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 setembre de 1992

Als meus pares i germans, que han tingut la sort que no els atabalés amb la meva tesi.

A la Vivien, que es mira les meves especulacions amb curiositat divertida.

A la Gemma Rigau, que m'ha donat àmims i suport a les verdes i a les madures.

A la meva família GB i altres animals, amb el desig que convergim sense estavellar-nos.

Al meu ordinador Ethelred, que, sense tenir el cor dur ni finestretes, sempre m'ha servit fidelment.

All data in this thesis are real, but they have been idealized in order to avoid the gibberish that may plague linguistic corpuses. People cited are also real, but they should not be held guilty for my interpretation of their theories. The views and theories introduced have mostly been formulated after careful and patient thought, but they should not be attributed any pretentious claim to truth. Copies of this thesis might possibly be subject to the condition that you pay the xerox owner.

"In the event of this thesis being brought to the cinema, I am pleased to give my thanks in advance to all the people that will take part in the film. Their collaboration will be highly esteemed because of the risk they will face of fatally tarnishing their prestige."

(plagiarized from Pere Calders (1986) **Ronda naval sota la boira**, Edicions 62, Barcelona, 1992, p. 9)

Table of contents

Acknowledgements.....ix
List of Abbreviations and Glosses.....x

Chapter 1 - Introduction.....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 - Burzio's Generalization and EPP.....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 Teory.....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 - AGR and Subjects.....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 - Infinitival Constructions.....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 - Preverbal Subjects in NSLs.....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, Hálldor A. Sigurdhsson, Isabel Herrero, Joana Louro, Kjel Åke Gunnarson, Lea Nash-Haran, Liliane Haegeman, Patrick Sauzet, Sabine Iatridou, Sten Vikner, Tali Siloni, Ur 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 abreviations 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	Binding Domain
BG	Burzio's Generalization
BT	Binding Theory
CLLD	Clitic Left Dislocation
DP(s)	Determiner Phrase(s)
EA(s)	External Argument(s)
EL(s)	Eragtive Language(s)
EPP	Extended Projection Principle
FC(s)	Functional Category/ies
IOS(s)	Infinitive(s) with an Overt Subject
IOS-INV(s)	IOS(s) with an Inverted subject
ISH	Internal Subject Hypothesis
non-NSL(s)	non Null Subject Language(s)
NSL(s)	Null Subject Language(s)

Glosses:

SE Reflexive particle not being a clitic (at least not a clitic of the Romance type) (e.g., Italian se, German sich, etc.).

SELF A type of crosslinguistically recurrent of word which, among other uses, is typically a member of reflexive DPs (German selbst, Italian stesso, Catalan mateix, etc.).

SI- 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.[1] 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 hypothesis 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.[2] My opinion is that ideally the theory should say little about it, because

---

[1] 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)).

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

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 behavior in INFL-to-COMP contexts.[1] 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.

---

[1] 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 for the facts essentially consists in a reinterpretation of the facts 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.

On the other hand, 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, as far as it misses some generalizations on subject inversion and expletives. Finally, I present a binding-theoretical approach to the nature of inverted subjects which will afford 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 been also developed in some way in Zagona (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 (1).a)) or as an adjunct to the VP (as in (1).b)):

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

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

a) locality in θ-marking: all Arguments of a verb, even the EA, are projected and θ-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.[4]

b) a more straightforward notion of θ-position: there are no A-positions being optional θ-positions (as the specifier of INFL was in previous hypothesis). θ-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

---

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

as a super-set of the set of θ-positions), but even so, there is no A-position being an optional θ-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. Such a kind of theory 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.

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 would be close to the ISH in (1).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).[5] 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

---

[5] 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.

subjects which are EAs will occupy a position peripheral to VP, while inverted subjects being internal Arguments will appear inside the VP.* 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.

b) Word order with respect to the V° cannot be telling, since V° 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.

---

* 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 θ-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.

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 is senseless). 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 condueix cap cotxe$_i$ el seu$_i$ propietari
Not drives no car the his owner
intended reading: 'No car is driven by his 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 keep on the safe side, I prefer to adopt the view 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$_i$ [$_{VP}$ e$_i$ [$_{VP}$ eats bananas]]
b. John$_i$ is [$_{AP}$ e$_i$ [$_{AP}$ apish]]
c. John$_i$ is [$_{NP}$ e$_i$ [$_{NP}$ a monkey]]
d. John$_i$ is [$_{PP}$ e$_i$ [$_{PP}$ 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° 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 (4) 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.[7]

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

(5) a. Avui s'encarregarà dels nens en Joan
Today will-take-charge of-the children the Joan
'Today JOAN will take 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 (6)) and Locatives not being the first Argument -(7)- seem to allow free word order w.r.t. the inverted subject (perhaps with a slight preference for the PP-inverted-subject order):

---

[7] 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 -(8)- also allow free word order, although perhaps the order inverted subject-adjunct is slightly preferred:*

(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

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

---

* 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.

(8) 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 it 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).[9]

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 (10)), the preferred word order is locative-subject, contrary to what we would expect:

---

[9] 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 to-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

(11) 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 (11) and (12) 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 thing 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 θ-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 (16) is similar to that 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).[19] 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 contend 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 interpretation of the Focus element.

---

[19] 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.

As argued in Kuno (1972), the element interpreted as Focus has what we 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.[11]

Now consider the sentences in (17) 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 to such a conflict, subject inversion will be alright with

---

[11] 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.

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 (20):

(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 (23), it is very thin:[12] in fact neither of the options in (23) should be very natural if asserted out of a context. No such

[12] The Italian speakers I asked found no such contrast.

contrast will appear in (24), which is a more felicitous sentence from the pragmatical point of view:

(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 poin 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 (25).b) (for (25).a)):

(25) a. I saw JOHN
b. John$_x$ [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 contend 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 bois 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 would 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 (28).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 (28).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 (29):

(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 is, as far 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 (31) 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 that 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 which AGRP dominates TP, as in Belletti (1991). The structure of the sentence is as in (34):

(34) $[_{AGRP} \cdots [_{TP} \cdots [_{VP} \ ] \ ] \ ]$

where other possible FCs (such as NEGP and Object-AGRP) are omitted.[13] 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 (34) 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°-movement. The general fact is that NSLs seem to exhibit longer V°-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°-movement.

Specifically, two facts will naturally follow from our theory:

---

[13] 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°, which is the highest FC in the hierarchy. The facts in both (35).a) and (35).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 Null case 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 rather 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 be one of resumptive strategy. 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° which assigns Nominative in both the agreement and government strategy. In Chapter 5, however, we will propose that T° is the basic Nominative-assigner by government, and that AGR° assigns government-Nominative only by combining with T°. AGR°, 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 θ-role.

The biconditional in (1) can be split into two different implications:

(2) a) External θ-role -> Accusative
b) Accusative -> external θ-role

In fact, (2).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 Case-less.

Still, (2).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$_{Acc}$-works/sleeps
'He does not work/sleep for eight hours'

c. Els cent metres els corre fàcilment
The hundred meters them$_{Acc}$-runs easily
'He easily runs the hundred meters'

(2).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, (2).b) strongly undermines the standard Case-theoretical account for DP-movement, namely, that a DP moves to obtain Case: it follows from (2).b) that a DP will fail to have Accusative Case precisely in the cases where DP-movement to obtain another Case is available,[1] the EA being absent. Consider the following D-structures:

(4) a. John [$_{VP}$ melt the ice] (transitive)

b. - [$_{VP}$ be melted the ice] (passive)

c. - [$_{VP}$ melt the ice] (unaccusative)

(5) a. J. [$_{VP}$ believe [$_{IP}$ the ice to VP]] (ECM)

b. - [$_{VP}$ be believed [$_{IP}$ the ice to VP]] (ECM-passive)

c. - [$_{VP}$ seem [$_{IP}$ the ice to VP]] (raising)

In the a. examples we have verbs assigning an external θ-role: as predicted by BG, Accusative Case is available. In the c. examples there is no external θ-role, while in the b. examples the external θ-role is not projected (at least as a visible DP).

[1] Abstracting away from PRO, if it does not obtain Case. In fact we will assume it does.

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, (2).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] -> DP-movement
- θ-criterion: DP-movement -> [-Ext. θ-role][2]
- Burzio's Gen.: [-Ext. θ-role] -> [-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.

Another question that should be addressed concerning BG is its universality. Are all languages subject to BG? If we take the

[2] Within the Internal subject hypothesis, NP-movement requires absence of external θ-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.

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)		External Argument	Internal Argument
ELs	Trans.	Ergative Case	Absolutive Case
	Unacc.	-	Absolutive Case
non-ELs	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.[3] 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.

---

[3] 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 (2).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 (8) 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$_{nom}$ sank t$_{acc}$

This first approach cannot work without further developments. First of all, we want to exclude cases such as (10):

(10) *The boat$_i$ to sink t$_i$-*acc* 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 (11), where the expletive satisfies the EPP and the internal Argument remains in situ and receives Accusative:

(11) *There developed [the problem]$_{acc}$
(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 θ-roles, subject to the condition they have only one Case. So both (*a man, t*) in (12).a) and (*there, a man*) in (12).b) would be instances of CHAINS:

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

In both CHAINS in (12) there is a single θ-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.*

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

---

* We could alternatively define (14) 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 notion of 'heading a CHAIN' is not different from the notion of 'being member of a CHAIN'.

internal Argument enters an INFL-CHAIN is excluded in some way[9]); 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 θ-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$_{Nom}$ developed t_{Acc}
b. *There$_{Nom}$ developed the problem$_{Acc}$

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.

Both examples in (16) 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:

---

[9] 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.

(18) a. *It is strange there to develop the problem$_{Acc}$

b. *It is strange the problem to develop t$_{Acc}$

c. *It seems there to have developed the problem$_{Acc}$

d. *It seems the problem to have developed t$_{Acc}$

In (18) 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 (14). 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:

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

(14) *Extended Projection Principle:*

Spec of INFL has to be a member of a CHAIN.

(15) A CHAIN must contain one and only one θ-role.

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

(19) Only the head of a maximal CHAIN can be Case marked.

Notice that (15) and (17) are standard assumptions for Chains (therefore they should also be for CHAINS). (13) is a theoretically desirable result on the grounds of simplicity and

generality.* The extended projection principle in (14) is also a natural development once we adopt a theory of CHAINS.

Notice also that (14) in conjunction with (15) 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 (15) 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 θ-role: a CHAIN headed by PRO needs a θ-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
- null expletive-inverted subject CHAINS (Italian: *pro* ha telefonato *Gianni* '*pro* has telephoned Gianni').

---

* 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).

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 θ-marks it, that indefinite DP's in these constructions are restricted to θ-positions, (see (20).a)), and specifically to θ-positions which are θ-marked by the verb which assigns partitive (see (20).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 [$_{SC}$ llibres oberts]
Not leave books open
'Don't leave (any) books open'
b. Aquesta política manté [$_{SC}$ gent en atur]
This policy keeps people in redundancy
'This policy keeps some people redundant'[7]

[7] 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 (21) 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 θ-marking the SC subject (i.e., the SC subject receives a compositional θ-role). If so, Belletti's contention that Partitive Case is an inherent (hence θ-associated) Case is not falsified.

There are two other cases where the indefinite subject of a Small Clause could possibly not be θ-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):[4]

(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

[4] For these cases another constraint seems to be at work: individual-level SCs do not allow a partitive.

like a Case at all, if Case is conceived as a merely formal entity of no interpretative import: Partitive forces the DP receiving it to be indefinite.* 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 (17) above):

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

(23) and (24) together allow the three only cases of expletive CHAINS which are attested:

---

* 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.

- 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 in Spec of INFL) and the indefinite is assigned Partitive. The two Cases do not conflict according to (23) 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 a kind of 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 (28); 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.[10] But they seem to falsify the claim, essential to the theory above, that there is always

---

[10] 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.

an INFL-CHAIN containing an Argument, unless we assume that oblique Arguments (such as the Dative in (28) 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 getantzt wurde
that yesterday danced was

Cardinaletti (1990) assumes that the kind of verbs in (28) 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 (28). 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 (28) and (29). This, however, amounts to admitting that expletive-Argument CHAINS do not have a uniform pattern. In the case of (28), it amounts to admitting 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 -(29)-, 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 amounts to loosening the notion of CHAIN a great deal.

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 (19), repeated here:

(19) 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 (19) has to be merely stipulated.

A way to derive (19) that comes to mind is the following: there is a kind of requirement to the effect that Nominative is 'preferred' to Accusative. So, to exclude (16) above, repeated here:

(16) a. *The problem developed t_{Acc}
b. *There developed the problem$_{Acc}$

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 (19) from a Case-hierarchy strategy.

Another set of theoretic 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*$_i$ came *a man*$_i$
LF: A *man*$_i$ came t_i

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 (33):

(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 (34) for an S-Structure like (32), 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) [$_{DP}$ there [*many* 1. 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 are*n't* here

For many speakers (35) cannot be interpreted with *many* having narrow scope, while the narrow scope interpretation in (32) 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 (36), 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.[11]

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

Now, let us 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 &

---

[11] 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.

Sportiche's papers (1988)/(1990). Let us call I-subject (suggesting internal subject) the Argument in this basic position which is coindexed with AGR. Let us assume that inverted subjects in Italian are Arguments in this basic position, as we proposed in Chapter 1, hence they are I-subjects.

Let us pose the following question: What is the nature of I-subjects with respect to the features {±pronominal} and {±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 and sentential subjects:

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

		English	Italian
- 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 (37) 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 (37) 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 (38)):

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

(39) differs only in three respects from the table of values in (37):

(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? Let us start with prediction (40).a). Consider a sentence like (41):

(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.[12] But in fact, constructions such as (41) share some properties with inverted subjects in Italian. (41) has an 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.[13]

The obvious objection one can pose to the contention that *himself* is the I-subject in (41) 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 (41), namely (43):

(43) En Joan ho ha fet *ell (mateix)*
The Joan it-has done he (SELF)

---

[12] 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.

[13] 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.
Parallelly, 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 emphatic 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.

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

(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:"

(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 really theoretical term being I-subject which is neutral w.r.t. both the features [±anaphoric] and [±overt].

" **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) [$_{DP}$ El ministre en persona] i [$_{DP}$ 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.

Now let us consider (40)-b). The prediction is that anaphoric I-subjects are not possible in Catalan. This prediction is also apparently fulfilled:

(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 (41)/(46) really significant? According to (38), 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:[15]

[15] 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.

(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.)[16]
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.)[17]
G. it-has done he (SELF) / SE SELF

e. O Joao fé-lo *ele (mesmo)/*si mesmo* (Portugu.)
The J. did-it he (SELF) SE SELF

f. Ion a scris *el insusi* acest proiect (Romanian)
I. has written he SELF this project

g. o jánis to káni o *idjos /*o eaftós tu* (Greek)
the J. it-did the HE-SELF/ the self of-his

h. Jon, ordea, *berak /*bere buruak* egin du (Basque)[18]
J. instead HE-SELF/ his self done has

---

[16] In modern colloquial Occitan **mateis** has been replaced in most dialects by the French borrowing **même**. This is a merely lexical fact.

[17] 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, 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.

[18] **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.

In all the above examples, the reflexive forms are unacceptable as emphatic I-subjects.[19] The acceptable emphatic I-subjects in (47) are possible in contexts where they are not bound (preverbal subjects, dislocated positions, complement positions).[20] 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:

---

[19] 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 either provide clear examples relevant for the theory, 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.

[20] 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

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)

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.[21] 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:

(52) a. *O Yanis*$_i$ theli i Maria na voithisi *ton idhio*$_{i/*j}$
The J. wants the M. that helps him-himself
'Yanis wants Maria to help *him himself'*
(Greek, Iatridou (1986:767))

b. *John*$_i$ told *Bill*$_j$'s sister that *he himself*$_{i/*j}$ 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 (52), 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 contextual 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

[21] See Bickerton (1987) for English he himself; Iatridou (1986) for Greek o idhios.

antecedents can happen to be present in the sentence and even c-commanding the emphatic/logophoric element).[22]

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:

(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/méme*, *el insusi*) or even the bare pronominals (Catalan *ell*, Occitan *el*): then 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.

---

[22] 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.

Non-NSLs should instead have emphatic anaphors in constructions equivalent to the ones in (47), as predicted by the generalization in (38). Let us consider the following examples:[23]

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

a. John did it *himself* (English)

b. Hans hat es *selbst* gemacht (German)[24]
   H. has it SELF done

c. Hans har gjort det *selv* (Danish)[25]
   Hans has done it SELF

d. John har gjort det *själv* (Swedish)[26]
   Hans has done it SELF

e. Jón gerdhi thetta *sjálfur* (Icelandic)
   J. did that $SELF_{NOM}$

f. Jean l'a fait *lui(-même)* (French)
   J. it-has done he (himself)

---

[23] 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.

[24] 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'

[25] 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.

[26] Norwegian behaves exactly like Swedish.

The analysis for the above examples is far from trivial. We will see, however, that under a reasonable interpretation it supports the generalization in (38): 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 element in the Romance NSLs (*lui-même*). But French strong pronominals function as anaphors,[27] 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_{arb}, etc.), the strong reflexive is *soi(-même)* (see (56).a)), and correspondingly the emphatic I-subject is *soi -même* (see (56).b)):

---

[27] 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.

(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'

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 (38) 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 (58) 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 (54).a/.b/.c) is a SELF element[28] left floating by DP-movement to Spec of AGR. So the structure of (59).a) would be (59).b) (and the same could be claimed for the other two languages):

(59) a. John has done it himself

b. $[_{AGRP}$ John$_i$ $[_{AGR'}$ has done it $[_{DP}$ t$_i$ himself]]][29]

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*,

---

[28] 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**).

[29] 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).

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 (60), floating is relevant for interpretation, in that it allows 'referential emphasis'. The *ec* accompanying the floating element is, however, a 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 (38), 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,[30] 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'

[30] 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.

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 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:[31]

---

[31] Even for **lui-même**, its floating status could be questioned in cases 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

(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)*

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.

* * *

(38) 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.[32] So, while English allows both (64).a) and b), Catalan and NSLs (which have SELF elements adjoined to DP's) only allow (64).c), not (64).d):

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.

[32] 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.

(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. **En Joan* ho farà *mateix*
The J. it-will-do SELF

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

(65) a. *Juan mismo* lo hará /**Juan* lo hará *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* fé-lo /**Ele* fé-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 empathic 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 insusi* is apparently the same as Italian *lui stesso* (he SELF), it is not equivalent: the SELF element is actually *el insusi* in cases like (66).a):

(66) a. [$_{DP}$ Ion el insusi] (cf.: **Ion insusi)
Ion himself
b. Ion a scris el insusi acest proiect
I. has written *el insusi* this project

Thus *el insusi* in (66).b) could be analyzed as $[_{DP}$ t el insusi], so that Romanian would not have reflexive I-subjects (si insusi) 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 insusi* 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 $[_{DP}$ 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 [$_{DP}$ *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).[33] 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° raised to INFL° and the post-verbal subject left adjoined to VP.[34]

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

---

[33] 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.

[34] 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.

(69) 'Amala Yusuf/huwa/*nafsuhu 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:

(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 $SELF_{Nom}$
b. Karaha Yusuf nafsahu
Hated Y. $himself_{Acc}$
'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. Suppose we 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,[19] *lui* when used as an emphatic I-subject precedes the object, while *lui stesso* follows it:

(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 ring a bell 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/* $[_{DP}$ 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 θ-position. Perhaps for head-initial languages all θ-positions are to the right of V, and pre-VP I-subjects (as in MSA, Romanian, Spanish) are always derived by short movement:

---

[19] G. Longobardi's judgements were clear cut.

(73) AGR°_{i} ... (I-subject) [$_{VP}$ 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).

On the other hand, 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 (74) as possible I-subjects (some of them only have (74).a) and (74).b)), whereas NSLs have the options in (75) (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 *lui-même*]
J. has done the work		SELF

c. overt anaphor:

Jean a fait le travail		$[_{DP}$ *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 (75) 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. 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 contend that the I-subject, once coindexed with AGR, has to obtain Case from AGR, either from AGR° (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° in NSLs or Spec of AGR in non-NSLs.

c) The link between expletive and Argument was problematic for Binding Theory, and 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 katuv '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 (4).a) are excluded because AGR is not coindexed with any DP or CP, sentences such as (4).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 [$_{CP}$ that...]

Let's call the DP/CP coindexed with AGR the I-subject. Let us review the principle in (1) to see how it works. First of all, (1) 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 (1) 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 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 -(8).a)-, while no controller is available in the unaccusative structure -(8).b). But the implicit Argument cannot control a complement clause -(8).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.[1]

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 θ-role is relevant is due to the independently assumed fact that the EA is projected in a more prominent position.

(1) 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

---

[1] 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).

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.

(1) 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 (1), 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 (1) would not be fulfilled, and the output sentence would be ungrammatical. Suppose, though, that (1) is parameterized as in (9), 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 getantzt 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.[2]

The present account, based on parameter (9), gives a unitary explanation for the existence of impersonal passives and the class of verbs in (11) in the same language: languages lacking impersonal passives do not have verbs of this kind. Icelandic

---

[2] 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.

would be another instance of such a correlation.' 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 (11) 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 (11) (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.

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 (1) is conceived for) and Ergative Languages (ELs). I will only consider one such case of ergativity, which is perhaps not most typical: Basque.

' 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.

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[4]). 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.[5]

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:

(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)

---

[4] 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.

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

The b) option would be the unmarked one and is equivalent to (1) 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 (12).

The obligatory character of AGR-coindexation in (12), which accounts for Burzio's Generalization when b) is taken, would account, when option a) is taken, for a 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 (12)-option a), to a designated position (object), which has thus 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, as far 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
Lat-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 (11) in general, it seems that these languages set the other parametric option in (9). 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'.[6] 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 (1) 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:

---

[6] 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 is Catalan, the CPs are likely the 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 (15), 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 (15) 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.[7] 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 would be nothing but a theoretical artifact used to our convenience.

The solution we propose to this puzzle 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.[8] We will see that this distinction is plausible and makes the correct distinctions.

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

---

[7] Recall that the existence of quasi-Arguments is crucial for our account: without quasi-Arguments, all weather verbs would be problematic.

[8] 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:*

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

It is clear that (18) is synonymous with (17).a) and not with (17).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

* 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.).

its CP to 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 neat, probably because the interpretation of *semblar*-2 stands in an relation of subset 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 is clearly means 'mere appearance' situation. Alternatively, *semblar-2* (= *semblar* with Accusative clitic) cannot be used to make confirming reply:[10]

(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:

(23) It looks as if ...

[10] Gemma Rigau pointed out to me the relevance of these examples.

In (23) 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 *semblar*-2 type involve a quasi-Argument roughly denoting 'appearances'.

Since *semblar*-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 *semblar*-2 can have a dative controller:

(24) a. Em$_i$ sembla PRO$_i$ estar somiant (Catalan)
Me-seems to-be dreaming
'I have the impression of being dreaming'

b. Me$_i$ parece PRO$_i$ estar soñando (Spanish)
Me-seems to-be dreaming

c. Mi$_i$ sembra di PRO$_i$ éssere in un sogno (Italian)
Me-seems of to-be in a dream

The verb in (24) is clearly *semblar*-2 as far as it conveys the 'mere appearance' meaning.[11]

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 *semblar*-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.

---

[11] I cannot explain why the English equivalent of semblar-2 ('to look/sound/etc.') cannot involve control:
*It looks to me to be dreaming

Nevertheless, *semblar*-2 does not allow a full DP (or CP) as an alternative to the (purported) quasi-Argument:

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

So the reason the quasi-Argument is the only option for *semblar*-2 has to be another. My suggestion is the following. Consider *semblar* when used without a clausal complement, as in (26):

(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 (26) 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 plogui
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 plogui*):

(28) Sembla $[_{SC}$ *quasi-A* $[_{CP}$ que plogui]]

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.[12]

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':[13]

---

[12] 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.

[13] As far as in the English translation of ((30).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 ((30).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:

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

(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 above hypothesis: the way we presented the contrast in (17), 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: (17).a) can be used in the 'mere appearance' interpretation. So the contrast is half-way:

- semblar-1: indicative
- semblar-2: subjunctive (or indicative)[14]

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 (31).a) and (.b) would be cases of *semblar*-2:

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 ((30).a) and (i) it would be milder.

[14] 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 ((17).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, (31).b) appears to more strongly suggest that 'raining is a mere appearance' than (31).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 (32).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 (32).b)):[15]

---

[15] 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 plogui
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 plogui)
Seems (, that it rains)

The same happens with another verb (*resultar* 'to turn out 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 build 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.[16] And in fact, predicates of this kind can be paraphrased as adverbial adjuncts ('apparently' for *semblar*-1, 'in fact' for *resultar* are close paraphrases).[17] There are other cases of CPs that cannot be pronominalized or dislocated. One is the following, in Catalan:

---

[16] The idea that raising verbs are adjunct predicates is proposed by Hernanz (1982), Rothstein (1983) and Torrego (1989).

[17] 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 (36), *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)

Semblar-1	Semblar-2
'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
= *seem* (or *seem*-1)	= *look/sound*/etc. (or *seem*-2)
similar verbs: *resultar*, *diu que*.	

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.

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.[18]

[18] In those dialects, it is disappearing even in the complement CP of optative verbs:
Voglio che vieni
I-want that you-come-IND

On the other hand, the finite CP Argument of *sembrare* can be dropped or dislocated, in contrast with Catalan (and Spanish) (34):

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

Since (39) 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 (1)/(9) 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 (1):

(1) 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, (1) 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 (1) predicts which A-Chains will be formed and what length they will have.

The standard theory of A-movement 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 constraints the DPs to move (or form expletive CHAINS) and the other constraints filter the undesired cases (super-raising, movement to a θ-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, (1) clearly overlaps with the other constraints on A-movement: the way it is formulated (1) 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 (41) 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 (1), 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 (1) 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 always can assume that the ECP and (1) 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 (1) 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 (42).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 (42).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 (42).a) can adjoin to the embedded VP as a first step, and then move on:

(43) $[_{VP}$ t' $[_{VP}$ read t]]

In the relevant structure (43), *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 (42).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 [$_{VP}$ 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 θ-role transmission takes place between their members: for the θ-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.[14]

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 θ-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-

[14] 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.

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 (1), 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 Burzios's Generalization. I do not contend that, given that (1) 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 (1), 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-rasing 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.[20] So there is a clear contrast between A- and A'-ECP violations:

(45) a. ?**Which book* does it seem that e was leafed e through

b. ***This book* seems that it/e was leafed e through

[20] In fact some speakers accept them: so it could be that for 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.

Summarizing so far, even if principle (1) is redundant with an ECP account for A-movement restrictions, the fact that the latter is not sufficient without some further assumption makes (1) a possible candidate to cover the gap. Still, if the ECP and other principles proved sufficient to constraint A-movement, we could try to reduce the power of (1) to avoid redundancy.

One way of doing so could be not to stipulate the prominence requirement on the DP/CP, so that (1) would be reduced to (46)

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

(46) 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 (46) 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 (47). Even the requirement that the I-subject must be in the c-command domain could be given up, since A-movement or (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 (46), (47) or (48) may be problematic for our account for German impersonal verbs, which was based on the parameter (9):

(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)*.

If we parameterize (46) or (47) 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 (49).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 (49).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 (1) and (9), leaving as an open question whether they could be simplified as (46), (47) or (48).

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. Remember that we cannot stipulate that Nominative is preferred to Accusative because Nominative is not the only Case option for a subject. What we want to make sure 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 contend 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 contend 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,[21] in the sense that it has to be able to display a complete range of phi-features: if AGR° 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°-clitics and, eventually, become part of the AGR° morphology: this is the standard explanation for Northern Italian dialects' evolution. 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° affix.

To implement this idea, let's assume the principles in (50).a) and (50).b) and the parameter in (50).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).[22]

c. AGR°/Spec of AGR is the AGR-identifier of AGR.

Suppose that when a language has a rich AGR° morphology, the first option in (50).c), which would be the unmarked one, is taken. This would be the case in NSLs. When AGR° is morphologically poor, the second option of parameter (50).c) has to be taken. Suppose we assume that:

---

[21] Or in all languages having agreement processes: perhaps languages like Chinese and Japanese could be characterized as completely agreement-less, 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).

[22] 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).[33]

This implies that, in English, some DP must appear in Spec of AGR. This is indeed the apparent situation in non-null-subject languages: 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 (52):

(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°, as a rich AGR-identifier, can assign Nominative directly to its I-subject. In non-NSLs, on the other hand, Nominative can only

---

[33] 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, (51) is an almost self-evident statement.

be assigned to the Spec of AGR. To derive this result, let us assume the following principles:

(53) AGR° 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 (54) 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°, 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 PRO also has Case).

In Italian, AGR° 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 as a provisory stipulation):[24]

---

[24] (55) is clearly at variance with Baker's (1985) Government Transparency Corollary.

(55) If AGR° combines with a head X by incorporation, it has the same governing capacities as X as far as Case assignment is concerned.

(55) 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 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 (55). In sum, V to AGR raising is a necessary condition for AGR° to be able to assign Case to its I-subject.[25] 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 a same idea to account for V-raising in infinitives in NSLs.

For convenience, we will call Nominative assignment in accordance with (55) *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

[25] 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.

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° 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 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 significative 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$_i$ [$_{VP}$ t_1 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° 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 $SELF_{NOM}$

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 (50).c): the I-subject is coindexed with AGR and, since AGR° 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.

(59).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 (59).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.[26]

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.:

- a full DP or pronominal (subject inversion).

- a variable, which accounts for the absence of that-t effects.[27]

- a *pro*, which is licenced by the AGR identifier, which recovers its content.

---

[26] 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^{max}, 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.

[27] 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.

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.

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 [±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 sjalv* (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 Joao já nao 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 (64), whereas in non-null-subject contexts an anaphor is preferred, as shown in (65):[24]

(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-were yourself in-days these

b. ??'An xoshev she-'ata lo 'ata be-yamin 'elu
I think that-you you-were 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

---

[24] Thanks to Ur 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

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 (66):

(66) a. John$_i$ is [$_{SC}$ 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$_i$ ... is [$_{SC}$ e$_i$ himself] ...
b. En Joan$_i$ ... és [$_{SC}$ e$_i$ 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 (67), 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 (67).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-reference 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 (68), 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):

(69) Greek:

a. O Kóstas$_{Nom}$ ine [$_{SC}$ t_{Nom} o kalíteros ipopsífios$_{Nom}$]
Kostas is the best candidate

b. Theorí [$_{SC}$ ton Kósta$_{Acc}$ ton kalítero ipopsifio$_{Acc}$]
He-considers Kostas the best candidate

German:

c. Dein Bruder$_{Nom}$ ist [$_{SC}$ t_{Nom} ein guter Mensch$_{Nom}$]
Your brother is a good person

d. Ich finde [$_{SC}$ deinen Bruder$_{Acc}$ einen guten Mensch$_{Acc}$]
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 (61)/(62) 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 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.[29]

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 eaftós tu*
The J. not is anymore the self his

where *o eaftós 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:

(72) $[_{DP}$ o $[_{NP}$ eaftós $[_{DP_{Gen}}$ 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

[29] 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.

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) [$_{DP}$ [$_{NP}$ [$_{DP_{gen}}$ 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