

2. Preverbal subjects in Romance null subject languages

2.1. Introduction

Central to Alexiadou & Anagnostopoulou's (A&A) (1998) Minimalist analysis of pro-drop is the assumption that all preverbal subjects in NSLs are clitic left dislocated (CLLD) in an A'-position. This is a direct implication of the assumption that the EPP is satisfied by pronominal agreement morphology in NSLs. Given the fact that the EPP is widely considered a principle of UG, the onus is very much on A&A to substantiate their claim that its mode of satisfaction is parameterised, and that NSLs, consequently, lack a preverbal subject position. This becomes all the more apparent, when we consider the following statement by Hulk and Pollock (2001:3):

“There is a consensus among both traditional and generative grammarians that the canonical surface word order of the Romance languages is subject-verb-object”.

For A&A's analysis to be falsified for a given language, it need only be proven that: (i) the verb raises to I in said language, and (ii) at least some preverbal subjects in said language do not occupy an A-bar position. Before such an endeavour can be undertaken, it is therefore necessary to assess the position of finite verbs in Romance NSLs and to define more clearly what is meant by the term CLLD, distinguishing it from other kinds of doubling found in Romance.

2.2. Overview of the chapter

Section 2.3.1 gives an overview of the evidence that verbs raise to a position in I in all of the Romance NSLs. Section 2.3.2 provides a clear definition of CLLD, including its relevant syntactic properties. With reference to these properties, section 2.4 examines the evidence in favour of the implicit claim that *all* preverbal subjects in Romance NSLs share *all* the properties of CLLD. This will include evidence from adverb placement, wide scope readings, the behaviour of indefinites, Italian dialect clitics and bound variable readings. Section 2.5 will discuss the counter evidence to A&A's claim: non

topic-like subjects, lack of topic-island effects, minimality effects, acquisition, the lack of *ad sensum* agreement and the position of null subject pronouns.

Based on this wealth of strong counter-evidence, it will be concluded that, although many preverbal subjects are CLLD in Romance NSLs, it cannot be maintained that *all* preverbal subjects are. As A&A's analysis requires this stronger claim to be true, it is concluded, in section 2.6, that an analysis based on a parameterised EPP simply cannot be maintained, in their terms. It will be suggested that the data argue in favour of an approach where agreement is the spellout of uninterpretable features, valued and deleted during the course of the derivation.

2.3. Background clarifications

2.3.1. Verb raising: Pollock 1989

Pollock (1989), drawing on work by Emonds (1978) argued that contrasts between French and English regarding the position of negation, adverbials and floating quantifiers such as those in (1)-(6) can be accounted for by positing a parametric difference regarding verb movement in the two languages:¹

(1) a. Jean (n') aime **pas** Marie.
 Jean neg likes not Marie

 b. *Jean (ne) pas aime Marie.
 Jean neg not likes Marie

(2) a. *John likes **not** Mary.

 b. John does **not** like Marie.

¹ In French the preverbal negative particle *ne* is assumed to occupy a different position to that occupied by true negation *pas*. Pollock (1989:366:fn 3) provides a diachronic justification for this assumption.

- (3) a. Jean embrasse **souvent** Marie
Jean kisses often Mare
- b. *Jean **souvent** embrasse Marie
Jean often kisses Mare
- (4) a. *John kisses **often** Marie
- b. John **often** kisses Marie
- (5) a. Mes amis aiment **tous** Marie
My friends like all Marie
- b. *Mes amis **tous** aiment Marie
My friends all like Marie
- (6) a. *My friends like **all** Mary
- b. My friends **all** like Mary

His basic insight is to propose that negation, adverbials and floating quantifiers occupy the same position in both languages and that the differences in word order are due to the fact that the finite verb raises into the IP domain in French, but not in English. Finite auxiliaries in English, on the other hand, *do* raise:

- (7) a. I have **not** seen Mary.
- b. *I **not** have seen Mary.
- (8) a. I have **often** wondered whether...
- b. ??I **often** have wondered whether...

(9) a. They have **all** attempted to complete the task.

b. ??They **all** have attempted to complete the task.

Pollock (1989:383) further proposes, based on evidence from the position of infinitivals in the two languages, that IP is split into two separate heads; T and AGR, in the following configuration:²

(10) Subject T **not/pas** Agr **Adv** [VP ...V ...]

The order of T and Agr has been disputed in the literature, with the general consensus being that Agr c-commands T, as proposed by Belletti (1990), and adopted by Chomsky (1991). This would give us the following configuration in French/English:

(11) Subject Agr **not/pas** T **Adv** [VP ...V ...]

² This was designed to account for the fact that infinitivals appear to undergo optional 'short verb movement' to a position below negation but above certain adverbs commonly taken to mark the edge of VP:

- (1) Ne pas manger de chocolat est une honte
CL neg to-eat of chocolate is a disgrace
- (2) *Ne manger pas de chocolat est une honte
CL to-eat not of chocolate is a disgrace
'To not eat chocolate is a disgrace.'
- (3) À peine parler l'italien après cinq ans d'étude est une honte
hardly to-speak Italian after five years of study is a disgrace
- (4) Parler à peine l'italien après cinq ans d'étude est une honte
to-speak hardly Italian after five years of study is a disgrace
'To hardly speak Italian after five years of study is a disgrace.'

The position was first proposed, therefore, merely as an alternative landing site for verb movement.

Applying Pollock's basic methodology, it would appear that, in all the Romance NSLs, the finite verb raises to a position in IP. However, given a degree of language-internal variation, one can maintain the idea that adverbials are only adjoined to maximal projections (as A&A crucially do), as long as (i) the IP is split into two functional heads, and (ii) the same adverbial can be adjoined in more than one position, depending upon semantic interpretation (following Svenonius 2002b).³

2.3.1.1.1. Italian

Belletti (1990) argues that emphatic elements such as *mai* (never) in NSLs, occupy the same position as *pas* in French, (adjoined to TP in our terms). The negative particle *non*, on the other hand, is head adjoined to the finite verb according to her analysis.⁴ Consider the following examples:

- (12) Maria non è **mai** uscita (Italian)
Maria not is never left
'Maria has never left.'

³ The assumption that adverbials are adjoined is by no means a given. Here, I assume that adverbials are adjuncts to maximal phrases, as A&A's argument crucially relies on this assumption. A similar stance is also argued for in Svenonius (2002) and Ernst (2002). Note, however, that Cinque (1999, 2004) presents an alternative view whereby adverbials are merged as the agreeing specifiers of functional heads. In his terms, the necessity for adjacency between the subject and finite verb is wholly dependent upon the position to which the verb (optionally) raises, as the inflectional field is made up of numerous functional heads which optionally house adverbials. As this analysis renders A&A's argument immediately invalid, I put it to one side here for the sake of exposition. Note, however, that there are strong arguments in favour of this position (cf. Cinque 2004). Note also that Cinque's analysis is perhaps closer to Pollock's own analysis than the approach taken here, which is very much a development of A&A's claim. However, as what is at stake here is ultimately the status of preverbal subjects, the question of optional adverbial placement/verb movement is, in a sense, extraneous. What is clear is that in all cases, the verb raises into the I-domain.

⁴ As such, *non* is syntactically similar to French *ne*, rather than to *pas*.

- (13) Maria parlava **pur/ ben/ già** /**sempre** di lui
 Maria spoke indeed /indeed /already /always of him
 ‘Maria indeed/already/always spoke of him.’
 [Belletti (1990:33-39)]

Given these facts, it would appear that finite verbs and auxiliaries raise to Agr (or whatever the highest functional head in IP is) in Italian, as was the case in French:

- (14) [_{AgrP} subject **verb** [_{TP} (mai/adv) T [_{vP} (adv) <subject> [_{VP...dO iO}]]]]⁵

For this reason, all adverbs *always* follow finite verbs and auxiliaries, as both maximal adjunction sites (TP and vP) are lower in the hierarchy.

2.3.1.1.2. Spanish

The status of Spanish (and EP) is less straightforward. Bok-Bennema (2001) argues, however, that the verb raises to the lower functional position in Spanish (our T). She arrives at this conclusion based on the following evidence. In Spanish, the finite verb must always precede manner adverbs, which according to Cinque (1999), occupy a low structural position:

- (15) *El chico *lógicamente* contestó la pregunta
 The boy *logically* answered the question
 ‘The boy answered the question logically’
- (16) El chico contestó *lógicamente* la pregunta
 The boy answered logically the question
 [Bok-Bennema to appear, citing Ojea López (1994:400)]

⁵ The bracketed structure assumes that adverbs are adjoined, however I remain agnostic as to whether this is really the case.

However, with other, higher, adverbials, word order is less rigid:

- (17) Juan **inmediatamente** cerró la puerta
 Juan immediately closed the door
 'Juan immediately closed the door.'
- (18) Juan cerró **inmediatamente** la puerta.
 Juan closed immediately the door
 'Juan immediately closed the door.'
- (19) Ese alumno **siempre/ya** asiste a mis clases
 That pupil always attends to my classes
 'That pupil always attends/is already attending my classes'
- (20) Ese alumno asiste **siempre/ya** a mis clases
 That pupil attends always to my classes
 'That pupil always attends/is now attending my classes'

[examples from Bok-Bennema, citing Ojea López 1994:409]

In examples (17)-(20), the verb can either precede or follow the adverb, triggering a subtle difference in meaning.⁶ Bok-Bennema assumes, in line with the spirit of an adjunction approach, that this variability is due to optionality in the placement of these adverbials, rather than optional verb movement. In Spanish, then, the verb always raises to the low inflectional head, meaning that adverbs can be merged before or after that head to TP or vP:

- (21) [_{AgRP} subject Agr [_{TP} (adv) **aux/verb** [_{vP} (adv) <subject> [_{VP}V]]]]

Crucially, in English adverbs can *never* surface between the finite verb and the direct object:

⁶ Bok-Bennema does not discuss this difference in meaning.

- (22) *John closed immediately the door.
- (23) *That pupil attends always my classes.
- (24) *That pupil attends already/now my classes.

This is captured if English also displays optionality in adverb placement but does not allow verb raising to T. It is further confirmed by the behaviour of finite auxiliaries, which seem to raise to T, the lower functional head, rather than the higher one, as Pollock suggests (cf. Kayne 1989):

- (25) John (??immediately) had (immediately) closed the door.
- (26) That pupil (always) has (always) attended my classes.
- (27) That pupil (already) has (already) attended my classes.

In these instances, too, the finite auxiliary has raised to T and as such adjunction sites are available preceding and following the verb (TP and vP).

2.3.1.1.3. European Portuguese

In EP, as is the case in Spanish, many adverbs can optionally precede or follow the finite verb (Costa 1996, 2004). Costa and Galves (2002) analyse these facts in a similar way to Bok-Bennema, positing a short verb movement to T:

- (28) O Pedro (provavelmente) viu (provavelmente) a Maria.
 The José probably saw probably the Maria
 'José probably saw Mary.'

Negrão et al. (2003) show that adverbs which follow the finite verb have a different interpretation to those that precede the finite verb in EP, suggesting that these different positions correlate to different semantics as Svenonius (2002) suggests (cf. Costa 1996, 2004 for discussion):

- (29) O João **sempre** foi a Paris de comboio.

(33) *My friends like **all** Mary.

(34) My friends have **all** left.

In summary, there appears to be good evidence in Spanish, Italian and EP that the finite verb raises out of vP to some head in the I-domain. In EP and Spanish this movement targets a low position (T), whereas in Italian (and French) it targets a high position (Agr). This has the direct implication that, where a subject appears preverbally, it must have moved out of vP in Romance NSLs. The possibility remains open, however, that post-verbal subjects might occupy the vP-internal position.⁹

2.3.1.2. Objections to the split IP

It is necessary to mention here, that, while the idea of split functional fields (cf. Rizzi 1997, Cinque 1999) is not overly controversial in mainstream Minimalism, the notion of an *Agr* head is more so. Many scholars have taken up the idea of a split IP, arguing that it is necessary in order to account for the wide range of empirical variation seen in natural languages (cf. Rizzi 1987, Cardinaletti 1997a, Poletto 2000, Ledgeway and Lombardi 2005 among others). In a number of Italian dialects, subject clitics are thought to instantiate the head *Agr* (cf. Rizzi 1986b, Brandi and Cordin 1989, Poletto 2000).

Ledgeway and Lombardi (2005) give evidence from Standard Italian and Southern Italian dialects to show that a more intricate split IP is required in order to capture differences in possible word orders between adverbs clitics and verbs.

⁹ Ian Roberts (p.c.) points out an alternative interpretation of these facts: verbs might not actually raise into the I-domain in Ibero-Romance, but may only raise to little 'v', assuming that adverbs mark the edge of (big) VP. In this case, verbs in English might not move at all, hence the obligatory adjacency between verb and object. This analysis is appealing but makes it difficult to account for the word order facts on Spanish and EP, as we shall see in chapter 3. For this reason, I assume for now that verb movement in Ibero-Romance targets the I-domain, leaving this alternative analysis open to future research.

Cardinaletti (1997a:47) likewise, argues that overt preverbal subjects in Romance NSLs occupy a higher position in a split AGR system.

However, there have also been serious theoretical objections to the existence of an Agr head. After proposing Agr_S and Agr_O in Chomsky 1991, Chomsky (1995) argues that both agreement heads are, in fact, redundant if held up to Minimalist scrutiny. The argument is based on the following fundamental Minimalist assumption:

“...the principles of UG involve only elements that can function at the interface levels; nothing else can be “seen” in the course of the computation...”

[Chomsky (1995:225)]

This in turn means that:

“Postulation of a functional category has to be justified, either by output conditions (phonetic and semantic interpretation) or by theory-internal arguments. It bears a burden of proof, which is often not so easy to meet.”

[Chomsky (1995:240)]

Chomsky claims that the other functional heads he has assumed (e.g. T, C and D,) arguably have “semantic properties” (Chomsky 1995:240), whereas Agr does not. Given this, his main objection to the head Agr is that:

“[T, C, D] have Interpretable features, providing “instructions” at either or both interface levels. Agr does not; it consists of -Interpretable features only,”

[Chomsky (1995:349)]

It is instinctively true that T, C and D differ from Agr in that they encode semantic information regarding tense, clause type and definiteness/referentiality. Agr arguably lacks this semantic encoding. However, the conclusion that Agr has no PF representation is motivated purely by theory-internal assumptions, based on the postulation that verbs enter the derivation fully inflected, a position which is later abandoned by Chomsky himself.

Chomsky (1995:238) points out that there are at least three possible ways for verbs to get inflected, given Minimalist assumption:

- i. Tense and phi-features might be assigned to the verb as it enters the numeration, as with case and phi-features on nouns.
- ii. Alternatively the verb might acquire these features by raising to T or Agr.
- iii. A third option discussed by Chomsky is that the verb might reach the phonological component uninflected, “the PF form resulting from interaction with functional elements within the phonological component”(Chomsky 1995:238).

He then goes on to assume that (i) is the correct analysis, leaving open the possibility that languages might differ with respect to the choice of (i), (ii) or (iii). Crucially, the choice of (i) means that the verbs enter the derivation fully inflected. The role of Agr_s then is thus greatly diminished. Under these assumptions, Agr_s serves only to force a subject position above spec TP:

“Agr is nothing more than an indication of a position that must be occupied at once by overt operations.”

[Chomsky (1995:351)]

If, rather, (iii) is true and Agr encodes PF-visible information, which is responsible for the insertion of agreement morphology in the phonological component, then this objection is weakened.¹⁰ There are, after all, many LF-uninterpretable lexical items in natural languages which have only a PF form, but which are crucially responsible for case assignment. Many prepositions arguably have a similar status:

(35) The destruction **of** the city **by** the barbarians.

¹⁰ This is true at least in languages with overt agreement morphology. It is not as easy to maintain for languages which lack agreement (e.g. Chinese). See Speas (2006) for a discussion.

In many languages, it seems that there is empirical reason to believe that IP must be split into at least two heads. The labels which these heads should have is, in a sense immaterial. I will adopt the labels Agr and T for reasons of convention. Where the distinction is not required I will simply use the generic term spec IP to refer to ‘the’ preverbal subject position.¹¹

2.3.2. Clitic Left Dislocation (CLLD)

In order to assess whether preverbal subjects display the properties of CLLD in Romance NSLs, it is first necessary to clarify what these properties are. CLLD refers to the process whereby XPs which function as topics, denoting ‘old information’, are base-generated in the following configuration:¹²

(36) XP_1 [...CL₁+V...]

I will abstract away from the exact peripheral location of this XP, remaining agnostic as to whether it occupies a position in CP (cf. Rizzi 1997) or is adjoined to IP (cf. Barbosa 1995).

Object CLLD, attested in Italian, Spanish, French and European Portuguese (EP), has specific structural properties, as first discussed by Cinque (1990). Firstly, it is worth remarking that, unlike clitic doubling, CLLD is not subject to ‘Kayne’s Generalisation’, whereby doubled elements must be case marked by a preposition (Kayne 1975).¹³ In this

¹¹ Chomsky himself acknowledges that there are potential empirical reasons to posit Agr_s, notably in Icelandic, which instantiates Transitive Expletive Constructions (TEC) (Chomsky 1995:349). However, the fact that German permits TECs in matrix clauses only suggests that they might involve the C system.

¹² The clitic can be proclitic or enclitic depending on language specific requirements.

¹³ Spanish and Romanian allow clitic doubling structures (CLD) where the object remains in situ and is doubled by a preverbal object clitic (Dobrovie-Sorin 1990):

(9) Lo hemos visto a Juan
 CL-ACC have:1pl seen A Juan
 ‘We have seen Juan.’

much it is *not* predicted that preverbal subjects will be limited to languages licensing clitic doubling (CLD) (e.g. Spanish and Romanian, and some Southern Italian dialects). Furthermore, the fact, as noted by Cinque (1990:58), that “CLLD can occur at the front of virtually any subordinate clause” (correctly) admits the possibility of preverbal subjects

- (10) I-am văzut pe Ion
 CL-ACC-have:1pl seen A Ion
 ‘We have seen Ion.’

This kind of structure is ungrammatical in Standard Italian, French and European Portuguese:

- (11) *L’ abbiamo visto Gianni
 CL-ACC have:1pl seen Gianni
 (12) *Nous l’avons vu Jean
 CL-ACC have:1pl seen Jean
 (13) *Temo-lo visto o João
 have:1pl-CL-ACC seen the João

Kayne (1975, and subsequent work) has attributed this property to the availability of the personal marker (*a/pe*) which is able to assign Case to the direct object in Spanish and Romanian respectively. Italian, French and Portuguese all lack such a marker and hence disallow CLD. There are many problems with Kayne’s generalisation. Firstly as Dobrovie-Sorin (1990) points out, this analysis does not really explain why CLD with indirect objects introduced by a preposition (which is presumably equally able to assign Case) is also ungrammatical in standard Italian, French and Portuguese, though not in some southern Italian dialects. Compare Spanish (14) where the indirect object is doubled by a dative clitic and the ungrammatical equivalents in (15)-(17):

- (14) Le di un libro a Juan [Spanish]
 CL-DAT gave:1s a book to Juan
 (15) *Gli ho dato un libro a Gianni [Italian]
 CL-DAT have:1s given a book to Gianni
 (16) *Je lui ai donné un livre à Jean [French]
 I CL-DAT have given a book to Jean
 (17) *Dei-lhe um livro ao João [EP]
 Gave-CL-DAT a book to João
 ‘I gave a book to John.’

Whatever it is which permits CLD in Spanish and Romanian, it seems to be separate from the NSP. A&A’s assertion that post-verbal subjects behave like clitic doubled subjects is therefore not unproblematic, as clitic doubling is not a property of all Romance NSLs.

in both main and embedded clauses. Finally, it seems that any “maximal phrase” (PP, AP or DP) can be CLLD (Cinque 1990:57) as long as the corresponding clitic exists in a given language.¹⁴ Note that the corresponding clitic for subject CLLD must be considered to be a combination of agreement morphology + PRO/*pro* in Barbosa’s terms and the weaker version of A&A’s analysis respectively, as both analyses maintain traditional theta-theory. In the stronger version of A&A’s hypothesis, as discussed in chapter 1, the agreement morphology itself acts as the corresponding clitic.

2.3.2.1. A base-generated dependency?

A number of the properties of CLLD led Cinque (1990) to assume that it is a base generated dependency, despite the fact that, on the surface, it appears that CLLD displays some of the properties associated with movement. For instance, Cinque (1990:59) observes that “there is obligatory connectivity between the [CLLD] phrase and the IP-internal position [in null subject languages- MS]”. This means that binding conditions appear to apply to the position of the clitic rather than that of the fronted XP:

- (37) A se stessa_i/*lei_i Maria_i non ci pensa. [Italian]
 about herself/her Maria not CL-there thinks
 ‘Maria is not thinking about herself.’

Moreover, CLLD is sensitive to strong, but not weak, island constraints (Cinque 1990:59):¹⁵¹⁶

¹⁴ Spanish, unlike French and Italian, does not have locative (*y/ci*) or partitive (*ne/en*) clitics and so is more limited in the kind of CLLD structures it can form. EP has a clitic-like locative ‘*la*’ but no partitive. Spoken BP, which has lost clitics, displays only LD similar to that seen in English, with full resumptive pronouns (cf. Raposo & Kato 2005).

¹⁵ Or at least CLLD with non-DPs is (Raposo 1996):

- (18) O capitão_i, encontrei [um negociante que lhe_i telefonou ontem]. [EP]
 the captain, (I) met the businessman that called 3pcl yesterday
- (19) *Ao capitão_i, encontrei um negociante que lhe_i telefonou ontem. [EP]
 To-the captain, (I) met the businessman who called 3pcl yesterday
 ‘The captain, (I) found [the businessman that called him yesterday].’

- (38) * A casa, lo abbiamo incontrato [prima che ci andasse]
at home cl-him have:1pl found before that cl-there went:3s
Lit. 'At home, we found him before he went there.'

However, CLLD also displays additional properties which would be unexpected were it to be derived via movement. Cinque (1990:66) notes that CLLD elements do not reconstruct, and as such always take wide scope over IP internal elements:

- (39) kapjo pedito to eksetase kathe kathigitis [Greek]
Some child CL-ACC examined every professor
'There is some child such that he was examined by every professor.'
[A&A (1998:505)]

Moreover, CLLD does not license parasitic gaps in Italian, EP, Spanish or French (cf. Cinque 1990:60 for Italian; Kato and Raposo 1994, Kato and Raposo 2005 for EP):

- (40) *Esses documentos, eu queimei-os depois de ler EC
Those documents, I burnt-CL-THEM after of to-read
Lit. 'Those documents, I burnt them after reading.'
- (41) *Ces documents, je les ai rangés, après avoir lus EC
Those documents, I CL-THEM have put away after of to-read
Lit. 'Those documents, I burnt them after reading.'

Cinque assumes that islands and binding are really restrictions on representations, concluding that CLLD is a base-generated dependency.¹⁷ This idea is picked up by A&A

¹⁶ Focus movement, on the other hand, is subject to both strong and weak islands.

¹⁷ Subsequently, this has been questioned by many. Raposo (1996) proposes that CLLD involves covert movement, and thus disallows PGs while being subject to island constraints. Belletti (2005), following Cecchetto (1999, 2000) proposes that in all doubling structures the clitics and the doubled XP are merged as a single constituent in argument position and are separated during the

as it provides a potential analysis of preverbal subjects in NSLs. If the verbal morphology + *pro* has the status of a (clitic) pronoun, then preverbal subjects might be base-generated in a peripheral position, either adjoined to IP or in the CP-layer, in the manner in which CLLD objects are.¹⁸

2.3.3. Relevant properties of CLLD

In order to use the properties of CLLD as a diagnostic for the status of preverbal subjects in Romance, we must first establish the properties of the former. The following sections describe the remaining properties of CLLD which will prove relevant in an empirical assessment of A&A's proposal.

2.3.3.1. Melodic pause

According to Raposo (1996:1), and others, the element doubled by a clitic appears at the left of the clause, forming "a clearly distinct melodic phrase, set off from the rest of the clause by a pause (represented by the comma)":

- (42) [DP Cet homme], je ne le connais pas [French]
This man I NEG CL-him know NEG
Lit. 'This man, I don't know him.'

This claim has been challenged by Solà (1992:268), who claims that "any clitic left-dislocated element can be pronounced without any special pause or phonological clue possibly differentiating it from what would be a "true" non-dislocated subject". It might be the case, then, that this pause is optional rather than obligatory, and therefore not a

course of the derivation. I will ignore these alternative analyses here as it is not clear how they fit in with A&A's analysis of null subjects.

¹⁸ If *pro* is the doubled element, as is assumed by A&A (in the weaker version of their theory) then the relationship between the peripheral subject and *pro* is really that of left dislocation (LD). This has slightly different properties as also discussed by Cinque. However, as the main objections raised to a CLLD account also apply to a LD account I will not dwell on this distinction here.

useful heuristic for the assessment of the status of preverbal subjects. I therefore omit this property from the discussion, pending further phonetic investigation.

2.3.3.2. Failure to reconstruct

Recall that one of the reasons that Cinque claims CLLD is a base-generated dependency is because CLLD elements fail to reconstruct in terms of scope (see example (39)). This makes the prediction that all preverbal subjects will have obligatorily wide scope in Romance NSLs. Again, this prediction is relatively easy to test.

2.3.3.3. No non-referential CLLD

Raposo (1986) and Kato and Raposo (2005) have confirmed Cinque's (1990) claim that CLLD with a resumptive clitic is not possible with robustly non-referential QPs or bare NPs in Italian, Spanish, EP or French. Examples here are taken from Raposo (1986):

(43) *Livros do Tintim, li-os antes de adormecer. [EP]
Books of-the Tintin read:1S-CL-THEM before of sleep:INF-1S
Lit. 'Books about Tintin, I read them before going to sleep.'

(44) *Poucos candidatos, ouvi-os falar na televisão. [EP]
Few candidates heard-them speak on-the television
'Few candidates, I heard them speak on the television.'

This is also the case with negative QPs in Standard French, Spanish, Italian, Catalan and EP:

(45) *Personne, je ne l' ai vu(e) [French]
Nobody I neg CL-ACC have seen
'I haven't seen anybody.'

Rizzi (1986b) provides a (stipulative) reason for this ungrammaticality, which appears to tie it in with the observations of Montalbetti (1984, 1986):

(46) A pronoun cannot be locally bound by a (non lexically restricted) quantifier.

Thus for some reason, perhaps a universal principle, pronouns cannot be locally bound by QPs in NSLs. CLLD creates an environment where a QP is first-merged in an A-bar position and so cannot leave behind a variable if quantifier raised. The basic intuition, which seems to be empirically accurate, according to Montalbetti's findings, is that "pronouns cannot function as primary variables, and can acquire variable status only parasitically, through binding from licit primary variables" Rizzi (1986b:395). I return to these facts below. If it is the case that all preverbal subjects are CLLD, then the ban on non-referential CLLD makes clear predictions about the behaviour of non-referential subjects. In A&A's terms, there should be no preverbal QP subjects in Romance NSLs.¹⁹

2.3.3.4. Recursivity

Cinque (1990:58) claims that there is "no theoretical limit" to the number of elements that can be CLLD.

(47) Di vestiti, a me, Gianni, in quel negozio, non mi ce ne
Some clothes, to me, Gianni, in this shop, not CL-ME CL-HERE CL-OF
ha mai comprati.
has never bought
Lit. 'Clothes, for me, Gianni, in this shop, he's never bought them here.'

This means that the presence of another CLLD element should not block the presence of a preverbal subject. Likewise if the only preverbal subjects are CLLD, then it is predicted that there should be no fixed subject position with respect to other XPs in the clause.

2.3.4. Summary

¹⁹ Barbosa's analysis, on the other hand, allows for the possibility that some preverbal subjects are A-bar moved to a preverbal position, making slightly different predictions. I return to this below.

CLLD elements display a coherent set of syntactic and semantic properties. For the purposes of this thesis, I will not propose an analysis of CLLD, but, rather, will accept Cinque's (1990) claim that it is a base-generated dependency. Rather, in the following sections I will be concerned with the accepted properties of CLLD, which I will use as diagnostics in an assessment of the status of preverbal subjects in Romance NSLs. If, as predicted by A&A, *all* preverbal subjects in NSLs are CLLD then *all* these preverbal subjects should display unambiguous wide scope, free positioning with respect to other fronted elements and a restriction to referential XPs. In what follows, it will be shown that none of these predictions hold. Moreover, it will be argued that Barbosa's (1995) weaker claim that some preverbal subjects are A-bar moved, and some CLLD, is also difficult to maintain.

2.4. Evidence that preverbal subjects are CLLD

Given that Romance NSLs are usually considered canonical SVO languages (Hulk and Pollock 2001:3), and that preverbal subjects in Romance NSLs are not separated from their predicate by a pause or melodic break, the burden of proof certainly lies with those claiming that *all* preverbal subjects are CLLD. We shall see shortly that this appears to be the claim made by A&A (1998), though it is not that made by Barbosa (1995). A&A, drawing on work by Barbosa (1995), give the following pieces of evidence that preverbal subjects in Romance are always CLLD:

- *An adverb can intervene between the subject and the verb;*
- *QPs and indefinites in preverbal subject position have unambiguous wide scope;*
- *In Catalan, post-verbal overt pronominal subjects can be bound, unlike preverbal pronouns.*

Barbosa (1995) gives additional evidence that referential subjects are CLLD from:

- *Italian dialect subject clitics.*
- *Indefinites*
- *Enclisis patterns in EP*

In the following sections, I consider each of these arguments on its individual merits in order to assess the validity of the claim that *all* preverbal subjects are CLLD. Note that under a different analysis it might be the case that many preverbal subjects are CLLD, though it need not be that case that all are.

2.4.1. Adverb placement

The first empirical argument provided by A&A relates to the fact that certain adverbs can intervene between the subject and the finite verb. A&A show that, in Spanish and Greek, this is possible (48)&(49), whereas the same is not true in French (50):

(48) Juan ya se fue. [Spanish]
Juan already self went
'Juan has already left.'

(49) O Petro xtes meta apo poles prospathies sinandisa ti Maria
Petro yesterday after from many efforts met Maria
Lit. 'Petro, yesterday, after a great deal of effort, met Maria.' [Greek]

(50) Jean *(déjà) veut (déjà) s'en aller [French]
Jean already wants already self-of go
'Jean already wants to leave.'

A&A assume that adverbs can only be adjoined to the maximal level. Given this fact, the grammaticality facts in (50) are explained if the subject is the specifier of the head housing the inflected verb, as no potential adjunction site intervenes between a head and its specifier. In the same way, the grammaticality of (48)&(49) is unexpected if they share the same structure, but easily explained if the preverbal subjects in (48)&(49) occupy an external CLLD position. A&A claim that this distinction is due to the status of Spanish and Greek as NSLs, and of French as a non-NSL.

2.4.1.1. The problem of Italian

This is, indeed, an interesting contrast between French and Spanish. However, the data pattern different for the other Romance NSLs. As we saw briefly in section 2.3.1, in Standard Italian, ‘positive adverbs’ of the ‘already’ type always occur after the verb, as is the case in French.

- (51) Maria *(già) parlava già di lui
Maria already was-speaking already of him
‘Maria was already speaking about him.’

For this reason, I argued that in both French and Italian the finite verb raises to Agr.

Sentential adverbs, on the other hand, can surface between the subject and the finite verb in Italian (Belletti 1990:39-42):²⁰

- (52) Gianni probabilmente telefonerà alle 5
Gianni probably will-phone at-the 5
‘Gianni will probably phone at 5 o’clock.’

Belletti (1990:43) analyses (52) as involving a left peripheral subject and a high sentential adverb. However, unlike A&A, she claims that this is a “derived structure”. In other words, Belletti makes the point that in Italian, subjects can appear *either* in the canonical preverbal subject position, *or* in a topicalised, left peripheral, position (which we can identify with Cinque’s CLLD position). This is not to say that subjects need *always* occupy a left peripheral position. Evidence that this might, indeed, be the case comes from the behaviour of non-referential subjects, which cannot be CLLD:

- (53) *Nessuno probabilmente telefonerà alle 5
Nobody probably will-phone at-the 5
‘Nobody will probably phone at 5 o’clock.’

²⁰ Note that if we were to adopt Cinque’s (1999) universal hierarchy of adverbs then several possible subject positions would necessarily become available.

- (54) NESSUNO probabilmente telefonerà alle 5
Nobody probably will-phone at-the 5
'NOBODY will probably phone at 5 o'clock.'

Belletti shows that 'nobody' cannot precede the sentential adverb unless it bears contrastive focus. In Cinque's (1990:61) terms, contrastive focus is the result of focus-movement to an A-bar focus position in the C-layer.²¹ It seems that *nessuno* cannot occur in the preverbal subject position before the adverb *probabilmente*. However, note that 'nobody' can appear preverbally without contrastive focus, where no adverb is present:

- (55) Nessuno telefonerà alle 5
Nobody will-phone at-the 5
Nobody will phone at 5 o'clock.'

This suggests that the subject in (55) occupies the canonical subject position, whereas that in (54) and (52) does not:

- (56) [IP Nessuno telefonerà [alle 5]]
(57) [CP NESSUNO[IP Gianni [IP probabilmente [IP *pro* telefonerà [alle 5]]]]

The fact that Italian, an NSL, patterns with French, rather than Spanish and Greek is problematic for A&A's argumentation. However, once we assume that verb movement targets a higher functional head in French and Italian, for whatever reason, then the data are unproblematic.

²¹ Cinque (1990) gives convincing evidence to differentiate focus-movement from base-generated CLLD, showing that the former but not the latter licenses parasitic gaps, and is sensitive to weak islands.

2.4.1.2. Spanish, European Portuguese and English

The fact that many adverbs can intervene between any kind of subject and the finite verb in Spanish and European Portuguese can be attributed to the fact that verb movement targets a lower position in the I-domain, as discussed in section 2.3.1, rather than to their status as NSLs:

(58) O Pedro/ninguém provavelmente viu a Maria. [EP]
The Pedro/nobody probably saw the Maria
'Pedro/nobody probably saw Maria.'

(59) Pedro/alguien ya ha visto a Maria. [Spanish]
Pedro/someone already has seen A Maria
'Pedro/someone has already seen Maria.'

It would seem that, while these facts do fall out from an analysis based on the idea of a split IP, they do not fall out from an account like that proposed by A&A. In fact, A&A essentially predict (58)-(59) to be ungrammatical, contrary to fact, as they involve non-referential QPs, which cannot be CLLD.

A&A themselves note that their assumptions are problematic given the data in some languages. For example English, as they note in footnote 12, also allows certain adverbs to intervene between the subject and the auxiliary element commonly held to be in I, as first discussed by Pollock (1989):

(60) [_{TP} Eric probably/really/never [_T has met Mary]

(61) There probably won't have been a riot.

(62) Nobody ever will complete the task.

Examples (60)-(62) seem to be fine with a certain intonation, where the auxiliary is stressed. A&A argue that the adverbs in (60)-(62) could be parenthetical, however, this does not appear to be the case, at least in (61). On the other hand, the data fall out from

an analysis in which finite auxiliaries in English raise to the low T position, as proposed in section 2.3.1.1.2.

In conclusion, the adverb placement facts seem to have more to do with differences in the target of verb raising than with the NSL vs. non-NSL opposition. Adverbs can be adjoined/merged between Agr and T and between T and v. In Italian and French the finite verb raises to Agr, and thus nothing can intervene between the subject and adverb, except where subjects are dislocated. In Spanish and EP, on the other hand, the verb targets T and so adverbs can precede or follow the verb. In English, auxiliaries also target the low T position.

2.4.2. Failure to reconstruct?

The second piece of evidence given by A&A in support of their claim, (this time drawing on an idea proposed by Barbosa 1995), comes from the scope of preverbal QPs and indefinites. Recall Cinque's observation that CLLD objects receive an obligatorily wide scope reading, failing to reconstruct at LF. A&A claim that in Greek and Spanish, preverbal QPs also always take wide scope over other quantifiers in the clause, suggesting that they are also CLLD. In post-verbal position, on the other hand, A&A claim that the subject can have either narrow or wide scope. They give the following examples from Greek:

(63) Kapios fititis sitihiothetise kathe arthro [wide scope only]
 Some student filed every article
 'Some (particular) student filed every single article.'

(64) Sitihiothetise kapios fititis kathe arthro [ambiguous]
 Filed some student every article
 a. 'Every article was filed by some student (or other).'
 b. 'Some (particular) student filed every single article.'

(65) Kapjo pedi to eksetase kathe kathigitis [wide scope only]
 Some child cl-ACC examined every professor
 'Some child is such that every professor examined that child.'

A&A argue that the effect observed in Greek is also true of Spanish. It is likely that this claim is based on evidence from the quantifier ‘todo’, which seems to prefer narrow-scope:

- (66) Algún estudiante sacó prestado todos los libros
Some student took borrowed all the books
a. ??‘All the books were borrowed by some student.’ [narrow scope subject]
b. ‘Some (specific) student borrowed all the books.’ [wide scope subject]

A weaker effect of this kind can be observed with English ‘all’, which strongly prefers narrow scope:

- (67) A student borrowed all the books [narrow scope of ‘all’ preferred]
(68) A student borrowed every book [ambiguous]

In fact, the Spanish facts are quite complex, and judgements seem to be unstable. Suñer (2002), for example, refutes A&A’s claim for Spanish, claiming that preverbal subjects are ambiguous in Spanish, just as they are in English:

- (69) Algún estudiante sacó prestado cada libro
Some student took lent each book
a. ‘Each book was borrowed by some student (or other).’ [narrow scope subject]
b. ‘Some (particular) student borrowed each book.’ [wide scope subject]

Suñer claims that the more natural reading for (69) is (a) where the preverbal QP has narrow scope in relation to the object QP. Suñer does note, however, that both readings are available. Again, this might be dependent on the choice of quantifiers in the sentence. In English too, the quantifier ‘each’ prefers a distributive reading, which

appears to correlate with narrow scope of the subject (reading (a)) in a pragmatically neutral context:²²

- (70) A student borrowed each book
a. 'Each book was borrowed by some student (or other).'
- b. '?Some (particular) student borrowed each book.'²³

Likewise in EP, it is possible to construct examples in which the preverbal subject has ambiguous scope, though in many cases wide scope is preferred for preverbal QP subjects:²⁴

- (71) Nenum estudante leu os patru livrus
-

²² Note, moreover, that Dobrovnie-Sorin (1990) points out that, in Romanian, preverbal subjects are uniformly ambiguous between wide and narrow scope readings:

- (20) Fiecare profesor va examina zece elevi
Each teacher will examine ten students
a. Each teacher will examine ten (possibly different) students.
b. There are ten students x, such that each teacher will examine x.

- (21) Toți judecătorii cunosc doi avocați
All judges know two lawyers
a. All the judges know (at least) two (possibly different) lawyers.
b. There are two lawyers x, such that all the judges know x

- (22) Toți copiii au văzut un actor celebru.
All children have seen an actor famous
a. For every child x, x knows at least one/a famous actor
b. There is one famous actor x, such that every child knows x

²³ Note however, the discussion of wide scope indefinites in section 2.4.3.

²⁴ Costa (2004:23) calls into question the very validity of the assumption that CLLD elements should fail to reconstruct, whereas A-moved subjects should not:

“As it is well-known, A-bar movement obligatorily reconstructs, unlike A-movement. If it is assumed that the preverbal subject is in Spec,IP, it is expected that there is no reconstruction of the subject, while a left-dislocated complement may reconstruct.”

This just goes to emphasize how shaky scope tests are in relation to the status of preverbal subjects. However, it is the case that, in English, A-moved subjects have ambiguous scope.

No student read the four books

- a. 'There is no student x such that x read the four books.'
- b. 'There are four books x such that no student read x.'

The same also appears to be true of preverbal subject QPs in Italian:

- (72) Qualcuno studente ha letto tutti i libri
Some student has read all the books
- a. 'All the books were borrowed by some student (or other).'
 - b. Some (particular) student borrowed all the books.'

It seems that outside Greek, there is little empirical support for A&A's claim that all preverbal QPs get a wide scope reading, and hence behave like CLLD elements.²⁵ At least in Spanish, Italian, EP and Romanian, the preverbal subject position is often potentially ambiguous in scope. In fact, scope appears to be dependent upon pragmatic/semantic factors relating to the specific quantifiers used.²⁶ The behaviour of indefinites, which requires further clarification, is discussed in the next section.²⁷

2.4.3. Indefinites: non-referential preverbal subjects

Barbosa (1995:36-37) gives evidence from EP, to suggest that subject indefinites always take wide scope over modals. She describes a context where someone is "being reminded of the items that [it is] necessary to submit in order to apply to a certain job".

²⁵ Even in Greek, the judgements reported by A&A have been disputed by Roussou & Tsimpli (2006).

²⁶ Consider the following sentences discussed by Rizzi (1982:123), taken from Jackendoff (1972):

(23) Many arrows didn't hit the target. Many > neg *neg>many

(24) The target wasn't hit by many arrows. Many > neg neg>many

Whereas (24) seems to be ambiguous as regards the scope of the two quantifiers, (23) can only receive the interpretation whereby *many arrows* takes wide scope over negation.

²⁷ One potential problem here is that, according to Barbosa (1995), robustly non-referential QPs should behave differently from other QPs. This is because the former are fronted via A-bar movement, whereas the latter are CLLD. I return to this point below.

In such a context, a preverbal indefinite subject is fine in English and French, but not in Catalan, Spanish, Italian or EP:

- (73) A letter of recommendation is required.
- (74) Une lettre de recommandation est requise. [French]
a letter of recommendation is required
'A letter of recommendation is required.'
- (75) #Una carta de recomanació és necessària. [Catalan]
a letter of recommendation is required
- (76) És necessària una carta de recomanació.
is required a letter of recommendation
'A letter of recommendation is required.'
- (77) #Una carta de recomendación es necesaria. [Spanish]
a letter of recommendation is required
- (78) Es necesaria una carta de recomendación.
is required a letter of recommendation
'A letter of recommendation is required.'
- (79) #Una lettera di raccomandazione è necessaria. [Italian]
a letter of recommendation is required
- (80) È necessaria una lettera di raccomandazione.
is required a letter of recommendation
'A letter of recommendation is required.'
- (81) # Uma carta de recomendação é necessária. [EP]
a letter of recommendation is required
- (82) É necessária uma carta de recomendação

is required a letter of recommendation

'A letter of recommendation is required.'

[all examples from Barbosa (1995:36-37)]

The same facts hold with respect to negation. In the same context, the following examples from French and Spanish illustrate the situation in non-NSLs vs. NSLs:

(83) Une lettre de recommandation n'est pas nécessaire. [French]
A letter of recommendation not is not necessary
'A letter of recommendation is not necessary.'

(84) # Una carta de recomendación no es necesaria. [Spanish]
A letter of recommendation not is not necessary

(85) No es necesaria una carta de recomendación.
Not is not necessary a letter of recommendation

As Barbosa points out, these effects fall out from an account which argues that all preverbal subjects are CLLD, on a par with that discussed in the previous section.²⁸ An account which argues that these preverbal subjects are A-moved requires additional machinery to account for the English/French vs. NSL distinction. As it stands, this is indeed a strong argument in favour of A&A's position.

However, matters are slightly more complex than they seem. Although Barbosa, like A&A, phrases her argument in terms of scope, there is reason to believe that this is not what is at stake in the above examples. In fact, Ruys (1992) has disputed the claim that indefinites can actually take *only* wide scope over other quantifiers. Consider the following example:

²⁸ Interestingly, this is not what Barbosa herself actually proposes. Rather she claims that preverbal subjects can either be CLLD or A-bar moved. As A-bar movement is usually assumed to permit reconstruction, it becomes unclear that this result is actually predicted by Barbosa's own proposal. However, A&A's more restrictive account certainly predicts these effects.

- (86) A student ate every cake.
a. All of the cakes were eaten by some student (or other).[narrow scope subject]
b. Some (specific) student ate all the cakes. [wide scope subject]

In actual fact, (86)(b) is just one possible way of satisfying the truth conditions of (86)(a). If one specific person ate all the cakes then it is also true that all of the cakes were eaten by someone. In this sense (86)(b) entails (86)(a), and as such cannot be seen as a distinct 'meaning' of the sentence in truth conditional terms. In fact Ruys argues that more generally, a wide scope reading of an indefinite entails a narrow scope reading. Thus, although a wide scope reading for 'somebody' might be pragmatically preferable, logically, it is indistinct from the narrow scope reading. The opposite is not true: (86)(a) does not entail (86)(b). It is not the case that if all the cakes were eaten by someone that the same specific person must have eaten them. In this sense a strict narrow scope reading is genuinely distinct from an ambiguous reading. Given this, the robust empirical contrast observed by Barbosa requires an alternative explanation. I argue that it can be attributed to a kind of definiteness effect.

2.4.3.1. Definiteness effects in Romance

2.4.3.1.1. Weak and strong readings in English

Milsark (1977) first pointed out that a distinction must be made between strong and weak quantifiers, independent of the definite/indefinite distinction. Only weak quantifiers are permitted in existential constructions in English:

- (87) There are some/two/many/no cats in the garden.
(88) *There is every/neither cat in the garden.
(89) *There are most/not all/both cats in the garden.

While some quantifiers (e.g. *every*, *neither*, *most*, *not all*, *both*) are unambiguously strong, others (e.g. *some*, *two*, *many*, *no*) are ambiguous between strong and weak readings. In example (87), as a result of the existential *there*, only a weak existential reading is available for these quantifiers:

- (90) There is an *x*, cat (*x*) & in the garden (*x*)
(91) There is an *x*, two (*x*) & cat (*x*) & in the garden (*x*)
(92) There is an *x*, many (*x*) & cat (*x*) & in the garden (*x*)

However, in examples like (93), these quantifiers are ambiguous between weak and strong readings:

- (93) Two cats are in the garden.
a. There is an *x*, two (*x*) & cat (*x*) & in the garden (*x*)
b. Two of the cats are in the garden.

While the weak reading for quantifiers is uniformly existential, the strong reading can either be partitive, specific or generic depending on the quantifier involved.²⁹ Quantifiers such as *many*, *some* and *two* receive a partitive interpretation on a strong reading (i.e. some of the *x*, many of the *x*, two of the *x*). Indefinites (e.g. *a cat*), on the other hand, get a specific reading and bare plurals get a generic reading.

Note that the strength of preverbal quantifiers also interacts with the status of the predicates with which they combine (cf. de Hoop 1992 building on work by Davidson 1967, and Carlson 1977). Individual-level predicates force a strong reading of a preverbal subject, whereas stage-level predicates permit either an ambiguous reading, or a weak reading of the subject. Compare the following as an illustration:

- (94) Two cats are black (individual-level = strong only)
(95) Two cats are purring (stage-level = ambiguous)

With bare plurals, on the other hand, stage-level predicates force a weak reading:

- (96) Cats are black (individual-level = strong only)
(97) Cats are purring (stage-level = weak only)
-

Thus in English, quantifiers in post-verbal position receive a weak reading, whereas in preverbal position, they are either ambiguous, weak, or strong, depending on the predicate and the internal properties of the quantifier.

2.4.3.1.2. Weak and strong readings in Romance

While it seems to be the case that non-focussed post-verbal subjects (where they are permitted) also receive a weak interpretation in Romance, the status of preverbal subjects is more complex.³⁰ For unaccusative verbs, with indefinites, it seems to be the case that preverbal subjects receive a necessarily strong reading, as reported by Pinto (1997:185) for Italian:³¹

(98) E' arrivato un linguista.
Is arrived a linguist
Lit. 'There arrived a linguist.'

(99) Un linguista è arrivato.
A linguist is arrived
'One (of the) linguist(s) arrived.'

Sheehan (2004) shows that this is also true of Spanish, suggesting it might be a property of Romance NSLs more generally.

I argue that the data from Barbosa (1995) given above is another example of this definiteness effect. Consider that Barbosa's examples become felicitous in EP, Spanish, Italian and Catalan where the preverbal indefinites get a partitive reading. Thus Barbosa purports that (100) becomes felicitous in the following context:

³⁰ See chapter 3 for a discussion of the distribution of post-verbal subjects in Romance.

³¹ Presentational verbs are those which allow either VS or SV word order in out-of-the-blue contexts. For a full discussion, see chapter 3.

Romance NSLs, as is the case in English. It is beyond the scope of this thesis to provide a full account of definiteness effects in Romance. For now, suffice it to say that, at best, the behaviour of indefinite subjects in Romance NSLs do not bear on A&A's analysis, and at worse, the former serves to refute the latter, illustrating that not all preverbal subjects can be CLLD. In chapter 3 I will show that some, though not all, of them fall out from an alternative theory of pro-drop, coupled with facts about case and the availability of expletive subjects.

2.4.4. Trentino and Fiorentino

Barbosa takes the clitic and agreement patterns of Fiorentino and Trentino as evidence that preverbal subjects are CLLD. These Italian dialects differ from Standard Italian in that they have subject clitics, and because post-verbal subjects trigger (i) only partial agreement on the verb and (ii) an expletive subject clitic:

(103) Le tu' sorelle l' hanno telefonato iere
The your sisters CL-Fempl have:3pl telephoned yesterday
'Your sisters telephoned yesterday.'

(104) L' hanno telefonato iere
CL-Fempl have:3pl telephoned yesterday
'They telephoned yesterday.'

(105) Gl' ha telefonato le tu' sorelle
CL-Exp has:3s telephoned the your sisters
'Your sisters telephoned.'

Barbosa notes the fact that preverbal subjects, like null subjects, and unlike post-verbal subjects, trigger strong agreement in Fiorentino and Trentino. Taking Fiorentino and Trentino to be like Irish, where overt subjects trigger no agreement and null subjects trigger rich agreement, Barbosa claims that these preverbal subjects are CLLD, co-occurring with a post-verbal *pro*.

Interestingly, Rizzi (1986b) cites the same data as evidence for the opposite claim. In his terms, post-verbal structures involve a null expletive in spec IP, whereas null subject utterances and preverbal subjects both involve a referential subject in spec IP. Referential subjects trigger full agreement, whereas expletive subjects trigger weak agreement, hence the fact that *pro* patterns with preverbal rather than post-verbal subjects. In fact, both of these interpretations of the data seem valid on their own terms. For this reason, this data cannot be taken as evidence in favour of either analysis.

Moreover, as Barbosa points out, the CLLD analysis of preverbal subjects rightly predicts that subject extraction will always take place from a post-verbal position. Again, the Fiorentino and Trentino data support this idea. The following is from Fiorentino but the same holds for Trentino:

(106) Chi gli è venuto con te? [Fiorentino]
 Who CL-Exp is come with you?
 Who has come with you?

(107) Quante ragazze gli è venuto con te?
 How-many girls CL-Exp is come with you
 'How many girls came with you?'

In (106)-(107) the subject clitic is an expletive and the verb agrees only in person with the extracted wh-subject. Barbosa, like Rizzi, (1986b) and Brandi and Cordín (1989) takes this as evidence that subjects are extracted from the post-verbal position.³³ Barbosa

³³ Rizzi (1982:124) shows that Italian quantifier raising is also subject to that-trace effects:

(25) Non pretendo che tu arresti nessuno
 Notrequire that you arrest nobody
 'There is no x such that I require that you arrest x.'

(26) Non pretendo che nessuno ti arresti
 Notrequire that nobody you arrest
 *'There is no X such that I require that x arrest you.'

(27) Non pretendo che sia arrestato nessuno.
 Notrequire that be arrested nobody

points out that Rizzi is forced to use further machinery in order to derive these facts. In Rizzi's terms the 'normal' A-position of overt subjects is preverbal and so additional machinery is needed to enable extraction from a post-verbal position (i.e. the merging of a null expletive in spec IP). Barbosa, on the other hand, derives this fact as a natural consequence of her analysis. The only A-position for subjects is the post-verbal one; therefore all extraction must proceed from this position. The fact that A&A and Barbosa's accounts seem to predict some of the correlations attributed to the NSP makes their approach highly appealing. However, ultimately, if their account fails to stand up to empirical scrutiny, it must be modified/rejected, despite its obvious appeal. In chapter 3 I will show that the word order facts across Romance do not support the idea that the post-verbal position is the only A-position in Romance NSLs. Thus, although it seems to be true that extraction proceeds from the post-verbal position, this post-verbal position need not be the only A-position. As long as another element serves to satisfy the EPP, then A-movement to spec IP, and that-trace effects, can be avoided (cf. Rizzi 1982, 2004).

2.4.5. Bound variable subjects

A related argument given by both Barbosa and A&A in favour of their approach comes from the behaviour of post-verbal subjects in Romance NSLs. Montalbetti (1984, 1986) showed that in Romance NSLs, preverbal overt subjects cannot be interpreted as bound variables unless linked to a bound pronominal variable (i.e. not a trace). The relevant data are as follows:

- (108) Muchos estudiantes piensan que ellos son inteligentes
- a. There is a group of many students x , such that x thinks x is intelligent
 - b. *Many (x) such that x is a student and x thinks x is intelligent

'There is no x such that I require that x be arrested.'

Under A&A's account these effects might be attributed to the fact that preverbal QPs are CLLD and thus have 'frozen scope'. In Barbosa's approach, on the other hand, where QPs are A-bar moved, it is not clear why this contrast should obtain.

(109) Muchos estudiantes piensan que *pro* son inteligentes

- c. *There is a group of many students *x*, such that *x* thinks *x* is intelligent
- d. Many (*x*) such that *x* is a student and *x* thinks *x* is intelligent

In A&A's terms this effect can be explained by the fact that all overt preverbal subjects are CLLD topics and, as such, cannot be bound, but only co-referential. Analyses assuming an active EPP need to provide an alternative explanation for these facts.

Moreover, A&A's analysis also predicts that the default A-position for subjects is the post-verbal one in Romance NSLs.³⁴ Interestingly, there is some evidence to suggest that overt subjects in this position can actually be bound. Barbosa cites data attributed to Solà (1992), who claims that, in Catalan, post-verbal, unlike pre-verbal, subject pronouns, can get a bound reading:

(110) Tots els estudiants₁ es pensen que ells₁ aprovaran.

all the students think that they passed

- a. All the students believe that they as a group will pass. (mutually encouraging)
- b. *For all the students it is true that *x* thinks *x* will pass. (egotistical)

(111) Tots els jugadors₁ están convençuts que guanyaran ells₁

all the players are persuaded that will-win they

- a. All the players believe that they as a team will win (football)
- b. For all the players it is true that *x* thinks *x* will win. (tennis)

Interestingly, these post-verbal overt subjects are ambiguous, receiving either a co-referential, or a bound reading. This contrast falls out if post-verbal overt pronominals are in an A-position, whereas preverbal overt pronominals are CLLD.

³⁴ In fact the distribution of post-verbal subjects in Romance is highly complex and in no way falls out directly from an analysis like that proposed by A&A. In chapter 3 I provide an empirical survey of possible word orders in the major Romance languages.

Interestingly, speakers differ as to their acceptance of the bound reading with post-verbal subjects. Judgments are rendered difficult by the subtle semantic distinctions required. A more robust contrast comes from the grammaticality of overt pronouns with robustly non-referential quantifiers. Thus Barbosa (1995) claims that the same point can be made for EP: post-verbal overt pronouns behave like null pronouns with respect to binding. In this case, because ‘nobody’ is too robustly non-referential to support a co-referential reading, an overt non-bound pronoun results in obligatory disjoint-reference, a co-referential construal with ‘nobody’ is impossible:

(112) *Nenhum aluno₁ disse que ele₁ falaria com ela
no student said that he would-talk with her
'No student said that he would talk to her

(113) Nenhum aluno₁ disse que falaria ele₁ com ela.
no student said that would-talk he with her.
No student said that he would talk to her.

Similar effects also hold in Italian:

(114) *Nessuno₁ ha detto che lui₁ l'avrebbe fatto
Nobody has said that he it-would-have done
'Nobody said that he would have done it.'

(115) Nessuno₁ ha detto che l'avrebbe fatto lui₁
Nobody has said that it-would-have done he
'Nobody said that he would have done it.'

Spanish behaves similarly. Where it clearly receives focus, native speakers accept (somewhat marginally) a bound reading:

(116) ?Nadie dijo que quisiera él, (y no su
Nobody said that want:subj he and not his
mujer), una manzana.
wife an apple

'Nobody said that he, and not his wife, wants an apple.'

Superficially, at least, this contrast seems to be explained by A&A's analysis.³⁵ Note, however, that the CLLD approach fails to account for the more complex fact that overt pronouns *can* be bound if linked to a bound null subject in Spanish (Montalbetti 1984, 1986):

- (117) Muchos estudiantes₁ dijeron que *pro*₁ piensan que ellos₁ son inteligentes
Many students said that they think that are intelligent
'Many students (x) said that x thinks that x is intelligent.'

This is even more striking where the matrix subject is a bare quantifier:

- (118) Quién₁ dijo que *pro*₁ piensa que él₁ es inteligente ? [Spanish]
who/nobody said that (he) thinks that he is intelligent
'Who/nobody said that he thinks he's intelligent?'

Under the CLLD approach, the binding of 'ellos'/él should always be blocked, as both pronouns are not 'arguments' in the true sense. However, according to Montalbetti (1986), (117) and (118) are fully grammatical with a bound reading. The same is true of the pronoun contained in the relative clause in (119):

³⁵ However, under the weaker version of A&A's proposal, and Barbosa's approach, clauses with a CLLD overt preverbal subject also contain a null *pro* in the post-verbal position. The fact that this *pro* cannot be bound still needs to be ruled out. Barbosa (1995) proposes that under Higginbotham's (1980) linking theory, only arguments can be bound. In any case, binding of the post-verbal *pro* across the CLLD pronoun is presumably ruled out under Relativised Minimality, the pronoun presumably acting as an intervener. More difficult to explain, however, is why null pronouns unlike overt pronouns cannot receive a co-referential reading in such cases. Some kind of restriction, inverse to that proposed by Montalbetti, is still required to capture the facts:

(28) A null pronoun in the scope of a quantifier cannot receive a co-referential reading.

In this sense, Montalbetti's generalisation does appear to be empirically inadequate and in need of amendment.

- (119) Nadie pensó [que [las fotos [que él tomó]] probarían [que *pro* estuvo ahí]].
Nobody thought that the photos that he took would-prove that was there
'Nobody thought that the photos that he took would prove that he had been there.'

I will not go into the full details of Montalbetti's theory, except to say that he accounts for the grammaticality of (119) from the fact that the pronoun in the relative clause does not c-command the embedded null subject, nor vice versa. For this reason, *pro* can be bound by the matrix QP and *él* linked to the bound variable *pro*.

In fact, Barbosa herself provides an alternative explanation for the fact that post-verbal, unlike preverbal subjects can be bound. She shows that, in EP, focused preverbal pronouns can also be optionally bound:

- (120) Nenhum aluno disse que **só ele** falaria com ela. [EP]
No student said that only he would-talk with her
'No x, where x is a student, x said that only x would talk to her.'

A similar fact appears to also be true of Spanish and Italian:

- (121) Nadie dijo que tan sólo él quisiera una manzana [Spanish]
Nobody said that only he want:subj an apple
'Nobody said that only he, (and not his wife), would like an apple.'

- (122) Nessuno_i ha detto che soltanto lui_i l'avrebbe fatto [Italian]
Nobody has said that only he it-would-have done
'Nobody said that only he would have done it.'

In her terms, this is due to the fact that foci are fronted under A-bar movement, whereas Topics are base generated in a non-argumental position. It stands to reason, therefore, that foci can be bound, being true arguments, whereas topics cannot. It will be argued in chapter 4 that this analysis carries over onto a deletion account where overt topics involve pronominal doubling whereas focus involves the failure to delete.

2.4.6. Summary

Although there is suggestive evidence to support Barbosa and A&A's claim that all preverbal subjects are CLLD in Romance NSLs, the evidence does not stand up to scrutiny. While it is certainly the case that preverbal subjects are CLLD on some occasions in Romance, it need not be the case that they are always CLLD. In fact, in the following section, it will be illustrated that there are several reasons to believe that they *cannot* always be CLLD.

2.5. Evidence that not all preverbal subjects are CLLD

The fact that post-verbal subjects get a pragmatically marked interpretation in all Romance NSLs except Spanish is often taken as an argument that Romance NSLs are canonically SVO, as this word order is pragmatically neutral.³⁶ However, as Barbosa (1995) rightly points out, a theory cannot be based entirely upon intuitive notions of markedness. However, there are also a number of other reasons to question the idea that all preverbal subjects in Romance NSLs are CLLD. Evidence of this kind comes from:

- *Differences in frequency*
- *Difference in pragmatics*
- *Non-topic-like preverbal subjects*
- *Robustly non-referential QP subjects*
- *Italian weak pronouns*
- *Non topic-like subjects*
- *Lack of topic –island effects*
- *Lack of Minimality effects*
- *Acquisition*
- *Lack of ad sensum agreement*
- *The fact that pro patterns with preverbal subjects*

³⁶ I return to these facts in chapter 3.

In the following sections, I review the arguments against A&A's analysis, concluding that the strongest version of their claim cannot possibly be maintained. I also consider Barbosa's (1995) slightly different proposal that preverbal subjects can be *either* CLLD *or* A-bar moved. I examine Barbosa's evidence for this position, based on clitic placement in EP, and conclude that it is not conclusive. Occam's razor therefore forces us to assume that the relevant phenomenon is A-movement in such circumstances. That Romance NSLs should allow preverbal A-subjects falls out from the assumption that they too have an active EPP of the move-XP type.

2.5.1. Differences in number of tokens

Costa and Adragão (2004) cite data from child acquisition of EP, to suggest that preverbal subjects are much more frequent than CLLD objects. They take this as evidence that not all preverbal subjects can be CLLD. However, studies of CLLD in French, where overt pronominal doubling occurs, reveal that subject CLLD is far more common than any other type of CLLD. Duarte (1995) cites the study by Barnes (1986), based on a *corpus* of colloquial French, which shows that 81% of the dislocated structures found were subject CLLD, and that the structure is not associated with a "special intonational contour", contrary to common claims:

- (123) a. Non, non, on est obligé de plonger.
No no one is obliged to dive
'No, no, you're obliged to dive.'
- b. On était obligé, le mec il m' a poussée!
One was obliged the guy he me has pushed'
'I was obliged alright, the bloke pushed me in!'

[Barnes (1986:217)]

In total, 62% of these tokens involved an apparent CLLD subject pronoun and, of these, 69% were the first person singular *moi*. This would seem to suggest that the special 'topic-like' status of subjects makes them a more common target for CLLD, than other arguments/adjuncts.

However, there is an alternative explanation for these findings. Barnes (1986:218) herself suggests that the subject pronouns have become agreement markers in spoken French, taking the place of agreement morphology on the verb. Building on this idea, Roberge (1990) has proposed that, in the colloquial language, French subject pronouns are really subject clitics of the type seen in Northern Italian dialects. If this is the case then these clitics have the status of agreement in colloquial French and are thus expected to be obligatory. Given these facts, the French data cannot be taken at face value, as what appears to be subject CLLD might actually be merely a subject and agreement marker. One possible problem with such an analysis is the fact that doubling of this type is not obligatory in French, even with the first person. A possible explanation for this fact comes from Barnes' further claim that 1st person singular occurrences without *moi* are only found where the speaker is the topic under discussion, suggesting that these might be instances of topic drop of the English type:

(124) (I) Might go out later

(125) (Moi) je vais peut-être sortir plus tard
I cl-1s go perhaps go-out more late
'(I) might go out later.'

This analysis, and even the empirical findings on which it is based, have been disputed by other researchers, and it seems, in fact, that French subject pronouns probably have a different status in different dialects of spoken French. Duarte (1995:105) cites a study by Blanche-Benveniste (1993), suggesting that only 10-13% of tokens in natural speech are made up of doubled structures of the type described above. This is clearly a matter for future empirical research: as it stands, the evidence is not conclusive.

While it may be the case that subjects make more obvious topics than objects cross-linguistically, Costa and Adragão's frequency argument is suggestive. Coupled with other facts about preverbal subjects in Romance NSLs, it serves as evidence that not all preverbal subjects are CLLD.

2.5.2. Pragmatic differences

CLLD serves to bring to the foreground old or given information in the universe of discourse, setting it as a 'point of reference' against which the main predication of the sentence is then construed (adapted from Raposo 1996:6). If all preverbal subjects are CLLD then we expect them to have the same information-packaging possibilities as other CLLD elements. Costa (2001) shows that in answer to sentence focus questions in EP, where all sentence constituents are focussed, "in transitive and intransitive structures, SV(O) word order is obligatory [and] left-dislocation is impossible":

- (126) a. O que é que aconteceu?
The what is that happened?
'What happened?'
- b. O Pedro partiu o braço.
the Pedro broke the arm
- b'. #Partiu o Pedro o braço.
broke the Pedro the arm
- b''. #O braço, o Pedro partiu-o.
the arm, the Pedro broke it

This is unexpected if preverbal subjects are CLLD. In this context the subject does not represent old or given information, but rather is part of a wide focus sentence. I return to these facts in chapter 3. For now, suffice it to say that preverbal subjects in Romance NSLs are not always 'topics', or 'old information'.

Cardinaletti (1997a) cites the following example from Anagnostopolou (1994) to suggest that in Italian, CLLD also has certain pragmatic effects, which are not shared by preverbal subjects. As an illustration, she gives the following data from English:

- (127) I saw a film about Wim Wenders₁ last night. The director₁ was presented as a rather introvert person.
- (128) I saw a film about Wim Wenders₁ last night. The director₂ was a young man from Spain.

Whereas in (127) *the director* refers back to Wim Wenders, in (128) *the director* “involves constructing a new referent” which is indirectly related to a context salient DP Cardinaletti (1997a:44), in this case the director of the film about Wim Wenders. According to her, in Italian, CLLD subjects can *only* pick up the more obscure reference (129). Non-CLLD subjects can potentially refer to either Wim Wenders or the new referent (130):

(129) Hanno premiato un film su Wim Wenders₁. Dopo la
Have:3pl rewarded a film about Wim Wenders. After the
proiezione, il regista_{1/2}, il premio, l’ha ricevuto dal ministro.
show the director the award it has received from-the minister.
‘A film about Wim Wenders has been awarded a prize. After the show, the
director accepted the prize from the Minister.’

(130) Ho visto ieri alla Biennale un film
Have:1s seen yesterday at-the Biennale a film
su Wim Wenders₁. Dopo la proiezione, il regista_{1/2} ha ricevuto un
about Wim Wenders. After the show, the director has received a
premio alla carriera
prize to-the career
‘I saw a film about Wim Wenders yeaterday at the Biennale. After the show, the
director accepted a prize celebrating his career.’

Cardinaletti therefore argues that the preverbal subjects in (130) cannot be CLLD as we would not expect the referent ‘Wim Wenders’ to be available. Further research is required in this area to check that the same facts hold for the other Romance NSLs.

2.5.3. Quantificational preverbal subjects

We have already seen that object CLLD is not possible with robustly non-referential QPs. Similar facts are true of subject left dislocation in French and BP:³⁷

(131) *[Ninguém]_i ele_i gosta de chorar [BP]
Nobody he likes of cry:inf
'Nobody likes crying.'

(132) *Personne, il n' aime pleurer. [French]
Nobody he not likes to-cry
'Nobody likes crying.'

However, preverbal non-referential QPs are the norm in Romance NSLs:

(133) Nadie quiere ser político
Nobody wants be:inf politician
'Nobody wants to be a politician.'

QPs can also occur post-verbally under the same conditions that other subjects can (see chapter 3). In such circumstances they must be doubled by preverbal negation:

(134) No lo cree nadie
Not it believes nobody
'Nobody believes it.'

³⁷ Note, however, that what looks like CLLD is possible with indirect objects:

(29) A nadie le dio Juan un beso
To nobody CL-HIM gave Juan a kiss
'Juan gave nobody a kiss.'

I will return to examples like this in chapter 3, where it will be argued that these non-referential QPs are fronted via movement to spec IP, rather than via CLLD. As such the clitic in (29) is not due to CLLD, but is rather a form of agreement between the element in spec IP and I. For further details see chapter 3.

A&A (1998:508) acknowledge that their analysis predicts that subject elements which typically cannot be left dislocated should never appear preverbally. They argue, however, that QPs, contrary to popular belief *can* actually be CLLD, giving the following evidence from Italian:

(135) Qualcuno (*lo) troveranno
 Someone him will-find:3pl
 ‘They will find someone.’ [existential:weak]

(136) Qualcuno lo troveranno
 Someone him will-find:3pl
 ‘They will find some specific person.’ [specific:strong]

A&A take (136), where an apparently non-referential QP appears preverbally as evidence that non-referential QPs can actually be CLLD. They claim that the specific reading associated with (136) is due to the referential use of the bare quantifier, which is itself a result of the fact it is CLLD (see section 2.4.3 above for details). The problem with this claim is that preverbal subject QPs do not always display any ‘referentiality’. It is not clear, for instance, how ‘nobody’ can have a referential reading of any kind, and yet it can surface in preverbal position.

Barbosa (1995) takes a slightly different approach to QP subjects, proposing that this group of non-referential QPs undergo A-bar movement to spec IP. She justifies this claim with complex data regarding the placement of clitics in EP. Let us take a brief excursion here to review these facts.

2.5.3.1. Portuguese QPs

Barbosa (1995) uses evidence from the placement of object clitics in EP to argue that referential subjects are CLLD, whereas non-referential subjects are A-bar moved to a preverbal position. She first observes that CLLD triggers enclisis in EP:

(137) Esse livro, a Maria deu-o ao Luis.
 This book the Maria gave-CL-IT to-the Luis

Lit. 'This book, Maria gave it to Luis.'

- (138) *Esse livro, a Maria o deu ao Luis.
This book the Maria CL-IT gave to-the Luis

Wh-movement of an object or adjunct, on the other hand, always triggers proclisis:

- (139) Quem o viu?
who CL saw
'Who saw him?'

- (140) *Quem viu-o?
who saw-CL-HIM

Referential preverbal subjects trigger enclisis, but non-referential QPs trigger obligatory proclisis:

- (141) A Maria comeu-o
The Maria ate-CL-IT

- (142) Ninguém o comeu
Nobody CL-IT ate
'Maria/nobody ate it.'

Barbosa (1995) takes this as evidence that referential subjects are CLLD, whereas quantificational subjects are A-bar moved to a preverbal position, at least in EP, triggering proclisis. Seen in these terms, Barbosa makes a strong case for differentiating between A-bar moved and CLLD subjects.³⁸

³⁸ Although further stipulation is required in order to correlate proclisis with the former and enclisis with the latter.

However, saying that it is A'-movement which triggers proclisis is a simplification of the truth. The presence of certain quantificational adverbs or negation also triggers proclisis, irrespective of the type of subject, (for a break down see Rouveret 1989, 1999 and Madeira 1994):

(143) O José já / não o comeu
The José already/NEG CL-IT ate

(144) O José nunca o comeu
The José never CL-IT ate

(145) O José as vezes come-o
The José sometimes eats- CL-IT

In some cases, this varies depending on the position (and subsequent meaning) of the adverb:

(146) O José sempre o come
The José always CL-IT eats
'José will eat it anyway.'

(147) O José come-o sempre
The José eats-CL-IT always
'José always eats it.'

Moreover, in embedded CPs, proclisis, and not enclisis is obligatory in all cases:

(148) Eu acho que te vi ontem
I think that CL-YOU saw yesterday
'I think I saw you yesterday.'

Barbosa (1996) suggests a solution to these problems by suggesting that the syntax interacts with the prosodic component. More specifically, she claims that EP clitics head a '*prosodic subcategorisation frame*' of the following kind:

(149) *_[intP cl X]

The constraint in (149) simply states that a clitic in EP cannot appear in PF at the edge of an intonational phrase. She claims that embedded CPs do not constitute separate intonational phrases and so proclisis in examples like (148) is expected. The problem is that it is difficult to see how a preverbal referential subject forms a separate intonational phrase to the finite verb it agrees with, even though it is widely accepted that no intonational break is required between the two (regardless of whether it is CLLD or not). This is especially problematic if we are also to assume that an embedded CP necessarily forms an intonational phrase with matrix material, skipping the embedded subject, which again, heads a separate intonational phrase. In short, Barbosa's account fails to define what exactly an intonational phrase is, and as such risks circularity.

It is beyond the scope of this thesis to provide an analysis of clitic placement in EP. EP clitic- placement is demanding of a principled explanation, however, it is not obvious that this analysis should depend upon the CLLD/A-bar moved status of preverbal subjects. The distribution of enclisis and proclisis superficially appears to suggest that there is a difference in status between QP and referential subjects; however, patterns of clitic placement with certain adverbs and in embedded contexts belie the idea that this is merely a structural difference regarding the position of preverbal subjects. Rather it appears to be crucially linked to the presence of quantification of any kind in matrix clauses (cf. Rouveret 1989). Moreover, the fact that the contrast is neutralised in embedded contexts suggests that it is an effect of interaction with the CP if anything, rather than a simple contrast between QP vs referential subjects.³⁹

³⁹ Ian Roberts (p.c.) suggests to me that this might be a V2 effect, with enclisis being the result of V movement to C. Despite the obvious appeal of this kind of approach, it is not without problems. An account must still be given of the contexts in which V2 is blocked, and proclisis ensues. In a sense, this debate does not effect the A-status of preverbal subjects in EP in any case, as SVO is still the unmarked word order in embedded clauses.

2.5.4. Weak pronouns

Cardinaletti (1997a) claims that in Italian not all overt preverbal subjects are in an A'-position. She shows that the weak pronouns *egli & essa*, only surface in preverbal position (150)-(151), and cannot be CLLD as objects (152), or subjects (153), unlike strong pronouns and R-expressions:

(150) Egli/lui/Gianni ha aderito
 He/he/G has adhered

(151) Ha aderito Gianni/lui/*egli
 has adhered He/he/G
 'He has adhered.'

(152) *Egli/lui/Gianni Maria non l' ha appoggiata
 Him/him/Gianni Maria not him has supported
 'Him, Maria has not supported.'

(153) *Essa questo problema non lo spiega
 It this problem not it explains
 Lit. 'It, this problem, doesn't explain.'

This data seems to suggest that not *all* preverbal subjects are CLLD, weak pronouns at least cannot be. This in turn gives evidence both in favour of the fact that the preverbal subject position is an A-position, and that the post-verbal position is not straightforwardly so. As we shall see in chapter 4, the post-verbal subject position is closely connected to focus.

2.5.5. Subjects present for disambiguation

Cardinaletti (1997) has shown that where verb forms are ambiguous in NSLs, pro-drop is limited to certain persons. In the present subjunctive of Italian, 1st, 2nd and 3rd person singular verb endings are all syncretic. Whereas 1st and 3rd person permit pro-drop, 2nd person requires an overt pronoun in out of the blue contexts:

(154) Che possa riuscirci non è chiaro
 That can:subj manage-there not is clear
 'It isn't clear that I/*you/he can manage it.'

(155) Che tu possa riuscirci non è chiaro
 That you can:subj manage-there not is clear
 'It isn't clear that you can manage it.'

Similarly, in the imperfect subjunctive of Italian, 1st and 2nd person singular verb agreement morphemes are identical. In this case, again, the 2nd person singular requires an overt pronoun:⁴⁰

(156) Che potessi riuscirci non è chiaro
 That could:subj succeed-there not is clear
 'It isn't clear that I/??you could manage it.'

This suggests either that Italian pro-drop is syntactically defective in the subjunctive mood, or that pro-drop, even in NSLs, is subject to pragmatic constraints. Clearly the overt pronoun in (155) is not a CLLD topic, but rather is merely required for

⁴⁰ Luigi Rizzi (personal communication) claims that the contrast here is stronger than that described by Cardinaletti. According to his judgment, while a null subject is perfectly fine with the imperfect subjunctive as long as the subject is contextually salient, the same is not true in with the present subjunctive. In the present subjunctive, where the syncretism is three way, it is virtually impossible to have a null 2nd person pronoun, even where there is a strong contextual salience. He provides the following example, where it is not possible to drop the embedded pronoun:

(30) #So che hai provato ma non è facile che possa riuscirci
 Know:1s that have:2s tried but not is easy that can succeed-there
 'I know that you've tried but it's not going to be easy for you to succeed,'

Given these facts, it might be the case, after all, that the number is syncretisms in a given paradigm is of importance after all (c.f. chapter 4 for a discussion). I leave this matter open for future research.

functional/pragmatic reasons to differentiate between ambiguous verb-forms. As such it is simply the phi-features of the 2nd singular pronouns which are required in (154)&(156).

Cole (2000) argues that the same thing is more generally true of Romance, where ambiguity arises. These preverbal subjects have no special discourse status as topics or foci:

(157) María y yo llegamos a casa. (yo) encontré las llaves...
Maria and I arrived at home (I) found the keys...

(158) María y yo llegamos a casa. *(yo\ella) tenía las llaves...
Maria and I arrived:1pl at home had:1s/3s the keys.

Cole claims where agreement identification fails, NSLs have recourse to discourse pro-drop strategies (i.e. they look for a single salient discourse topic), and where these fail, require an overt subject. A CLLD approach to overt preverbal subjects fails to allow for the presence of overt preverbal subjects which do not receive a topic reading. These examples seem to be genuine examples of pronouns used as pronouns in NSLs.

2.5.6. Islandhood

As pointed out by Gutiérrez-Bravo (to appear:7, attributed to Goodall 2001), there is evidence that preverbal subjects in Spanish are not fronted topics. In Spanish, clauses with fronted topics are islands for extraction, whereas clauses with preverbal subjects are not:

(159) *A quién crees [que el premio se lo dieron]?
to whom you-think that the prize DAT-CL ACC-CL they-gave
(Lit. 'Who do you think that the prize they gave to?')

(160) A quién crees [que Juan le dio el premio]?
to whom you-think that Juan DAT-CL gave the prize
'Who do you think that Juan gave the prize to?'

[Example from Goodall (2001)]

Again, an alternative approach, under which the preverbal subject occupies an A-position does not face this problem. If the preverbal subject raises to an A-position then the embedded sentence in (160) has a different status to that in (159). Assuming that it is the presence of an A-bar element in the left periphery of (159) which blocks A-bar extraction of another, more embedded element, it follows that the preverbal subject in (160), as an A-element will not block A-bar extraction in the same way. If, on the other hand, the preverbal subject is treated as an A-bar element, as proposed by A&A and Barbosa, then this contrast is very difficult to account for.

2.5.7. Recursivity

Costa appears to abstract away from Cinque's distinction between base-generated CLLD A'-subjects and moved A'-subjects, assuming all preverbal topics to have the same base-generated status.⁴¹ This means that he concentrates on showing that preverbal subjects in EP behave differently from preverbal topics. He gives the following set of evidence that not all preverbal subjects in EP are in an A'-position. Preverbal subjects in EP do *not* give rise to:

- *A' minimality effects (following Ambar, 1992)*
- *Multiple topicalisation effects*
- *Changes between topics*

Thus whereas a topicalised indirect object cannot intervene between a moved wh-element and its trace (161), (in certain environments) a preverbal subject can, (162):⁴²

(161) *Que livro₁, **ao Paulo**, deste t₁?

⁴¹ He gives evidence attributed to Duarte to show that preverbal topics without resumptive clitics are not derived via movement (Costa 2001:98).

⁴² Note that, as discussed in chapter 3, bare wh-words require subjacency between the wh-phrase and the finite verb/auxiliary. This is not the case with wh+N phrases.

What book, to-the Paulo, gave:2sg
'Which book did you give to Paulo?'

- (162) Que livro₁ o João leu t₁?
What book the João read
'Which book did João read?'

Costa, following Ambar (1992), takes this as evidence that the preverbal subject is not in an A-bar position, assuming the ungrammaticality of (161) to be due to Minimality.

Costa (2001) shows, likewise, that EP subjects behave differently from objects with respect to topicalisation. Whereas only one complement can be preposed to a preverbal A'-position (163)-(164), a topicalised complement and a preverbal subject can felicitously co-occur (166):

- (163) Esse livro, deram ontem à Maria.
That book gave:3pl yesterday to-the Maria
'That book, they gave to Maria yesterday.'

- (164) ??Esse livro, à Maria deram ontem
That book to-the Maria gave:3pl yesterday
Lit. 'That book, to Maria, they gave yesterday.'

- (165) Esse livro, leu o Paulo ontem
That book read the Paulo yesterday

- (166) Esse livro, o Paulo leu ontem
That book the Paulo read yesterday
'That book Paulo read yesterday.'

Similarly, in relative clauses preverbal subjects are fully grammatical, but fronted topics result in ungrammaticality:

- (167) Já li o livro que o João ofereceu à Maria.

Already read:1s the book that the João offered to-the Maria
'I've already read the book that João gave to Maria.'

- (168) *Já li o livro que, à Maria, lhe ofereceu ontem o João.
Already read:1s the book that to-the Maria her offered yesterday the João
Lit. 'I've already read the book that to Mary, João gave yesterday.'
[example from Costa & Duarte (2002)]

In her defence, Barbosa (2001) points out that the distinction to be made is more complex than subject vs. non-subject. She proposes that CLLD is subject to a 'thematic hierarchy' whereby subjects occupy a lower position than CLLD objects, and experiencer objects because of their theta-role occupy a lower position than theme objects. Experiencer objects, like subjects, are therefore fine at the front of relative clauses and questions:

- (169) Vi hoje a casa que, à Maria, mais (lhe) convém comprar.
Saw:1s today the house that to-the Maria more her is-convenient to-buy
'Today I saw the house that for Maria is the most convenient to buy.'

- (170) Que discos, ao João, mais (lhe) agradará receber?
What records to-the João more him will-please to-receive
'What records will John be most pleased to receive.'

An alternative to Barbosa's thematic hierarchy' is to assume, as Masullo (1993) has argued, that experiencers behave like subjects in syntactic terms. This is supported by the possible presence of the clitic in (169)-(170), which is a kind of agreement. In this case then the subject vs. non-subject asymmetry is maintained, and Costa's claim remains valid.

2.5.8. Acquisition

Costa and Adragão (2004) (C&A) point out that A&A and Barbosa's accounts of null subjects both predict that the child should start producing preverbal subjects at the time at which he/she acquires CLLD. Before this, all subjects should be null or should surface

in the post-verbal position, given the generally accepted fact that children acquire V-to-I movement very early on (cf. Gonçalves 2001). On the other hand, if spec IP is actually an A-position, then we do not necessarily predict the production of preverbal subjects to cluster with “the emergence of other fronting strategies” (C&A 2004:2). In actual fact C&A (2004:4) report that VS word order accounts for only 7% of 1060 utterances containing an overt subject (child aged between 31 and 43 months). Of these small number of VS tokens 79% are made up of unaccusative verbs, which merge their subjects in object position, and which allow post-verbal subjects in unmarked contexts in adult grammars (C&A 2004:6). They also show that children do not suffer from a pragmatic deficit, contra Wexler (1998), and in support of De Cat (2002). From this, C&A conclude that (i) the preverbal position is unmarked in EP (ii) not all preverbal subjects are CLLD (iii) the post-verbal position is marked, the subject being focussed. I will return to this point in chapter 3.

2.5.9. Lack of *ad sensum* agreement

Suñer (2002) gives evidence from *ad sensum* agreement (as well as scope and reconstruction) to suggest that many preverbal subjects in Spanish must be in an A-position. Following Bosque (1999), Suñer shows that true CLLD subjects in Spanish allow optional *ad sensum* agreement:

(171) El jurado, María nos aseguró que estaban presionados
 The jury:ms María us assured:3s that were:3pl pressured:mpl
 ‘The jury, Maria assured us that they felt pressured.’

In (171) the predicate ‘be pressured’ has 3pl agreement. Its subject ‘the jury’ is grammatically singular but notionally plural. Agreement therefore reflects semantic number rather than syntactic number. This *ad sensum* agreement is not possible with all preverbal subjects, notably those directly adjacent to the finite verb:

(172) El jurado *estaban presionados /estaba presionado
 The jury were pressured was pressured
 ‘The jury felt pressured.’

One possible way to explain this is to assume that ad sensum agreement is only possible where the subject is CLLD.⁴³ This leads Suñer to conclude that preverbal subjects cannot *always* be the result of CLLD in Spanish.

2.5.10. The position of the null pronoun

An analysis such as that of A&A (1998) predicts that null subjects should share properties with post-verbal subjects, rather than preverbal subjects. In actual fact the opposite appears to be the case. Cardinaletti (1997a) argues that *pro* occupies a preverbal position based on its typological status as a weak pronoun. Drawing on Cardinaletti and Starke (1999), and making a comparison with French, she claims that *pro* patterns with deficient 'il-type' pronouns, rather than with the strong 'lui-type' in its distribution and semantics. Thus *pro*, like 'il', can refer to non-human referents, can act as an expletive or quasi-referential subject, or as an impersonal subject. Cardinaletti goes on to show that weak pronouns in French, unlike strong pronouns, have a very restricted distribution, being limited to the immediate preverbal position:

(173) *Il, il est parti.

He he is left

(174) Lui, il est parti.

Him he is left

(175) *Il est parti, il.

He is left, he

(176) Il est parti, lui.

He is left, him.

⁴³ Note that this is not the case in British English, in which sentences such as the following are common:

(31) The police think that he has escaped

(32) The jury find him innocent

'He has left.'

Cardinaletti therefore states that the null hypothesis is that the same distributional restriction will be true of the weak pronoun *pro* in Italian.

Indeed, as Cardinaletti (1997) notes, both Rizzi and Burzio have previously pointed out empirical evidence that *pro* patterns with preverbal, rather than post-verbal subjects. Null subjects are barred from existential constructions in Italian, which require a post-verbal subject:

(177) *Io ci sono alla festa
I CL-LOC am at-the party

(178) Ci sono io alla festa
CL-LOC am I at-the party

(179) *Ci sono alla festa
CL-LOC am at-the party
'There's me at the party.'

On the other hand, floating quantifiers are possible with preverbal subjects or *pro* in Italian, but not with post-verbal subjects:

(180) [Tutti i soldati] sono andati via
All the soldiers are gone away

(181) [I soldati] sono tutti andati via
The soldiers are all gone away

(182) Sono andati via [tutti i soldati]
Are gone away all the soldiers

(183) *Sono tutti andati via [i soldati]
Are all gone away the soldiers

- (184) Sono **tutti** andati via
 Are all gone away
 'All the soldiers left.'

There is thus empirical evidence in Italian, to suggest that *pro* occupies the same structural position as preverbal subjects. The same appears to be true of Spanish (185)-(189), and Portuguese (190)-(194):⁴⁴

- (185) [**Todos** los chicos] se han comprado un coche [Spanish]
 All the boys CL-SELF have bought a car

- (186) [Los chicos] se han comprado **todos** un coche
 The boys CL-SELF have bought all a car

- (187) Se han comprado un coche [**todos** los chicos]
 CL-SELF have bought a car all the boys

- (188) *Se han comprado **todos** un coche [los chicos]
 CL-SELF have bought all a car the boys

- (189) Se han comprado **todos** un coche
 CL-SELF have bought all a car
 'All the boys bought themselves a car.'

- (190) [**Todos** os soldados] foram embora [EP]
 All the soldiers went away

- (191) [Os soldados] foram **todos** embora
 The soldiers went all away

⁴⁴ Existential constructions are formed differently in Spanish and Portuguese and so the first test cannot be applied.

(192) Foram embora [todos os soldados]
 Went away all the soldiers

(193) *Foram todos embora [os soldados]
 Went all away the soldiers

(194) Foram todos embora.⁴⁵
 Went all away

Assuming preverbal subjects to be in an EPP-driven subject position, instances of floating quantifiers can be attributed to the movement of the subject and subsequent stranding of the quantifier in the original subject position. This correctly rules out the possibility of a low subject coupled with Q-float. As such, it is implied that null subjects must occupy a preverbal position as they also permit Q-float. On the other hand, if preverbal subjects are CLLD, and thus base-generated in a peripheral position, it becomes harder to explain the distribution of floated quantifiers. In this case, it remains unexplained why null subjects and preverbal overt subjects both permit Q-float, whereas overt post-verbal subjects do not.

Further evidence comes from a large number of Italian dialects, in which, similarly, there is reason to believe that *pro* occupies a preverbal position (contra Barbosa 1995). Preverbal subjects trigger full agreement, unlike post-verbal subjects, which typically fail to agree in person in the 3rd person. For example in Ancona, post-verbal 3rd person plural subjects do not trigger plural agreement, as Cardinaletti (1997a) shows:

⁴⁵ Note that examples such as the following are also possible in EP:

(33) [Os soldados todos] foram embora
 The soldiers all went away
 ‘All the soldiers left.’

I argue, in chapter 3, that the quantifier *todos* is not stranded here but is actually a post-nominal modifier in the DP.

(195) Questo, lo fa sempre i bambini.
This, it does always the children

(196) *Questo, i bambini lo fa sempre.
This the children it does always

(197) Questo, i bambini lo fanno sempre.
This the children it do always
'THIS, the children always do.'

In the same context, a null subject would have to trigger plural agreement. Cardinaletti shows that this is not simply due to functional reasons as even where a floating quantifier makes it clear that the subject is plural, plural agreement on the verb is also required:

(198) Questo, lo dimostrano tutte
This, it demonstrate all:FPL

(199) *Questo, lo dimostra tutte
This it demonstrates all:FPL
'THIS, they all demonstrate.'

All this strongly suggests that the null subject does not occupy the post-verbal position. In many different Romance NSLs then, there appears to be good evidence that null subjects occupy a preverbal subject position, and even that they cannot surface post-verbally, unlike lexical and (some) pronominal subjects.

2.6. Summary

Given the fact that many Romance NSLs appear to display SVO word order and given the common generative assumptions about universal grammar, the burden of proof lies with those claiming that *all* preverbal subjects are CLLD, rather than simply A-moved to satisfy the EPP. In this chapter, I have examined the evidence for such a stance and

concluded that it is by no means persuasive. Although, in many cases preverbal subjects are undoubtedly the result of CLLD, (with a null pronoun in spec IP), it cannot very easily be maintained that *all* instances of preverbal subjects are. In fact in many cases, it seems far less problematic to assume that the subject occupies the canonical preverbal subject position. In the following chapter, we will see more evidence for such a stance from word order possibilities across Romance.