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

Since this is obviously a case of subject orientation, the facts are not expected. H. Sigurdhsson (p.c.) pointed out to me that perhaps the problem is that *sig* cannot be used because it lacks a nominative form. For all the other germanic languages we have been implicitly assuming that *sig selv/sich selbst*/etc. are Nominative forms in the Anaphoric Copulative Constructions. Then, why is Icelandic reluctant to use *sig* as Nominative?

I think there is a notable difference between Icelandic and Mainland Scandinavian, English or French: in the latter, there are some Nominative pronominals that can only be used in Spec of AGR position:[10]

(75) a. *He* is sick
b. It is *him/*he*
c. *Him/*he*, he's a liar

In these languages, the so-called Nominative pronouns are special forms confined to Spec of AGR. The forms usually taken to be non-Nominative are, I assume, unmarked. So let's assume in these languages, by extension, even reflexive forms would be unmarked, and therefore can be used as Nominative forms unproblematically. We could say that in some languages having unmarked forms is the unmarked case.

The situation is quite different in Icelandic: Case morphology is not confined to pronominals and it is quite rich. Therefore, taking *sig* as Nominative would contradict the general pattern, namely that DPs have no neutral Case-forms. As for German, which also has Case morphology across the board, we have to assume that this language does have neutral Case-forms (at least for *sich selbst*). In fact, German Case-morphology is much poorer than in Icelandic. So, unlike in Icelandic, the German

[10] I assume that Normative pressure concerning ((75).b) is forcing an ungrammatical construction.

sich selbst element does not distinguish between Nominative and Accusative:

(76) German:

a. Hans ist *sich selbst* nicht mehr
   H. is SE $SELF_{Nom}$ not more

b. Hans sag *sich selbst*
   H. saw SE $SELF_{Acc}$

Icelandic:

c. Jón er ekki lengur *hann sjálfur*
   J. is not longer he $SELF_{Nom}$

d. Jón meiddi *sig sjálfan*
   J. hurt SE $SELF_{Acc}$

Therefore, the problematic case in Icelandic could be assumed to be due to morphology: no Nominative or neutral form being available for *sig*, the Nominative form *hann* has to be used as a suppletion for *sig*.

Even if there are some problematic cases, I think that the contrast shown in (61)/(62) is significant enough not to disregard the issue.

2.3.3. On the Status of Binding Theory

We have proposed a definition of Binding Domain which is intended to cover two empirical phenomena (the [±anaphoric] status of I-subjects and of copulative predicates) that, as far as I know, had not been addressed in the literature thus far and were not even considered relevant for Binding Theory.

Binding Theory has often been conceived of as a means for accounting for co-reference restrictions between DPs having

independent θ-roles:[31] in the above proposal, however, I-subjects do not have independent θ-roles from their (possible) preverbal antecedents or, in the case of copulative clauses, the post-copular element is not a referring entity (it is a predicate). So they are not cases of co-reference in a reasonable sense of the term.

Then, if the above approach is on the right track, it strongly suggests that Binding Theory is a purely formal device which blindly extends beyond the scope of co-reference.

A second issue to be addressed concerns the empirical complexity of Binding Theoretical facts: recent research on a variety of languages has shown that Binding Theory is much more complex than early studies about English and similar languages suggested (one has to face complexities as long distance binding, subject (anti-)orientation, logophoricity, etc.).[32] Our reformulation of Binding Theory does not say anything about these issues, and one might suspect it is too naively tied to a simplistic view of classical Binding Theory.

Although I admit that a more comprehensive approach to Binding Theory is necessary, I think the present proposal has several advantages:

- as we pointed out, it is neutral w.r.t. the standard cases of Binding Theory: the additional requirement of a Case-position for A in the determination of the Binding Domain for A does not affect objects or oblique Arguments, for it is implicit in standard accounts that for them the Case position is the position

---

[31] See for instance Reinhart & Reuland (1991) for such a view.

[32] See for instance Hestvik (1990) and Vikner (1985) for Scandinavian, Everaert (1986) for Dutch, Koster & Reuland (1991) for long distance anaphora.

where they stand at S-structure."[33] Clitics (especially in clitic doubling structures) could be a case where this is not true, but I assume that clitics involve head Chains and head Chains as such are not subject to Binding Theory.

- it seems to me that, in the field of anaphora, there are two privileged, 'more central', unmarked cases: these are *subject-oriented* anaphors and *clause-bound* anaphors. The binding-theoretical account advanced above for I-subjects and anaphoric copulative constructions clearly belongs to the field of subject-oriented clause-bound anaphors.

- The definition of Binding Domain we proposed does not make reference to accessible SUBJECTS or to Complete Functional Complexes (see Chomsky (1981)/(1986-b)). Since any FC (including CP) can be a Binding Domain, the traditional problem of excluding examples like:

(77) *I think that myself am sick

is solved provided we can assume that C° is an appropriate governor for this definition.[34] I think that Accessible SUBJECT is a tricky notion, and that the notion of CFC cannot be used once we assume the Internal Subject Hypothesis. We will discuss this issue in the next chapter in connection with PRO.

---

[33] Chomsky's (1992) proposal of movement of the object to an AGR-Obj specifier does not challenge this idea: an object will be anaphoric w.r.t. the external Argument in a domain where the object has a case position, namely Spec of AGR-Obj.

[34] A potential problem could be:
??They want very much for each other to be happy
If acceptable, CP does not count as a Binding Domain. Perhaps the rather acceptable status of this sentence would be due to the possibility for for each other to be interpreted as a benefactive controlling the infinitive.

So, the present proposal, although not intended to provide any new insight into the standard cases of BT, seems to fit into it without problems.

2.4. The Position of pro and Floating Quantifiers

In the above account, it is crucial that *pro* is able to occur as an I-subject, contrary to the standard assumption that subject *pro* is only licenced in the Spec of INFL/AGR. In the present account, pro-drop is in some sense an epiphenomenon: the I-subject can be null in all languages: it will be *pro* in NSLs and a null anaphor in non-NSLs. So Taraldsen's (1980) original idea that AGR morphological richness is the key to pro-drop is here expressed in some indirect way: richness makes AGR° the AGR-identifier and this leads, through Binding Theory, to the existence of a null pronominal as I-subject. Therefore, we cannot assume that *pro* is licenced in Spec of AGR.

Rizzi has adduced two pieces of evidence in favor of the preverbal position for *pro* (the argumentation is reproduced in Roberts (1989) and Roberts (1991-a)). One of them is expressed by the following paradigm (from Italian):

(78) a. Essendo stanco, Gianni è andato via.
Being tired G. is gone away
'Being tired went away'

b. Essendo stanco, è andato via.
Being tired is gone away
'Being tired, he went away'

c. *Essendo stanco, è andato via Gianni.
Being tired is gone away G.
'Being tired, JOHN went away'

Rizzi's argument, adapted in terminology, is as follows: the gerund subject is a PRO which has to be controlled by a preverbal subject, which can m-command it. This is why the postverbal subject in (78).c), which cannot m-command the gerund, does not licence the controlled PRO and the sentence is ungrammatical. Since both the pre-verbal subject in (78).a) and the null subject in (78).b) licence the gerund, the conclusion is that *pro* is preverbal.

Catalan (and Spanish) behave in a similar way:[35]

(79) a. Estant tan cansat, en Joan se n'ha anat al llit
Being so tired the J. SI-off-has gone to-the bed
'Being so tired, Joan went off to bed'

b. Estant tan cansat, se n'ha anat al llit
Being so tired SI-off-has gone to-the bed
'Being so tired, he went off to bed'

c. *Estant tan cansat, se n'ha anat al llit en Joan
Being so tired SI-off-has gone to-the bed the J.
'Being so tired, JOAN went off to bed'

There are two main objections to Rizzi's account for the facts. Firstly, if we adjoin the gerund to a superordinate sentence, we obtain the following paradigm:

---

[35] Other non-finite sentences, such as temporal infinitives in Catalan (and Spanish) (see Rigau (1992) for an analysis of this construction), manifest the same paradigm as (79):

(i) A l'arribar, (en Joan) se'n va anar al llit
At the-to-come J./pro went off to bed
'When Joan/he arrived, he went to bed'

(ii) *A l'arribar, se'n va anar al llit en Joan
At the-to-come went off to bed J.

Rigau convincingly argues that these infinitives do not involve PRO. If so, the facts cannot be explained as a problem of control, as we will see.

(80) a. Estant cansat, és millor que en Joan se'n vagi
Being tired is better that the J. goes away
b. Estant cansat, és millor que se'n vagi
Being tired is better that goes away
c. *Estant cansat, és millor que se'n vagi en Joan
Being tired is better that goes away the J.

The examples in (80) are parallel in acceptability to the ones in (79) except that the gerund is in a position where c-command (or m-command) by the preverbal subject of the embedded clause is not possible. So the claim that c-/m-command is a necessary condition to licence the gerund PRO cannot be maintained. There must be some other account for the fact that the postverbal subject does not licence the gerund.

In order to preserve Rizzi's analysis, one could assume that control takes place at some level where the gerund is adjoined to the embedded sentence, the surface structure in (80) being a result of gerund raising. This level, however, cannot be D-structure, since at that level the subject of *anar-se'n* 'go away' does not m-command the gerund anymore than the postverbal subject, on the assumption that *go away* is an unaccusative verb and its subject is an object at D-structure.

But independently of this problem, there are more basic empirical problems for the control account. If we replace the gerund by a finite adjunct clause, the distribution of acceptability does not vary:[16]

---

[16] The paradigm in (81) has a correlate in Italian:
(i) a. Come era stanco (Gianni) è andato via
As he-was tired (G.) is gone away
b. *Come era stanco è andato via Gianni
As he-was tired is gone away G.
Paradigm (80) also has an Italian correlate. However, since gerunds are a little literary, dislocation of the gerund to an upper clause gives slightly awkward results:
(ii) a. Essendo stanco è meglio che (Gianni) vada via
Being tired is better that (G.) goes away

(81) a. Com que estava cansat, en Joan se n'ha anat
As he-was tired the J. went away
b. Com que estava cansat, se n'ha anat
As he-was tired went away
c. *Com que estava cansat, se n'ha anat en Joan
As he-was tired went away the J.

In (81) no controlled PRO is involved, so the paradigm has to be explained in an alternative way. I think the facts in all the preceding c. examples, where the postverbal subject cannot co-refer with the subject of the adjunct clause, can be accounted for as cases of WCO. Since the post-verbal subject is interpreted as Focus, and Focus triggers WCO effects, as can be seen in (82) (see Chomsky (1981)), the post-verbal subject cannot be co-referent with a pronoun it does not c- (or m-)command:

(82) *His$_i$ mother loves JOHN$_i$

In fact, in Catalan Rizzi's paradigm has a correlate in cases where we are not dealing with a subject, but a dislocated/clitic object:[37]

b. *Essendo stanco è meglio che vada via Gianni
Being tired is better that goes away G.
For paradigm (81) if we dislocate the adjunct clause to an upper clause the acceptability results do not change in either Catalan or Italian.

[37] Italian gerunds seem to be subject oriented, and so this paradigm cannot probably be instantiated in Italian.

(83) a. *PRO* estant malalt, a *en Joan*$_i$, *l'anirem* a veure
being ill to the J. him-we'll-go to see
'Since he's ill, we'll visit Joan'

b. *PRO* estant malalt, *l'anirem* a veure
being ill him-we'll-go to see
'Since he's ill, we'll visit him'

c. ?**PRO* estant malalt, anirem a veure *EN JOAN*
being ill we'll-go to see the J.
'*Since he's ill, we'll visit *JOAN*'

Both (83).a) and (b.) cluster together as opposed to the non-dislocated Focus object, which triggers WCO. The only difference between this paradigm and Rizzi's paradigm is that the object, as opposed to the postverbal subject, need not be Focus (so that (83).c) improves if *en Joan* is not Focus.

So the paradigm adduced by Rizzi proves irrelevant for the position of null subjects: in (83).b) we could not claim there is a preverbal empty object c-commanding the gerund.

In fact there are cases where pre-verbal subjects have co-reference restrictions which do not cluster with null subjects. Consider:

(84) a. El cotxe d'*en* *Joan*$_i$, *ell (mateix)*$_i$, no el condueix
The car of-the J. he (SELF) not it-drives
'*Joan*'s car, *he himself* never drives'

b. ?*El cotxe d'*en* *Joan*$_i$, no el condueix (-> *pro*$_i$)
The car of-the J. not it-drives
'*Joan*'s car, *he* does not drive'

c. El cotxe d'*en* *Joan*$_i$, no el condueix *ell*$_i$
The car of-the J. not it-drives he
'*Joan*'s car is not driven by *him himself*'

In (84) *pro* contrasts with overt subjects, which behave alike. I think the explanation for that paradigm could be based on the idea that (in some languages) a null subject always involves a topic, whether overt or not. If it is overt it can appear in several places (it can adjoin in a higher or lower position, even in superordinate clauses). A topic cannot bind an R-expression. In a strict sense, the dislocated element *ell mateix* in (84).a) does not c-command the other dislocated element *el cotxe d'en Joan* containing the co-referential R-expression (*en Joan*), because *ell mateix* appears to the right of the other left adjoined element: the first branching node including the inner adjunct will not contain the outer adjunct.

Now suppose null topics always have maximal scope. Then (84).b) would be excluded for the same reason as:

(85) *Ell mateix$_i$, el cotxe d'en Joan$_i$ no el condueix
He SELF the car of-the J. not it-drives
'*He does not drive Joan's car'

where *ell mateix* illicitly binds *en Joan*. I leave the issue here.

The other empirical argument presented by Rizzi is based on the following paradigm:

(86) a. I soldati sono andati tutti via
The soldiers are gone all away
'The soldiers have all gone away'
b. Sono andati tutti via
Are gone all away
c. *Sono andati tutti via i soldati
Are gone all away the soldiers

The Floating Quantifier (*tutti* 'all') can co-occur with both preverbal subjects and null subjects, but not with post-verbal subjects. This would suggest, according to Rizzi, that null subjects pattern with preverbal subjects. In more theoretical terms, we could say that a FQ is licenced by being in the c-command domain of a referential subject. This would be the case in (86).a) and even in (86).b) if we assume there is a preverbal *pro*, but not in (86).c), where the referential subject is in inverted position and only an expletive *pro* appears in preverbal position.

In fact, we can redefine the pattern in (86) by saying that FQs can appear in contexts where there is null I-subject (*pro* in this case): both (86).a) and (86).b) would have, in our theory, a *pro* as I-subject; (86).c), instead, would have *i soldati* as I-subject.

In Chapter 1 we assumed that FQs in languages like Catalan cannot be the result of movement of the element they are adjoined to at D-structure, for some of them do not form a possible constituent with this element. Now we have a theoretical reason for this fact: there is no A-movement of the I-subject in NSLs.

Putting these ideas together, we can assume that FQs in NSLs are elements adjoined to *pro*. Then (86).a/.b) would have the structure:

(87) (I soldati$_i$) sono andati [$_{DP}$ tutti *pro*$_i$] via

(86).c) is ungrammatical because there is no *pro* I-subject for the FQ to adjoin to. In fact, if the quantifier adjoins to the I-subject, the result is:

(88) Sono andati via tutti i soldati
Are gone away all the soldiers

which, of course, is not a case of Quantifier Floating.

This analysis, however, faces important problems for the present theory. For a start, FQs do have a freer distribution than I-subjects: they can precede the complements of the verb (whereas I-subjects in general follow the complements in Italian or Catalan). In English overt I-subjects (*himself*) and FQs (*all*) even have complementary distribution (the former are VP-final and the latter are VP-initial).

Another problem is related to emphatic I-subjects: while inverted subjects do not allow FQs, emphatic I-subjects do allow FQs:[38]

(89) Els nois s'han fet *tots* el llit *ells (mateixos)*
The boys SI-have made all the bed they (SELVES)
'The boys have *all* made their bed *themselves*'

If we want to maintain that emphatic elements like *ells (mateixos)* are I-subjects, then the above claim that FQ's in NSLs are adjoined to a *pro* I-subject cannot be true, if *pro* has to occupy the I-subject position too. In fact FQs seem to have a freer distribution than I-subjects (they can precede the verbal complements). On the other hand, as we pointed out in Chapter 1, FQs are likely to occupy derived positions, rather than the I-subject position.

I cannot provide any clear solution for the problem raised by (89). I will only suggest two possible approaches. We have seen above that a sentence having a null subject can be claimed to involve a null topic, detectable in cases like (84).b). We have also seen that the null topic is only present when there is no overt topic: in (84).a), the overt topic *ell (mateix)* prevents a null topic from appearing. We said that null topics are

[38] In fact, some speakers do not easily accept these constructions. See above for parallel facts with clitic doubling.

licenced when there is a null subject. Suppose we assume that null topics are licenced whenever a preverbal (non-Focus-fronted) subject is possible: if the preverbal subject appears (wherever it is dislocated) it acts like a topic. If there is no preverbal subject, then a null topic appears. Now the contexts allowing preverbal subjects are, in our theory, either a null I-subject or an emphatic I-subject. Since these are the contexts allowing preverbal subjects to appear, we can assume that these contexts also trigger a null topic to be present if there is no preverbal (overt) subject.

Now, we could assume that a FQ has to have an antecedent c-commanding it, and that the antecedent can be a null topic. This is only a vague approach. We leave the question open of how FQs are generated and interpreted unexplained. Something in the theory has to guarantee that FQs are clause-bound, which does not follow from the licensing condition of being bound by a (possibly maximal scope) topic.

Notice that this approach is close to Rizzi's contention that there is a preverbal *pro* in his example (86).b): we also propose there is an empty category. The difference is that the empty category is not in Spec of INFL (AGR), but rather has maximal scope.

Another solution could consist in adopting Rigau's (1988) proposal about *strong pronouns* in Romance languages. According to her, emphatic I-subjects would be strong pronouns. Strong pronouns, in her theory, are not in A-positions. Then, a possible account for (89) would be:

- FQs are generated in A-position and, in coherence with our approach, in NSLs they are adjoined to a *pro*.

- since emphatic strong pronouns are not in A-position, FQs and emphatic strong pronouns can cooccur.

- FQs, even if generated in A-position can/must move to some higher position (which accounts for their freer distribution).

- emphatic strong pronouns, even if not in A-position, are subject to Binding Theory, and therefore abide by the generalization about the [±anaphoric] status of I-subjects, even if they are not I-subjects themselves.

The problem with emphatic (pseudo-)I-subjects is similar to the problem of pronominal clitic-doubling in Catalan or Spanish:

(90) El veig a ell
Him-I-see to him
'I see HIM'

Clitic doubling is also at the basis of Rigau's claim that strong pronouns are not in A-positions. Clitic doubling pronouns, like emphatic (pseudo-)I-subjects, can cooccur with a FQ:[»]

Context: I hand the new school books to *the parents* of the small children, but...

(91) els grans, *els* dono *a tots* el llibre *a ells (mateixos)*
the older, them-hand-I to all the book to them (SELFs)
'the older ones, I hand the book to them all personally'.

In both (89) and (91) the (emphatic) strong pronouns would be occupying a -θ-position, thus leaving the object/I-subject position free for the FQ (independently of whether the FQ then moves to some higher position).

So we seem to draw back to the traditional assumption, which we crucially challenged before, that emphatic anaphors and pronominals are not in an A-position, without giving up the crucial idea that these elements are subject to Binding Theory. In fact, both clitic doubling strong pronouns and emphatic (pseudo-)I-subjects, even if not in a θ-position, are not far

[»] Speakers not readily accepting (89) do not accept (91) either.

removed from it: their word order distribution is nearly the same as, respectively, standard objects and inverted subjects, and only more subtle evidence, such as their co-occurrence with FQs, leads to reconsideration of the theory. One possibility is that these elements are adjoined to their respective θ-positions, in a structure like (92):

(92) $[_{DP}\ [_{DP}$ *pro* (FQ)] *strong pronoun*]

In (92):

- the *pro* would be licenced by the clitic (if clitics licence a *pro*) or by AGR° (in NSLs), depending on whether the DP is in object position or in I-subject position,

- the strong pronoun is an adjunct carrying emphasis. It is subject to the same BT constraints as the *pro* it is adjoined to.

- the FQ can be moved to some other position, provided it is in the c-command scope of, respectively, AGR° or the clitic.[40]

This is a rather speculative and little explanatory approach. I will abstract away from this digression in the remainder of the thesis.

2.5. On the Nature of AGR Requirements

In section 1., we advanced several rules requiring that AGR must be coindexed with some DP or CP Argument, and on the one hand AGR must have an AGR-identifier which must provide the I-subject with Case, on the other hand:

---

[40] French cases like:
(i) Il faut tout que je face
It needs everything(FQ) that I do
'I have to do everything'
are atypical cases of long distance FQ. I cannot say anything about the issue.

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

(94) a. AGR must have an AGR-identifier.
b. X can be an AGR-identifier iff X is rich in phi-features (number and person).
c. AGR°/Spec of AGR is the AGR-identifier of AGR.

(95) AGR° can optionally assign Nominative Case to Spec of AGR by agreement or to some other position under government.

(96) The I-subject must receive Case from its AGR-identifier.

Within the framework of Principles and Parameters, the obvious question is whether we should postulate any such specific rules at all. This theoretical framework should optimally consist of principles and parameters of a very general and pervasive nature, and rules affecting one single category as AGR looks at odds with such a desideratum.

Two considerations, however, can be put forward in defense of these rules. One is that, even if category specific, they are not by any means language specific: they seem to hold in a pervasive way across a good deal of languages. So they cannot be considered mere ad hoc theoretical devices to account for highly idiosyncratic facts. If not genuine principles of UG, they are at least good candidates to be theorems of the grammar.

I think the above set of rules can be interpreted as a specific case of a set of the general constraints that Chomsky (1986-b) dubs Licensing and Full Interpretation. In Chomsky's view, Licensing is a condition usually holding of two items that somehow need each other in order to be fully interpreted: Operator and variable, predicate and Argument, etc.; not only do the two elements have to cooccur, they also must stand in a proper relation.

In the same spirit, we could conceive that AGR and I-subject are two elements that need each other: AGR has to be licenced by having an I-subject, and the I-subject needs Case."

The idea I want to pursue is the following: once AGR is present in a language, it has to be licenced by having an I-subject, providing Case to it, and being rich in features. The question is: why has AGR to be present?

We could assume that AGR is not present in all languages: Japanese and Chinese would possibly be languages lacking AGR (see Fukui & Speas (1986)). These languages challenge Taraldsen's original idea that null subjects are dependent on rich AGR-morphology. There are, on the other hand, languages having no overt AGR-morphology and nevertheless being closely similar to the Indo-European languages we have considered, namely Scandinavian languages. If the idea that Chinese/Japanese lack AGR is to make sense to account for their exceptional behaviour, then we should ask why Scandinavian languages do not take the negative setting for the [±AGR] parameter. This parameter should have some trigger for one or the other value.

We could then argue that the trigger for the positive value is the presence of phi-features in the language. As argued by Fukui & Speas (1986), Japanese (and Chinese) seem to lack phi-features altogether.

This approach is, however, problematic in one sense: Japanese, which has overt Case morphology and is a Nominative-Accusative language, seems to abide by Burzio's generalization as far as its Case array is concerned. Since our account for BG is based on AGR, we cannot adopt the view that Japanese (and probably Chinese) lacks AGR.

" The latter idea could be challenged: we could claim that the Case filter (or visibility requirements) are an epiphenomenon, due to the fact that AGR (and other FCs) have to discharge Case (thanks to Jeff Runner for this suggestion). I will no pursue the issue.

We will suggest another possibility. *All languages have AGR.*[42] Not all languages have phi-features. Then we could reformulate our rule (94) as:

(97) a. AGR must have an AGR-identifier.
b. X can be an AGR-identifier iff X is rich in the phi-features which are present in the language.
c. AGR°/Spec of AGR is the AGR-identifier of AGR.

Since Japanese (and Chinese) have a null set of phi-features, condition (97).b) can be fulfilled in a trivial way. AGR° is vacuously rich in Japanese, and it can licence *pro* as an I-subject.

Summing up, the rules we postulated for AGR can be conceived as belonging to a set universal licensing principles:

a) AGR must be licenced by having an I-subject (as expressed by (93)) and providing this I-subject by Case as expressed by (95) and (96));

b) since AGR consists in features, it must also be licenced by being able to display a sufficiently rich set of features (as expressed by (94)). This will allow the I-subject to be an empty element whose content is recoverable from by the AGR-identifier (in a trivial way if the language has no phi-features).

All these proposals are highly speculative, and there is a good deal of vagueness in them. The main difference with other instances of licensing principles is the strictly formal nature of the licensing principles for AGR. A licensing principle for, say, Operator-variable structures is more deeply rooted in interpretation and, therefore its existence seems to be a matter

---

[42] In Chapter 4 we will propose that some infinitival constructions do not have AGR: if so, the correct claim would be that all languages have AGR in finite sentences.

of necessity if language has to be an instrument for carrying meaning. If, however, the above proposal proves to be on the right track, it constitutes an argument for the autonomy of syntax, whose formal principles would be independent from the constraints on well-formed semantic structures.

2.6. Inverted Copulative Constructions

There seems to be an interesting correlation between the NSL status and the possibility of having an inverted agreement pattern of some copulative constructions of the type (examples from Catalan):

(98) a. El president sóc jo
The president am I
'*I* am the president'
b. *The president am I/me

We cannot simply say that (98).a) is a case of subject inversion, because then the preverbal subject would have to be a dislocated element without a resumptive clitic. Actually, nominal predicates can dislocate, and then a clitic is used as a resumptive element:

(99) El president, no *(ho/el) sóc pas
The president, not it/him-am at-all
'The president, I am certainly not'

So the clitic is obligatory in (99) and impossible in (98), which suggests *el president* is not a dislocated in (98).a).

On the other hand, the possibility of (98).a) seems tightly correlated with the NSL status: Old French had the equivalent of (98).a) and lost it as it lost its NSL status.

In order to derive these facts, we need some previous theoretical assumptions. First, I adhere to the view, defended by several authors (Andrea Moro, Giuseppe Longobardi) that all copulative constructions contain a predicate, and there are no 'equative' constructions. Therefore, *el president* in the above examples has to be the predicate element, for pronoun (*jo* 'I') cannot plausibly be a predicate. The underlying structure for the above examples would be:

(100) ser... [$_{SC}$ jo el president]
To-be I the president

Both Moro's and Longobardi's proposals essentially say that, when the predicate is a DP, either the Argument or the predicate are allowed to raise to Spec of AGR. Plausible and simple though this idea is, it does not immediately explain why:

- agreement takes place with the postverbal DP.
- these examples are only allowed in NSLs.

I think our theory on AGR-identifiers is well equipped to say something on the way of an explanation for these facts. We proposed that Spec of AGR is not the AGR-identifier in NSLs. So Spec of AGR is not required to be filled. In fact, as we will argue in Chapter 5, it appears to be a position with much looser requirements for the DP filling it than in non-NSLs. In any event, the fact that the predicative DP raises to it does not interfere with AGR-identification of Nominative assignment.

In non-NSLs, raising of the predicate to Spec of AGR would not abide by the requirements that there must be an AGR identifier and that this AGR-identifier has to transmit its Case to the I-subject, which implies it has to form an A-Chain with it. Therefore, non-NSLs cannot have DP-predicate raising: French lost it as soon as it became a non-NSL.

In conclusion, even if we cannot provide a precise explanation for why the predicative DP can raise to Spec of AGR, we can explain why it cannot in some non-NSLs.

We cannot explain why, in this type of construction, the pre-verbal DP can be dropped and the postverbal DP (which would be the I-subject) cannot:

(101) a. Sóc jo (el president)
Am I (the president)
'It's me (the president)'
b. *El president sóc
The president am

Probably (101).b) is excluded because this kind of construction is precisely used to focalize the subject, and hence it cannot be dropped. It is not clear why the preverbal DP can be dropped if it is not recovered in content by AGR.

3. Indefinite I-subjects

So far, the predictions are that postverbal subjects can be:

- [-anaphoric] in NSLs.
- [+anaphoric] in non-NSLs.

Both kinds of languages, however, freely admit indefinite in post-verbal object position:

(102) a. There came a man
b. Viene un uomo
Comes a man

Within the theory sketched above, the question is: why are indefinites able to occur post-verbally without violating BT in

English? Recall that for NSLs the existence of indefinite subjects is not a problem anymore than definite inverted subjects are: they both would be licenced as far as they are both [-anaphoric] I-subjects. Indefinite I-subjects, however, are uniformly acceptable independently of the null/non-null-subject contrast. So they should be licenced independently of how inverted subjects are licenced in NSLs.

Since Binding Theory plays a crucial role in our account of subject inversion, so that only NSLs allow [-anaphoric] I-subjects, we will exploit the idea that something allows indefinite DPs to escape from BT effects. We will address the question in section 3.2.

In section 3.1., we will address another important question: how are indefinite DPs Case-marked? If we adopt Belletti's (1988) hypothesis that indefinite DPs receive Partitive Case, then an obvious problem arises in connection with our rule of AGR coindexation, since we crucially assumed that the DP AGR is coindexed with could not be inherently Case-marked. We will refine the notion of inherent Case in a way to allow Partitive Case to be simultaneously structurally Case marked.

3.1. Partitive Case

We assumed that, for languages such as English or Catalan, at D-structure AGR has to be coindexed with an non-(inherently)-Case-marked DP (or CP). Reference to inherent Case-marking was crucial in order to prevent a DP other than an EA or an object to become the I-subject.

In Chapter 1, we noticed that Belletti's hypothesis of Partitive Case provides an interesting means of excluding sentences like:

(103) a. *I consider girls interesting (existential reading)
b. *There seem girls to have come

which would be excluded because Partitive Case can only be assigned to positions which are θ-marked by the head assigning Partitive.

If we assume that indefinite DPs in object position are assigned Partitive Case, then they should not be candidates for AGR to coindex with, according to our rule of AGR coindexation, repeated here:

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

However, it is crucial for our theory that Partitive DPs may be possible I-subjects. Otherwise, sentences like:

(105) There are many children

would have no I-subject, and should be excluded, as (104) is obligatory.[41] On the other hand, it is plausible that the indefinite DP in (105) is the subject in some sense, since it agrees (in number) with the verb in English and many other languages.

---

[41] For this reasoning to hold, we must exclude the possibility that the expletive itself becomes the I-subject (thanks to Jeff Runner for pointing this out to me). We can obtain this result by assuming that:
a) The expletive is directly generated in Spec of AGR.
b) The notion of c-command relevant for (104) is strict c-command: AGR° does not c-command its specifier.
Alternatively, we could stipulate that the I-subject has to have 'semantic content'.

To solve the problem, we will assume that partitive Case differs from other inherent Cases in that Partitive alone is not sufficient to Case-licence a DP. Suppose we postulate that:

(106) Partitive has to be realized as Structural Case at S-structure.

If we want to maintain that Partitive is an inherent Case, then it must not be one relevant for (104). It would anyway be a defective Case, which has to be supplemented by a structural Case. It is not clear why (106) should hold. There is, though, some evidence for its empirical validity. Consider the following pair of examples:

(107) a. There have developed typhoons here
b. John developed theories on that issue

Suppose both indefinite DPs in (107) are assigned Partitive at D-structure, and some structural Case at S-structure. As we remarked, there is some evidence that Nominative is assigned to the indefinite DP in (107).a), which would naturally account for the fact that there is agreement between the verb and the indefinite DP. As for (107).b), it can be naturally assumed that the indefinite DP receives Accusative Case at S-structure. There is some evidence from Romance languages pointing to that conclusion. Consider the following paradigm of Causative constructions (examples from Catalan):

(108) a. Li fa menjar les patates
Him-DAT makes eat the potatoes
'S/he makes him eat the potatoes'
b. El fa menjar
Him-ACC makes eat
'S/he makes him eat'

The generalization is that in Romance (and many other languages') causative constructions the subject of the infinitive clause is assigned Dative if the infinitive has an Accusative Case-marked complement; when the infinitive does not have any Accusative complement, then its subject is assigned Accusative.

According to Baker (1983), the generalization could be expressed as follows: the complex formed by the causative verb and the infinitive can at most assign one Accusative Case; if that Case is required by the object of the infinitive, then the subject of the infinitive has to receive Dative; if not, it is the subject of the infinitive that is assigned Accusative.

Now consider the following example:

(109) Li fa menjar patates
Cl-DAT makes eat potatoes
'S/he makes her/him eat potatoes'

which minimally differs from (108).a) in that the object of the infinitive is indefinite. If it is indefinite, it has been assigned Partitive (it is enough for our argument that this is at least a possible option). If Partitive did not additionally require a structural Case, as we are claiming, the CP *patates* would not spend the Accusative Case which the complex of verbs can afford, so the subject of the infinitive could be assigned Accusative, contrary to fact:

(110) *El fa menjar patates
Him-ACC makes eat potatoes
'S/he makes him eat potatoes

So, it is reasonable that Partitive Case is not sufficient by itself to Case-licence a DP. We could then redefine (104) as (111):

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

where a DP is Case-saturated if the Case it bears is sufficient to licence it, Partitive Case-marked DPs at D-structure not being so.

Chomsky's (1986-b) characterization of inherent Case requires that inherent Case has to be realized at S-structure, so it could appear that our special characterization of Partitive Case as a special inherent Case additionally requiring a structural Case can be accommodated within this general view.[*]

I will not develop the question further. It suffices for the present purposes that Partitive Case, if at all an inherent Case, is different from other inherent Cases in that it additionally requires structural Case, so that (111) is an accurate rule for AGR coindexation. It could turn out, as we pointed out in Chapter 1, that Partitive is not a Case, but

---

[*] I think, however, that Chomsky's proposal cannot be trivially adopted nowadays. First of all, Chomsky's assumption that inherent Case has to be 'realized' does not involve structural Case-assignment. Secondly, the facts Chomsky's theory is intended to cover are basically related to Genitive case. Since it is crucial in his account that Genitive is both assigned and realized in the government domain of N, the theory should be carefully revised in the light of the Determiner Phrase hypothesis, if we are to accommodate it to present day common assumptions.

rather a special interpretative option for DPs. For convenience, though, I will keep using the term *Partitive Case*.

3.2. Indefinite DPs and phi-features

Now let us address the other problem raised by indefinite subjects, namely that they are possible I-subjects in non-NSLs, which otherwise do not allow [-anaphoric] I-subjects.

A solution to this problem can be formulated in the following terms. Let us assume that:

(112) a. Partitive DPs do not have person features.[43]
b. Partitive DPs may/may not have (grammatical) number features.
c. AGR and Spec of AGR need not agree in number features when person features are not present.
d. A binds B if A c-/m-commands B and A and B share some phi features.

The parameter in (112).b) is intended to account for the variation languages seem to exhibit with respect to agreement with an indefinite:

---

[43] This is also assumed in Rigau (1991). Rigau also assumes that Person Agreement and Number Agreement are different functional categories. I think this idea could be adopted within the present theory, but it would require some careful elaboration, for it is crucial for the present account to work that AGR° is coindexed with the I-subject, contrary to what Rigau assumes for partitive constructions: at least Person-Agreement is not coindexed with the partitive in her view. Here I contend that there is always coindexation, which may be devoid of content if the binding features are absent.

(113) Agreement in number:

a. There are children in the garden (Standard English)

b. Es *sind* Kinder in den Garten (German)
   It are children in the garden

c. Ci *sono* bambini nel giardino (Italian)
   there-are children in-the garden

d. *Arriben* turistes (Standard Catalan)
   Arrive tourists

(114) No agreement:

a. Il y a des enfants au jardin (French)
   It there-has of-the children in-the garden
   'There are children in the garden'

b. *Arriba* turistes (North-Western Catalan)[46]
   Arrives tourists

c. There*'s* children in the garden (Coll. English)

(112).b) could be characterized as follows. Suppose indefinite DPs involve a (possibly empty) D° which may or may not inherit the number features of its complement NP.[47] If it does, then we have Partitive DPs cum number. Otherwise, Partitive DPs do not have number features.[48]

---

[46] See Rigau (1991), where an explanation is provided for the contrast between languages showing verb number-agreement with the indefinite and languages with no such agreement.

[47] Actually, it is not crucial for the present purposes that we have an empty D°: it could as well be an empty quantifier, if we assume indefinite quantifiers and determiners are different FCs, as proposed by many authors with several implementations (Cardinaletti & Giusti (1991), Rigau (1991))).

[48] French would be a language where Partitive DPs do not have number features. We could argue that this is a related to the fact that a preposition usually precedes the NP in French Partitive DPs:

(i) Je mange beaucoup de pommes
    I eat many of apples
(ii) Je ne mange pas de pommes

Let us consider the following simplified D-structure:

(115) AGR developed $\{_{DP}$ typhoons $\}_{Part}$

In (115) AGR is coindexed with the Partitive DP. In a non-NSL where the Partitive DP does not have number features, AGR and the Partitive do not share any phi-features. So, when the AGR-identifier (namely Spec of AGR) is filled by an (expletive) DP, this DP will not either share any features with the Partitive DP. According to (112).d), there is no binding relation between the AGR-identifier and the Partitive, so BT does not force the I-subject to be [+anaphoric] (and in fact prevents it from being so), as is otherwise the case with non-NSLs.

Suppose, on the other hand, that the partitive DP has number features. Then the prediction is that in (115) AGR° and the indefinite DP will share number features. If the AGR identifier (= Spec of AGR) shared these features, then a Binding relation would stand between Spec of AGR and the I-subject, and the indefinite DP would be excluded as a BT 3rd principle violation. However, non-NSLs of this kind can have recourse to the (possibly marked) option admitted in (112).c), namely that AGR° and Spec of AGR do not agree in number features, so, again, no BT violation ensues if the Partitive remains in place.

In the preceding account, it is not clear why some non-NSLs allow null expletives (or even null quasi-Arguments). Perhaps the residual character of AGR-identifiers in indefinite I-subject constructions, where the AGR-identifier does not display any features, allows for it to be dropped.

I Neg eat not of apples
and this prepositions blocks number inheritance by the determiner. However, other languages having similar prepositions allow number in Partitive, so this account cannot be trivially correct:

(iii) Catalan: moltes de pomes
many-fem-pl of apples(-fem)-pl

Another possibility could be developed. Suppose non-NSLs can take the marked option of having AGR° as the AGR-identifier in constructions where I-subject does not have person features (i.e., it is Partitive). Suppose this option is subject to the constraint that AGR° is rich enough to recover the content of the I-subject: if it is the case that the I-subject has number features, then AGR° has to be rich in number features, i.e., it has to at least distinguish between 3rd-singular and 3rd plural. From this we could derive the following descriptive generalization:

(116) If a non-NSL does not show overt *there*-expletives, the verbal morphology distinguishes between 3rd singular and 3rd plural.

Recall that the implication does not hold in the other direction: English distinguishes between 3rd-sng and 3rd-pl (only in the present tense, though). German, Icelandic or Brazilian Portuguese are all languages abiding by this generalization.[*]

Many questions are left open here:

a) Why is Partitive restricted to certain verbs in English (There came a man/*There broke a glass)?

b) Why is Partitive allowed for non-objects in some languages (German)?

---

[*] Some Brazilian dialects do not have the 3rd-sng/pl morphological distinction anymore and, as far as I know, they still have null expletives (thanks to Cristina Figueiredo for pointing this out). Maybe we could simply say that Brazilian Portuguese (or at least these dialects) has Partitive DPs without number features, which implies that AGR° is vacuously rich to licence a Partitive DP.

b) Why do some of the languages allowing null indefinite-expletives also allow null quasi-Arguments?[50] Are quasi Arguments also lacking person features?

I cannot properly address these issues.

4. CP I-subjects and Small Clauses

Thus far, the theory we have presented essentially consists of the following components:

- a rule of AGR indexation that coindexes AGR with some DP or CP, which becomes the I-subject.

- some principles of Case theory to the effect that the I-subject has to receive Case from the AGR-identifier.

We have contended that CPs are candidates for becoming I-subjects, but we have said nothing about whether they also are assigned Case. Since Stowell (1981), a widely accepted hypothesis has been that CPs do not accept Case. For the present theory it is essential that I-subjects obligatorily obtain Case from their AGR-identifier, in order to explain how Accusative is never assigned to an I-subject.

So if I-subjects obligatorily obtain Case from their AGR-identifier, we seem to be compelled to the conclusion that CP I-subjects receive Case too, contrary to Stowell's hypothesis.

In any event, it is clear that CPs show a distribution which is clearly different from that of DPs. These are some of the essential facts:

- object CPs do not have to be adjacent to the verb:

(117) I said the other day that...

---

[50] Rizzi (1986) proposes there is a gradation expletive -> quasi-Argument -> full-Argument, and that if pro is licenced in one of the options, it is also licenced in the preceding options. See next section.

- CPs cannot occupy the subject position of sentences and Small clauses:

(118) a. *Does that John comes bother you?
b. *I Consider that John may come impossible

- (in many languages) CPs cannot be the complement of a preposition and do not require 'of' insertion:

(119) a. *I talked about that...
b. I am sure (*of) that...

I think the above evidence is rather compelling. So, we will adhere to Stowell's claim that CPs cannot be assigned Case. Before proceeding, let us consider another aspect of argumental CPs that we have not addressed thus far.

The fact is that argumental CPs can (and sometimes have to) be 'doubled' by an expletive:

(120) a. *It* is evident *that*...
b. I consider *it* evident *that*...
c. I can accept (*it*) *that*...

Suppose that the expletive in the above examples is linked to the CP in some way, and that it is the expletive that occupies the θ-position at D-structure and is Case-marked at S-structure. In other words, suppose that it is the expletive that fulfils all the requirements of being an Argument, except that it transmits its θ-role to the extraposed CP.

Suppose that the linking device between the expletive and the CP is coindexation. In the preceding section we have provided a means of preventing BT effects in structures like:

(121) *There* came a *man*

by assuming that coindexation does not imply binding if there are no phi-features shared. Similarly, we can assume that CPs are not sensitive to being coindexed and c-commanded by their *it*-expletive, because they have no phi-features. That CP I-subjects cannot have phi-features is suggested by the fact that coordinating them does not give plural agreement:

(122) It is/*are well-known that ... and that ...

Coordinated preverbal CPs can trigger plural agreement:

(123) That ... and that... are two well known facts.

Since, as we and many authors argue, these CPs are not in subject position (they would be dislocated), we could assume that the real subject (some empty category) is a resumptive element having plural features.

In fact expletives linked to CPs are exceptional in that they are the only ones occupying θ-marked positions, as shown in (120).b/.c). So let's propose there are two kinds of expletives:

- expletives of the *there*-type, which merely fulfil the formal requirement of filling the AGR-identifier in (some) non-NSL and transmitting Case to the I-subject. They are not θ-marked, since they do not play any role in interpretation.
- expletives of the *it*-type, which are linked to a CP and behave like Arguments (they are Case- and θ-marked) except for the fact that they transmit their θ-role to the CP.

We cannot claim that CPs are always associated to an *it*-expletive, because there are obvious counterexamples:

(124) I think (*it) that...

The distribution of CPs and *it*-expletives can be accounted for on the following assumptions:[51]

(125) a. CPs cannot be (directly) Case-marked. CPs can escape Case-marked by having recourse to:
- extraposition.
- being associated to an *it*-expletive.

b. CPs cannot occupy specifier positions at any level of representation.

According to (125).a) two options are available for the CP complement of *accept*:[52]

(126) a. I accept t_i $[_{CP}$ that... $]_i$ (extraposition)
b. I accept it_i $[_{CP}$ that... $]_i$ (*it*-expletive)

According to (125).b), only the *it*-expletive option is available for subject CPs:

(127) a. *It* is evident *that*...
b. I consider *it* evident *that*...

Now consider NSLs. We have claimed that *there*-expletives do not exist in NSLs. We also claim that no expletive exists in subject inversion constructions. This is a natural assumption

---

[51] This proposal is inspired on ideas in Authier (1991), who deals with the contrast between French and English CP-expletives. Here we will not go into a detailed discussion of the issue.

[52] Only the extraposition option is available for the CP complement of other verbs like say or think. The explanation for these facts could lie on the different status of subcategorized CP complements.

under the view that such kinds of expletives are merely formal AGR-identifiers, and are not necessary in languages where Spec of AGR is not he AGR-identifier. *It*-expletives, on the other hand, are of a different nature, as we have argued. Since they fulfil the important role of linking CPs to A-positions, they should be in principle universally available.

Rizzi (1986) argues convincingly that null expletives exist in V-governed positions in Italian. The essential of his argumentations goes as follows. Some languages (such as Italian) allow null objects in a way Modern English does not. Rizzi claims that such null objects are *pro*:

(128) a. Questo conduce *pro* a concludere quanto segue
This leads to conclude what follows
b. This leads *(people) to conclude the following

See Rizzi (1986) for an explanation of how such a *pro* is licenced. One appeal of Rizzi's characterization of the facts is that it predicts an interesting correlation: languages allowing null objects as in (128).a) also allow null expletives as in (129).a); languages not allowing null objects (as English, see (128).b)) do not allow null expletives either (see (129).b):

(129) a. Gianni ritiene *pro* probabile che Mario venga
b. John considers *(it) probable that Peter comes

Early Modern English (until the XVIIIth century) allowed both null objects and null expletives, so the correlation is highly plausible, as it holds of as closely related languages as early Modern English and present Modern English.

In the light of these facts, then, we conclude that:

- *there*-expletives are an exclusive property of (some) non-NSLs.

- *it*-expletives are universally available. They are null in contexts where null pronominals are allowed.

With the above assumptions, we can address the problem of CPs and obligatory Case for I-subjects: whenever a CP is coindexed with AGR, there must be an I-subject to which AGR provides Case. Since the CP itself cannot be assigned Case, it will have the option of being linked to an *it*-expletive, which does not transmit Case to its associated CP.[33]

One question we could address is where the CP linked to the expletive is attached to. Non-dislocated sentential CP I-subjects seem to occupy a VP-final position. When a CP is a subject of Small Clauses (as in (129)), it also occupies a Small-Clause-final position. The underlying place of attachment of the CP should not be far removed from the underlying position occupied by the *it*-expletive it is linked to. Therefore, the fact that the CPs in (129) are steadily Small-Clause-final seems to suggest that this is also the basic position for the *it*-expletive.

In other words, we suggest that:

- Small clauses have some FC structure which provides the specifier for Small-Clause subjects preceding the predicate.

(130) I consider $[_{XP}$ John $[_{X'}$ $[_{AP}$ intelligent]]]

- The underlying position for Arguments of a predicate is always to the right of the predicate (this is true for any

---

[33] Our proposal, then, is that it-CP CHAINS are the only type of CHAINS which are available in all languages: there-Partitive CHAINS are only used in non-NSLs, and null-expeltive/inverted-subject CHAINS do not exist.

lexical category being a predicate, in both clauses and Small Clauses).

- Therefore, the underlying position for *it*-expletives (and their CP, which stands nearby) is to the right of the predicate even in Small Clauses: this is why, even if the *it* expletive moves to a FC specifier, the CP remains in SC-final position.

(131) I consider [$_{XP}$ it$_i$ [$_{X'}$ [$_{AP}$ strange t$_i$ that...]]]

- If this analysis for Small clauses is correct, we predict other types of I-subjects, beyond CPs, to appear to the end of the Small Clause. The following example suggests this is on the right track:

(132) - John is too stupid to do the job
- Why don't they resort to Bill?
- Because they consider Bill stupid himself

(133) They consider [$_{XP}$ Bill$_i$ [$_{X'}$ [$_{AP}$ stupid [$_{DP}$ t$_i$ himself]]]]

The Catalan equivalent of (133) is not well-formed:

(134) ?*Considero en Joan estúpid ell mateix
I-consider the J. stupid he SELF

In fact, Small Clauses are not Null-Subject structures even in NSLs like Catalan, so the prediction would be that they should not have [-anaphoric] I-subjects. They do not have [+anaphoric] I-subjects either:

(135) **Considero en Joan estúpid (si) mateix
I-consider the J. stupid (SE) SELF

I leave the issue open.

5. French Stylistic Inversion and Northern Italian Dialects

There is one case of subject inversion that does not easily fit into our theory: French Stylistic Inversion:

(136) a. Où (dis-tu qu') est allé Jean?
Where (say-you that) is gone J.
'Where do you say John went'

b. La personne avec qui (je crois que) viendra Jean
The person with who (I believe that) will-come J.
'The person with whom I think that Jean will come'

c. J'espère que vienne Jean
I-hope that come-SUBJ J.
'I hope Jean will come'

There is a general agreement that it is essentially of the same nature as subject inversion in Romance NSLs (see Kayne & Pollock (1978) for an initial characterization). If French is a non-NSL, I-subjects should be, according to the present theory, [+anaphoric]. There is a fact, however, that looks consistent with the present approach: when there is stylistic inversion, no preverbal subject is present. Standard accounts would say that only null expletives licence subject inversion. In our alternative theory not involving null expletives in subject inversion we should say that in French Stylistic inversion, since there is no preverbal subject, there is no binder for the I-subject and therefore it can be and has to be [-anaphoric].

One aspect of stylistic inversion which is not easy to deal with is the characterization of the contexts allowing it. It is triggered by Wh-movement and subjunctive mood. Concerning the former, one cannot say it is licenced by a [+Wh] COMP, because it can be licenced in the clause where the Wh- has been extracted

from even if this clause is not [+Wh] itself (see (136).a/.b) with the parenthetical part).

One way of characterizing the licensing contexts in a unitary way is the following. Suppose we assume that stylistic inversion is licenced *whenever the clause involving it is non-assertive*. Clauses containing a Wh-gap are non-assertive, and so are clauses in subjunctive mood. How could we relate this to (stylistic) subject inversion?

We have crucially assumed that I-subjects in NSLs are directly assigned Case by AGR° through Chain-government, which is made possible if V-raising to AGR° takes place. French, in fact, has V-movement to AGR° (V-movement to the top INFL category, in Pollock's (1989) terms). Therefore, if French does not have subject inversion in the general case, it is only because it is not a NSL, not because the requirement of V-movement to AGR is not met. Let us tentatively make a rather speculative proposal in this connection.

Suppose that what in fact makes French a non-NSL is not that AGR° is not intrinsically rich enough to make French a NSL, but rather that AGR° cannot manage to govern its I-subject through Chain-Government. Suppose the reason is that there is something between AGR° and the I-subject blocking government. Suppose this blocking element is absent in non-assertive sentences. Let's implement the idea.

French (like English and many non-NSLs) has the negative particle *pas* (Cf. English *not*, German *nicht*, etc.) below the inflected verb. This is what led Pollock to assume NEGP is a FC placed between T and AGR. Suppose this is right. Suppose, however, that this category is not NegP, but ΣP, as Laka (1990) suggests, Σ being a FC which includes Negation and Affirmation. Suppose we assume that Σ also contains the feature [±assertion], and that [+assertive] Σ has some blocking effect. If [-assertive] Σ did not have this blocking effect, then AGR° would be able to

Chain-govern its I-subject and French will possibly behave like a NSL in [-assertive] sentences.

The above suggestion is only tentative and faces several problems:

- it is not clear at all what the blocking effect of [+Assertive] Σ could be: it does not block V-head movement, for there is long V-movement in French assertive clauses. It only would block Chain-government.

- in [-assertive] clauses, French is not actually a NSL: it allows subject inversion but not null subjects.

In any event, it is significant that non-NSLs often have post-verbal negative particles, which suggests that Σ intervenes between AGR and T, while NSLs often have preverbal negative particles.[54] The correlation seems to have some significance: languages losing the NSL-status often shift from preverbal negation to postverbal negation: this happened in the transition from Old English to middle English and in the transition from Middle French to Modern French (see Pollock (1989)).

The fact that French [-assertive] clauses can have subject inversion, but not null subjects, suggests that the two facts need not correlate. Let us rephrase the facts in our theoretical terms.

We proposed that when AGR° is rich enough, it is the AGR-identifier and then:

a) it can directly Case mark the I-subject.

b) the I-subject can be a null pronominal (in non-NSLs it can only be a null anaphor).

---

[54] Zanuttini's (1991) typology of languages concerning negation is based on the same observation. Actually, the correlation with the (non-)NSL status is not strict: there are Italian dialects with post-verbal negation which are NSLs; and Brazilian Portuguese is a non-NSL having preverbal negation. We will speculate on the parameterization of the relative hierarchical position of ΣP in Chapter 5.

Suppose that a) is a necessary but not sufficient condition for b) to hold. French [-assertive] clauses would be a case where a) but not b) holds.

There is another well-known case where something similar happens: some Northern Italian dialects (see, e.g., Rizzi (1982), Brandi & Cordin (1989)). In these dialects, the agreement morphology in a strict sense is not rich enough to allow null subjects. But, as initially proposed by Rizzi (1982), subject clitics are part of AGR and make AGR rich enough to allow null subjects. Now, in some of these dialects (Paduan -Rizzi (1982), Trentino and Fiorentino -Brandi & Cordin (1989)-) subject clitics are required to allow null I-subjects, but they are not necessary (and in fact not possible) to allow an overt I-subject:

(137) Paduan (Rizzi (1982)):

a. (Giorgio) *(el) vien
   G. he-CL comes

b. (*El) vien Giorgio
   He-CL comes Giorgio

Like in the case of French Stylistic inversion, this suggests that the requirements for null subjects are stronger than the requirements for subject inversion. In our terms, the requirements for *pro* I-subjects are stronger than the requirements for overt [-anaphoric] subjects. The former require some minimal richness (as the one displayed in French AGR-morphology or Northern Italian strict AGR-morphology). The latter requires a full range of AGR-distinctions (which is simply not possible in French, and possible by resorting to clitics in Northern Italian dialects).

In order to capture these facts, we should refine our parameter for the AGR-identifier, repeated here:

(138) a. AGR must have an AGR-identifier.

b. X can be an AGR-identifier iff X is rich in phi-features (number and person).

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

(138).b) mentions richness in number and person as necessary for AGR° to be the AGR-identifier. In fact, if we look at the Italian dialects mentioned above, subject inversion involves person agreement, but not number-agreement (see Brandi & Cordin (1989:fn 10):[55]

(139) *6-person paradigm with a pronominal inverted subject:*

Trentino		Fiorentino	
vegno	mi	e vengo	io
te vigni	te	tu vieni	te
ven	elo/ela	e viene	lui/lei
vegnim	noi	si vien	noi
vegní	voi	vu'venite	voi
ven	lori/lore	e vien	loro

In (139) the AGR-affixes+clitics makes all person/number distinctions except for 3rd-sng/3rd-pl, which are not distinguished. On the reasonable assumption that 1st-sng is a different person from 1st-pl ('we' is not simply the plural of 'I'), and similarly 2nd-sng is a different person from 2nd-pl, so that only 3rd-sng and 3rd-pl are really the same person (or non-person), we can interpret (139) as indicating than subject inversion in those dialects involves person agreement but not number agreement (the number agreement for 1st and 2nd persons being parasitic on the person: e.g. the 4th person 'we' is inherently plural, as opposed to the 1st person 'I'). Null I-

[55] Modern Standard Arabic behaves the same as these dialects.

subjects, instead, require both person and number agreement, as we saw. Thus, we could reformulate (138) as (140), and then add (141):

(140) a. AGR must have an AGR-identifier.
b. X can be an AGR-identifier iff X is rich in *person* phi-features (five distinctions: 1st, 2nd, 3rd, 4th and 5th).
c. AGR°/Spec of AGR is the AGR-identifier of AGR.

(141) A null I-subject (be it [±anaphoric]) requires an AGR-identifier being rich in person and number features (six distinctions).

Let us assume that the Northern-Italian dialects under consideration choose the AGR°-option in (140).c), because the AGR-morphology, with the help of AGR-clitics, makes AGR° rich enough (with at least 5 distinctions). But then these dialects use AGR-morphology+clitics in a parsimonious way: they use all 6 distinctions to abide by (141) when the I-subject is null (a null pronominal), but they use only 5 distinctions when the I-subject is overt.

The fact that inverted subjects seem to require Person features (except for the special Case of French Stylistic Inversion), does not support Rizzi's (1986) view that expletives only require formal licensing (not feature identification), if, in consonance with Rizzi's view, subject inversion involved an expletive. In fact, Rizzi considers three levels of feature requirements for *pro*: person and number (referential *pro*), only number (quasi-Argument *pro*) and no requirement (expletive *pro*). We see, however, that:

- subject inversion requires person features (and in some languages even number). We contend that there is no expletive involved.

- null 'there'-expletives require number (e.g. Brazilian, German) or no feature, depending on the language (see section 3.2.).

- null 'it-CP' expletives do not apparently require features (at least in V-governed position, see Rizzi (1986)), but

- both null 'it-CP' expletives and null 'there'-expletives (and quasi-Arguments) seem to be favored in in non-NSLs at least distinguishing 3rd singular from 3rd plural (Brazilian Portuguese, Icelandic, German). So it seems that number features are important to licence null expletives.

So Rizzi's three-level distinction appears not to be straightforwardly confirmed by cross-linguistic evidence.

One tentative alternative proposal could be the following:

a) Null *there*-expletives do not exist in NSLs.

b) Null *there*-expletives require number features (i.e., a morphological distinction between 3rd-sng and 3rd-pl) (German, Brazilian Portuguese, Icelandic).

c) Null Quasi-Arguments require number features and some additional condition (which is met by Brazilian Portuguese and Icelandic but not by German).

d) Subject inversion requires person features (in our technical terms, AGR° can be the AGR-identifier only if it is rich in person features).

e) Referential *pro* requires number and person.

This is only a tentative approach, conceived solely on the basis of finite clauses. For infinitival clauses, see next chapter. What the present approach shares with Rizzi's (1986) proposal about null pronominals is the idea that, beyond requirements on formal licencing, empty pronominals are subject to requirements on *recovery of content*, which can be more or less stringent (even possibly vacuous) depending on the nature of the

pronominal. We propose, however, a reduction of the empty pronominals available.

French does not fit into the theory yet: in [-assertive] contexts, where overt [-anaphoric] I-subjects are allowed, it is not the case that 5 distinctions are provided by the verbal morphology (and subject clitics cannot be used, because they are not AGR-clitics, but merely phonological clitics). I cannot provide an account for this fact. Two possibilities could be pursued:

- French Stylistic Inversion is some marked option in which (140).b) is relaxed.

- (140).b) should be relaxed for all languages, so that AGR° can in principle be the AGR-identifier in French (and *a fortiori* in Northern Italian dialects and other Romance NSLs) but not in English or Germanic Languages, where not even [-assertive] sentences allow subject inversion.

All the preceding proposals in this section are highly speculative and only tentative. For convenience, I will continue to use (138) in the remainder of the discussion.

There are, in addition, some remaining problems which appear to be even harder to account for within the hypotheses advanced here. One is Old French, where Null Subjects and Subject Inversion are restricted to V-second contexts, i.e., root contexts for the most part (see Adams (1987)). The other is Corsican: in spite of the fact that this language has rich AGR morphology, it only behaves like Italian in root contexts. In embedded contexts a subject clitic is required for both Null Subjects and subject inversion (see Agostini (1956)):

(142) a. Quand'*(ellu) canta...
When he-CL sings
'When he's singing...'
b. Quand'*(ellu) canta Petru...
When he-CL sings P.
'When Petru is singing...'
c. Quand Petru canta...
When P. sings

Our theory the null/non-null subject contrast is based on morphological richness and does not say anything about the root/embedded distinction. It is therefore unable to deal with the issue as it stands. I think, however, that the key to account for these facts lies precisely in morphological richness: both Old French and Corsican have more syncretisms in the agreement paradigm than other Romance NSLs (in the case of Old French, spelling often showed distinctions which had no pronounced counterpart). This seems to suggest that 'slightly' empoverished paradigms can still give birth to a restricted form of Null-subject-hood, only available in root clauses.

6. A-dependencies and Minimality

Our characterization of I-subjects can be summarized as follows:

- in NSLs, a dependency is created between AGR° and an I-subject, which ultimately implies that the I-subject will directly receive Case from AGR°.

- in non-NSLs, the same dependency is created, but since AGR° itself is too poor, Spec ofy AGR has to be filled by a DP.

We have tacitly assumed that in the latter case, the DP in Spec of AGR forms a Chain with the I-subject. If Chains are formed by movement in the standard case, this suggests that the

I-subject itself moves to Spec of AGR, leaving an empty anaphor as I-subject at S-structure. Is this result correct?

For the Germanic languages, we saw that, even when the I-subject is not empty, a trace can always be claimed to be present, as exemplified in (143), where the non-empty I-subject is in fact a floating element:

(143) a. John$_i$ did it t_i
b. John$_i$ did it [$_{DP}$ t_i himself]

In French, however, the I-subject cannot be claimed to always involve an empty anaphor:

(144) Jean l'a fait lui
J. it-has done he
'John did it himself'

As we saw, in (144), the I-subject *lui* cannot be analyzed as a floating element cooccurring with a trace as it does not form a possible constituent with the preverbal subject. Then no trace is involved in this Case. So the dependency between *Jean* and *lui*, even if similar to a Chain, is not a standard case of Chain (nor is it a standard case of CHAIN either, since there is no expletive).

We have not presented any account for why French allows such a non-standard kind of Chain formation while Germanic languages do not. We will try to derive this contrast from some independent linguistic facts.

Kayne (1987) assumes that French has Object Agreement. Although he is not very precise as regards the characterization of this Agreement, it could be naturally conceived, within the present widespread conception of Functional Categories, as Functional Category to which the (participial) verb raises.

The present day multiplication of hypotheses about Functional Categories after Pollock (1989) initial proposal raises a fundamental question about their universality. One possible view is that all of them are universal, although possibly not morphologically realized in some languages. According to this view, English would have Object-Agreement, which, as opposed the French one, would have no morphological manifestation. An alternative, and perhaps more realistic view, is that only those FCs exist in a language which have some morphological manifestation. In this view, English would not have object Agreement.

Of course, an intermediate hypothesis is possible: some FCs are universal (whether they have overt manifestations or not) and others are not (and will be present only in languages where they can be detected by the morphology). Chomsky's recent hypotheses (See Chomsky (1992)) postulate that Object-Agreement belongs to the universal type of FC, since it is the universal means of Accusative-assignment. In fact, this hypothesis is at odds with of our theory on Accusative, as far as we claim that Accusative Case is in principle available independently of lexical idiosyncrasies.

I think, however, that French Object-Agreement and Chomsky's abstract Object-Agreement need not be assimilated as a single concept. French object Agreement is only operative in a restricted class of sentences, namely the ones involving a moved object of some kind (an object clitic, a Wh-moved object or a A-moved object). To account for this limited distribution, Kayne assumes that the Specifier of Object Agreement in French is not a Case position. Since Accusative assignment in French is quite independent of the Object Agreement restricted paradigm, we are led to the conclusion that, if a universal Object agreement is to be postulated, it has nothing to do with French overt Object Agreement.

We will ignore for the moment the possibility of non-overt Object Agreement, and assume that only French has an Object Agreement FC. I come back to the issue below.

Suppose that Specifier of Agreement FCs counts as an A-position, as would be predicted by Rizzi's (1991-b) theory. If we adopt Relativized Minimality (RM), then the existence of an Object Agreement poses an immediate problem for the Internal Subject Hypothesis. Since the Spec of O-AGR counts as an A-position, raising of the external Argument to Spec of Subject-AGR will skip such an A-specifier, which is forbidden in RM. A French sentence like (145) should violate RM:

(145) Jean$_i$ les$_j$ a [$_{O\text{-}AGRP}$ t_j repeintes$_v$ [$_{VP}$ t_i [$_{VP}$ t_v t_j]]]

Since (145) is well-formed, the prediction is incorrect. To solve this problem without giving up RM, let us assume the following UG options.

Suppose that, whenever Spec of (Subject) AGR has to be filled because it is the AGR-identifier, the unmarked way of supplying the required DP is for the I-subject itself to raise to this position. However, when a language (such as French), has A-specifiers intervening between the I-subject and Spec of (Subject) AGR, then this option will not be available in sentences where Object Agreement is present, namely, sentence with compound tenses. In this case, a marked option can be adopted by which a DP is directly inserted in Spec of AGR and it is coindexed to the I-subject in a resumptive-like way. Since French is forced to take this option, then overt I-subjects in French need not involve an empty anaphor, as is the case in (144). In Germanic languages, on the other hand, no Object Agreement being present, the unmarked option is taken and Spec of AGR is always filled by movement, which accounts for the floating character of overt I-subjects.

In Germanic languages, both overt I-subjects and FQs would be elements left floating by A-movement. We argued in section 2.4. that in NSLs FQs need not form a possible constituent with the preverbal subject. We argued this is because in NSLs no A-Chain is formed between the preverbal subject and the I-subject, so there is no reason why FQs should be analyzable as elements left behind by movement. Now, for French we assume there is a Chain, but this Chain is not (always) a movement Chain, but rather a resumptive-strategy Chain. Then the prediction is then that in French FQs are not necessarily left behind by movement and therefore need not form a possible constituent with their preverbal subject. Specifically, *tous les deux/trois* 'all the two/three' (= 'both'/'the three of them') are FQs that do not form a possible constituent with their antecedent:

(146) a. *Les enfants /ils* sont allés *tous les deux* au cinéma
The children/they are gone all the two to-the movies
'The children/they have both gone to the movies'
b. *Tous les deux les enfants / *tous les deux ils
All the two the children all the two they
'Both the children' / 'They both'

For NSLs having object agreement with a pattern similar to the French one (Italian, some Catalan dialects), the minimality problem does not arise, since we assume that those languages never involve I-subject raising.

Let us briefly speculate on two issues that the above approach raises. One is the existence of resumptive A-Chains, as we could call the Spec-AGR/I-subject dependencies in French cases like (144). How do they escape being subject to the ECP? In fact, a representational point of view (which is always worth keeping an eye on) would not easily distinguish between resumptive and movement A-Chain.

It seems, however, that the distinction between movement dependencies and resumptive dependencies is an irreducible complexity: in the case of A'-dependencies it would be hard to reduce one to the other.[54] It is clear as well that resumptive strategies are more permissive than movement strategies. Then it seems that it could be interesting to extend the distinction to A-Chains, provided we manage to put some constraints on the availability of the options (such as the markedness strategy we propose).

Another issue to consider is Chomsky's Object-AGR. As we said, the existence of universal Object-AGR is in principle a welcome possibility for a theory claiming that Accusative is always available in principle. However, Chomsky's proposal raises several problems. One is that it is considerably theory internal and empirically under-determined: for many languages there is little evidence that object agreement exists. If objects receive Accusative by moving to Spec of Object-AGR, then we could possibly expect some head initial languages to have the order Object-Verb (i.e., languages having overt object movement and covert V-movement). This unattested possibility can be excluded somehow, but then the hypothesis has little predictive power for comparative syntax.

Perhaps part of Chomsky's theory could be kept without conflicting with RM. In fact, as far as I know, Chomsky's recent proposals have been built as an alternative to RM, and it is not surprising that the two theories conflict. However, sometimes it happens that alternative research strategies are not pointing at the same theoretical domain and are not, therefore, theoretically incompatible. It could turn out, for instance, that Chomsky's Spec of Object-AGR position is not an A-position as far as RM is concerned (for instance, we could try to define, in the spirit

---

[54] See Cinque (1990) for extensive discussion on the issue.

of Rizzi (1991-b), (non-θ) A-positions as Specifiers of FC showing overt agreement.[57] Object-AGR would never be involve overt agreement in many languages.

Another possibility we could contemplate is that Object-AGR assigns Accusative by government, and that then its Specifier is not (necessarily) projected, so that no problem for RM arises. I leave the issue here.

7. Summary

In this chapter we have advanced a theory for deriving Burzio's Generalization and characterizing the status of the Null Subject Phenomenon. Concerning the former, we have proposed some parameters for accounting for impersonal constructions in languages like German on one hand, and for Ergative languages on the other hand.

Concerning the Null Subject phenomenology, his theory tries to minimize the differences between NSLs and (non)-NSLs: all languages have some form of null subjects and subject inversion, the difference lying in the [±anaphoric] character of both.

Our theory crucially relies on the properties of AGR, which we claim are universal (*modulo* some parameters concerning Ergative Languages and the presence of phi-features in a given language), and can be conceived of as instances of licensing principles.

This formulation led us to revise Binding Theory and Case Theory. We have provided some independent motivation for our revision of Binding Theory (anaphoric copulative constructions).

---

[57] In that sense, English null C°, which is a manifestation of agreement in Rizzi's theory, would be 'overt' agreement in the sense that its obligatory null form in the relevant cases is in 'overt' contrast with the optionally overt form (that).

The strong requirements AGR imposes under our characterization of the facts led us to consider whether the ECP should be the appropriate means of constraining A-movement. We noticed that there is potential redundancy between the ECP account and our account, and that we could eliminate this redundancy by reducing the power of our rules for AGR.

We have also addressed a variety of empirical problems (inverted Copulative constructions, French Stylistic Inversion, Northern Italian dialects, etc.) and theoretical problems (Minimality, Chain formation, etc.) in a rather speculative and often inconclusive way: it could not be otherwise, given the wide range of implications the present thesis is involved in.

Chapter 4
Infinitival Constructions

In the preceding chapter we have developed a theory on how AGR determines what is the I-subject in a clause, in order to derive Burzio's Generalization and explain the contrast between NSLs and non-NSLs w.r.t. subject inversion. We claimed that AGR morphology plays a central role in determining a good deal of cross-linguistic contrasts.

If this is correct, then our account of infinitival clauses cannot be a trivial extension of our theory for finite clauses, for non-finite clauses have the central property, in many languages, of not showing any AGR morphology. So, two possibilities come to mind: either non-finite clauses have a radically different behavior w.r.t. the phenomena discussed in the previous chapter, or morphology is not so crucial as we claimed in accounting for those phenomena.

Our proposal will be that neither situation is exactly true: although non-finite clauses have more restricted possibilities, they are in many essential respects similar to finite clauses, because, on the one hand, they have alternative means of recovering AGR content apart from morphology and, on the other hand, they are subject to some *parallelism* principles w.r.t. the finite clauses in the same language. In other words, the speaker recovers the lack of information in non-finite clauses from both UG and some parametric options fixed on the basis of finite clauses, ultimately, from the richness of AGR in finite clauses.

1. Some Properties of Infinitival Constructions

A mainstream of generative analyses of infinitives is based on the hypothesis that infinitives are full clauses differing from finite clauses only in their impoverished inflectional content. The highly restricted possibilities for subjects in infinitives would be derived from the weak character (or perhaps absence) of inflectional content. The standard analysis since Chomsky (1981) assumes that infinitival INFL can neither govern nor Case-mark its subject, so that either this subject has to be PRO (which need not be governed nor Case-marked) or obtains Case independently of the infinitival INFL (in ECM constructions, 'for'-infinitives or raising constructions).

Concerning PRO, its restricted distribution has been made to derive from the postulation that it is a [+pronominal, +anaphoric] DP. Since BT requirements on [+pronominal] and [+anaphoric] elements are contradictory, PRO has to escape such requirements by being ungoverned, hence having no Binding Domain. Such a theory faces two main problems:

a) Since PRO escapes all binding requirements, it remains a mystery why its reference requirements are so highly restricted: either it is controlled or it receives arbitrary interpretation, the choice not being free in most cases. We will not propose any interesting solution for this fact, but we will contend that it is not possible to derive control from other modules of the grammar (such as Binding Theory).

b) Given a theory of Visibility as defined in Chomsky (1986-b), if PRO does not receive Case, it should not be licenced as an Argument.

Kayne (1991) presents a hypothesis that avoids at least the former problem: PRO is always governed (at least at LF) and therefore it is subject to binding requirements, so that control is reduced to BT. We will argue that this positions is untenable.

Kayne does not say anything about the possibility that PRO may receive Case, but his theory could be extended in that direction more easily than standard ones. We will discuss Kayne's proposal in detail in section 1.3.

1.1. Parallelism between Finite and Infinitival Constructions

Infinitives are like finite clauses except for their lack of explicit content in INFL features, from which their dependent character in both temporal interpretation and subject possibilities results. Essential for our concerns is the fact that I-subjects in infinitives show the same restrictions as in finite clauses, as far as the generalization we proposed in Chapter 2, repeated here as (1), is concerned:

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

The following examples show that (1) holds in infinitives the same as in finite clauses:[1]

(2) <u>English, French, German:</u>

a. John decided [to do it *himself/*him*]

b. C'est mieux [de le faire *soi-même/*on /*lui(-même)*]
   It is better to it-do SE-SELF/ one/ he (-SELF)
   'It is better to do it oneself'

c. Hans beschloss [es *selbst/*er* zu machen]
   H. decided it SELF / he to do
   'Hans decided to do it himself'

---

[1] In controlled infinitives, NSLs allow pronominals as I-subjects, but not R-expressions. We will derive this fact from our theory of control.

(3) Catalan, Spanish, Italian:

a. En Joan prefereix [de fer-ho *ell (mateix)/*si mateix*]
   The J. prefers to do-it he (SELF) / SE SELF

b. Juan prefiere [hacerlo *él (mismo) /*si mismo*]
   J. prefers to-do-it he (SELF) / SE SELF

c. Gianni preferisce [di farlo *lui (stesso) /*se stesso*]
   G. prefers to do-it he (SELF) / SE SELF

Given that our account of the distribution of I-subjects crucially relies on the Case position for the I-subject and, on the other hand, PRO-infinitives are standardly assumed to contrast with finite clauses as far subject-Case is concerned, the facts in (2)/(3) do not trivially follow from the assumptions we made to account for (1) in finite clauses.

Suppose we assumed that PRO does not have Case. Then the A-Chain containing the I-subject would not have a Case position, and no Binding Domain would be definable for the I-subject, according to our definition of Binding Domain, repeated here:

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

If no Binding Domain is definable, no prediction is made concerning the [±anaphoric] character of the I-subject. This is an undesirable result, since the infinitives show exactly the same behavior as finite sentences, as we see in (2)/(3), and this should not be a matter of accident. In addition, what we called reflexive copulative constructions, whose behavior we derived from the same definition of Binding Domain behave exactly the same in finite and infinitival constructions:

(5) a. He tried [to be *himself/*him* again]

b. On *doit* essayer [*d'être soi-même/*on/*lui(même)*]
One must try to be SE-SELF / one/he (SELF)

c. Er versuchte [wieder *sich selbst/*er* zu sein]
He tried again SE SELF / he to be
'He tried to be himself again'

(6) a. Intentava de tornar a [ser *ell (mateix)/*si mateix*]
He-tried to return to be he (SELF) / SE SELF
'He tried to be himself again'

b. Intentaba volver a [ser *él (mismo)/*si mismo*]
He tried to-return to be he (SELF) / SE SELF

c. Non riusciva ad [essere *lui (stesso)/*se stesso*]
Not managed to be he (SELF) /*SE SELF
'He didn't manage to be himself'

So we are led to the conclusion that infinitival I-subjects have a Case position, and that the Case position is, for as given language, the same as that of the finite clauses.[2] More specifically, infinitival I-subjects inherit Case from Spec of AGR in non-NSLs, while they are directly Case-marked in NSLs. Since this is a necessary requirement for our theory to be extendable to infinitives, we have to build some plausible theory that achieves this result. As far as the theory of visibility or, more generally, any theory pointing to uniform Case requirements for all Arguments is to be welcome on the grounds of simplicity, our proposal will have some independent plausibility.

[2] We will propose that this is not the case for some languages such as Occitan and Sardinian.

1.2. V-movement in Infinitives

Pollock (1989) shows that verb movement (for lexical verbs) in infinitives is shorter than in (French) finite sentences:

(7) a. Jean ne pense *pas toujours* au futur
J. ne thinks not always of-the future
b. Ne pas (toujours) penser (toujours) au futur...
ne not (always) to-think (always) of-the future.
c. *Not to always* think of the future...

Pollock's theory derives this fact from the 'weak' or 'poor' character of Tense in non-finite sentences. Since all Romance infinitives show no Tense morphology on the infinitival verb, the prediction should be that infinitives in these languages do not allow long verb movement, contrary to fact: in Italian, Spanish and Catalan, infinitival verb movement is apparently as long as finite verb movement.[3]

(8) a. Non ama *più* Maria (Italian)
Not loves *anymore* M.
b. Non (*più*) amare *più* Maria...
Not to-love *anymore* M.

[3] If not longer: we will argue that enclisis in infinitives should be analyzed as extra verb movement (our proposal will differ from Kayne (1991), who argues for a very short kind of extra movement). See Belletti (1991) for the idea that elements like più, mai, etc. occupy a position similar to French pas, plus and other negative elements. On the other hand, there are NSLs with short V-movement in infinitives. We will address the question in section 1.3.

c. No diu *mai* la veritat (Catalan)
Not tells *never* the truth
d. No (*mai*) dir *mai* la veritat...
Not to-tell *never* the truth

e. No està *nunca* cansado (Spanish)
Not is *never* tired
f. No (*nunca*) estar *nunca* cansado...
Not to-be *never* tired

Thus, there seems to be a correlation between the Null-Subject status of the language and the possibility for the infinitival verb to raise to a high (the highest) functional category: AGR in Belletti (1991)'s theory and our own. Another fact, which is likely to be parasitic on the former, is the possibility of clitic climbing. Kayne (1989) argues that clitic climbing (and also long V-movement in infinitives) is due to the strong character of INFL in NSLs languages.

The problem is how to express the correlation between the strong character of AGR in finite clauses and the purported strong character of AGR in non-finite clauses. In infinitival clauses, AGR is not apparently strong in Italian or Spanish, as far as morphology can tell us. To simply stipulate that infinitival AGR is strong because finite AGR in the same language is strong appears to be a mere stipulation.

What I want to propose is the idea that what extends from finite to infinitival clauses is a parameter value. But before proceeding, let us consider the facts considered in Kayne (1991) concerning V-movement and clitic placement in infinitives, and the proposal Kayne presents to account for the facts.

1.3. Kayne's Proposal

Kayne (1991) shows that in Romance languages there is a correlation between enclisis in infinitives and the possibility of having interrogative infinitives with a COMP particle of the type 'if', which he convincingly argues is a C° particle:

(9) *Italian:*

a. andarci / vederlo (enclisis)
   to-go-*there* to-see-*it*

b. Non so se andarci
   Not know-I if to-go-there
   'I don't know whether to go or not'

(10) *French:*

a. y aller / le voire (proclisis)
   there-to-go it-to-see

b. *Je ne sais pas si y aller
   I ne know not if there to-go
   'I don't know whether to go or not'

Kayne argues that this is a genuine correlation holding of many Romance languages, and that it does not correlate with the (non) NSL status of the language.[4] So Catalan, Spanish, and some Italian dialects (Piedmontese, Milanese and Paduan) are like Italian, while Occitan, Sardinian and some Italian dialects (Gardenese) are like French. As for English (and languages having short V-movement in infinitives in general), they would pattern with French in not allowing 'if'-infinitives:

---

[4] Kayne argues that the NSL status is a necessary but not sufficient condition for a language to have enclisis and thus 'if'-infinitive constructions.

(11) a. No sé *si* anar-*hi* (Catalan)
Not know-I if to-go-there

b. No sé *si* hacerlo (Spanish)
Not know-I if to-do-it

c. *Sabi pas se *hi* anar (Occitan)
Know-I not if there-to-go

d. *No'isco si andare (Sardinian)
Not know-I if to-go

e. *I don't know if to go (English)

Kayne's analysis of these facts accounts for both the (im)possibility of 'if'-infinitives and the enclisis phenomenon. Concerning the latter, he assumes that clitic attachment is (at least in Romance) *left*-adjunction to a functional head, for principled reasons having to do with morphological headedness. Therefore, whenever a clitic is strictly left-adjacent to the verb, the structure would be:

(12) $[_{I^\circ}$ cl $[_{I^\circ}$ V°]]

where I° is some functional (inflectional) head. In fact, the general condition Kayne imposes on clitic attachment is that clitics must attach to a functional head, whether it contains the moved V° or is silent. However, he follows Baker (1985) in assuming that the head the clitic adjoins to cannot contain a trace. A consequence of this assumption would be that, since verb movement is cyclic for principled reasons, the clitic cannot adjoin to a functional head lower than the one where V° has moved, because this would imply that the clitic is adjoined to a head containing a trace:

(13) *$[_{X^\circ}$ V°] ... $[_{Y^\circ}$ cl $[_{Y^\circ}$ t]]

where X° and Y° are functional heads and *t* is a necessary intermediate trace of V°-movement. The reverse situation, where the clitic is in a higher functional head than V° is not excluded:

(14) $[_{X^\circ}$ cl X°] ... $[_{Y^\circ}$ V°]

This would be the structure for languages having the clitic preceding the infinitival verb but not necessarily adjacent to the verb (Occitan, earlier French).

Since both V° and the clitic are heads, why is it that the clitic can move non-cyclically (giving (14)) while V° cannot (as far as (13) is not allowed)? Kayne's solution is based on the idea, developed in Kayne (1989), that for NSLs infinitival INFL (or some of its members) is strong enough to l-mark and void some potential barriers,[5] so that long head movement is allowed to some extent. However, V°-movement is subject to a further constraint: it has to pick up the affixal functional heads. This would explain the contrast (13)/(14): in (13) the verb has to move through Y° to merge with this affixal head, while the clitic in (14) meets no similar requirement.

Since cyclic movement (as expressed by the HMC) is not a matter of principle, but is rather derived (from the ECP and affixation requirements), there might be structures where V° skips the clitic position without violating any principled requirement. Kayne argues that there is one such structure. Suppose the clitic moves to a functional head I°, left-adjoining to it (as it has to); suppose I° has no content, so that V° is

[5] Specifically, in Kayne (1989) VP would be the potential barrier that is voided by INFL° l-marking. Although Kayne (1991) does not address the question, at least some of the 'heirs' of INFL after the INFL splitting hypothesis should be l-marking elements for the hypothesis in Kayne (1989) to be extendable to the proposal in Kayne (1991).

not required to move to it to pick up any affixes; suppose finally the verb adjoins to the I' projection, giving:

(15) $[_{I'}\ V^{\circ}\ [_{I'}\ [_{I^{\circ}}\ cl\ I^{\circ}\ ]\ ]\ ]$

(15) does not possibly violate the ECP, since I' is not a minimality barrier (in the sense of Chomsky (1986-a)) because it does not exclude V°; it does not violate any affixation requirement as far as I° is content-less. Therefore it is a possible structure. Kayne claims that (15) is actually the structure for enclisis in Romance infinitives: the apparent fact that the clitic is right adjoined to the (functional head containing the) verb is due to the verb having skipped the position of the clitic. Kayne assumes that I° is T°, which is likely to be content-less in infinitives. On the other hand, the proposal that clitic attachment (in Romance) is left adjunction to a functional head can be assumed to hold without exception. And the very short span by which V° out-raises the clitic accounts for the apparent right-adjunction of the clitic to the verb. For more details on this proposal, see Kayne (1991).

The existence of a structure like (15) is only partially derived in Kayne's account: since it involves long clitic movement, it should be restricted to NSLs (as argued in Kayne (1989)), but being a NSL is not a sufficient condition for allowing (15), as far as there are NSLs not having enclisis in infinitives (Occitan, Old French, Sardinian). In any case, Kayne develops a proposal that makes the existence of such a structure highly appealing, in that it allows for an explanation of the generalization we mentioned above, namely that only languages having enclisis in infinitives allow 'if'-infinitives (recall the contrast (9)/(10)).

Kayne's account is based on several assumptions. One is that 'if' is a C°-particle that, contrary to empty infinitival C°, is

able to govern the specifier of IP, which is occupied by PRO in control infinitives.[5] If we simply assume that PRO cannot be governed (due to the PRO-theorem), then we explain that 'if'-infinitives are not allowed in French or English. We do not thus far explain, however, why languages having enclisis allow 'if'-infinitives, since the existence of the structure (15) does not trivially bear on the PRO-theorem problem.

In order to account for the facts, Kayne adapts a proposal by Chomsky (1986-b) on the definition of Binding Domain. Chomsky wanted to account for cases where pronominals and anaphors are not in complementary distribution, such as:

(16) a. *They* like *their* pictures
b. *They* like *each other's* pictures

Chomsky's proposal is that the definition of Binding Domain is sensitive to the [±anaphoric] character of the element whose Binding Domain is determined. Without going into the details of Chomsky's technical definitions, the essential idea is that the Binding Domain for a pronominal/anaphoric element has to fulfil

---

[5] Kayne crucially assumes that prepositions preceding (controlled) infinitives in Romance, as in French (i), are not in C°:

(i) Jean essaie de comprendre

for they do not have the same effects as 'if'. I agree on this point, and this will be crucial for my account too. However I think the preposition de (and its Romance counterparts) cannot be in Spec of CP, as Kayne assumes. On the one hand, since a specifier cannot be occupied by a head, we would be dealing with a complement-less PP. On the other hand, if de is in Spec of CP, de-infinitives would be like Wh-islands, contrary to fact:

(ii) Quand$_i$ as-tu essayé [de venir t$_i$]
When have-you tried de to-come
Cf. *Quand$_i$ m'as-tu demandé [où Jean allait t$_i$]
When me-have-you asked where J. went

We could assume that de is outside CP. This is not in fact incompatible with Kayne's or my theory, nor is it crucial to either, provided de is not in C°.

the condition of virtually allowing for the pronominal/anaphor to be respectively free/bound. In more concrete terms, this means that X is Binding Domain for an anaphor only if X contains an A-position c-commanding the anaphor; since pronominals do not have to (in fact cannot) be bound in their Binding Domain, there is no requirement of virtual binding for pronominals.

With these assumptions, the facts in (16) can be explained. Consider whether the NP containing the genitive is a possible Binding Domain for the pronominal or anaphor in this genitive position. Suppose this NP fulfils that part of the definition of Binding Domain which is common to pronominals and anaphors, namely:

(17) B is a Binding Domain for A iff B is the minimal CFC containing A and a governor of A.

For the anaphor (*each other*), however, there is the further requirement that the Binding Domain has to contain an A-position c-commanding it. Since the anaphor is in the specifier of the NP, i.e., the highest specifier in this NP, such a requirement is not met, and the Binding Domain will be the next CFC up, namely the whole sentence, where the anaphor will be correctly bound. For the determination of pronominal's Binding Domain, there is no such requirement, so the NP itself can be the Binding Domain, allowing the pronominal (*their*) to be free in it.

PRO, like genitives, can be assumed to occupy the highest specifier of the IP. Since IP is a CFC, Chomsky's proposal can easily be made to bear on the PRO distribution. Since Chomsky's proposal allows for the determination of the Binding Domain to give different results depending on the pronominal/anaphor status of the element in question, then the simultaneous pronominal/anaphoric status of PRO need not lead to the PRO-theorem. Specifically, if a CFC XP contains PRO, a governor of

PRO and no A-position c-commanding PRO, then PRO as pronominal will have XP as its Binding Domain, where it will be free as required, but XP will not be the Binding Domain for PRO as anaphor: the Binding Domain will be the whole IP next up, where it will be bound by the controller in the standard case.

Now consider the structure of an 'if' infinitive in English or French, where it is excluded:

(18) $[_{C^\circ}$ if $]$ $[_{IP}$ PRO $[_{I'}$ I° $]$ $]$

Since there is no governor of PRO inside IP, IP cannot be the Binding Domain for PRO. In fact, the governor for PRO is C° (containing 'if'), but since C° is outside the IP CFC, the Binding Domain has to be extended to the next CFC, namely the superordinate clause, where PRO, although being correctly bound by the controller as an anaphor, is incorrectly bound as a pronominal. Therefore the 'if'-infinitive is correctly excluded by the second BT principle.

Now consider the structure of an 'if'-infinitive construction in a language having infinitive enclisis. In these languages, V° is adjoined to I':[7]

(19) $[_{C^\circ}$ if $]$ $[_{IP}$ PRO $[_{I'}$ V° $[_{I'}$ I° $]$ $]$ $]$

In (19), Kayne argues, C° cannot govern PRO because there is a closer governor, namely V°, which creates a minimality effect. Therefore IP is a CFC containing a governor for PRO (namely V°). For IP to be the Binding Domain for PRO as pronominal nothing else is required, and PRO as pronominal is correctly free within IP. On the other hand IP does not fulfil the virtual binding

[7] V° adjunction to I' would take place in all infinitives in these languages independently of whether an (en)clitic is present or not: enclisis would only be a manifestation of the phenomenon.

requirements of PRO as anaphor, so the Binding Domain for PRO-anaphor is extended to the superordinate clause, where it is licitly bound by the controller.

So NSLs having infinitival enclisis allow for PRO to be governed without violating BT. At S-structure, the other set of languages do not have a governed PRO. Since, however, the hypothesis that PRO may be governed and, therefore, subject to BT is a natural way of accounting for the strict referential constraints on PRO (it is controlled or arbitrary), Kayne proposes that the governed status of PRO is universal (i.e., PRO never escapes BT). In languages where it cannot be licitly governed at S-structure, it is governed at LF, where V-raising would create a structure similar to (19).

Kayne's analysis is appealing in several respects:

- It accounts for an interesting generalization concerning enclisis and the existence of 'if'-infinitives.* This generalization could hardly be a matter of accident and, even if it was, it would pose a problem for learnability as far as, for languages not having 'if'-infinitives, the learner would need negative evidence in order not to generalize Wh-interrogative infinitives to 'if'-infinitives on the basis of finite clauses, where both Wh- and 'if'-interrogatives are possible.

- It sets a plausible basis for reducing control to BT, a desirable result in view of the hitherto poorly understood phenomenon of control.

- it accounts for (the possibility of) enclisis in a highly principled way, on the basis of the ECP and affixation constraints in V-raising, as well as morphological-headedness constraints on clitic attachment.

---

* Portuguese is a potential problem for the empirical generalization, as we will see above.

However, Kayne's proposal is far from crystal-clear in many respects, which we will consider in turn.

One question Kayne's proposal raises concerns his crucial claim that V° adjoins to I'. There are two possible problems with this proposal. One is whether adjunction of X° to Y' should be allowed at all. The other problem is that Kayne's analysis of proclisis vs. enclisis analyses the former as having the clitic more closely attached to the (functional head containing the) verb than the latter. There is some evidence pointing to the opposite way (see Benincà & Cinque (1990)). Since, however, these two possible objections are extrinsic or peripheral to Kayne's discussion, we will not pursue them here. We will concentrate on intrinsic problems Kayne's theory cannot escape facing.

There is a problem that is essentially connected to Kayne's proposal, although he does not explicitly address it. Kayne seems to tacitly assume that the only means of having PRO governed inside IP in structures like (18)/(19) is having the governing head adjoined to I'. Let us see why this tacit assumption is necessary. Kayne analyses Sardinian as a language having infinitival V-raising to T° (= I° in (18)), i.e., to the functional head the clitic is attached to. Since this language does not allow 'if'-infinitives, we must imply that V in I° cannot govern PRO while V adjoined to I' can. Otherwise, Sardinian would allow 'if'-infinitives. Or would it not? Let us consider the possibility that a head governs its specifier.

Kayne seems to emphasize the idea that V-adjoined-to-I' plays the role of blocking government by C° by minimality: so, Kayne could argue, I°, even if able to govern its specifier PRO when filled by V°, would not block government by C°. But notice that the government requirement in the definition of Binding Domain (as in (17)) is not a requirement of exclusive government: if I° governs PRO in its specifier, that is sufficient for IP to become a potential Binding Domain (and an actual one for PRO as

pronominal). So, to the extent we allow a head to govern its specifier, and we assume that functional head containing V° is a governor, then it is immaterial whether C° governs PRO or not whenever PRO is governed by I° containing V. So Sardinian, for which Kayne assumes V raises to the functional head whose specifier contains PRO, would be predicted to allow 'if'-infinitives, contrary to fact.

An option for solving this problem could be that heads do not govern specifiers, so that adjunction to I' is the only means for V to govern PRO from inside IP.

But this is not easily tenable either. Consider again Chomsky's original proposal. If heads do not govern specifiers, we would have to assume that in a case like:

(20) *They* like [$_{NP}$ *their* books]

the genitive pronoun (*their*) has to be governed by a head adjoined to N' (or D' if we adopt the DP hypothesis) in order to prevent the verb (*like*) from governing the genitive and enlarging the Binding Domain to the whole clause. Put in general terms, Kayne's proposal, although designed to account only for the PRO distribution, leads to the conclusion that any case of non-complementary distribution of pronominals and anaphors involves X° adjunction to Y'.

If we go further into the consequences of this proposal, even more basic problems appear. Consider a structure like:

(21) **John* doesn't remember if [$_{IP}$ *himself* actually won]

If the 'if' C° is a governor of the Spec of IP in infinitives, it is likely to be so in finite contexts as well. Therefore, the IP in (21) should not count as a Binding Domain unless, again, we assumed that the some element adjoined to I'

blocks this government. In this case, to impair things, the element in question would be empty, in fact radically empty if affix lowering takes place at S-structure, as in Chomsky (1988).[9] Kayne crucially assumes for independent reasons that an anaphor has to abide by BT requirements at S-structure, so that (21) could not be rescued at LF.

To try to rescue Kayne's hypothesis from both (20) and (21) there are some ways-out to try. We could assume that in genitives the *'s* particle (which would be morphologically irregular for *their*) is actually a D° element adjoined to D', which would elegantly account for its enclitic character.[10] To avoid the problem in (21), we could simply assume Rizzi's (1989-b) idea of the 'Anaphor Agreement Effect', by which anaphors in subject position of tensed clauses would be excluded on independent grounds. However no similar account is possible for the grammaticality of (22), which would be incorrectly ruled out because of BT second principle, in parallel with (21) being incorrectly ruled in:

(22) *John* does not remember if $[_{IP}$ *he* actually won]

In conclusion, although Kayne unproblematically accounts for the impossibility of 'if'-infinitives in short infinitival V-movement languages, the account is less clear for enclisis languages, and it potentially runs into problems for Sardinian,

---

[9] Notice that (21) is a problem for Chomsky's (1986-b) proposal as well: he has to assume (and he does quickly in passing) that AGR counts as a virtual antecedent without being a possible antecedent, in order for the IP to be a Binding Domain for **himself** without possibly being one in which **himself** can be bound.

[10] We would assume, as in Fukui & Speas (1986), that 's is the head of D° being a genitive Case assigner.

where it is not clear why 'if'-infinitives are not possible, and, besides, for cases like (20), (21) and, more conclusively, (22).

If we assume that I° can govern its PRO specifier then Sardinian ought to allow 'if'-infinitives as we argued. Suppose we assumed that Sardinian infinitival V-movement, although very 'long' (almost as long the Italian one) does not however reach the I° whose specifier contains PRO, but rather stops a step short of it. Then, the Sardinian status would be essentially the same as the French one: an empty I° is unable to govern and the 'if' C° really matters. Although this position is tenable for Kayne, it undermines the cruciality of the V-adjoining-to-I' proposal: we could simply assume that long movement is movement to I°, (which for some reason implies enclisis), and that suffices to prevent C° to be the exclusive governor of PRO. In languages with short movement (and now Sardinian is not relevantly different from English, French or Occitan) I° is empty and C° is the exclusive governor. So the idea of adjunction to I' is, at best, less motivated than Kayne claims it is.

There is another problem of a rather speculative nature. In both Chomsky's and in Kayne's proposal (implicitly in the latter), the notion of CFC plays an essential role in the definition of Binding Domain: since only a CFC can be a Binding Domain, whenever a CFC (such as the infinitival IP) fails to be a Binding Domain (because the governor is outside it or it does not satisfy virtual binding requirements) then the Binding Domain switches to the whole next IP up. The CP or VP immediately dominating the IP are not possible candidates simply because they are not CFCs. At the time Chomsky formulated his proposal, he assumed a still fairly simple structure for the sentence (the S'/S category) and, specifically, he did not consider the internal subject hypothesis at all: (NP,S) was the θ-position for at least external Arguments, so S (=IP) was clearly the minimal

constituent containing all the θ-positions of a clause (at least in agentive clauses), thus a CFC.

The moment the internal subject hypothesis is assumed, however, things are not so simple: it could be argued that VP"" (in Koopman & Sportiche's (1988) sense) is the (minimal) CFC. Similar problems would arise in connection to the DP hypothesis if we generate nominal θ-positions inside the (strict) NP. In order to preserve Chomsky's results, there is an obvious solution: we define CFC as the minimal constituent containing all the θ-positions of a predicate plus the A-specifiers of the FCs locally dominating it. If the specifier of the highest IP (or the genitive specifier of DP) are A-positions, then the definition will give CFCs which coincide in essence with Chomsky's and Kayne's proposal: the highest IP for clauses and DP for nominals.

This simple solution is, however, possibly too simplistic: the notion of A-position is now not an obvious one anymore, and both Rizzi (1991) and Chomsky (1992) are trying to define it in a derivative way. At least Rizzi's characterization (which allows Spec of CP to be an A-position under certain conditions) could be problematic. I will not speculate further on these ramifications.

Crucial for the present thesis is, however, the fact that we postulate that, at least in the set of NSLs considered, I-subjects (which are the most prominent θ-positions in a clause) do not form an A-Chain with Spec of IP, then there is no plausible motivation for extending the CFC from VP"" to IP. In fact, our definition of Binding Domain basically restricts the candidates by requiring them to be functional categories:

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

In all hypotheses about FC structure which simultaneously assume the internal subject hypothesis, FCs (as the various members of IP or DP) form a 'shell' around the predicative lexical categories (resp. VP^{Max} and NP). Therefore requiring the Binding Domain to be a FC amounts to ensuring that a Binding Domain will always contain the whole CFC (if the latter is defined on the basis of θ-positions only).

There is another problem Kayne himself points out: if enclisis is the manifestation of close adjacency between V and the clitic without both forming a constituent, then the prediction is that V-cl cannot move as such. Since AUX-to-COMP (as characterized in Rizzi's work) has enclisis on the AUX, this could only follow from the accidental fact that the clitic happens to (left) adjoin to C^{0} and the AUX adjoins to C'. Although not an impossible accident, it is rather suspicious that enclisis is preserved in AUX-to-COMP.

Kayne points out an additional problem: European Portuguese allows both enclisis and proclisis in infinitives. Nevertheless, it allows 'if'-infinitives. Kayne's proposal would be strongly confirmed if Portuguese speakers allowed 'if'-infinitives *only* in the enclisis option, which is not the case: there seems to be no contrast at all between the two options, both allowing 'if'-infinitives with total naturalness. The alternative account of Kayne's data I will propose is, I think, less in trouble with Portuguese than Kayne's, even if this is at the cost of being less restrictive.

Let me point out a final problem. Kayne's attempt to reduce control (and arbitrary PRO would also be a form of control) to Binding Theory (BT) is appealing, for control is thus far one of the modules of grammar that has remained most obscure of all. However it is well-known that the mysterious nature of control is hard to reduce to the standard locality conditions in BT. So in (24), (40) there is an unexpected minimal contrast between PRO

and genitive (the latter being akin to PRO in not showing complementary distribution between pronominals and anaphors);

(24) a. **Each other's* accusation triggered many problems for *them*
b. *PRO* to accuse Nixon triggered many problems for *them*

The problematic cases are numerous and diverse. To mention another two, both Wh-movement and dislocation may bleed or feed BT possibilities while control remains immune to both. More importantly, there are cases of dislocation where control and BT seem to meet contradictory requirements:

(25) Votar *en Joan*, em sembla que ni *ell mateix*
To-vote-for the J. me-seems that not-even he SELF

s'imagina que t'ho proposessis
guesses that you-intended
'Vote for *Joan*, I think not even *he himself* could guess you intended to"

In (25), (41) BT requires *en Joan* not to be c-commanded by *ell mateix*, whereas the controlled PRO in the dislocated infinitival clause would have to be c-commanded by the subject of *t'ho proposessis* 'you intended to' if control is to be reduced to binding. In whatever level of representation one requirement is met, the other is not.

In addition, I have serious doubts that control can be reduced to purely structural conditions: at least for typical control verbs (*hope*, *convince*, etc.) there seem to be irreducible semantic or lexical factors (consider, e.g., *persuade* and

promise) on the determination of the controller.[11] A sad conclusion though it may be, control seems to remain a poorly understood field of the grammar. What can easily be determined on a structural basis is only the controllee, not the controller. Our proposal, as the standard ones, will capture this fact.

Summing up, Kayne's hypothesis appears problematic for at least the following reasons:

- it challenges the structure preserving hypothesis concerning adjunction: a head is adjoined to a X'.

- it predicts a contrast in degree of attachment in enclisis vs. proclisis which is the opposite of what independent morpho-phonological evidence suggests.

- it implicitly assumes that a head does not govern its specifier, a problematic assumption for standard accounts of Binding Domain as Governing Category, unless we generalize X°-adjunction-to-Y' to all the parallel cases.

- it crucially relies on the notion of CFC in a way that is not trivially adaptable to present-day assumptions on clause structure.

- it represents V+enclitic as not forming a constituent, thus predicting that V-cl cannot move.

- Portuguese is a potential counter-example.

- control cannot be easily reduced to BT.

---

[11] There is a recent proposal by Larson (1991) which tries to derive whether the object or the subject is the controller from purely structural (c-command) conditions. Although plausible for some basic cases, the proposal is bound to adopt rather prolix assumptions to cover only the English data. In addition, the minimal contrast he postulates between objects c-commanding and not c-commanding the infinitive should optimally be derived from some Universal Theta Assignment Hypothesis, for they are unlikely to be learned and, in addition, object/subject control status seems to be relatively uniform for synonymous verbs across languages.

Perhaps it is not advisable to simply dismantle a proposal with such a basic appeal and perspicuous insights as Kayne's, on the basis of technical problems: these problems might disappear as a deeper understanding of the facts develops. However, at least one of the problems, namely the pervasive consequences of taking the proposal of X°-adjunction-to-Y' seriously, which has dramatic effects on cases like (20) and (21), does not seem to be a side-problem and casts serious doubts on the first glance plausibility of the theory.

The alternative account we will present is, admittedly, less ambitious, for it gives up some of the promising achievements in Kayne's theory, such as reducing control to BT.

2. The AGR-identifier of Infinitives

In the preceding chapter, we adopted a parameter that determines which is the AGR-identifier for a (finite) clause, repeated here as (26):

(26) AGR°/Spec of AGR is the AGR-identifier of AGR.

We assumed that AGR° is the unmarked value, and that languages having poor AGR-morphology were forced to adopt the marked value (Spec of AGR).

Suppose (26) is a parameter which is set once for all types of clauses in a given language. The trigger for setting the parameter would be finite clauses, but the value chosen would relevant for infinitival ones in the unmarked case.[12]

That the AGR identifier must be uniform across sentence-types (finite/infinitival) is a necessary assumption for the theory above to work. The reason is that the [±anaphoric]

[12] Occitan and Sardinian will be claimed to be marked languages in this connection. See below.

character of I-subjects does not seem to vary from finite to infinitival clauses, as we have seen:

(27) a. Gli dispiace dover farlo *lui (stesso)*
Him-dislikes to-have-to do-it he (SELF)
'He dislikes to have to do it himself'
b. It bothers him to have to do it *himself*

As for R-expressions, they are not allowed as I-subjects in controlled infinitives in NSLs, but this will follow from the assumption that control involves a PRO which binds the I-subject, and R-expressions cannot be bound.

2.1. Control

To make clear what we want to arrive at, let us advance the following idea, which we will try to motivate as we proceed:

(28) In the unmarked case, controlled infinitives have the same AGR-identifier option (AGR°/Spec-AGR) as finite sentences, in a given language.

Let us see what (28) predicts for Italian and English (we assume they are 'the unmarked case'). Consider the D-structures in (29):

(29) a. John wants $[_{AGRP}$ AGR° to come DP]
b. Gianni vuole $[_{AGRP}$ AGR° venire DP]
G. wants to-come

In both cases, AGR is coindexed with the DP. Consider first the English case construction. In this case it is Spec of the infinitival AGR which is to be filled to be the AGR-identifier.

Let us assume that this specifier is occupied by PRO (whose precise status we will discuss below). Let us assume that, contrary to what is standardly claimed, PRO can (and has to) receive Case, in order to abide by the Visibility Condition or some equivalent requirement. Let us assume that it is AGR° which assigns Nominative to PRO by agreement.[13] Suppose that infinitival AGR° can assign a Nominative only to PRO (a kind of weak Nominative that is specific to PRO).

Given the above assumptions, PRO is the Case-position for the I-subject, and the whole IP (= AGRP) is the Binding Domain for the I-subject. The I-subject will then be anaphoric, as desired.

In the Italian construction, on the other hand, it would be AGR° itself which would be controlled. To obtain this result we could tentatively assume that:

- AGR° obtains person features by control, and becomes rich.

- so V raising to AGR is allowed (assuming, in the spirit of Pollock (1989) that only heads having some content allow V° to move to them).

- AGR° has to assign Nominative directly to the I-subject, by government, which requires its combining with T.

- so T has to raise to AGR, and it will if cyclic V-movement to AGR takes place.

- as in finite sentences, the Case position for the I-subject will be the I-subject itself, so that the Binding Domain

---

[13] We will crucially assume that heads do not governs their specifiers, so that, as we assumed for finite sentences, Nominative is assigned to Spec of AGR by agreement, not by government.

I think that the assumption that heads do not govern specifiers it is a tenable position in general. In addition, it subsumes much of the canonical-government proposals in earlier literature: non-canonical government is impossible in the general case because it would be government of a Specifier.

for the I-subject will exclude Spec of AGR, and the I-subject will be [-anaphoric].

This is a first characterization of the facts. Let's address more specific questions. One is that in NSLs like Catalan or Italian, the emphatic I-subject in controlled infinitives is pronominal, but it cannot be an R-expression. So to say that I-subjects are [-anaphoric] is not enough, since the [-anaphoric, -pronominal] option is excluded:

(30) a. Gianni vuole { venire lui/pro }
G. wants to-come he/pro
b. *{Lui} vuole [venire Gianni]
(He) wants to come G.

We could simply assume that the R-expression *Gianni* in (30).b) is excluded because it is bound by the controller in the upper clause (*lui*). Although this account is correct, it is not sufficient. As we pointed out above, there are cases where the controller seems not to c-command the infinitive at any level of representation. This is the case with dislocated controlled infinitives (see the discussion in section 1.3., where we conclude that certain controlled infinitives cannot be c-commanded by the controller at any level of representation).

We will assume that the impossibility of R-expressions as subjects of controlled infinitives is due to a violation of the 3rd principle of BT, but the offending binder will not be the controller itself, but PRO. Let us assume that PRO is always present in control structures (and in cases of PRO_{Arb} as well). We will later elaborate on the nature of PRO. Let's take it for granted that it is necessarily present in both NSL and non-NSL control structures.

How can this be accommodated within the above assumption that (in the unmarked case) NSLs have a controlled AGR°, while

non-null subject languages have a controlled Spec of AGR? At first glance, this looks at variance with the assumption that both types of languages involve a controlled PRO, which is in Spec of AGR. In order to solve this apparent contradiction, let us make the following assumptions:

- PRO is a necessary element of control (and PRO_{arb}) structures: only PRO can be controlled (perhaps PRO_{arb} is controlled by an empty Argument or operator).

- PRO cannot be governed for the familiar reasons (we adhere to the PRO theorem).

- principle (28) requires (or, being a markedness principle, favors) AGR° as the controllee in NSLs. If this option is to be fulfilled, the only means it can be is that PRO adjoins to AGR°. Let us assume that PRO is a maximal projection, but that only its head (call it PRO°) has real content, so that adjunction of this head to AGR° actually implies placing the controller inside AGR°.

Under this view, NSLs would have, in the unmarked Case PRO° adjoined to AGR°, thus fulfilling (28). Now consider the following infinitival structure:

(31) $[_{AGRP}\ [_{PRO}\ PRO^{\circ}\ ]\ [_{AGR'}\ AGR^{\circ}\ ]\ ]$

Suppose PRO is to adjoin to AGR°. We have assumed that a head does not govern its specifier. Therefore, movement of PRO° into AGR° would violate the ECP. Therefore, with such a structure, PRO° could never adjoin to AGR° in order to make AGR° to controllee element.

Suppose, however, that above AGRP, but below CP, there is an intermediate functional category (XP) (which we will try to characterize later) to which AGR° can move, giving:

(32) $[_{X^{\circ}}\ AGR^{\circ}_{i}\ ]\ [_{AGRP}\ [_{PRO}\ PRO^{\circ}\ ]\ [_{AGR'}\ t_{i}\ ]\ ]$

In this configuration, PRO° is allowed to move to the X° head containing AGR° without violating the ECP. Therefore, (32) is a structure that allows for the controllee to end up in a head position which at the same time contains AGR°. This would be the case for NSLs as Catalan or Italian.

Can (32) be somehow motivated? I will assume that enclisis, which is typical of NSLs, is a manifestation of the structure (32): whenever AGR° (which contains V°) raises to X°, enclisis is manifested. So we are in a position to account for enclisis and the empirical facts discussed in Kayne (1991). Let us see how.

In order to motivate our account for the enclisis facts, we will concentrate on two questions: How is 'longer' V-movement related to enclisis? and, What are the effects of adjoining PRO° to AGR° (or a head containing AGR°).

2.1.1. Enclisis

Let us first consider a well known case of enclisis not concerning Romance infinitives: French (complex) verb/subject-clitic inversion. An interesting and recent account for this phenomenon is Rizzi & Roberts (1989). The literature on this topic agrees on the point that V-INFL movement to COMP is a crucial factor for this construction. What concerns us here is the following fact: once the V-INFL raises above INFL into C°, something forces the clitic subject to cliticize to C°. I think this can follow from the clitic's need for a host, but there is no obvious reason why the clitic does not *left*-adjoin to C° instead of *right* adjoining. Whatever the account is, we could take it to have a rather pervasive nature. The situation, then, would be:

- there is a designated landing site where clitics *left*-adjoin: AGR° for both object clitics in Romance and subject clitics in French and Northern Italian dialects.

- whenever the content of this head moves a head up, the clitic is left without an appropriate host and it has to cliticize into the new host position: in this case (for whatever reason) the clitic right-adjoins to the upper host.

Now, if French (complex) inversion involves longer V-movement than in assertive clauses, and this longer movement is not a minimal further step as in Kayne's (1991) account for enclisis, but rather a one-head-up-more step, is there evidence that Romance enclisis also involves further head-to-head movement? I think there is some evidence, at least in Catalan and Spanish (I will be presenting the Catalan examples; Spanish equivalents behave identically as far as I know).

One way of measuring the length of the further-step movement in enclisis constructions is to consider word-order phenomena. If enclisis involves V-INFL moving one more head up w.r.t. proclisis, then this extra movement will skip over possible specifiers or adjuncts of the maximal projection in between:

(33) $[_{X^0}\ V\text{-}AGR_i\ ]\ [_{AGRP}\ adjunct\ [_{AGRP}\ Spec\ [_{AGR'}\ t_i\ ]\ ]\ ]$

so that the order will be adjunct-V or Spec-V in proclisis constructions and V-adjunct or V-Spec in enclisis constructions.

I think there is evidence of precisely this kind. In Catalan and Spanish, there are several kinds of adverbs (Catalan *sempre* 'always', *mai* 'never', *ja* 'already'/'yet', *encara* 'still', etc.) which usually precede the verb in finite sentences.[14] In fact, when they precede the verb, they must be adjacent to it:

[14] They can also follow the verb with a slightly lower degree of naturalness. This does not affect the argument below.

(34) Sempre/mai /ja /encara (*en Joan) hi va
Always/never/already/still (the J.) there-goes
'Joan always/never/already/still goes there'

So these elements are likely to occupy the specifier of the head where the verb stands in finite sentences.[15] In fact, the generalization is that these adverbs can precede the verb in precisely the constructions having proclisis: finite sentences (except affirmative imperatives, but including negative imperatives), and cannot in the constructions having proclisis: infinitives, gerunds and affirmative imperatives. In the latter case, they obligatorily follow the verb:

(35) a. infinitives: (*ja) anar-hi (ja) avui
(*already) go-there (alr.) today
'to already go there'
b. gerunds: (*ja) anant-hi (ja) avui
(*already) going-there (alr.) today
'already going there'

[15] Remember that in NSL Spec of AGR is not necessarily filled by the subject DP, so there is no problem if we assume this specifier is Spec of AGR. Ja 'already/yet' and encara 'still/yet' can precede the other adverbs in preverbal subject, giving marginal results:
(i) ?Ja sempre ve a classe
Already always comes to class
Possibly these two adverbs can be adjuncts to AGRP in addition to being specifiers.
In fact, we will assume that Spec of AGR is filled by PRO in control structures, but the facts in (35) are independent of control (they show up in non-controlled infinitives and imperatives).

c. aff. imperatives: (*ja) ves-hi (ja) avui!
(*already) go-there (alr.) today
'do go there right today'

d. neg. imperatives: (encara) no hi vagis (encara)!
(yet) not there-go (yet)
'Don't go there yet'

The acceptability judgements are clear cut. Especially significant is, I think, the case of imperatives: a minimal contrast such as affirmation/negation involves a change from enclisis to proclisis, and correlatively, a sharp change in the possibility for those adverbs to precede the verb. A natural and simple explanation is that, since enclisis involves an extra step in head-to-head movement, the adverbs in the specifier of the lower head shift in word order w.r.t. the verb once the verb moves a head up. This account is not possible within Kayne's hypothesis, where the extra movement is adjunction to next I', since this does not predict any word order change w.r.t. specifiers or adjuncts.

I will not develop in depth the question of what the upper head in this enclisis account is. In the imperatives it could be C°: as evidence for this we have the fact that affirmative imperatives strictly forbid COMP material preceding them, whereas negative imperatives do not:

(36) a. (**Que) ves-hi!
(That) go-there!
'Do go there!

b. (Que) no hi vagis!
(That) not there-go!
Don't you go there!

In infinitives and gerunds the upper head for extra-movement could be respectively a 'nominal' and 'adverbial' FC head. At least for infinitives, it cannot be COMP, because COMP can be filled by the 'if' particle as we have seen. Let us assume it is a FC intervening between CP and AGRP. We will call it NOM.[16] This head is probably not available in languages without enclisis. This can be a good explanation for the strongly nominal character of infinitives in languages like Italian, Catalan and Spanish, where it is used with a preceding definite article in some constructions (a. examples), and, besides can be readily used as a nominalization (b. examples):[17]

(37) Catalan:

a. A l'arribar
   At the-to-arrive
   'In arriving'

b. El desvetllar-se de la natura
   The to-awake of the nature
   'The wakening of the nature'

---

[16] Kayne (1991) suggests that there is an INFN ('infinitival') functional category which is specific to infinitives. I agree on this point in proposing NOM. What is odd about Kayne's proposal is that INFN is rather at the bottom of the clausal FC-hierarchy. If this extra FC is to be responsible for the categorial specificity of infinitives (their nominal behavior, for instance), it is more natural that it is the shell containing the other, sentence specific, FCs.

[17] Actually, even Sardinian, which has proclisis in infinitives, hence, in our terms, no movement of AGR° to NOM°, readily allows for infinitives as nominalizations. Perhaps Sardinian really has this nominal head but does not use it. After all, the nominal (or adverbial) character of English gerunds does not imply long verb movement. So the existence of an extra nominal FC would be a necessary but not sufficient condition for verb extra-movement (amounting to enclisis) in infinitives.

Spanish:

a. Al llegar
   At-the to-arrive
   'In arriving'

b. El cantar de los pajaritos
   The to-sing of the birdies
   'The singing of the birdies'

Italian:

a. L'aver (lui) affermato questo
   The-to-have (he) stated this
   '(His) having stated this'

a. Lo svegliarsi della natura
   The to-wake of-the nature
   'The wakening of the nature'

Raposo (1987-a) argues, that Romance Infinitives' nominal properties account for the apparent fact that they have to be Case-marked.

Even if, as I said, I cannot explain why enclisis and not proclisis takes place when there is extra head movement, at least we can generalize the phenomenon to the case French (and Northern Italian) (complex) inversion, both being cases of extra head-movement which force the proclitic of the lower head to become an enclitic of the upper head. Perhaps the reason of the change of directionality of cliticization is a matter of diachronic change, having to do with the 'easiest' way a new generation can reanalyse the parents data when reanalysis takes place: for instance Old French subject inversion in interrogatives had nothing to do with cliticization even when the subject was a pronominal; when cliticization started to be active, the easiest way for the new generation to accommodate facts was to assume right-adjunction of the subject clitic to C°. Perhaps similar accounts could be given for Romance clitics, assuming their

original position was less fixed (subject to the Tobler-Mussafia law -see Benacchio & Renzi (1987)). I will not elaborate on the matter.

Since the enclisis/proclisis facts are not predicted from a synchronic point of view, Portuguese is not necessarily a counter-example: since in this language clitics have had a development which is rather different from the one in other Romance languages (there is nothing similar, in the latter, to Portuguese clitic-order alternations in main clauses being sensitive to negation and preverbal quantification), it might simply be that extra-movement to NOM° does not necessarily correlate with enclisis because (maybe because of a different reanalysis process in the history of Portuguese). This conclusion is too loose to be of great interest. I think that the co-occurrence of proclisis and enclisis could be the key to a more accurate analysis of the facts. It might also be that enclisis is the core option.[18] I leave the issue here.

In the next section we will see how the above assumptions interact with the PRO-theorem.

Interrogative infinitives in Romance NSLs present a problem which also appears in finite embedded interrogatives. Consider the following finite paradigm from Catalan (which is representative of all Romance NSLs as far as I know):

---

[18] Joana Louro (p.c.) pointed out to me that enclisis in standard Portuguese requires some morpho-phonological alternations which are not used in the colloquial version of enclisis and that are rather annoying. So it might be that:
- enclisis is the unmarked option, but proclisis is also allowed as a marked variant.
- to speak Portuguese 'correctly', you have to learn cumbersome rules affecting enclisis.
- so speakers shift to proclisis to avoid both cumbersome school grammar rules and speaking 'incorrectly'.

(38) a. Crec (*en Joan) que (en Joan) el veuré
I-think (the J.) that (the J.) him-will-I-see
'I think that Joan I'll see'

b. No sé (en Joan) quan (*en Joan) el veuré
Not know-I (the J.) when (the J.) him-will-I-see
'Joan, I don't know when I'll see (him)'

c. No sé (en Joan) si (en Joan) el veuré
Not know-I (the J.) if (the J.) him-will-I-see
'Joan, I don't know if I'll see (him)'

(38).a) shows that dislocated elements and clause adjunct adverbs cannot precede the complementizer *que* 'that', but only follow it. So they would be IP adjuncts. (38).b) shows that these elements cannot follow a Wh-word (i.e., cannot intervene between the Wh-phrase and the verb), which could suggest that there is V-INFL to COMP movement (as in English main interrogatives), but the fact that now these elements can precede the Wh-phrase seems to suggest that it is the Wh-phase which is lower (in some IP-specifier), not the V-INFL that raises. To make thinks worse, (38).c) shows that 'if' interrogatives allow the adjunct to both precede or follow the 'if' particle. In Chapter 5 we will provide some way of explanation for the facts.

What is to be noticed now is that infinitive interrogatives display a rather similar paradigm:

(39) a. No sé (en Joan) quan (*en Joan) visitar-lo
Not know-I (the J.) when (the J.) to-visit-him
'Joan, I don't know when to visit (him)'

b. No sé (en Joan) si (??en Joan) visitar-lo demà
Not know-I (the J.) if (the J.) to-visit-him tomorrow
'Joan, I don't know if to visit (him) tomorrow'

To the extent these data cast doubts on the idea that 'if' is a C° particle, they are a potential problem. I think, however, that these facts do not seriously challenge our analysis: It may well be that the assumption that Wh-elements and interrogative 'if' in Romance NSLs are in the same functional category as the one hosting *que* 'that' is too simple, and there is an intermediate projection. So provided 'if' is in the head of the FC immediately dominating NOM° (or AGR°), our account is tenable, whether this FC is CP or not.

2.1.2. The PRO Theorem

Before going into the discussion of the PRO-theorem effects, let us say something about what a theory on PRO and control has to minimally specify. We have argued that control cannot be reduced to Binding Theory in a positive way (i.e. with a theory predicting the reference of PRO as a case of BT, as is intended in Kayne (1991)). We will adhere, however, to the more traditional view that PRO's distribution is determined by BT in a negative way: PRO has to escape binding requirements by being ungoverned (or not having Case, see below).

Since control is, under this view, still a mysterious module of the grammar, nothing of great interest can be said about it. What I want to suggest, however, is that control can be made minimally an interesting phenomenon if we relate it to another phenomenon which is apparently akin in nature: subjunctive obviation.

For both control and obviation there have been authors trying to derive either from Binding Theory (Kayne (1991) and Picallo (1985), resp.). Both phenomena, however, are reluctant to such accounts. To mention a major problem, dislocation or movement of the infinitive/subjunctive does not have any effect on the control/obviation facts, contrary to what happens with

genuine BT facts. Recall our discussion with examples such as (24), (40) and (25), (41). Similar examples can be built for obviation (consider (42)):

(40) a. **Each other's* accusation triggered many problems for *them*

b. *PRO* to accuse Nixon triggered many problems for *them*

(41) Votar *en Joan*, em sembla que ni *ell mateix*
To-vote-for the J. me-seems that not-even he SELF

s'imagina que t'ho proposessis
guesses that you-intended
'Vote for *Joan*, I think not even *he himself* could guess you intended to'

(42) *Que votis *en Joan*, em sembla que ni
That you-vote-SUBJ the J. me-seems that not-even

ell mateix s'imagina que t'ho proposis
he SELF guesses that you-intend
'Vote for *Joan*, I think not even *he himself* could guess you intend to'

In (42) the dislocated subjunctive *que votis en Joan* is subject to the obviation constraint w.r.t. the subject of the embedded clause (this is the reason for the ill-formedness of the structure), in spite of the fact that, were it not for obviation, the example would be fine and it would involve crucial non-c-commanding of *en Joan* by *ell mateix* 'he himself'.

I will leave the question here, only suggesting that control and obviation are closely related phenomena which could constitute an autonomous module involving some sort of obligatory coindexation/anti-coindexation without recourse to c-command.

Since we assume that control is outside Binding Theory, the traditional assumption that PRO is an element having to escape BT seems reasonable.

Now let us try to account for Kayne's (1991) generalization that enclisis is a necessary (perhaps sufficient) condition for having 'if'-infinitives. Let us begin with languages not having enclisis (English, French, Occitan, Sardinian). Let us simply assume that in those languages a controlled infinitive has the following essential structure:

(43) $[_{CP}\ C^{\circ}\ [_{AGRP}\ PRO\ [_{AGR'}\ AGR^{\circ}\ ]\ ]\ ]$

PRO is not governed by AGR°: in English, French and Occitan, simply because AGR° is not filled and therefore is not intrinsically able to govern. For any language, however, there is a major reason: AGR° is not structurally able to govern PRO, because, we assumed, heads do not govern specifiers. If, however, C° is filled with an intrinsically possible governor (such as 'if'), then PRO is governed and there is a conflict with the PRO-theorem. So we adhere to Kayne's (1991) initial idea that the impossibility of 'if-infinitives in some languages is due to an illicitly governed PRO.

In languages having enclisis, however, the structure is the following (where NOM is the nominal head of infinitives, as we proposed above):

(44) $[_{NOM^{\circ}}\ AGR^{\circ}_{i}\ ]\ [_{AGRP}\ [_{PRO}\ PRO^{\circ}\ ]\ [_{AGR'}\ t_{i}\ ]\ ]$

where AGR° has moved to the next head.

We assumed that in this structure PRO° is able to adjoin to NOM°. Therefore:

a) Since AGR° is merged in the same head as PRO, it becomes rich in (control-provided) features. Thus it is the AGR-identifier, and can assign Nominative directly to the I-subject, in accordance with the principles we postulated in Chapter 3:

(45) AGR° can optionally assign Nominative Case to Spec of AGR by agreement or to some other position under government.
(46) The I-subject must receive Case from its AGR-identifier.

b) Hence (headless) PRO in Spec of AGR will not have Case, and *its Binding domain will not be definable*, according to our definition, which makes reference to a Case position. Therefore it is immaterial whether it is governed or not as far as the PRO-theorem is concerned (in (44) it would be governed by $[_{NOM°}$ AGR°]). In the present account, then, PRO can escape BT either by not being governed or by not having a Case position.

c) PRO°, on the other hand, once adjoined to a head, is not in an A-position and is not subject to BT.[19]

d) (headless) PRO in Spec of AGR is coindexed with AGR° (by Spec-head agreement) and AGR° is coindexed with the I-subject (by the very same reason it is in finite sentences). Then PRO A-binds the I-subject, with the result that the latter cannot be an R-expression (in the following example, the infinitive is dislocated, so that the reason for the BT-3rd Principle violation cannot be binding by the main subject):

---

[19] If this idea is correct, then it could have far-reaching consequences for other cases of head-movement of a pronominal (thanks to Carme Picallo for pointing this out to me): if, for instance, clitics are analyzed as DP heads adjoining to some FC, then, by the same logic of our reasoning about PRO°, no Binding Principle would apply to clitics, contrary to fact. On the one hand, however, it is not clear that clitic attachment is the output of head movement from the A-position. On the other hand, the Binding Theoretical properties of clitics are an obscure area at least in the case of reflexives (reflexive clitics often become passivizer or unaccusativizer morphemes across languages).

(47) a. Per això ell/*en Joan, no ho desitja pas
To-do this he/ the J. not it-wishes at-all
'To do this himself/*John, he does not wish'

b. [Per [$_{IP}$ t_{PRO} t]$_i$ [$_{VP}$ [$_{VP}$ això] ell_i/*$en\ Joan_i$]]

The Case position for the I-subject is the I-subject itself, thus the Binding Domain for the I-subject excludes PRO, and the I-subject can be (and has to be) a pronominal.

An important question arises here. Since we crucially derive enclisis with PRO°-adjunction, and PRO°-adjunction ultimately forces the infinitive I-subject to be pronominal, languages not having enclisis in infinitives will not behave like NSLs in infinitival constructions, even if they are NSLs in finite sentences. Specifically, we would expect those languages to have pronominal I-subjects in finite clauses and anaphoric I-subjects in (controlled) infinitives.[20] This would be the case for Occitan and Sardinian. Is this prediction borne out?

Quite disappointingly, both Sardinian and Occitan are languages using pronominals *qua* (strong) reflexives in an exclusive way:[21]

---

[20] Recall we assume that the parallelism principle (28) holds only 'in the unmarked Case'. Occitan and Sardinian would thus be marked in this connection.

[21] The Sardinian example comes from Jones (1990), where he explicitly states that there is no strong pronoun/strong anaphor lexical contrast. The same is true for Occitan as far as I know: even if literary Occitan has tried to retrieve the Medieval strong reflexive form (se (mateis)), it is never used in colloquial speech.

(48) Occitan:

Jacme$_i$ parla d'el$_{i/j}$ (mateis)

J. talks of-he (SELF)

'Jacme talks about himself/him (himself)'

Sardinian:

Gavini$_i$ l'at comporatu pro isse$_{i/j}$ (matessi)

G. it-has bought for him (SELF)

'Gavini bought it for himself/him (himself)'

So we cannot obtain confirmation for our prediction that these languages have a contrasting AGR-identifier strategy for finite/infinitival sentences, although, happily enough for our hypothesis, our prediction is not falsified either.

In fact, I think it is not an accident that languages not distinguishing strong pronouns/anaphors are the ones both being NSLs and having proclisis in infinitives: in the absence of evidence for the [±anaphoric] character of infinitival I-subjects, these languages have more unproblematically adopted proclisis, because it is more unlikely to conflict with the I-subject BT data of a possible earlier period where the parallelism finite/infinitival (i.e., the unmarked case) held. Or, the other way around, the fact that they are NSLs with proclisis has favored the adoption of neutral pronoun/anaphor elements in order to avoid pronominal/anaphor switch in finite/infinitive clauses: if the learner does not observe the switch in the parents' data, it would be easier for him/her to reinterpret pronominals as neutral pronominal/anaphoric elements than to adopt the switch practice. The idea would be that the learner undergoing diachronic change tries to build grammars that are the least conflicting possible with the parents data.

This idea extends to old French, which also was a NSL with proclisis: in fact, even modern French has neutral pronominal/anaphor forms, surely a residue of its NSL period.

Assuming that the existence of neutral anaphoric/pronominal elements is a marked option, the present theory provides a possible trigger for its existence in these languages: optimal data reinterpretation in diachronic change.

Summing up, we have provided an account for infinitival constructions that consists of the following assumptions:

- in the unmarked case, infinitives have the same AGR-identifier option as finite clauses for a given language.

- given that PRO is always the key element in control, and it is a maximal projection, it has only two options: either it is ungoverned (as required by its extraordinary status), or its head incorporates to the head containing AGR.

- the latter option requires the V-AGR head to move to an upper head in order to make the landing site for PRO°-movement a position governing the trace; and it allows AGR° to become the AGR-identifier, a favored option in NSLs.

- long V-movement to AGR° (and possibly NOM°) is only allowed as far as AGR° ends up being contentful; it is in addition required if the AGR-identifier is AGR° and has to Chain-govern the I-subject to assign it Case.

2.2. Raising

In Chapter 3, we crucially assumed that raising constructions are characterized as involving a non-CP infinitive:[22] this is why AGR in the upper clause is coindexed with an I-subject internal to the infinitive, this I-subject being the first DP/CP it c-commands. Suppose, nevertheless, that

[22] Recall we suggested that only verbs with an epistemic meaning are likely candidates to be raising verbs, at this could be the basis for accounting for their exceptional non-CP character or their complements: they would form a (semantically) mono-clausal structure with their complement.

the infinitival AGR has to abide by one of the options concerning the AGR-identifier: either it is AGR° or Spec of AGR. We provisionally assume the following parallelism principle:[21]

(49) In the unmarked case, raising infinitives have the same AGR-identifier option as finite sentences in a given language.

Consider the D-structures in (50), taking English and Italian as representative languages:

(50) a. AGR° seem $[_{AGRP}$ AGR° to have come DP]
b. AGR° sembra $[_{AGRP}$ AGR° esser venuto DP]

In both cases, both the main clause AGR and the embedded AGR (if there is one) are coindexed with the DP in the embedded clause, since it is the first DP or CP in the c-command domain of both. This means that the two AGR's end up coindexed.

For English and non-NSLs, both Spec's of AGR have to be filled by a DP to render both specifiers AGR-identifiers. This is the unmarked option in (49) and, in fact, the only one for the infinitival AGR°, given that the it is not rich to be an AGR-identifier itself.

The infinitival AGR° cannot assign Case to its specifier: we assumed that infinitival AGR° can only assign Case to PRO, and PRO cannot occur in this position because it is governed by the main verb, as standardly assumed. Since we are trying to derive Burzio's Generalization, we assume that 'to seem' verbs can in principle assign Accusative. So, in principle the Spec of the

---

[21] This principle has an obvious parallelism with (28) (which concerns control, see page 213). We will reduce them to a single principle. For the moment we distinguish them for convenience.

infinitival AGR could be assigned Accusative. However, since principles (45) and (46) force the main AGR-identifier (Spec of AGR) to transmit its Case to the I-subject, the result would be that the I-subject would end up obtaining both Accusative Case from the verb 'seem' and (Nominative) Case from the main clause Spec of AGR.

Therefore, the only option for the infinitival Spec of AGR is to be a non-Case-marked empty category and form a Chain with the Specifier of the main clause AGR, from which it will obtain Case.

Given the definition of Binding Domain, the main clause will become the Binding Domain for the I-subject in the infinitive, so that both the I-subject in the infinitive and the infinitival Spec of AGR can only be [+anaphoric]:

(51) a. John seems $[_{AGRP}$ to have done it $[_{DP}$ e (himself) $]]$

As for the Spec of AGR, its anaphoric behavior cannot be instantiated by the presence of an overt anaphor or floating *himself*:

(52) John seems $[_{AGRP}$ e/*himself to have done it]

This fact should be due to some restrictions on the distribution of emphatic elements. I'm not able to say anything interesting about this issue, except that emphatic subjects seem to be restricted to some positions.

Now consider Italian. In this case, it is AGR° which has to become an AGR-identifier, if the unmarked option is taken. For the infinitival AGR° to become the AGR-identifier, it has to be rich. As before, in raising constructions the infinitival AGR and the main AGR end up coindexed with the same I-subject and,

therefore, coindexed themselves. Suppose that, since the main AGR is rich, the infinitival AGR° may inherit phi-features from it and become rich itself.[24] If so, it will be the AGR-identifier and will have to assign, by the combination of principles (45) and (46), Case to its I-subject by Chain-government, provided the infinitival verb raises.

Since our Parallelism Principle (46) induces, for NSLs like Italian or Catalan, the main AGR° to assign Case to its I-subject, we have to assume that, in raising constructions in these NSLs, the main and the embedded AGR are reanalyzed as an 'extended' Government-Chain, consisting of the union of the two head Chains formed by V-movement up to AGR in both the main and the infinitival clause. The foot of one Chain (V°'s trace of 'seem') governs the head of the other, and in addition the two AGR°s are coindexed. So Chain extension is a very unnatural device in this case.

In fact, Torrego (1989) argues that in Spanish the infinitival verbal form incorporates into the main verbal form in raising constructions.[25] If so, no device of Chain extension

---

[24] We crucially assume that the main AGR is rich and its coindexation with the embedded AGR makes the latter rich. When the richness of the main AGR° is only obtained after PRO° incorporation (see section 2.1.2. above), we might wonder whether, since incorporation is an S-structure phenomenon, it comes 'too late'. In fact, 'seem' verbs have a degraded status in infinitival constructions in Catalan (and possibly other NSLs):

(i) ??Semblar estar borratxo no t'ajudarà
To-seem to-be drunk not will-help-you

(ii) ??Vull semblar estar borratxo
I-want to-seem to-be drunk

I cannot pursue this idea here.

Coindexation between main and embedded AGR takes place in subject controlled constructions as well, so that in this case we predict that both PRO° incorporation and AGR coindexation converge to the same effect.

[25] She claims that this incorporation takes place only at LF, on the basis of examples where some adverbs intervene between the two verbs:

(i) Parecía ayer haber muchos mosquitos

would have to be stipulated. Or, reversing the argument, in NSLs incorporation would be triggered by the need for Chain-extension.

Given our BT definitions, the Binding Domain for the I-subject will be the infinitival TP, and thus the I-subject will be [-anaphoric]:

(53) a. Sembra [averlo fatto *pro*/lui/Gianni]
Seems to-have-it done *pro*/he /G.
'He/Gianni seems to have done it'
b. Chi$_i$ sembra [averlo fatto t_i]
Who seems to-have-it done
'Who seems to have done it?'

In the above examples, one could argue, it is not clear that the inverted subject is inside the embedded clause: it could be as well right adjoined to the main VP. There is, however clear evidence that this is not necessarily so (see (54)) and even not possibly so (see (55)) (examples from Spanish):

(54) a. El libro, parece [$_{IP}$ haberlo comprado Juan en Londres]
The book seems to-have-it bought J. in London
'That book seems to have been bought by John in London'
b. ??El libro, parece haberlo comprado en Londres Juan
The book seems to-have-it bought in London J.

Seemed yesterday there-to-be many mosquitoes
'Yesterday there seemed to be many mosquitoes'

To my ear (for Catalan and my Spanish), the presence of an intervening adverb sharply degrades the acceptability of the sentence. so incorporation would be an S-structure phenomenon in Catalan.

(55) a. Parece [$_{IP}$ haberlo escrito Juan] por la letra[25]
Seems to-have-it written J. by the handwriting
'It seems to have been written by John to judge by the handwriting'

b. *Parece haberlo escrito por la letra Juan
Seems to-have-it written by the hndwrng. J.

What (55) shows, is a particular instance of a more general fact: constituents of the embedded clause cannot be extraposed to the main clause:

(56) a. Ha dit [$_{CP}$ que donaria el llibre als nois] avui
Has said that he'd-give the book to-the boys today
'She said that s/he would give the book to the boys today (S/he said it today)

b. **Ha dit [$_{CP}$ que donaria el llibre] avui als nois
Has said that he'd-give the book today to-the boys

(55).b) and (56).b) do not improve at all if the extraposed element (subject/dative) is heavy, (provided *avui* 'today' is actually interpreted as specifying the 'saying' time interval).

It is a trivial matter that the parallelism principles in (28) (for control infinitives) and (49) (for raising) can be reduced to a single principle:

(57) In the unmarked case, infinitives have the same AGR-identifier option as finite sentences in a given language.

---

[25] The constituent por la letra should be read as non-dislocated, the interpretation being then, roughly: 'It is the handwriting that persuades (me) that it has been written by Joan'

We assume that, at least for control, this is only an unmarked case option for there would be languages (Occitan and Sardinian) taking the marked option. Now, if we have a single principle, the prediction would be that if Occitan and Sardinian take the marked option in control, they take the marked option in raising infinitives too, namely, the AGR-identifier would be the Spec of AGR in raising infinitives and AGR° in finite clauses. We argued that Spec of AGR in a raising infinitive cannot be PRO because it would be governed by the 'seem' verb. What other options are there left for Occitan or Sardinian?

If it were an empty anaphor, as is the case in English, it could in principle, like in English, form a Chain with the main clause Spec of AGR. But the main Spec of AGR does not receive Nominative Case, because in finite clauses AGR° assigns case only through Chain-government, as we argued in Chapter 3. Therefore an empty anaphor is excluded. (Remember preverbal subjects in NSLs enter a resumptive pronoun strategy with the I-subject, and elements resumed by an expletive are not assigned Case themselves).

Since the Infinitival AGR° is rich by being coindexed with a main AGR°, it could conceivably assign Case to is Specifier by agreement. But since the main AGR-identifier (AGR°) has to assign case by forming a Government-Chain, there would be conflicting requirements on the way Case is assigned: a Government-Chain can only assign Case by government, and its foot cannot be an agreement-Case-assigner. If this (perhaps dubious) assumption is on the right track, it would leave Occitan and Sardinian with no option for the AGR-identifier (Spec of AGR) of raising infinitives. In fact, Sardinian does not have infinitival raising.[27]

---

[27] Jones' (1990) description of Sardinian, which is rather cautious in excluding non-genuine constructions (Italianisms) is categorical in this connection.

As for Occitan, it does apparently allow raising (example from Sauzet (1989)):

(58) Lo paire semblava la me voler donar
The father seemed it-me-to-want to-give
'The father seemed to want to give it to me'

However Sauzet presents evidence showing that Occitan has what he dubs 'pseudo-raising': cases where there is apparent raising out of a finite clause:

(59) Lo paire semblava que la me voliá donar
The father seemed that it-me-wanted-he to-give
'The father seemed to want to give it to me'

I will not go into the details of his analysis: what is essential is that pseudo-raising is an exceptional construction (at least within the Principles and parameters framework): Sauzet assumes that in (59) what 'raises' is a topic element. Therefore, we could make the following argument:

- Occitan does not allow raising constructions for the reason we mentioned above.

- for whatever reason (perhaps strong interference with French) it has acquired a construction (namely (58)) which looks like raising. But it is not subject raising, but rather 'topic raising' (pseudo-raising).

- once the speaker accepts pseudo-raising with infinitives (58)), the construction generalizes to finite sentences, which could not be predicted by a theory assuming that (58) is simply a standard case of raising.

So our claim that Occitan or Sardinian cannot have (standard) raising makes some sense: it predicts that either the construction is absent (Sardinian) or that, if apparently

present, it is not standard raising (and there is independent motivation for pseudo-raising with finite clauses). Since pseudo-raising is likely to be a marked option, we could exploit the idea that the trigger is a historical one: French pressure (which is very strong in Occitan areas) has forced a raising-looking construction into the language and speakers have interpreted it as pseudo-raising and then generalized it to cases where French does not allow raising (namely (59)).

In fact, 'pseudo-raising' is a widespread phenomenon (a fact that could cast serious doubts on the accuracy and cross-linguistic significance of our theories for 'standard' raising). Basque is another language which has some sort of pseudo-raising:

(60) a. Haiek dirudite hauteskundeak galdu dituztela
They seem election-the lost have-that
Lit.: 'They seem that (they) have lost the election'
b. Bush-ek dirudi Perot-ek hauteskundeak irabaziko ditu
Bush seems Perot election-the to-win has
Lit.: 'Bush seems that Perot will win the election'

(60).b) clearly shows we are not dealing with raising, since the main clause subject does not even bind an Argument inside the embedded clause. In fact, our characterization (in Chapter 3) of Basque as having a designated position for AGR-coindexing would make 'true' raising impossible in this language, for it is essential to raising that the main AGR may be able to coindex with whatever DP is the most prominent one in its c-command domain.

In Chapter 3 we assumed that our principle of AGR-coindexation could be too powerful in that they would exclude super-raising without any need for the ECP (see the section 1.4 in Chapter 3). Then we considered the possibility of reducing the

power of our principle of AGR-coindexation. One way or another, super-raising is not a problem in the present theory.

Our account for raising in NSLs having proclisis does not involve the formation of an A-Chain, so there is no trace in the Spec of the infinitival AGR. Rizzi (1982-a) (cited in Burzio (1986:206) adduces evidence in favor of the existence of such a trace. Since a trace has to be (properly) governed, the infinitive cannot be moved, dislocated or, in general, appear outside the governing domain of the 'seem' verb, for then the 'seem' verb would not govern the trace (I adapt the Italian data to Catalan):

(61) *És [*t* estar cansat], que en Joan sembla.
Is to-be tired that the J. seems
'*It's to be tired, that Joan seems'

Cf. with:

És [*PRO* estar cansat], que en Joan tem
Is to-be tired that the J. fears
'It's being tired that Joan fears'

In our account, if there is no trace, what is the cause for the ungrammaticality of (61)? If our suggestion is correct that an extended Government-Chain is formed by the union of the V-movement Chains of the main and infinitival clauses, then it is reasonable that the foot of one (sub-)Chain has to govern the head of the other. We even suggested that, in Spanish or Catalan, the lower verb incorporates into the main verb: that would be an even stronger reason for the ill-formedness of this construction.

And even we could suggest that an AGR-identifier has to c-command its I-subject at all levels. Notice that in control structures the main clause AGR is not the AGR-identifier of the embedded I-subject, the control relation being of another nature.

2.3. ECM and 'for' Infinitives

In these constructions, the infinitival Spec of AGR receives Case from the head governing it, according to the standard analysis:

(62) a. I ... [$_{V^0}$ *believe*] [$_{IP}$ him [$_{I'}$ to be there]]
b. [$_{C^0}$ *for*] [$_{IP}$ him [$_{I'}$ to be there]] ...

ECM constructions are exactly like raising constructions except that the main AGR is not coindexed with an I-subject inside the infinitive, because there is a preferred candidate in the main clause (e.g., the Experiencer Argument of *believe*). In both, the infinitival AGR is poor and cannot become rich by any means. Therefore, only Spec of AGR can possibly be the AGR-identifier. Since the infinitival AGR° is not able to Case mark (it can in principle only Case-mark PRO, which is excluded by its being governed by the main verb or 'for'), Spec of AGR has to receive Case from that upper governor.

The present theory predicts that ECM and 'for' infinitives should not be possible for NSLs taking the unmarked option by which infinitives have the same AGR identifier as finite sentences: the infinitival AGR° would have to be the AGR-identifier, which it could not, being irreparably poor. Case-assignment to the Spec of the infinitival AGR would be useless, this position not being the AGR-identifier.

The claim that NSLs taking the unmarked option (for infinitival AGR-identifiers) have no ECM might have a counterexample in classical Latin.[28] In this language, however, infinitives cum Accusative are not clear cases of ECM, as far as

[28] If it is the case that classical Latin was a language taking the unmarked option for infinitives, of which I have no evidence.

the presence of the Accusative appears to be quite independent of the main verb's lexical characteristics and Case properties. So we have (examples from Maraldi (1983)):

(63) a. Dicitur eos venisse
Said-is them to-have-come
b. Manifestum est eum abisse
Evident is him to-have-gone

In standard accounts, we do not expect a passive verb or a copula to assign Accusative. In the present theory any verb can assign Accusative, but:

- either the infinitival is a CP (and therefore becomes the I-subject) and the main verb cannot govern the infinitival subject in Spec of IP,

- or the infinitival is an IP and the infinitival subject becomes the main clause I-subject, and then it will be Nominative.

In whatever theory, it seems that Latin infinitives cum Accusative are not ECM, but some other construction. I have nothing to say about the issue.

Raposo (1987-b) proposes that some verbs take an IP complement in Portuguese, but, if this is correct, the resulting structure does not allow ECM (it would be one of the types of inflected infinitive): in our terms, because, as we argued, assignment of Case to Spec of AGR is pointless in such a language. Thus our conclusion that NSLs taking the unmarked option for infinitives cannot have ECM is not challenged as far as I know.

Since, we assumed, Occitan and Sardinian take the marked option of having infinitival Spec of AGR as the AGR-identifier, we predict ECM is possible in principle for these languages,

which is not apparently the Case[29]. In fact, nothing forces a language to have ECM (or raising): they are anyway marked options. I do not know what might trigger (if anything other than positive evidence during acquisition) the presence of ECM or raising in a language.

French is another language not having ECM in general, except for the fact that the Wh-extracted version of an ECM construction is allowed (see Kayne (1981), Rizzi (1982-b)):

(64) Le garçon que je croyais être intelligent
The boy that I believed to-be intelligent

Even Italian has this construction:

(65) Il ragazzo che ritenevo essere intelligente
The boy that I-believed to-be intelligent

Verbs allowing this construction are epistemic verbs and verbs of saying (see Rizzi (1982-b:78). Since these verbs are, unlike English *believe*, control verbs in the general case, these exceptional constructions seem to require some analysis expressing their exceptional status, which does not necessarily

---

[29] Sardinian allows infinitival complements without control, as we will see in section 2.4., but the overt subject is Nominative and has the distribution of an inverted subject
As for Occitan, it does have apparent cases of ECM (see Sauzet (1989)):
(i) Pensava las vacas manjar son sadol
He-believed the cows to-eat their fill
Sauzet argues that these cases are to be analyzed as involving a PRO controlled by a topic adjoined to CP, since, unlike in English, the absence the DP las vacas gives control by the subject 'he', and, on the other hand, the DP (las vacas) can be adjoined to a finite CP:
(ii) Pensava las vacas que manjavan son sadol
He-believed the cows that they-ate their fill
in which case the topic binds a pro. In any event, the infinitival (apparent) subject is not Accusative.

threaten our claim that Italian cannot basically have ECM. What is essential here is that Spec of IP is not Case-marked,[30] for it is not an AGR-identifier and, therefore, not a Case-position. The proposal in Kayne (1981) and Rizzi (1982-b) that in (64) and (65) the main verb Case-marks the trace of the Wh-phrase in COMP does not contradict this claim.

French, even if allowed to have ECM, does not happen to. I leave the question here.

2.4. Infinitives with an Overt Subject

What I will call infinitives with overt subjects (IOSs) should be clarified: we have seen that both control and rasing infinitives do have overt I-subjects of a [±anaphoric] restricted nature (depending on the (non-)NSL status) and, for NSLs, of a [±pronominal] nature (depending on the whether there is control or raising). I will use the term IOS to name only those infinitives that allow an overt subject in a way not predictable from control or raising. Romance languages show a variety of them:

- AUX-to-COMP infinitives in Italian and Portuguese:[31]

(66) Ritengo non esser lui in grado di farti niente
I-believe not to-be he in a position to do-you nothing
'I believe him not to be able to do anything to you'

---

[30] It is apparently Case-marked in AUX-to-COMP infinitives, but this is an exceptional construction (see section 2.4.). As Rizzi (1982-b) points out, even if there is a correlation between the verbs allowing (65) and the ones allowing AUX-to-COMP, the latter construction is significantly more marked (high literary speech-level). In section 2.4.2. we will challenge the view that AUX-to-COMP involves Case-marking of the Specifier of AGR.

[31] See Rizzi (1982-b)/(1986) and Raposo (1987-b).

(67) Penso terem os deputados trabalhado um pouquinho
I-think to-have-3pl the deputies worked a little-bit
'I think the MPs finally worked just a little bit'

- Non-subcategorized prepositional infinitives in Catalan and Spanish (examples from Catalan):

(68) a. Aquesta habitació és per jugar-hi els nens
This room is for to-play-there the children
'This room is for the children to play in it'
b. A l'arribar la Maria, vaig fer-li el dinar
At the to-arrive the M. I cooked-her the meal
'Upon Maria's arrival, I cooked a meal for her'

- Dislocated infinitives (in colloquial Catalan and Spanish):

(69) Anar-hi en Joan, no em sembla pas la solució
To-go-there the J. not me-seems at-all the solution
'John going there, I don't think it's the solution'

Interestingly, Occitan has more or less the same possibilities, but with preverbal subjects:

(70) a. Aquesta cambra es per los dròlles jugar
This room is for the children to play
b. En Joan arribar/arribant, farem lo trabalh
In J. to-come/coming we'll-do the job
(71) a. Joan far tot lo trabalh, m'estonarià.
J. to-do all the work me-would-surprise

- Non-controlled complement infinitives in Sardinian:[32]

(72) Non keljo a vénnere tue
Not I-want to to-come you
'I don't want you to come'

The Sardinian case is specially interesting (and puzzling): what makes the infinitive in (72) a non-controlled infinitive is the presence of the preposition a 'to' before it. Without this preposition, the infinitive is obligatorily controlled (see (73).a)). When the non-controlled *a*-infinitive has no overt subject, its subject is interpreted as arbitrary ((73).b)):

(73) a. Non keljo vénnere
Not I-want to-come
'I don't want to come'
b. Non keljo a vénnere
Not I-want to to-come
'I don't want someone/anyone to come'

Several proposes, which I will not discuss here, have been provided to account for IOS. Rizzi (1982-b) deals with AUX-to-COMP. He proposes that an Aux in non-finite COMP acquires some capacity to assign Nominative. Raposo (1987-b) makes a much more restrictive claim (AGR has to be Case-marked itself to be a Case assigner). Although he makes several interesting predictions (he restricts AUX-to-COMP to NSLs, and predicts which contexts allow it), some of his crucial proposals on the precise nature of c-selection by verbs are far from obvious, and, in addition, his proposals on Portuguese are not easily extendable to AUX-to-COMP in Italian and even less, I think, to IOS in Romance. For IOS in

[32] See Jones (1990).

Spanish, Fernández Lagunilla (1987) simply states that Nominative is assigned by default, an attractive idea only if we can provide means of excluding the default option for the cases when it is not possible, which she does not even attempt to do. Rigau (1992) accounts for the postverbal position of subjects in Catalan and Spanish prepositional temporal IOSs (see (68).b)) not in terms of Case theory but in terms of tense interpretation (a weak T° has to incorporate to the temporal preposition preceding these infinitives). Independently of this, Nominative is assigned -and *pro* is licenced- by an abstract AGR. Abstracting away from *pro*, this idea is much in the spirit of Reuland's (1983) proposal for English gerunds with an overt subject. Galves (1991) revises Raposo's (1987-b) proposal in an interesting way, but it is still a theory basically conceived for European Portuguese.

Some authors assume it is AGR which assigns Nominative in IOS (Raposo (1987-b), Galves (1991), Rigau (1992)); others (Fernández-Lagunilla (1987), Hernanz (1992), Delfitto (1990), Belletti (1991)) assume AGR does not take part in Nominative assignment in some non-finite sentences. Except for the case of Portuguese inflected infinitives, in which AGR is obviously present, I will propose that AGR is not present in most types of IOS, and I will contend that this is not at variance with what happens in finite clauses.

The proposals we will introduce here are highly speculative and far from precisely established. They will however lead to important qualifications to the theory sketched so far. Whether our speculations are on the right track or not, we cannot, I think, ignore the issue or treat it in a independent way from what we have assumed so far.

We will classify IOS's in two groups, assuming the distinction is of theoretical relevance: IOS's where, we will argue, the overt subject is an I-subject (the cases in Catalan,

Spanish, Sardinian) and IOS's where the overt subject is in Spec of AGRP (AUX-to-COMP, Occitan).

2.4.1. IOS with an Inverted Subject

Let us assume that the IOS's where the subject follows the verb without necessary adjacency V-subject are cases of subject inversion, in our terms, cases involving an overt [-anaphoric] I-subject. In Catalan, where inverted subjects are VP-final (with the qualifications we introduced in Chapter 1), IOS's preferentially show VP-final subjects:

(74) Fer (???en Joan) la feina (en Joan), em sembla...
To-do the J. the work the J. me-seems
'That Joan does the work, well, it seems to me...'

In Spanish, instead, where the VSO word-order is a possible form of subject inversion in finite clauses, VSO word order in IOS is quito usual.[33]

Let us assume that the cases we have seen in (68), (69), (72) and (73).b) are cases of IOS with an inverted subject (*IOS-INV*). (73).b) shows another phenomenon: Sardinian IOS-INV allow a null subject with arbitrary interpretation. Catalan and Spanish IOS-INVs show an even more surprising fact: the subject can be null and fully referential (see Rigau (1992)):

---

[33] As we will see in the next chapter infinitives with a VSO word order are predicted to be possible in principle. In fact, this example is not as unacceptable as a parallel case with a finite clause. Still, the prediction is that Spanish allows VSO in IOS with no problem, while Catalan marginally allows it for other reasons. This is in accordance with the spirit of the Parallelism Principle (in whatever version, see below) we are proposing.

(75) En acostar-m'hi, em va mirar provocativament
In approach-me-there me-looked-at provocatively
'As I approach him/her, s/he looked at me provocatively'

The null subject in is not controlled and is fully referential. We will assume, with Rigau (1992), that this null subject is a *pro*. We will also assume that the arbitrary null subject in Sardinian IOS-INV in (73).b) is a *pro*.

There are at least two questions our theory should address:

(76) a. How is Nominative assigned to the I-subject in IOS-INV?
b. What makes referential *pro* possible in IOS-INV?

Let us start with (76).a). In the present theory, the most natural assumption is that Nominative is assigned by government directly to the I-subject, in a way similar to Nominative assignment in finite clauses. This is a desirable prospect if we want to account for the fact that IOS's with an inverted subject are, as far as I know, restricted to NSLs, which have subject inversion in finite clauses.

In Chapter 3 we contended that direct Case assignment to the I-subject is a consequence of AGR° being rich and becoming thus the AGR-identifier, plus the proposal that it is the AGR-identifier which has to provide the I-subject with Case: if the AGR-identifier is AGR° it has to Case-mark the I-subject by head-governing it. In this Chapter we have contended that infinitival AGR° can end up being the AGR-identifier if some configuration allows it to be rich: either PRO° incorporation (in control structures) or, in raising structures, coindexation with the upper AGR (possibly involving incorporation). In IOS's, however, there is no apparent non-*ad-hoc* device by which the infinitival AGR° could end up being rich.

Let us first speculate on some theoretical basis for accounting for the facts which is consistent with the present theory. Our account for finite AGR in Chapter 3 is based on the assumptions that AGR is present and has to be licenced (by having an I-subject, providing Case to it and being rich in features). In this chapter we have developed the idea that AGR is also present in infinitives and, therefore, there must be some means to make the AGR satisfy the same requirements, provided everything follows from control, raising or ECM legitimate structures.

Suppose, however, we assume that, since infinitives are morphologically silent in AGR content, AGR is only optionally present. Then no infinitive has to have an AGR in principle. What determines the presence of AGR when it is morphologically silent? Let us explore some possibilities.

In the case of lexically determined control, infinitives would be forced to have AGR because control would be a universally available grammatical option, which 'tries to apply' whenever possible. This idea could be related to the Elsewhere Condition, which essentially states that regular processes apply automatically unless they are blocked by the existence of more specific/irregular processes (e.g., in morphology, a regular verbal form is used unless an irregular form exists).

Suppose control is a universal option, in fact the most 'regular' one possible for infinitives, and applies whenever possible, unless a more specific/idiosyncratic option exists (for instance, an idiosyncratic subcategorization specification for 'believe' ([__IP]) in English). Then the existence of AGR in lexical control structures is forced by the obligatory application of control: control involves PRO; PRO can only appear

in Spec of AGR;[14] an element in Spec of AGR has only two options to be licenced, in accordance with our previous assumptions: either it becomes an AGR-identifier or it is resumed by an independently Case-marked I-subject (see next chapter). Putting aside the second option, the conjunction of the above assumptions gives the result that AGR has to be present in control structures.[15]

In the case of raising and ECM, since they are marked options, no problem of necessity (but rather one of mere availability) arises: these structures are allowed, not required, by UG, and will only exist if some lexical idiosyncratic (pattern of) specification(s) is learned (I think ECM and raising pose no problem for being learned on positive evidence). Since raising or ECM always involve infinitival complements (not, for instance, adjuncts), if raising or ECM configurations are not used in a language, then lexical control takes over (e.g., 'believe' is a

---

[14] It is far from obvious that theories being as rich in structural positions as the Split-INFL hypothesis, or Larson's (1988) VP-shell theory, can manage to confine PRO to Spec of AGR. Two solutions come to mind: a) PRO (or PRO°) is inherently an AGR-identifier; b) PRO can appear in other positions, but then it is not available for control.

[15] It is not obvious that the Elsewhere Condition, which has been used in phonology and morphology, should be relevant for syntax. I think it is at variance with Chomsky's (1988) implementation of the idea of economy of derivation, in that he explicitly states that regular/universal options take preference over irregular/idiosyncratic options. In fact Chomsky's proposal is intended to deal with subtle, theory internal problems, while the EC could be argued to be relevant for syntax in more obvious cases (e.g., if a language has object agreement morphology, it has to use it, and give up what is likely to be the more universal option of not using it). In addition, the idea of economy need not be tied to Chomsky's universal-over-particular constraint. For instance, Roberts (1991-a) uses a notion of economy based on the length of derivations.

Since, however, Chomsky is concerned with the economy of derivations in a very subtle sense, it could turn out not to be at variance with the Elsewhere Condition if some subtle distinction could be made between the field of application of the EC and economy. I leave the question open.

control verb in languages not having ECM[36]) or the construction simply does not exist.

So the existence of AGR in complement infinitives is always forced to be present by control or simply allowed, to give raising/ECM structures. In some adjunct infinitives, however, neither lexical control nor raising is a possible option. Therefore, there is in principle no necessity by which the infinitive should have AGR. PRO_{Arb} or dative control is still an available option in most cases, but not always a necessary one, in that it is not subject to fixed lexico-semantic control requirements. Therefore, absence of AGR would give the result that no principle of licensing concerning AGR applies.

If this is the situation, the prospects are not highly promising yet within the present theory: we are in the middle of nowhere, for we have given up the key element we had recourse to in order to account for distribution of subjects: AGR. It would be little interesting to simply propose independent constraints that apply to subjects when there is not AGR, especially because the distribution of I-subjects in the above cases is obviously reminiscent of the one in finite, control and raising/ECM cases: I-subjects are strictly [-anaphoric].

Therefore, the optimal theory should try to characterize what is the minimal common factor between control/raising/ECM infinitives and IOS. In more technical words, if the preceding considerations are on the right track, we should determine what is common between AGR-ful and an AGR-less infinitives that bears on I-subject distribution. And in addition we should be able to derive the existence of IOS-INV from the NSL status.

Concerning this last desideratum, it could be the case that IOS with a [-anaphoric] I-subject is parasitic on the existence

---

[36] English verbs allowing both control and ECM (e.g. expect) would have [___ IP] only as an optional subcategorization idiosyncrasy.

on long V-movement in infinitives, which, crucially in our theory, allows Nominative assignment through Chain-government, and in turn would be parasitic on the existence of such a long movement in controlled/raising constructions, taking control to be a core case of infinitive that learners use to set parametric options in this field of grammar.

The idea then would be that once there is long V-movement, Nominative assignment by government is possible, even if AGR does not take part in the process. Our previous proposal concerning Case is that the AGR-identifier must provide Case to its I-subject. If it is Spec of AGR, it must transmit its Case to the I-subject. If it is AGR°, we claimed, it must Case-mark the I-subject.

Suppose we assume, *alternatively*, that the requirement is of the type: the AGR-identifier must participate in providing Case to the I-subject. If it is Spec of AGR, we have Case transmission. If it is AGR°, and therefore Case is provided by (Chain-)government, it must combine with a head that is *the actual Nominative Case-marker by government*: T°. Thus it is not only finite T° which is a Nominative assigner, as has been claimed so often: the fact that infinitival T° is not apparently able to assign Case would be due to the independent requirements on AGR, that must participate in the process, if AGR is (forced to be) present. If AGR is not present, then T° can assign Nominative by itself. Why can it not in non-NSLs? The idea would be that in NSLs the case where T° assigns Nominative by itself (in IOS) is not at variance with the case where T° combines with AGR°. Suppose we rephrase the parallelism principle (57) above as:

(77) In the unmarked Case, the I-subject in infinitives obtains Case in the *same* way as in finite sentences (where 'same' means involving sets of processes that are 'unifiable' in the set-theoretical sense).[37]

Then in NSLs taking the unmarked option, T°'s assigning Case by itself to the I-subject is the *same* option as combined assignment by AGR° and T° by Chain-government. In non-NSLs, instead, this is not true, for, in finite sentences, the AGR-identifier in Spec of AGR cannot combine with T° to provide it with case: instead, we have contended, it has to be Case-marked by AGR under agreement and then transmit its Case to the I-subject.[38]

There remain some problems: why does Sardinian allow complement infinitives not to be controlled when a preposition precedes them (see (72))? No similar blocking effect occurs in other languages having prepositional controlled infinitives. On the other hand Sardinian apparently contradicts our claim that lexical control is a strong and pervasive requirement that applies whenever possible. Sardinian, as we saw, is exceptional in another sense: in many cases of IOS, if we take out the overt

---

[37] An alternative (and only apparently simpler) formulation would be:

In the unmarked case, the I-subject in infinitives is Case-marked with the same option as in finite clauses, the options being: a) (Chain-)government; b) Agreement.

The trans-derivational and trans-structural character of this formulation does not change w.r.t. (77). I leave the question open.

[38] What we call parallelism principles should perhaps be called parameters in that they allow for variation. However, I do not know of any non-NSLs taking the marked option (i.e., having Nominative assignment under government in infinitives) If there is none, this asymmetry should obviously be captured. In any case that is why I prefer to keep to the term 'principle': as a parameter, it should have the two options freely available.

subject, the subject reference shifts to control or the PRO_{arb} interpretation can appear when there is no possible controller. Sardinian prepositional infinitives, however, take the arbitrary-existential interpretation steadily (see (73)).[39] Although it is clear what the speaker has internalized as a pattern (no preposition -> control/preposition -> no control), it is not clear how this pattern can have been developed.

In order to account for these facts in line with the above considerations, let us assume that in Sardinian some prepositions are lexically specified to block control (so that the Elsewhere Condition would take this more irregular option).

Let us try to summarize and integrate all the above considerations:

a) Control is the most regular option as far as the Elsewhere Condition is concerned: it applies if the structural conditions are met and there is no more irregular, language-particular option blocking its application.

b) Languages can have lexically determined options blocking control: verbs subcategorizing for IP instead of CP (raising/ECM); prepositions specified for non-control (as in Sardinian), etc. Of course, these language-particular options have to abide by the learnability problem: they have to be easily recognizable, and allowed by UG. Thus, for instance, the existence of ECM or raising has to be learnable (and I think it is, on the mere basis of hearing the constructions). The pattern in Sardinian (absence/presence of preposition <-> control/non-control) is, I think, not difficult to identify, for at least the prepositional IOS complement is readily identified.

c) Another language particular fact that the learner can easily identify is the existence of AGR-inflected infinitives

---

[39] Thus the Sardinian cases of IOS-INV we are considering are comparable to ECM verbs like English expect, which shift to control when the infinitival subject is null.

(see next section). In this case, the issue of whether AGR is present or not is not at stake: it is obviously present and poses no special problem of learnability. This option blocks control.

One idea contained in the above proposals is that the theory must not explain too much: whenever a construction is language specific and poses no obvious problem for being straightforwardly learned, it would be simply inadequate to try to directly derive its existence from principles and wide-scope parameter settings. Principles and parameter settings should only allow the construction, not determine it.

There is a remaining important problem concerning Sardinian: if Sardinian, as we argued, takes the marked option in the Parallelism Principle, then IOS are not expected at all. As for Occitan, it is certainly better behaved in this connection: as we saw in (70)/(71), IOS are (rather) constructed with preverbal subjects. As far as I know, IOS's with inverted subjects are marginal.

Now let us address question (76).b): what licences (referential) *pro* in IOS-INV? We have assumed, in the spirit of Rizzi (1986), that *pro* requires recovery of content. I think it would be senseless to assume that an abstract AGR can recover the content of *pro*, for this would void the traditional intuition that *pro* is allowed as far as its content is overtly expressed in a head.

The solution we will propose is similar in spirit to the one we have proposed for Nominative assignment in the absence of AGR. The idea is that:

(78) AGR, if present in a clause, must be rich enough to licence an empty I-subject (which will be [±anaphoric] depending on the (non) Null Subject status of the language).

(78) is a concise expression of what we have assumed thus far for AGR-licensing, except that we did not consider the possibility of AGR being absent. If it is absent in IOS-INV in a NSL, then the null I-subject being [-anaphoric] (*pro*) will not be required to be identified by AGR, and some other device (such as context recovery) takes place.

The idea advanced here about 'free' *pro* is obviously a simplification: we cannot account for why *pro* can be fully referential in some IOS-INV (e.g., Catalan and Spanish prepositional infinitives) and only arbitrary in others (e.g., Sardinian non-controlled complements). A comprehensive account should even account for cases of null objects (if they are indeed *pro*, as in Rizzi's (1986) proposal about Romance).

I think the account for the referentiality of 'free' *pro* should be expected to be derived from independent modules concerning context recovery (referential *pro*) and unselective Binding by sentence operators (arbitrary *pro* -and PRO), which should interact to allow only the attested cases.[40]

2.4.2. IOS with a Subject in Specifier of AGR

Let us start with AUX-to-COMP, and assume, following Rizzi (1982-b) and subsequent work, that it involves a subject in Spec of INFL (= AGR).

AUX-to-COMP, as far as it is a NSL phenomenon, falls out of all the above proposals: here we are dealing with a subject which is not an I-subject: unlike I-subjects (inverted subjects) in Italian and Portuguese, subjects in AUX-to-COMP construction *may* appear right after the auxiliary. There is another outstanding difference: while all other cases of IOS are colloquial (often

---

[40] See Authier (1991) for an account of Romance null objects in terms of unselective binding.

substandard), AUX-to-COMP is a rather literary construction. Therefore we can take it to be a non-core construction.

In fact, I want to suggest that, in the spirit of the present theory, AUX-to-COMP can be treated as having much in common with IOS-INV. We tentatively will propose the following account:

a) In AUX-to-COMP movement of the verb to C° is triggered because T° moves to C°, for some reason (see below).

b) In AUX-to-COMP, Case assignment to the I-subject is carried out by Chain-government by T° (in combination with AGR°, see below):

(79) Ritengo...
[$_{C°}$ avere+T_i] [$_{AGRP}$ [$_{AGR}$ t_i] [$_{T°}$ t_i] fatto questo Gianni/lui]
to-have done this G. /he
'I believe Gianni/him to have done this'

c) As in finite clauses, once the I-subject is licenced, it can act as a resumptive element licensing a DP in Spec of AGR:

(80) (Peter had it done by a lawyer, but...) ritengo...
[$_{C°}$ averlo+T_i] [$_{AGRP}$ Gianni [$_{AGR}$ t_i] [$_{T}$ t_i] fatto lui/*pro*]
to-have-it Gianni done he
'I believe Gianni to have done it (himself)'

In (80), *pro* is licenced without its content being recovered by a rich AGR. If we assume that AUX-to-COMP in Italian is AGR-less (just like IOS-INV in other languages), then this *pro* would not require content recovery. In fact, however, this is not a plausible approach: AUX-to-COMP in Italian, unlike IOS-INV in other languages, is a construction not allowing arbitrary or referential null subjects, and, anyway, it would be senseless to

assume that the *pro* in (80) is recovered from the context, for it is acts as a resumptive pronoun.

We will alternatively assume that in this case AGR° is *exceptionally* allowed to be enriched by the DP in its specifier (*Gianni*). This is certainly an unusual possibility, which we did not consider thus far:[11] this would be the exceptional point in our proposal, which would account for the marked character of AUX-to-COMP.

d) We want to account for why AUX-to-COMP is restricted to auxiliaries, copulas and modals. We can assume that the reason for this movement is that T° has to move to C° because it is selected (by a factive verb, for instance) or required for the interpretation of the construction (gerunds). Suppose this selection or requirement has some further restriction: T° which is allowed to move to COMP is one having the special property of *not selecting* a VP in its base position: in compound-tense clauses, as well as in copular structures and modal-verb clauses, T° does not select a VP, but rather, respectively, a participle (which is a FC of some sort), a non-verbal predicate (a Small Clause) or an infinitival (if modal-verb clauses are mono-clausal and modals subcategorize for an infinitival). It is crucial to this account that Auxiliaries, copulas and modals do not head their own VP (which would be a complement of T°), but are rather generated under a FC: perhaps under T° itself. If this last suggestion was correct, we could even reformulate the theory another way: only verbal forms which are generated under T° can raise to C° to satisfy selectional restrictions involving T°.

In fact, there is a more appealing possibility: let us assume that the T° which can move to COMP in AUX-to-COMP has to be one not selecting an event:

---

[11] In fact there is circularity in this procedure: pro resumes a preverbal subject which enriches AGR°, which in turn licences pro.

- Modals select a proposition which would take the form of an infinitive.

- As for compound-tense auxiliaries, even if the clause contains an eventful VP, the auxiliary does not carry the Tense interpretation for the event itself, but rather the tense interpretation for the reference time (*At five o'clock*$_1$, John *had-T*$_1$ already *finished-T*$_2$), which is not eventful itself.[12]

- As for copulas, it is harder to contend that they do not select an event, if the predicate is stage-level (see Kratzer (1988)), unless we assume that some (meaningful) FC intervenes between the copula and the stage-level Small Clause.

These are rather speculative considerations. I leave the issue here.

We can assume that AUX-to-COMP in Portuguese is licenced in a similar way. The obvious difference between AUX-to-COMP in Italian and in Portuguese is that in the latter AGR is morphologically present. Our theory predicts, if nothing else is said, that inflected infinitives are essentially like finite clauses as far as AGR and subjects are concerned. In fact, as shown by Raposo (1987-b), there are many restrictions on the distribution of infinitival overt AGR. We cannot address the issue here. The optimal situation for our theory would be that these restrictions are the result of the interaction of factors other than AGR, such as:

- the AUX-to-COMP trigger (which we have assumed is ('event-less') T° raising to COMP).

---

[12] See, e.g., Giorgi (1992) for an implementation of the idea that in compound tenses the Reference time and the Event time are expressed in different positions in the syntactic structure.

- incompatibility between (lexical) control and overt AGR: it must be the same restriction which prohibits PRO in finite sentences.

As for the Occitan examples in (70)/(71), they seem involve a subject in Spec of AGR, not an I-subject. Since we have assume that Occitan takes the marked option in the parallelism principles (i.e., Occitan infinitives have Spec of AGR as the AGR-identifier), these facts are not at odds with the present theory. These examples would be similar to English gerunds with an overt subject. I will not address the issue of what licences non-finite clauses having an overt Spec of AGR (hence an AGR not involving control, raising or ECM).

* * *

The above contention that T° is the basic Nominative Case-marker by government leads to another speculation: the way we have formulated Nominative assignment through agreement, T° cannot take part in this process, for this manner of assignment does not require T raising, manifested as long V-movement (English is an instance of language without obligatory long-V-movement in finite sentences). Therefore, Nominative through government and Nominative through agreement could actually be different Cases. It is noteworthy, in this connection, that non-NSLs tend to develop pronominal forms which show a Case form which exclusively occurs in Spec of AGR (e.g., in colloquial English *he* is not used outside Spec of AGR, contrary to Italian *lui* 'he', which can appear in postcopular position and in dislocated position). So forms like *himself*, French *lui*, which we take as Nominative when they are I-subjects, and must be non-Nominative in other cases, would be neutral forms: English (and

French) would have no Case distinction except for pronominal agreement-Nominative forms.[43]

Summing up, Nominative assignment works in the following way:

a) T° is the unmarked Nominative Case-marker: it is so in NSLs, which are the unmarked option: both in finite clauses and, with the unmarked parallelism option, also in infinitives.

b) AGR° is, to use a metaphor, 'jealous' of T°'s Nominative Case marking: if T° actually assigns Nominative (in the cases mentioned in a)), then, if AGR° is around (in finite clauses, control, and Raising, and AUX-to-COMP), it wants to take part (just like a jealous younger baby wants to take part in the older baby's game whenever the latter plays). Since AGR has a lot of restrictions for its own licensing, this ultimately will reduce the possibilities for subjects to control, raising and ECM (in the metaphor, the younger baby's taking part actually reduces the possibilities for the older one's games).

c) If AGR° is not around (in IOS minus AUX-to-COMP), then T° can assign Nominative alone, and the possibilities for subjects seem to surprisingly increase (the little baby is not pestering around).

d) In languages where T° never assigns Nominative (non-NSLs), AGR is the only Case assigner (by agreement).

---

[43] In Chapter 3 we commented on these facts in another sense: these Spec-of-AGR-only pronominal forms would be AGR-identifier forms, which, diachronically, tend to cliticize to AGR° to become (unmarked) AGR° AGR-identifiers. I think both ideas (Case singularity and AGR-identifier singularity) can converge in a natural way, perhaps one being derived from the other. The fact that these Case distinctions are exclusively pronominal suggests that the AGR-identifier singularity is more basic, for pronominals, unlike full DPs, are minimal sets of AGR-features.

3. Summary

In this chapter we have extended our theory in Chapter 3 to infinitival constructions. In fact our theory on the [±anaphoric] status of I-subjects forced us to make some assumptions on the existence and status of AGR in non-finite clauses. Although this extension from finite to non-finite is far from being trivial (and could possibly be implemented in other ways), I think it is has some plausibility in that it accounts for a range of facts that are thus far poorly understood: long infinitival verb movement in many NSLs and the existence of neutral pronominal/anaphoric forms in NSLs not having infinitival long V-movement; and it gives reasonable alternative account to Kayne's (1991) theory concerning clitics and PRO, avoiding the problems we noticed for this proposal.

Our treatment of Infinitives with an Overt Subject (IOS) is only tentative. The literature on the issue is fragmentary and far less developed than that devoted to other types of infinitives. I think the reason for this is that Principles and Parameters (and all the research stream leading to it since the 1960's) is, as far as infinitives are concerned, intrinsically feeble to account for IOS. Most research on infinitives has ignored IOS, perhaps because English lacks it, and this fact has possibly biased research trends. This is a typical situation in any empirical science, which is rooted in the necessarily accidental component of research.

Chapter 5
Preverbal Subjects in NSLs

1. Finite Sentences

There is an important question that remains unanswered in the above theory: what is the status of preverbal subjects in NSLs? If they are not required as AGR-identifiers, how are they licenced?

A reasonable position is that specifiers are not always filled: it would at least be difficult to contend they always are for any category and any well-formed structure (we keep neutral w.r.t. the issue whether non-filing implies non-projection). The obligatory filling of a given specifier should rather follow from principles and parameter settings. The above theory characterizes the requirements on the AGR category in a way that makes Spec of AGR obligatorily filled only for non-NSLs. So in a sentence like the following (Catalan):

(1) Ho ha fet (en Joan)
It-has done (the J.)
'JOAN/he did it'

Spec of AGR would be empty, at least when the I-subject is an R-expression (*en Joan*); when it is empty (*pro*), it is not logically impossible that Spec of AGR be filled by another empty category (possibly *pro*), but the null hypothesis is that it is not.

Now we have to answer at least three questions:

- Is Spec of AGR ever filled in NSLs? If it is:
- What is its status? More specifically:
- Is it filled only by elements coindexed with the I-subject?

We will address these questions in sections 1.1. and 1.2.

1.1. Dislocation vs. Specifier (of AGR)

A conceivable approach to the nature of preverbal subjects in NSLs is to assume that they are left dislocated DP's. There are various arguments in favor of this view. We will see that it is nevertheless too simplistic a view.

Before addressing the issue, let me say a word about word order. It has often been claimed that inverted subject constructions cannot be the 'basic word order' for Romance languages allowing this option, for they involve a Focus interpretation which is contextually and pragmatically marked. Therefore preverbal subject constructions (which are indeed pragmatically more neutral) have to be basic structures, and a basic structure will not involve dislocation. I think that these considerations are pointless: they involve a naive conception of transformational grammar that has been largely overcome: nowadays underlying levels of representation are abstract and complex representations which simply cannot be claimed to be more or less 'basic'. When I claim that inverted subjects are Case positions in NSLs I am not committing myself to any claim about what is more 'basic' or 'neutral': inverted subject constructions are simply an available option of the grammar which happens to be pragmatically or contextually marked (in that it usually involves Focus).

A basic argument in favor of dislocation for preverbal subjects in NSLs is a theoretical one, based on indeterminacy:[1] even the most radical theory contending preverbal subjects in NSLs are not dislocated must admit that they can be. So in a simple sentence like:

(2) En Joan no ha vingut
The J. not has come

would ambiguous between the dislocation and Spec of AGR status of *en Joan*. More or less explicitly, many authors have contended this ambiguity is only apparent: the dislocation version would imply both a phonological pattern (usually a rise-and-fall intonation) separating the preverbal subject from the rest of the sentence, and a special interpretation by which the preverbal subject is read as 'as for Joan' or something similar.

It is true that there may be a phonological clue for dislocated elements. What is not true, at least in Romance, is that the it is obligatory: any clitic left-dislocated (CLLD) element can be pronounced without any special pause or phonological clue possibly differentiating it from what would be a 'true' non-dislocated subject. Here are some examples CLLD (and the pattern generalizes to any CLLD element and to any Romance language as far as I know), where no pause or phonological clue obligatorily indicates dislocation:[2]

---

[1] This is one of the arguments wielded in Rosselló (1986).

[2] Normative grammarians do not like dislocation in formal speech, especially in written formal speech. Liberal ones, they allow it only in informal speech, and then a comma should be used. This might be at the origin of the belief on the existence of the pause that has gone unchallenged by many linguists.

(3) a. A en Joan li han robat la cartera
To the J. him-have robbed the wallet
'They stole the wallet from Joan'

b. Amb ell no hi parlo
With him not there-speak-I
'To him, I never talk'

So both preverbal subjects and CLLD elements have an only optional pause, which depends on the degree of emphasis (in fact, it is not a discrete sign: one can add more emphasis by overdoing the tone break or adding a pause, in a continuous way). It may be the case that with CLLD elements the (optional) dislocation pattern is used more often than with subjects. But if so, it would be natural: subjects are more often [+human] than internal Arguments; [+human] DPs are more often D-linked for obvious reasons; D-linked DPs need less emphasis to be introduced; therefore, by diffuse-logic *modus ponens*, subjects need less often the emphatic dislocation pattern.

As for the interpretative clue, it has often been contended that a true dislocated element cannot be used in a sentence which is a natural answer to 'What is happening?'. This test, apart from being a bit vague (it implicitly relies on pragmatic factors such as speakers' optimal cooperation) does not give clear results. All of the following examples in Catalan are a reasonable answer to 'What's happening?' (e.g., when someone arrives and feels something strange in the atmosphere):

(4) a. A en Joan li han robat la cartera.
To the J. him-have robbed the wallet
= (3).a)

b. Al nen l'ha mossegat una rata
To-the child him-has bitten a rat
'A rat bit the child'

c. A en Joan li han concedit una beca
To the J. him-have-they awarded a grant
'Joan has been awarded a grant'

d. A en Joan li ha vingut un atac de cor
To the J. him-has come an attack of heart
'Joan had a heart attack'

In fact, the non-dislocated versions of the above examples would be slightly unnatural as an answer to 'What's happening?' (they would be certainly acceptable in literary speech, for they would be the outputs of the avoid-dislocation correction effort').

We must point out the in (4) the CLLD element is human and definite. These factors are clearly relevant. So the following examples would not be appropriate replies to 'What's happening?':

---

[3] There is a long tradition in traditional grammars stating that SVO is the unmarked word-order, and it is a dearly recommended one when the grammar is prescriptive. This partly can be traced back to the XVIIIth century belief that French was the language of reason, because, among other trifles, it had the virtue of systematically expressing the 'natural order of thoughts' (i.e. 'agent-action-object') properly (i.e. SVO). If this sheer nonsense deserves any theoretical attention it is precisely a warning against being misled by it as far as it has become 'common sense' belief among literate people that SVO is the unmarked order.

(5) a. Una sargantana, l'he vista al jardí
A lizard it-have-I seen in-the garden
'One (of the) lizard(s), I saw in the garden'
b. De lingüistica, en parlarem després
Of linguistics of-it-will-we-talk later
'Linguistics, we'll talk about later'

But about the same happens with preverbal subjects, as the following are not appropriate replies either (perhaps they can be felt as more acceptable than the preceding because of normative pressure for SV(O) word order):

(6) a. Un roc ha caigut
A stone has fallen
'One (of the) stone(s) fell'
(not just: 'A stone fell')
b. Una cotxe ha passat
A car has gone-by
'One (of the) snake(s) went by'
(not just: 'A snake went by')

There seem to be other factors favoring neutral (i.e., 'What's happening?'-appropriate) dislocation or preverbal subjecthood: there is a gradation agent-dative-object-oblique going from most to least favoring. This favoring gradation is specially apparent with 'inverted' psych-verbs (It. *piacere*, 'to like' see Belletti & Rizzi (1988)), which most often have the Dative or Accusative Experiencer Argument as CLLD and the subject Theme as an I-subject. I will not pursue the issue. Suffice it to be the case that no criterion singles out, as far as we have seen, preverbal subjects as opposed to CLLD elements in a clear-cut way: both subjects and CLLD datives or objects have the same favoring conditions for (non-Focus-fronted) preverbal position,

even if one of the favoring conditions is possibly being an external Argument (then a subject).

There is a more solid criterion that has become a classical test for subjecthood as opposed to dislocation: only true subjects can be quantified. I think this criterion can be a good test. But I think too that it has many times been oversimplified. It is not enough to come up with a few examples (such as (6)) and conclude they are genuine cases of true subjects. On the empirical side, it may happen (and it does happen) that not all quantifiers behave alike. On the theoretical side, we know there are various kinds of quantifiers: apart from 'true' quantifiers, researchers have proposed subclasses such as D-linked quantifiers (Pesetsky (1982)), branching quantifiers (May (1985), Hornstein (1984)), indefinite DPs with a referential reading (Fodor & Sag (1982)), etc. I will not address the issue of a proper classification (these classes are not complementary), but many of the proposals coincide in characterizing some quantifiers as being 'less quantificational' than the 'true' ones. For convenience, let us call the former 'weak' quantifiers (D-linked, referential)[4] and the latter 'strong' quantifiers.

Now one possible expectation would be that only 'weak' quantifiers are possible preverbal subjects or CLLD elements. The expectation is more or less fulfilled (remember, however, what we said above about other factors, which might distort the results). So it is the case that in Romance NSLs preverbal indefinite subjects tend to be 'referential' (= 'a certain') or D-linked (= 'one of the'), while the purely existential interpretation is hard to obtain (see (6)). Bare indefinite DPs

---

[4] Cinque (1990) uses referential for what we call 'weak', assuming referentiality subsumes D-linking. In fact, it seems to me that with D-linked quantification what is really referential is the set over which the quantifiers range. But it is only a matter of terminology: 'weak'/'strong' are in fact vague terms used here to avoid being committed to specific theories of quantification.

are simply not allowed as preverbal subjects in these languages (while they are in English) -(7).a); Some negative and proportional quantifiers are not either -see (7).b/.c); (8) shows that the same restrictions hold for CLLD elements resumed by definite clitics:

(7) a. *Rocs cauen de la muntanya
Stones fall from the mountain
b. *Res ha passat
Nothing has happened
c. *Pocs estudiants han vingut
Few students have come

(8) a. *Estudiants els he vist
Students them-have-I seen
b. *Res ho ha fet
Nothing it-has done
c. *Pocs estudiants els ha convidat
Few students them-has invited

Spanish and Italian display about the same restrictions (examples from Italian; recall we are not dealing with Focus-fronted elements, which would be acceptable in (9), with a different intonation):

a) Peverbal subjects:

(9) a. *Studenti sono arrivati
Students are arrived
'Students have arrived'
b. ?*Pochi studenti sono arrivati
Few students have arrived
c. ?*Niente è successo
Nothing has happened

b) CLLD elements:

(10) a. *Studenti li ho visti
Students them-have-I seen
'I have seen students'

b. *Pocchi studenti li ho visti
Few students them-have-I seen
'I've seen few students'

c. *Niente l'ho visto
Nothing it-have-I seen
'I've seen nothing'

The reason these examples are not allowed is that these quantifiers are 'strong', and cannot be forced to a D-linked or 'referential' reading.

Similar considerations can be made of proportion quantifiers ('a few', 'most', 'many') and numerals: the D-linked or 'referential' reading prevails for preverbal subjects, while postverbal subjects are ambiguous:

(11) a. Molts estrangers treballen aquí
Many foreigners work here
'Many of the foreigners work here'
Not: 'There are many foreign workers here'

b. Aquí hi treballen molts estrangers'
Here there-work many foreigners
a) 'There are many foreign workers here'
b) 'Many of the foreigners work here'

As for universal quantifiers, there do not seem to be sharp occurrence restrictions.

There is, in addition, an interpretative constraint which holds true of any preverbal quantifier in Catalan (and, I assume,

in other Romance languages): they never have narrow scope w.r.t. a another quantifier inside their clause. Narrow scope for the subject quantifier has been reported to be less preferred in English and other languages, but I think that in Romance NSLs is not only less preferred: it is excluded.

The facts may have often been obscured because scope interactions can be interfered by the 'weak' status of one of the quantifiers: especially in the case of indefinite DPs with D-linked or referential reading, this reading is such that it gives the same extensional interpretation as the wide scope reading (consider, e.g., the sentence 'Everybody loves a certain woman', which is extensionally equivalent to 'Everybody likes a woman' with wide scope for 'a woman'). So, in the following example:

(12) Tot estudiant ha llegit tres llibres
Every student has read three books

we can conceive of three readings: two with scope interactions (wide scope for 'every student' and wide scope for 'three books') and one with no scope interaction, when the numeral is interpreted as 'some certain three books'. Since the wide-scope reading for 'three books' is extensionally equivalent to the non-scope reading, our claim that 'every student' cannot have narrow scope can only rely on the intuition that 'three books', when apparently wide scope, is 'referential'. So (12) cannot be conclusive w.r.t. our claim that 'every student' cannot have narrow scope.

There are however cases where the wide scope interpretation of a quantifier is not equivalent to its scopeless reading. Consider the scope interactions between two numerals in the following example:

(13) He enviat tres exemplars a quatre editorials
Have-I sent three copies to four publishing-houses
.
'I've sent three copies to four publishing houses'

In (13) both numerals are inside the VP, and we have the two relative scope readings ('3 copies and possibly 12 publishing houses' and '4 publishing houses and possibly 12 copies'), and in addition we have the non-scope reading, with 'just 3 copies and 4 publishing houses'. If, however, one of the numerals is a preverbal subject, as in the following example:

(14) Tres directors han dirigit quatre films
Three directors have directed four films

we have the scopeless reading and only one of the scope readings, namely the one with the preverbal subject having wide scope. As I said, even in English the wide scope reading has been reported to be preferred, but it is simply the only option in Catalan.[5]

Negative quantifiers[6] and cannot either have narrow scope when allowed as preverbal subjects or CLLD elements:

---

[5] There are actually some speakers who admit it, but most people steadily exclude it. My Italian reports are also less clear-cut.

[6] I use the term 'negative quantifier' for convenience. Zanuttini (1991) argues these elements in some Romance languages are negative universal quantifiers involving negative concord, and argues against a negative polarity item analysis. I can not address the question here. Only recall that in Romance NSLs these elements do not cooccur with the sentence negation particle when they are preverbal.

(15) a. *Preverbal subjects*:

Ningú coneix tots els llibres

Nobody knows all the books

a) 'There is no person knowing every book'

b) *'Every book is such that no person knows it'

b. *CLLD elements*:

A ningú el coneix tothom

To nobody him-knows everybody

a) 'There is no person being known by everybody'

b) *'Everybody is such that he knows nobody'

Similarly, existential and proportional quantifiers never have narrow scope in the following cases:[7]

(16) *Preverbal subjects*:

a. Un/algun metge es cuida de tots els malalts

A /some doctor takes-care of all the patients

b. Molts/la majoria d'estudiants llegeixen tots els llibres

Many /most students read all the books

[7] For some reason, existential quantifiers of the type 'some NP or other' and numerals like 'at least one' can have narrow scope no matter how higher up from the other quantifier they are:

(i) Some book or other, I think every student has read

(ii) At least one car, I guess everybody can afford

Since this narrow scope reading is not predictable from the general clause-boundedness constraint for scope interactions, we leave it as irrelevant. Therefore, in ((16).a) and ((17).a) this type of reading, which is usually obtained with a special intonation pattern, should be disregarded. Interestingly, ((17).a) only allows this reading if the preposition optionally accompanying the CLLD object is absent.

(17) *CLLD elements:*

a. (A) un/algun malalt el visiten tots els metges
(To) a /some patient him-visit all the doctors
'A/Some patient is visited by all the doctors'

b. Molts/la majoria de llibres els llegeix tothom
Many /most books them-reads everybody
'Many/most books are read by everybody'

So the quantification test for preverbal subjects gives no clear-cut result because quantification is not a trivially unitary concept in natural language. A minimal characterization of the facts could be the following:

- some indefinite DPs (bare NPs, non-D-linked non-referential existential quantifiers -such as 'few'- and some negative quantifiers not having the [+human] feature - e.g. 'nothing') are not possible as preverbal subjects, in the same way as they are not as CLLD elements.

- other quantifiers (negative elements, universal quantifiers) are possible preverbal subjects or CLLD elements, but they always have wide scope.

Given this state of affairs, an obvious possibility is to assume that:

- preverbal subjects in NSLs are dislocated elements resumed by a *pro* (as CLLD elements are resumed by a (definite) clitic).

- only 'weak' quantifiers can be dislocated (this is less evident for universal and negative quantifiers).

We will see that this hypothesis can not be maintained. Thus far we have not considered an additional criterion that can distinguish dislocated elements from elements in Spec of AGR: word order. If we assume that:

a) V moves to AGR° in NSLs (and we crucially assumed that);
b) there can be no maximal projections adjoined to X' (and specifically to AGR'): this is a reasonable constraint on X'-structure;

then it follows that no XP can intervene between a true subject in spec of AGR and the verb in AGR°. This criterion is useless as far as non-quantified subjects are concerned: there is no way to tell whether adjacency is due to the preverbal subject's being in Spec of AGR or simply to the accidental fact that no XP happens to intervene. But for quantified subjects the data are clearly revealing. At least some of the quantified preverbal subjects require adjacency (we will see examples directly).

So we seem to finally arrive at some conclusion about preverbal subjects: at least some of them are in a fixed position, which is likely to be Spec of AGR. But let us raise another question first: we have seen that the restrictions on preverbal subjects are the same as on CLLD elements. So it might be that even for CLLD elements there was an adjacency requirement when they are quantifiers of a certain type. And this is indeed the case. What follows is set of pairs of examples (each pair containing a CLLD example and a preverbal subject example), in a gradation from the most ill-formed to the best well-formed cases. For the examples with preverbal subjects we abstract away from Focus fronting, which is irrelevantly acceptable without the part in the parenthesis and with another intonation, as it would be with objects and other Arguments: for the latter, however, no clitic would appear.* The asterisk at the beginning means the

---

* I suspect that there are processes which are similar to Focus-fronting and nevertheless do not have the typical intonation and contrastive interpretation of typical Focus-fronting. So, in the judgements below I tried to disregard the acceptability of the preverbal subjects when a parallel (clitic-less) object fronting is available which intuitively has the same phonological, interpretative and stylistic flavor. For what is at stake here is if preverbal subjects are any different from

sentence is ill formed even without the part in the parenthesis; the asterisk inside the latter means that the part in the parenthesis impairs the sentence or makes it bad:

(18) a. 1. *Res (*en Joan) ho ha fet
Nothing the J. it-has done
2. *Res (*a en Joan) li ha passat
Nothing to the J. him-has happened
b. 1. *Cap (*en Joan) l'ha llegit
None the J. it-has read
2. *Cap (*avui) ha arribat
None today has arrived
c. 1. *Pocs amics (*aquest any) els ha convidat
Few friends this year them-has invited
2. *Pocs amics (*aquest any) m'han convidat
Few friends this year me-have invited
d. 1. ?*Cap paquet (*en Joan) l'ha enviat
No packet the J. it-has sent
2. ?*Cap paquet (*avui) ha arribat
No packet today has arrived
e. 1. A ningú (*aquesta vegada) l'han acceptat
To nobody this time him-have-they accepted
2. Ningú (?*aquesta vegada) ha aprovat l'examen
Nobody this time has passed the-exam
f. 1. A tothom (?*aquesta vegada) l'han acceptat
To ev.body this time him-have-they accepted
2. Tothom (?*aquesta vegada) ha aprovat l'examen
Ev.body this time has passed the-exam

---

clitic-resumed elements.

(18) g. 1. A qualsevol (???al Brasil) el poden robar
To anybody in Brasil him-can rob

2. Qualsevol (???al Brasil) et pot robar
Anybody in Brasil you-can rob

h. 1. A tots (??demà) els veuré a la reunió
To all tomorrow them-will-I-see at the meeting

2. Tots (??demà) aniran a la reunió
All tomorrow will-go to the meating

i. 1. A algú (?fa una estona) l'han tancat al lavabo
To sm.body a while ago him-have locked in-the bathroom

2. Algú (?fa una estona) s'ha tancat al lavabo
Sm.body a while ago himself-has locked in the bathroom

j. 1. A tots els alumnes (?demà) els veuré
To all the pupils tomorrow them-will-I-see

2. Tots els alumnes (?demà) vindran a veure'm
All the pupils tomorrow will-come to see-me

k. 1. A tots aquests alumnes (demà) els veuré
To all these pupils tomorrow them-will-I-see

2. Tots aquests alumnes (demà) vindran a veure'm
All these pupils tomorrow will-come to see-me

l. 1. La majoria d'aquests (en Joan) no els coneix
The majority of-these the J. not them-knows

2. La majoria d'aquests (a mi) no em coneixen
The majority of-these to me not me-know

m. 1. A molts d'aquests (avui) no els he vist
To many of-these today not them-have-I seen

2. Molts d'aquests (avui) no han vingut
Many of-these today not have come

(18) n. 1. A dos amics meus (avui) els han arrestat
To two friends mine (today) them-have arrested
2. Dos amics meus (avui) han tingut un accident
Two friends mine today have had an accident
o. 1. Alguns amics (al casament) no els hi convido
Some friends to-the wedding not them-there-invite-I
2. Alguns amics (al casament) no hi vindran
Some friends to-the wedding not there-will-come
p. 1. A un meu amic (l'altre dia) el van arrestar
To one my friend the other day him-arrested-they
2. Un meu amic (l'altre dia) va venir a veure'm
One my friend the other day came to see-me

Examples from a. to d. are unacceptable as preverbal (and even worse if not adjacent to the verb): they are all 'strong' quantifiers hardly interpretable as D-linked or 'referential'. They are all [-human]. From e. to j., they are acceptable only if adjacent to the verb (with various degrees of ill-formedness if they are not). The rest of the examples are fully acceptable even if there is not adjacency to the verb. The gradation can be clearly related to the 'strength'/'weakness' distinction. The more a quantifier is likely to be interpreted as D-linked or 'referential', the more it is acceptable as dislocated. I think that in fact the intermediate degrees of ill-formedness are actually not 'intermediate' in grammatical status, but rather in pragmatic acceptability: what is intermediate is the chances for the speaker to imagine a likely pragmatic context where D-linking or 'referentiality' are plausible. So for instance the 'referential' use of *algú* 'somebody' implies that the speaker is wanting to be enigmatic or unexplicit in referring to a person. It is more usual that, when one wants to refer to somebody the other person does not know, one minimally introduces a

descriptive restriction for the quantifier (e.g., 'a friend of mine', 'some guy I met the other day', etc.).

So, the conclusion seems to be that:

- 'Strong' [-human] quantifiers cannot be preverbal (unless moved by Focus fronting).

- 'Strong' [+human] quantifiers can be preverbal but they have to be adjacent to the verb, which suggests they are in Spec of AGR. They involve a resumptive clitic strategy if they are not subjects. If subjects, they would be resumed by *pro*.

- 'Weak' (D-linked or 'referential') quantifiers can be dislocated.

What is of our present concern is that there is a preverbal Specifier position (we assume for the moment it is Spec of AGR) that can be filled by [+human] strong quantifiers being resumed by a minimal (clitic or *pro*) pronoun. So our previous speculation that all preverbal subjects are all dislocated is false, but now we conclude their position is not one specific to subjects.[9]

Another fact which is worth considering is that the acceptable preverbal quantified elements having the V-adjacency requirement (examples from e. to g. above) are clause bound: if they appear in the upper clause they are unacceptable approximately to the same degree as if they are not adjacent (the following examples are identical to examples from (18).e) to g., except that instead of an intervening XP we have a superordinate clause segment *crec que* 'I think that'):

---

[9] In fact, if that position exists, nothing prevents non-quantified preverbal subjects from being there, unless we assumed it is an exclusively quantifier position. We will address this issue in the next section.

(19) a. 1. A ningú (*crec que) l'han acceptat
To nobody (I-think that) him-have-they accepted
2. Ningú (?*crec que) ha aprovat l'examen
Nobody (I-think that) has passed the-exam
b. 1. A tothom (?*crec que) l'han acceptat
To ev.body (I-think that) him-have-they accepted
2. Tothom (?*crec que) ha aprovat l'examen
Ev.body (I-think that) has passed the-exam
c. 1. A tothom (???crec que) el poden robar
To ev.body (I-think that) him-can rob
2. Tothom (???crec que) roba
Ev.body (I-think that) robs
d. 1. A tots (??crec que) els veuré a la reunió
To all (I-think that) them-I'll-see at the meeting
2. Tots (??crec que) aniran a la reunió
All (I-think that) will-go to the meating
e. 1. A algú (??crec que) l'han tancat al lavabo
To sm.body (I-think that) him-have locked in-the bathroom
2. Algú (?crec que) s'ha tancat al lavabo
Sm.body (I-think that) SI-has locked in-the bathroom
f. 1. A tots els alumnes (?crec que) els veuré
To all the pupils I-think that them-will-I-see
2. Tots els alumnes (?crec que) vindran a veure'm
All the pupils I-think that will-come to see-me

Finally, there is still another fact that has to be considered. 'Strong' [-human] quantifiers are not allowed as preverbal subject or clitic resumed elements in neutral sentences. But there is a factor that can rescue them in this

position: they are allowed when they are somehow construed with a postverbal contrastive Focus:[19]

(20) a. 1. ???Res ho ha fet EN JOAN
Nothing it-has done the J.
'Nothing has been done by Joan' (-> everything has been done by others')

2. ??Res li ha passat A EN JOAN
Nothing him-has happened to the J.
'(Many things happened to many people, but) nothing happened precisely to Joan'

b. 1. ?Cap l'ha llegit EN JOAN
None it-has read the J.
'(We have read many of these books but) none was read precisely by Joan'

2. Cap ha arribat AVUI
None has arrived today
'(Many have arrived, but) none arrived precisely today'

---

[19] To my ear, examples in each pair 1./2. (resp. clitic resumed elements and subjects) are equally acceptable. Some speakers prefer the examples with subjects. I assume this is due to the fact that these speakers are assigning some fronting analysis to these examples, which is impossible with the clitic-resumed elements, as far as fronting does not envolve clitics. On the other hand, there are kinds of fronting which are not Focus fronting, and do not envolve the typical phonological pattern of Focus fronting:

(i) A mi, res m'han dit.
To me nothing me-have told
'Me, I was told nothing whatsoever'

It would be this 'intonation-flat' fronting which makes the examples in 2. better for some speakers. In my dialect, this kind of fronting is not used.

(20) c. 1. ?A pocs assistents els havia convidat EN JOAN
To few people-present them-had invited the J.
'(Many people were were at the party, but) few had been invited by Joan'

2. Pocs amics m'han convidat A SOPAR
Few friends me-have invited to dinner
'(Friends have invited me to lots of things, but) few have invited me to have dinner'

d. 1. (?)Cap paquet l'ha enviat EN JOAN
No packet it-has sent the J.
'(Many packets were sent, but) none by Joan'

2. Cap paquet ha arribat PER CORREU
No packet has arrived by mail
'Packets arrived, but none arrived by mail'

To summarize, quantified subject- or clitic-resumed-elements are allowed if:

- they are 'weak' (then they can be dislocated)
- they are 'strong' [+human] (then they are in a Spec position).
- they are construed with a postverbal Focus elements (and then they are in this Spec position too).

The facts reported from Catalan extend to Spanish with a reasonable degree of approximation. As for Italian, the judgements I have been reported are less clear and systematic. Southern dialects[11] seem to be close to Catalan, with the

[11] Thanks to Giuseppe Longobardi for his judgements, comments and suggestions

difference that adjacency to the verb does not seem to play any role:[12]

(21) a. 1. *Niente (Gianni) lo ha fatto
Nothing (G.) it-has done
2. *Niente (a Gianni) gli è successo
Nothing to G. him-has happened
b. 1. *Nessuno (oggi) l'hanno letto
None (today) it-have-they read
2. *Nessuno (oggi) è arrivato
None (today) has arrived
c. 1. *Pochi amici (questo anno) li ha invitati
Few friends (this year) them-has invited
2. *Pocchi amici (questo anno) mi hanno invitato
Few friends (this year) me-have invited
d. 1. ?*Nessun pacco (Gianni) l'ha spedito
No package (G.) it-has sent
2. ?*Nessun pacco (oggi) è arrivato
No package (today)
e. 1. *Ognuno (questa volta) l'hanno accettato
Everyone (this time) him-they-have accepted
2. *Ognuno (questa volta) ha passato l'esame
Everyone (this time) has passed the-exam
f. 1. ?Nessuno (questa volta) l'hanno accettato
Nobody (this time) him-they-have accepted
2. Nessuno (questa volta) ha passato l'esame
Nobody (this time) has passed the-exam

---

[12] Thanks to Giuseppe longobardi and Giovanni Albertocchi for the data (which I take as roughly representative of resp. Southern and Northern varieties of Italian).

Italian has less unambiguously {+human} quantifiers than Catalan or English: so nessuno is both 'nobody' and 'none'; tutti is both 'everybody' and 'every one/all', etc. Nevertheless the [±human] interpretation is equally significant in the data below.

(21) g. 1. Tutti (questa volta) li hanno accettati
All (this time) them-they-have accepted
2. Tutti (questa volta) hanno passato l'esame
All (this time) have passed the-exam

h. 1. Chiunque (in Brasile) lo possono derubbare
Anybody (in Brasil) him-they-can rob
2. Chiunque (in Brasile) ti può derubbare
Anybody (in Brasil) you-he-can rob

i. 1. Tutti (domani) li vedrò alla riunione
All (tomorrow) them-I'll-see at-the meeting
2. Tutti (domani) andranno alla riunione
All (tomorrow) will-go to-the meeting

j. 1. ?*Qualcuno (pocco fa) l'hanno chiuso nel bagno
Somebody (just now) him-have locked in-the bathroom
2. Qualcuno (pocco fa) si è chiuso nel bagno
Somebody (just now) himself-has locked in-the bathm.

k. 1. ?Tutti gli studenti (oggi) li vedrò in classe
All the students (today) them-I'll-see in class
2. Tutti gli studenti (oggi) verranno a vedermi
All the students (today) will-come to see-me

In Northern dialects, adjacency seems to play a role. They differ from Catalan in that preverbal subjects appear to be more acceptable than other clitic resumed elements.