears in a higher position.

Infinitival Spec of AGR in non-NSLs is never available for an overt DP in control or raising, because it is occupied resp. by PRO or a null anaphor. It is in ECM and 'for' infinitives. In other cases it is conceivably available, but infinitives with overt subjects (i.e., infinitives not being either controlled or raising) are not attested in non-NSLs except for ECM. The only case of non-finite clauses having non-ECM-Case-marked overt subjects are gerunds. In the case of gerunds, they either use an alternative means of case marking (genitive, ECM) or they are adverbial:

- (38) a. John's loving Mary so much is a drag  
b. I'd prefer John hating Mary a bit  
c. Roddy accepted Eleine's dirty proposals, he being a confessed debauchee.

### 3. Summary

In this Chapter we have argued that:

a) Preverbal subjects in NSLs do not have the same status as preverbal subjects in non-NSLs. This is expected under our previous theory (which predicts only the latter to be AGR-identifiers, while the former will be *pro*/pronoun resumed elements). We have seen that preverbal subjects in NSLs are restricted for quantification and, in some of the languages (Catalan, Spanish) their position is not restricted to subject elements.

b) The reason for the restrictions on quantification must lie on some (not accounted for) restrictions imposed by the head. We have argued that this head, although being the host of AGR, is not exhaustively AGR. It also contains a 'meaningful' FC, which is subject to parametric variation ( $\Sigma P$  being the unmarked option taken by NSLs). Since Spec of  $\Sigma P$  is an appropriate host for interrogative Wh-elements, this is the place these elements move to. Both word order phenomena and lack of Subject-Aux inversion point to that direction.

Although our discussion is too brief to be conclusive, I think the ideas advanced are worth exploring quite independently of the precise and intricate theoretical assumptions in the previous chapters: in my opinion, the null hypothesis that Romance NSLs minimally differ from English (which has often been the departing point of study for many linguistic phenomena) should be taken more carefully, even if we aim at universal grammar.

## References

- Adams, M. (1987) *Old French, Null Subjects and Verb Second Phenomena*, Doct. Diss., UCLA
- Agostini, P.M. (1956) *L'usu di a nostra lingua*, Ajaccio, Corsica.
- Authier, J.-M. (1991) "V-Governed Expletives, Case Theory, and the Projection Principle", *Linguistic Inquiry*, 22-4, pp. 721-40.
- Baker, M. (1983) "Assume GF" and the Order of Morphemes", *MIT-WPL*, 5, MIT, pp. 53-103.
- Baker, M.C. (1985) *Incorporation: A Theory of Grammatical Function Changing*, Doct. Diss., MIT, Cambridge, MA.
- Baker, M.C. (1988) *Incorporation: A Theory of Grammatical Function Changing*, Un. of Chicago Press, Chicago, Illinois.
- N.e.t.
- Baker, M., K. Johnson & I. Roberts (1989) "Passive Arguments Raised", *Linguistic Inquiry*, 20-2, pp. 219-51.
- Belletti, A. (1988) "The Case of Unaccusatives". *Linguistic Inquiry*, 19-4, pp. 1-34.
- Belletti, A. (1991) *Generalized Verb Movement*, ms., Un. de Genève.
- Belletti, A. & L. Rizzi (1988) "Psych-verbs and Theta-Theory", *Natural Language & Linguistic Theory*, 6, pp. 291-352.
- Benacchio, R. & L. Renzi (1987) *Clitici Slavi e Romanzi*, Quaderni Patavini di Linguistica, Monografie 1, Un. di Padova & C.N.R..
- Benincà, P. & G. Cinque (1990) *Su alcune differenze fra enclisi e proclisi*, ms., Un. di Venezia.
- Bickerton, D. (1987) "He himself: Anaphor, Pronominal, Or...", *Linguistic Inquiry*, 18-2, Squibs and Disc., pp. 345-48.
- Bonet, E. (1989) *Postverbal Subjects in Catalan*, ms., MIT.
- Borer, H. (1986) "I-Subjects", *Linguistic Inquiry*, 17-3.
- Brandi, L. & P. Cordin (1989) "Two Italian Dialects and the Null Subject Parameter", in Jaeggli, O. & K.J. Safir (eds.) (1989) *The Null Subject Parameter*, Kluwer Academic Publishers,

SNLLT 15, Dordrecht, pp. 111-142

Burzio, L. (1981) *Intransitive Verbs and Italian Auxiliaries*, Doct. Diss., MIT.

Burzio, L. (1986) *Italian Syntax*, Reidel Publ. Co., Dordrecht.

Cardinaletti, A. (1990) *Pronomi nulli e pleonastici nelle lingue germaniche e romanze*, Doct. Diss., Un. di Padova.

Cardinaletti, A. & G. Giusti (1991) *Partitive ne and the QP-/hypothesis. A Case Study*, 2nd draft, Un. di Venezia.

Chomsky, N. (1981) *Lectures on Government and Binding*, Foris Publ., Studies in Generative Grammar 9, Dordrecht.

Chomsky, N. (1982) *Some Concepts and Consequences of the Theory of Government and Binding*, The MIT Press, Linguistic Inquiry Monographs 6, Cambridge, MA.

Chomsky, N. (1986-a) *Barriers*, MIT Press, Cambridge, Mass.

Chomsky, N. (1986-b) *Knowledge of Language. Its Nature, Origins and Use*, Praeger, New York.

Chomsky, N. (1988) *Some notes on Economy of Derivation and Representation*, ms., MIT.

Chomsky, N. (1992) *A Minimalist Program for Linguistic Theory*, draft, MIT.

Cinque, G. (1990) *Types of A'-Dependencies*, MIT Press, LI Monographs 17, Cambridge MA/London.

Contreras, H. (1987) "Small Clauses in Spanish and English", *Natural Language and Linguistic Theory*, 5-2.

Couquaux, D. (1981) "French predication and Linguistic Theory", in May, R. & J. Koster (eds.) *Levels of Syntactic Representation*, Foris Publications, Studies in Generative Grammar 10, Dordrecht, pp. 33-64.

Delfitto, D. (1990) *Generics and Variables in Syntax*, Tesi di Perfezionamento, SNS di Pisa, Italy.

Delfitto, D. & M. Pinto (1992) "How Free is 'Free Inversion'", in Delfitto, D., J. Schroten & W. Zwanenburg (eds.) *Recherches de Linguistique Française et Romane d'Utrecht*, XI,

- OTS, Utrecht, pp. 1-7.
- Everaert, M. (1986) *The Syntax of Reflexivization*, Foris Publ., Publications in Language Sciences 22, Dordrecht
- Fernández Lagunilla, M. (1987) "Los infinitivos con sujetos léxicos en español", in Demonte, V. & M. Fernández Lagunilla (eds.) *Sintaxis de las lenguas románicas*, El Arquero, Textos Universitarios, Madrid, pp. 125-147
- Fodor, J.D. & I.A. Sag (1982) "Referential and Quantificational Indefinites", *Linguistics and Philosophy*, 5, pp. 355-398
- Fukui, N. & M. Speas (1986) "Specifiers and Projections", *MIT-WPL*, 8, pp. 128-72.
- Galves, C.C. (1991) *Inflected Infinitive and AGR Licensing*, ms., submitted to *Linguistic Inquiry*, Remarks and Replies, Un. de Campinas.
- Giorgi, A. (1992) "Syntactic Structures and Temporal Representations", *15th GLOW-Colloquium*, Lisbon.
- Gràcia i Solé, M.L. (1986) *La teoria temàtica*, Tesi Doctoral, Un. Autònoma de Barcelona, Spain.
- Heim, I. (1990) *Introduction to Formal Semantics*, draft, MIT.
- Hernanz, M.L. (1982) *El infinitivo en español*, Publicacions de la UAB, Bellaterra, Spain.
- Hernanz, M.L. (1991) "Spanish Absolute Constructions and Aspect", *CWPL*, 1991, pp. 75-129.
- Hestvik, H.G. (1990) *LF-Movement of Pronouns and the Computation of Binding Domains*, Doct. Diss., Brandeis Un.
- Hornstein, N. (1984) *Logic as Grammar. An Approach to Meaning in Natural Language*, MIT Press, Cambridge, Mass.
- Iatridou, S. (1986) "An Anaphor Not Bound in Its Governing Category", *Linguistic Inquiry*, Squibs and Disc., pp. 766-72.
- Iatridou, S. (1988) "Citics, Anaphors, and a Problem of Coindexation", *Linguistic Inquiry*, Squibs and Disc., pp. 698-703.
- Jones, M.A. (1990) *Sardinian Syntax*, ms, Un. of Essex, a to be published at Routledge, London.

- Kayne, R. (1981) "On Certain Differences between French and English", *Linguistic Inquiry*, 12, 3, pp. 349-71.
- Kayne, R.S. (1987) *Facets of Romance Past Participle Agreement*, ms, MIT, November 1987.
- Kayne, R.S. (1989) "Null Subjects and Clitic Climbing", in Jaeggli, O. & K.J. Safir (eds.) (1989) *The Null Subject Parameter*, Kluwer Academic Publishers, SNLLT 15, Dordrecht, pp. 239-261.
- Kayne, R. (1991) "Romance Clitics, Verb Movement and PRO", *Linguistic Inquiry*, 22, 4, pp. 647-86.
- Kayne, R. & J.-Y. Pollock (1978) "Stylistic Inversion, Successive Cyclicity, and Move NP in French", *Linguistic Inquiry*, 9-4, pp. 595-621.
- Kenstowicz, M. (1989) "The Null Subject Parameter in Modern Arabic Languages", in Jaeggli, O. & K.J. Safir (eds.) (1989) *The Null Subject Parameter*, Kluwer Academic Publishers, SNLLT 15, Dordrecht, pp. 263-276
- Kitagawa, Y. (1986) *Subject in Japanese and English*, Doct. Diss., Un. of Massachusetts at Amherst.
- Koopman, H. & D. Sportiche (1988) *Subjects*, ms., UCLA.
- Koopman, H. & D. Sportiche (1990) *The Position of Subjects*, ms., UCLA.
- Koster, J. & E.J. Reuland (eds.) (1991) *Long Distance Anaphora*, Cambridge Un. Press, Cambridge, MA.
- Kratzer, A. (1988) *Stage-Level and Individual-Level Predicates*, ms., Un. of Mass. Amherst.
- Kuno, S. (1972) "Functional Sentence Perspective", *Linguistic Inquiry*, 3, 3, pp. 269-320.
- Kuroda, Y. (1988) "Whether we Agree or not: A Comparative Syntax of English and Japanese", in Poser, W. (ed.) *Papers on the Second International Workshop on Japanese Syntax*, CSLI, Stanford Un.
- Laka, I. (1990) *Negation in Syntax: On the Nature of Functional Categories and Projections*, Doct. Diss., MIT.
- Larson, R.K. (1988) "On the Double Object Construction", *Linguistic Inquiry*, 19-3.

- Larson, R.K. (1991) "Promise and the Theory of Control", *Linguistic Inquiry*, 22-1, pp. 103-39.
- Levin, B.C. (1983-a) *On the Nature of Ergativity*, Doct. Diss., MIT.
- Levin, B.C. (1983-b) "Unaccusative Verbs in Basque", in Sells, P. & C. Jones (eds.) *Proceedings of ALNE 13/NELS 13, GLSA*, Amherst 1983, pp. 129-144
- Llinàs Grau, M. (1988) *The Syntax of Some Verbal Sequences in Catalan*, M.A. Diss., University College, London
- Maraldi, M. (1983) "New Approaches to Accusative Subjects", in Pinsker, H. (ed.) *Latin Linguistics and Linguistic Theory*, Studies in Language Companion Series, Vol 12, J. Benjamins Pub. Co., Amsterdam/Philadelphia, pp. 167-76.
- May, R. (1985) *Logical Form. Its Structure and Derivation*, the MIT Press, Cambridge, MA.
- Montalbetti, M.M. (1984) *After Binding. On the Interpretation of Pronouns*, Doct. Diss., MIT.
- Moro, A. (1988) "Per una teoria unificata delle frasi copulari", *Rivista di Grammatica Generativa*, 13, pp. 81-110.
- Moro, A. (1991) "The Raising of Predicates: Copula, Expletives and Existence", XV MITWPL, MIT, Cambridge, MA.
- Moro, A. (1992) "The anomaly of copular sentences", *Un. of Venice Working Papers*, 8.
- Motapanyane, V. (1989) *La position du sujet dans une langue à l'ordre SVO/VSO*, Mémoire de Prédoctorat, Un. de Genève.
- Ouhalla, J. (1990) *The Syntax of head Movement: A Study of Berber*, Doct. Thesis, Un. College, London.
- Ouhalla, J. (1991) "Functional Categories and the Head Parameter", *14th GLOW Colloquium (on Heads)*, Leiden.
- Ortiz de Urbina, J. (1989) *Parameters in the Grammar of Basque. A GB Approach to Basque Syntax*, Foris Publ., Studies in Generative Grammar, 33, Dordrecht
- Pesetsky, D. (1982) *Paths and Categories*, ms, Doctoral Diss., MIT
- Picallo, C. (1985) *Opaque Domains*, Doct. Diss., City Un. of New York.
- Picallo, C. (1991) "Possessive Pronouns in Catalan and the Avoid Pronoun Principle", *Catalan*

WPL-1991, pp.211-34.

- Poletto, C. (1991) "L'ausiliare avuto e la proiezione sintattica di aspetto", *17o Incontro di Grammatica Generativa*, Trieste.
- Pollock, J.-Y. (1989) "Verb Movement, UG, and the Structure of IP", *Linguistic Inquiry*, 20-3, pp. 365-424.
- Raposo, E.P. (1987-a) "Romance Infinitival Clauses and Case Theory", in Neidle, C. & R.A. Núñez Cedeño (eds.) *Studies in Romance Languages*, Foris Publ., Publications in Language Sciences 25, Dordrecht, pp. 237-249.
- Raposo, E.P. (1987-b) "Case Theory and INF-to-COMP: The Inflected Infinitive in European Portuguese", *Linguistic Inquiry*, 18-1, pp. 85-110.
- Reinhart, T. & E.J. Reuland (1991) "Anaphors and Logophors: An Argument Structure Perspective", in Koster, J. & E. Reuland (eds.) *Long Distance Anaphora*, Cambridge Un. Press, Cambridge, MA.
- Reuland, E.J. & A.G.B. ter Meulen (eds.) (1987) *The Representation of (In)definiteness*, MIT Press, Current Studies in Linguistics 14, Cambridge, Mass.
- Rigau, G. (1988) "Strong Pronouns", *Linguistic Inquiry*, 19, 3, pp. 503-511
- Rigau, G. (1991) "On the Functional Properties of AGR", *Catalan Working Papers in Linguistics-1991*, pp. 235-60.
- Rigau, G. (1992) *La legitimación del sujeto en las construcciones temporales de infinitivo*, ms., Un. Autònoma de Barcelona, to appear in Proceedings of the 2o Coloquio de Gramàtica Generativa, Vitoria-Gazteiz, April-1992.
- Rizzi, L. (1982-a) "Comments on Chomsky's Chapter: On the Representation of Form and Function", in Mehler, J., E.T.C. Walker & M. Garrett (eds.) *Perspectives on Mental Representation*, Lawrence Erlbaum Ass., London, pp. 441-53.
- Rizzi, L. (1982-b) *Issues in Italian Syntax*, Foris Publ., SGG 11, Dordrecht.
- Rizzi, L. (1986) "Null Objects in Italian and the Theory of pro", *Linguistic Inquiry*, 17-3.
- Rizzi, L. (1989-a) *Relativized Minimality*, MIT Press, Linguistic Inquiry Monograph Series,



Cambridge, MA.

- Rizzi, L. (1989-b) *On the Anaphor-Agreement Effect*, ms., a Rivista di Linguistica, Un. de Genève.
- Rizzi, L. (1991-a) *Residual Verb Second and the Wh-Criterion*, ms. Un. de Genève.
- Rizzi, L. (1991-b) "Proper Head Government and the Definition of A-positions", *14th GLOW Colloquium*, Leiden, the Netherlands.
- Rizzi, L. & I. Roberts (1989) "Complex inversion In French", *Probus*, 1, 1, Foris Publications, pp. 1-30.
- Roberts, I.G. (1987) *The Representation of Implicit and Dethematized Subjects*, Foris Publ., Linguistic Models 10, Dordrecht
- Roberts, I.G. (1989) "Case Assignment Parameters and the History of Inversion", *12th GLOW Colloquium*, Utrecht, Holland.
- Roberts, I. (1991-a) *Verbs and Diachronic Syntax*, ms. Un. de Genève.
- Roberts (1991-b) Class-notes, Un. de Genève.
- Rosselló, J. (1986) *Gramàtica, configuracions i referència. Per una teoria alternativa del PRO-drop romànic*, Tesi Doctoral, Un. de Barcelona.
- Rothstein, S. (1983) *The Syntactic Forms of Predication*, Doct. Diss., MIT, Cambridge.
- Saccon, G. (1991) *A Semantic Condition on Agreement*, ms., LSA Meeting.
- Sauzet, P. (1989) "Topicalization et prolepse en occitan", *Revue des Langues Romanes*, XCIII, La structure de la proposition dans les langues romanes, Un. Paul Valéry/CNRS, Montpellier, pp. 235-273.
- Solà i Pujols, J. (1989) *Trets i categories lèxiques: moviment del verb als nusos inflectius*, Master Diss., Un. Autònoma de Barcelona.
- Speas, M.J. (1986) *Adjunctions and Projections in Syntax*, Doct. Diss., MIT.
- Sperber, D. & D. Wilson (1986) *Relevance: Communication and Cognition*, Harvard Un. Press, Cambridge? MA.
- Stowell, T.A. (1981) *Origins of Phrase Structure*, Doct. Diss., MIT.

- Taraldsen, K.T. (1980) *On the Nominative Island Condition, Vacuous Application, and the that-trace filter*, distributed by the Indiana Un. Linguistics Club.
- Torrego, E. (1989) *Experiencers and Raising Verbs in Spanish*, ms., Un. of Massachusetts Boston.
- Vikner, S. (1985) *Parameters or Binder and Binding Category in Danish*, Working Papers in Scandinavian Syntax 23.
- Zagona, K. (1982) *Government and Proper Government of Verbal Projections*, Doct. Diss., Un.of Washington.
- Zanuttini, R. (1991) *Syntactic Properties of Sentential Negation. A comparative Study of Romance Languages*, PhD Thesis, Un. of Pennsylvania.
- Zribi-Hertz, A. (1990-a) "NP lui-même", *Recherches Linguistiques*, XIV, Uni. de Metz/Ed. Kluicksieck, Paris. pp. 377-402.
- Zribi-Hertz, A. (1990-b) "Lui-même argument et le concept de 'pronom A'", *Langages*, 97, Larouse, pp. 100-127.