

A Case for an Agree-based Theory of Control

Pilar P. Barbosa

Universidade do Minho/CEHUM

1. Introduction

In a recent paper, Szabolcsi (2009) discusses evidence that there are languages where control and raising infinitives have overt subjects. Her descriptive generalization regarding this set of languages is the following:

- (1) The overt subjects of control complements can only be pronouns. The overt subjects of raising complements can be pronouns or lexical DPs.

Even though similar observations have been made before for Italian and Spanish (cf. Burzio 1986, Torrego 1996, Belletti 2005), Szabolcsi's paper is particularly interesting given that it covers a considerable number of languages and reveals the existence of at least three different typological patterns:

- (i) Languages that do not admit expressed subjects either in control or in raising complements; this is the case in English, German and French.
- (ii) Languages that allow for explicit subjects in raising and control complements: Hungarian, Italian, Romanian, Turkish, Brazilian Portuguese.
- (iii) Languages that allow for expressed subjects only in raising complements, like Russian, Finnish and (possibly) Hebrew.

In this paper, I will examine European Portuguese (henceforth EP) and I will show that it complies with (1). I will argue that the explicit subjects that occur in these infinitival complements in EP (as well as in Spanish or Italian) are genuine post-verbal subjects of the kind that is commonly attested in a standard consistent Null Subject language (NSL) and I will examine generalization (1) in light of the current debate between Agree-based *vs.* movement-based theories of obligatory control (OC). I will show that the evidence underlying (1) constitutes a strong case for the need for a non-raising approach to OC. In particular, there are differences between raising and control complements in EP that resist explanation under a movement theory of control.

A closer look at the sample of languages studied by Szabolcsi reveals that all of the consistent NSLs in her sample fall under Pattern (ii). Szabolcsi conjectures that the availability of null subjects might play a role but then rejects this hypothesis in virtue of languages such as colloquial Brazilian

Portuguese, which is no longer a consistent NSL and yet obeys pattern (ii). Nonetheless, even though the correlation between consistent *pro*-drop and Pattern (ii) cannot be stated as a two way implication, there appears to exist a one way implication in the sense that there is not a single consistent NSL in the sample that doesn't fall under pattern (ii). My goal in this paper is to present an explanatory theory of (1) that captures this one way implication.

This paper is organized as follows. In section 2 I introduce the relevant data. In section 3 I argue that the DPs that occur in these infinitival complements are genuine subjects and I discuss the relevance of the data examined for the current debate between movement-based *vs.* Agree-based theories of OC. In section 4, I present the analysis proposed.

2. The data

In this section, I introduce the data discussed in Szabolcsi (2009). I will use examples from EP (a language that the is not discussed by Szabolcsi) but the same facts obtain in Italian and Spanish, as reported in the paper cited.

2.1 Obligatory control complements

Consider the following example:

- (2) **Só ele** decidiu ir ao mercado.
 only he decided to-go to-the market
 'He is the only one who decided to go to the market'

(2) has the interpretation indicated in the English gloss, with the Focus operator having matrix scope. Example (3) below, with a post-verbal subject, is ambiguous and can have the two readings indicated in [a] and [b] depending on whether the focused DP is interpreted as having scope over the matrix verb (reading [a]) or over the embedded verb only (reading [b]). Following Szabolcsi, I will call each of these readings the High and Low readings, respectively:

- (3) Decidiu ir ao mercado **só ele**.
 decided to-go to-the market only he
 [a] 'He is the only one who decided to go to the market'
 [b] 'He decided for it to be the case that only he goes to the market'

The ambiguity of (3) can be easily explained under the assumption that it is structurally ambiguous: the focused pronoun can be the (post-verbal)

subject of the matrix clause (cf. (4a)) or it can be the (post-verbal) subject of the embedded clause (cf. (4b)):

- (4) a. [decidiu [PRO ir ao mercado] **só ele**]
 b. [*pro* decidiu [ir ao mercado **só ele**]

In example (5) below, the subject precedes the PP complement:

- (5) Decidiu ir **só ele** ao mercado.
 decided to-go only he to-the market

When (5) is uttered without a pause between the subject and the PP, it is not ambiguous and can only have the interpretation indicated in (3b). This is so because the only possible analysis of (5) is the one indicated below, where the focused pronoun is inside the embedded clause in a VSO structure:

- (6) *pro* decidiu [ir **só ele** ao mercado]

Now compare example (3) above with (7) below:

- (7) Decidiu ir ao mercado **só o João**.
 decided to-go to-the market only the John
 [a] 'He is the only one who decided to go to the market'
 [b] *'He decided for it to be the case that only John goes to the market'

(7) differs from (3) in that the focused DP is fully lexical rather than a pronoun. Curiously, the Low reading disappears in this case and the focused DP can only be interpreted with matrix scope. Under a standard, non-raising approach to OC, the reading in (7a) corresponds to the syntactic representation in (8a). The syntactic configuration that would serve as input for the Low reading (7b) is represented in (8b):

- (8) a. [decidiu [PRO ir ao mercado] **só o João**]
 b. **pro*_i decidiu [ir ao mercado **só o João**_i]

The configuration in (8b) is a standard Condition C violation and this is why the Low reading is not available. What this means, then, is that the subject pronoun in the Low reading is indeed in a low position, inside the embedded infinitival clause.

Note that the pronoun in the embedded clause need not be modified by a focus operator. It may occur by itself, as shown below:

- (9) Decidiu ir **ele** ao mercado.
 decided to-go he to-the market
 ‘He decided that he would be one to go to the market’

In (9) the pronoun is focused and the sentence has the reading indicated in the gloss (below, we will return to the issue of the interpretation of these post-verbal subject pronouns). As expected, (9) only has the Low reading when uttered without a break between the pronoun and the PP.

It is worth pointing out that these facts are not related to restructuring (cf. Rizzi 1978). The examples mentioned so far contain the verb *decidir* ‘decide’, which is not a restructuring verb. This is confirmed by the fact that it doesn’t allow clitic climbing, a wellknown test for restructuring. A restructuring verb like *querer* ‘want’ allows clitic climbing (cf. *Não quero vê-la* / *Não a quero ver* ‘I don’t want to see her’). *Decidir*, on the other hand, does not allow clitic climbing (cf. *Não decidi convidá-lo* ‘I didn’t decide to invite him’ / **Não o decidi convidar*).

Likewise, it is not possible to relate these facts to lack of a CP projection (cf. Costa 2004, Alexiadou et al. 2008)) given that these controlled pronominals can occur in infinitival complements introduced by an overt complementizer (cf. the EP example (10)). This fact has already been pointed out by Torrego (1996) for Spanish (cf. (11)) and Belletti (2005) for Italian (cf. (12)):

- (10) Não sabemos [se assinar **só nós** a carta ou não].
 not know.1stPl if to-sign only we the letter or not
 ‘We don’t know whether only us should sign the letter or not.’
 (11) No sabemos [si firmar **nosotros** la carta].
 not know.1stPl if to sign us the letter
 (12) Maria mi ha chiesto [di parlare **io** com Gianni].
 Maria me has asked DI to speak I with Gianni
 ‘Maria asked me to talk to Gianni myself.’

In sum, in obligatory control complements, the (noninflected) infinitival clause may have what look like explicit subjects, which are obligatorily controlled. As mentioned above, Torrego (1996) and Belletti (2005) reach similar conclusions for Italian and Spanish. Besides pronouns, the following two kinds of anaphoric expressions can also occur in OC constructions:

- *Partitive QPs containing a pronoun*
- (13) Pensamos falar alguns de nós / vários de nós / muitos de nós com ela.
 think.1stPL to-speak some of us / several of us / many of us with her
 ‘Some of us / several of us / many of us intend to talk to her’

- *Certain kinds of anaphoric collective DPs (cf. Torrego 1996)*

(14) Não sabemos como falar a turma toda / o grupo todo com ela.
 not know.1stPl how to-talk the class whole / the group whole with her
 'We don't know how the whole class / the whole group will talk to her'

2.2 Raising complements

Szabolcsi (2009) observes that, in raising complements, the distinction between pronouns and (non-anaphoric) lexical DPs is lost. Thus, the examples that follow allow for the Low reading regardless of the nature of the DP.

(15) *EP*:

- a. Não parece cantar **só eu** nesta gravação.
 not seem.1stSg to sing only I in this recording
 'It doesn't seem to be the case that only I sing in this recording.'
- b. Não parece cantar **só o João** nesta gravação.
 'It doesn't seem to be the case that only John sings in this recording.'

(16) *Italian*

- a. Non sembra cantare **solo io** su questo nastro.
 'It doesn't seem to be the case that only I sing in this recording.'
- b. Non sembra cantare **solo Gianni** su questo nastro.
 'It doesn't seem to be the case that only John sings in this recording.'

Since it is not easy to distinguish the High from the Low readings with the verb 'seem', Szabolcsi uses other examples with raising verbs and designs different scenarios that clearly distinguish the different truth-conditions associated with each reading. For lack of space, I won't review those facts here and I refer the reader to the paper. There, it is clearly established that the raising complements with an explicit subject are fully ambiguous regardless of the nature of the DP (a pronominal or lexical DP).

To conclude this sub-section, raising complements differ from control complements in that they do not require the subject in the embedded clause to be a pronoun or an anaphoric expression.

2.3 Discussion

The facts just presented are particularly interesting in light of the current debate between Agree-based vs. movement theories of OC. Under the view of Control as raising (Hornstein 1999 and subsequent works), one might be tempted to account for the facts just discussed in terms of *backward control* or *backward raising* (Polinsky and Potsdam (2002), whereby the subject of

the infinitival clause is the Spell-out of the copy of the raised subject (see also Alexiadou et al. 2008 for a slightly different implementation). As pointed out by Szabolcsi, however, this cannot be done in the case of the data discussed here for the following reasons. First, the embedded subjects do not have matrix scope; in particular, the High reading is unavailable whenever the focused pronoun precedes the object (cf. (5)). This is unexpected if the pronoun is the copy of a raised subject. Secondly, the contrast between raising and control complements regarding non-pronominal subjects is not easily explained. Recall that, with control verbs, only pronouns (or a restricted set of anaphoric collective DPs) are compatible with the Low reading (cf. (3) vs. (7)). In raising complements, any kind of subject allows for the Low reading. Earlier, we attributed this to a Condition C effect, but this account is lost under a backward control analysis, which would predict raising and control infinitives to behave alike with respect to the choice of the phonological shape of the pronounced copy.

3. Multiple subjects and emphatic pronouns

One other fact pointed out by Zsabolsci (2009) as well as Belletti (2005) is that these control and raising complements may take multiple subjects. Here we transcribe examples in EP (cf. (17)) and Italian (cf. (18)).

- (17) a. **Só o João** detestou resolver **só ele** o problema.
 Only the J. hated to-solve only he the problem
 ‘Only John hated it that only him solved the problem’
 b. **O João** acabou por resolver **ele** o problema.
 the J. ended up to-solve he the problem
 ‘John ended up solving the problem himself.’
 (18) a. **Solo Gianni** vuole andare **solo lui** a scuola.
 only Gianni wants to-go only he to school
 ‘Only Gianni wants to be the only one to go to school’
 b. **I ragazzi** risultarono aver riposto **loro** alla domanda.
 the boys turned out to have answered them to the question

These examples are strongly reminiscent of the constructions with “emphatic pronouns” discussed in Burzio (1986) among others. In all of the languages under discussion a pronoun (bare or modified by a focus particle) may co-occur with an explicit pre-verbal subject in a simple sentence. The pre-verbal subject may be a full DP or even a pronoun:

- (19) A Teresa / ela escreveu **só ela** o poema.
 the T. she wrote only she o poema
 ‘Teresa wrote the poem only by herself.’

- (20) a. A Teresa / ela escreveu **ela** o poema (ninguém a ajudou).
 the Teresa / she wrote she the poem (noone her helped)
 b. A Teresa / ela escreveu o poema **ELA** (ninguém a ajudou).
 ‘It was Teresa who wrote the poem, noone helped her.’
- (21) *Italian* (Belletti 2005)
 Gianni/lui verrà **lui**.
 Gianni/he will come he
 ‘John/he will come himself’
- (22) *Spanish* (Sanchez 1993)
 Pedro abrió **el** la puerta.
 Pedro opened he the door
 ‘Pedro opened the door himself’

In (19) the pronoun is modified by a Focus particle. In the remaining examples it appears by itself. As we can see in the EP example (20), the pronoun may precede the object (cf. (20a)) or follow it (cf. (20b)). In the latter case it must bear stress (as indicated by the use of capital letters). The reading obtained is the one indicated in the glosses, where the pronoun is focused and acquires an interpretation roughly equivalent to that of a cleft, as indicated in the gloss in (20). This interpretation can be paraphrased in English by sentences with the adjunct SELF anaphor. This is why these pronouns have been labelled “emphatic” by Burzio 1986.

In view of these examples, one might think that the pronouns in the embedded clauses in (19-22) are not genuine subjects, but rather some kind of anaphoric adjuncts. In the next section, we argue that this hypothesis is untenable and that the pronouns in bold in all of the examples mentioned above are genuine post-verbal subjects.

3.1. Evidence that emphatic pronouns are subjects

According to some linguists, emphatic pronouns are not real subjects but rather adjunct anaphors (cf. Piera 1987), or the phonetic realization of a trace (Burzio 1986). Rigau (1987), Sola (1992), Barbosa (1995) and Belletti (2005), however, claim that they are post-verbal subjects. In what follows, I present some of the arguments given in Barbosa (1995).

The idea that these pronouns are adjuncts (very much like the adjunct SELF anaphor *himself* in English) faces the problem that they do not have exactly the same properties as adjunct anaphors. In the first place, if emphatic pronouns are adjuncts, we should expect them to attach to any DP in the sentence, as happens with the English anaphor (cf. *I talked to the director himself*). However, this is not the case: emphatic pronouns cannot be attached to post-verbal subjects (cf. (23a)) nor can they be associated with

objects (cf. (23b)); in these cases, the complex SELF anaphor must be used (cf. 24):

- (23) a. *Apareceu o director ELE.
 appeared the director HE
 b. *Falei com o director ELE
 talked.1stSg to the director HE
- (24) a. Apareceu o director **ele próprio**.
 appeared the director himself
 ‘The director himself showed up’
 b. Falei com o director **ele próprio**.
 ‘I talked to the director himself’

(23a,b) are evidence that pronouns cannot occur as DP adjuncts. Moreover, (24a,b) show that EP does have a lexical counterpart to English *himself*, namely the complex SELF anaphor *ele/a-próprio/a*.

Another related property that distinguishes emphatic pronouns from complex SELF anaphors is that the former, though not the latter, are subject oriented. Thus, the emphatic pronoun in (25a) can only be anaphorically related with the subject; this restriction does not apply to the complex anaphor in EP (25b) or in English (26):

- (25) a. [A criança]_i foi felicitada pel[a professora]_k ELA_{i/*k}
 the child was congratulated by the teacher she
 b. [A criança]_i foi felicitada pel[a professora]_k ela própria_{i/k}
 the child was congratulated by-the teacher herself
- (26) The girl was congratulated by [the teacher]_i herself_i

Thus, we observe that (i) emphatic pronouns cannot appear as DP modifiers; (ii) the counterparts to the English adjunct SELF anaphor in EP are NOT the emphatic pronouns. These two observations taken together constitute a problem for the hypothesis that emphatic pronouns are adjuncts.

On the other hand, emphatic pronouns have the same distribution and interpretation as regular post-verbal subject pronouns. Thus, on apart with (20a,b) above we have (27a,b), where the post-verbal pronominal subject is interpreted as exhaustive focus and the sentence can be paraphrased by a cleft, as indicated in the glosses.

- (27) a. Escreveu **ela** o poema (ninguém a ajudou).
 wrote.3sg she the poem (noone her helped)
 b. Escreveu o poema **ELA** (ninguém a ajudou).
 ‘It was she who wrote the poem, noone helped her’

In this article, I assume the analysis of VSO/VOS alternations proposed in Ordóñez (1998) for Spanish, Costa (1998) for EP and Cardinaletti (1998) for Italian. According to this analysis, post-verbal subjects occupy their base position inside the vP. Since all of these languages arguably have V raising to T, the order obtained is VSO; VOS order is derived by object scrambling to a position above vP. Thus, (27a) is analysed as in (28a), and (27b) is analysed as in (28b):

- Em (28b) the pronoun is the mostly embedded constituent. Hence it is assigned Nuclear Stress by the Nuclear Stress Rule (cf. Cinque 1993, Zubizarreta 1998). This fact explains why the pronoun in that position must bear stress, as indicated above. Note that this effect obtains in (27b) as well as in (20b) above, with an emphatic pronoun. This, for us, is clear indication that the emphatic pronoun is subject to the same restrictions as post-verbal subject pronouns.

- a) The emphatic pronoun is the spell out of the copy of the moved subject (cf. Burzio 1986);
- b) The emphatic pronoun and the pre-verbal subject form a unique syntactic constituent, a “big” DP of the kind proposed by Torrego (1995), Uriagereka (1995), Cecchetto (2000) in order to account for Clitic Doubling in Romance (cf. (29)); part of the big DP, DP₂, moves to pre-verbal position and the remaining part (the pronoun) remains in the right-periphery of the clause; this is Belletti’s (2005) proposal¹:

(1) a. [[DP₂ *pro*] [_T escreveu [_{VP} [DP₁ *ela* [_{DP₂} ~~*pro*~~] o poema]]
wrote she the poem



(30) [[_{DP2} A Teresa] [_T escreveu [_{VP} [_{DP1} ela [_{DP2} ~~A Teresa~~]]] o poema]]

- c) The DP in preverbal position is not the thematic subject, but is rather a left-dislocated topic doubled by the pronominal subject in post-verbal position (Rigau 1987, Sola 1992, Barbosa 1995); in the representation that follows we assume that left dislocated topics are *base-generated* in the left periphery of the clause and are licensed by ‘rules of predication’ in the sense of Chomsky 1977: the clausal projection contains an ‘open’ position (the pronoun) satisfied by the entity referred to by the dislocated DP (cf. Demirdache 1992, Agnastoupoulou 1997):

(31) [_{DP_k} [_{TP} [_{V/v} [_{V/VP} ... *subject pronoun_k* ...]]]

(32) [A Teresa]_i [_{TP} escreveu [_{VP} ela_i [_{V'} ~~escreveu~~ o poema]]]]

One simple way of testing hypothesis (c) consists in seeing whether these “double subject” constructions are compatible with non-referential quantified expressions. As is wellknown, these negative quantifiers cannot in general be left dislocated. The following examples from French illustrate this fact:

- (33) a. Pierre il aime la musique.
 Pierre he loves the music
 b. *Personne il aime la musique.
 noone he loves the music

Given this, hypothesis (c) predicts that a non-referential quantified phrase should be incompatible with an emphatic pronoun, and, in effect, this prediction is confirmed (cf. (34-36).

(34) *Nenhuma criança escreveu ela o poema.

no child wrote she the poem

(35) *?Nessuno verrà lui.

noone will-come he

[Belletti 2005]

b. A Teresa decidiu [[PRO escrever [_{VP} [_{DP1} **ela** [_{DP2} ~~PRO~~]]] o poema]]
 the Teresa decided to-write she the poem

- (36) *Context*: The teacher worked and ... [Szabolcsi 2009]
 a. Gianni ha lavorato anche lui.
 Gianni has worked also he
 b. *Ogni ragazzo ha lavorato anche lui
 every boy has worked also he

These facts favor hypothesis (c) and constitute a problem for hypotheses (a) and (b). Recall that, under hypothesis c), the pronoun is the argumental subject and the dislocated DP is a base-generated topic. Since non-referential QPs cannot be left dislocated, the only configuration available for them is the one in (37), where the QP is the subject; hence there is no room for the subject pronoun:

- (37) [Nenhuma criança escreveu [_{VP} ~~nenhuma criança~~ escreveu o poema]]

Now consider hypotheses (a) and (b). If the emphatic pronoun is the spell out of the copy of the raised subject (hypothesis (a)), there is really no explanation for why this option is blocked in the case of a non-referential QP. The same reasoning applies to hypothesis (b). Under the “big” DP approach it is also not very clear why such a DP may not contain a non-referential QP given that it can be independently shown that non-referential QPs can be clitic doubled (recall that the “big” DP approach was specifically designed to deal with clitic doubling):

- (38) No le ablé a nadie. [Spanish (Torrego 1998)]
 not CL talked to nobody

Besides this, the “big” DP approach doesn’t really have a satisfactory answer to why there are no occurrences of DP₁ in subject or object position. Thus, (39a,b) are unattested (in order for such examples to become good, the DP *a Teresa* must be clearly detached from the rest of the sentence as an appositive):

- (39) a. *Ela a Teresa escreveu o poema.
 she the Teresa wrote the poem
 b. *Já falei com ela a Teresa.
 already talked.1stSg with her the Teresa

In view of these facts, I conclude that emphatic pronouns are not contained inside a “big” DP and are rather genuine subjects sitting in their base position.

3.2 Multiple subjects in raising and control complements

Turning now to the multiple subject constructions of raising and control complements, we note that they are remarkably similar to the mono-clausal examples with emphatic pronouns. Given these similarities, it is reasonable to conclude that the pronouns that occur in the embedded infinitival complements are also genuine (post-verbal) subjects.

In fact, all of the expressions that can occur as subjects of control infinitives can easily occur as post-verbal subjects in a simple clause triggering exactly the same kind of agreement morphology that is found in a control structure (for instance, collective DPs such as *a turma inteira* ‘the whole class’ co-occur with 1st person plural agreement in control as well as in simple clauses):

- (40) Chumbámos nós / só nós / só nós os dois / alguns de nós /
 flunked.1stPL we / only we / only we the two / some of us /
 / a turma inteira
 the class whole

Evidence that the control and raising complements with multiple subjects should be analysed along the same lines as the monoclausal examples with emphatic pronouns comes from the following interesting contrast in the distribution of non-referential QPs in raising vs. control complements. Unlike referring DPs (cf. (41a)), non-referential QPs may not occur in a “multiple” subject construction with a raising verb (cf. (41b)):

- (41) a. A empregada não apareceu e o hóspede acabou por
 the maid not showed-up and the guest ended up
 fazer ele o pequeno-almoço.
 to-do he the breakfast
 ‘The maid didn’t show up and the guest had to cook breakfast himself.’
 b. A empregada não apareceu, mas eu fui lá e
 the maid not showed-up but I went there and
 * nenhum hóspede acabou por fazer ele o pequeno-almoço.
 no guest ended up to-do he the breakfast
 ‘The maid didn’t show up but I went there and no guest turned out to prepare breakfast himself’

Our explanation of the contrast between (32) and (34) above carries over to the raising case: in (41a) the DP is a base-generated topic doubled by the

pronoun which is the “in situ” subject of the infinitival clause.

(42) [o hóspede [acabou por [fazer **ele** o pequeno-almoço]]

This configuration is not available in (41b) given that the non-referring expression *nenhum hóspede* ‘no guest’ cannot be a base-generated topic; the QP can only be merged as subject of the embedded clause and then raised to the matrix. For this reason, there is no position for the subject pronoun to be merged in the embedded clause.

Now, interestingly, control complements can take multiple subjects with a non-referential QP².

(43) Estou certa de que nenhum hóspede optará por fazer **ele**
 Am certain of that no guest will-choose by to-make he
 o pequeno-almoço todos os dias.
 the breakfast every the days
 ‘I am certain that no guest will choose to prepare his breakfast himself
 every day’

The contrast between (41b) and (43) can be easily explained under the assumption that the QP *nenhum hóspede* is merged as an argument of the matrix control verb. In this case, the pronoun is the subject of the infinitival clause and interpreted as a bound variable:

(44) [[nenhum hóspede]; optará [~~nenhum hóspede~~; optará [por fazer ele; o pequeno-almoço]]]

Now consider the alternative view that control is a special case of raising. Under this approach it is not at all clear how to explain the minimal pair formed by (41b) and (43). If the pronoun in (43) is the spell out of the raised subject then the ungrammaticality of (41b) is unexpected. Furthermore, the contrast between (41a) and (41b) is also left unaccounted for.

For this reason, I conclude that the paradigm discussed constitutes a

² Szabolcsi (2009) mentions the following Italian example, which minimally contrasts with (36b) above:

(i) *Context*: The teacher worked and ...
 Ogni ragazzo vuole lavorare anche lui
 every boy wants to work also he
 ‘Every boy wants it to be the case that he too works’

strong case not only in favor of a non-movement account of obligatory control but also in favor of the view that the pronoun is a subject of the infinitival complement. Our account of the difference between (41b) and (43) relies on the assumption that the pronoun is merged in the embedded clause as an argument and not as an adjunct. If it were an adjunct it should be able to co-occur with an explicit subject in (41b), contrary to fact³.

4. Analysis

4.1 Introduction

In her paper, Szabolcsi (2009) applies the above mentioned tests to a series of languages and detects the existence of three different patterns:

- (i) Languages that do not admit expressed subjects either in control or in raising; this is the case in English, German and French.
- (ii) Languages that allow for explicit subjects in raising and control complements; this is the case of Hungarian, Italian, Romanian, Turkish, Brazilian Portuguese (to which we can add EP now).
- (iii) Languages that allow for expressed subjects only in raising complements, like Russian, Finnish and (possibly) Hebrew.

Szabolcsi observes that the languages that fall under pattern (i) are non NSLs; the others admit some kind of subject *pro*-drop (Finnish has partial *pro*-drop and Russian has null expletives). On the other hand, all of the consistent NSLs studied by Szabolcsi are included in pattern (ii). The only exception is Brazilian Portuguese, a language that is undergoing a process of change toward becoming partial *pro*-drop (Rodrigues 2004). Even so, looking at Szabolcsi's sample, the correlation between consistent *pro*-drop and pattern (ii) can be stated as a one way implication, since there isn't a single consistent NSL in the sample that doesn't fall under pattern (ii). This one way implication is what I wish to explore here.

In this section, I start by presenting my assumptions regarding the syntactic properties that characterize consistent subject *pro*-drop (of the Romance type). Then, I show how this set of assumptions, combined with an Agree-based theory of Control, can be used to capture the three patterns described.

³ Torrego (1996) maintains the idea that the pronoun is an adjunct in these cases, but stipulates that the pronoun can only be an adjunct of a null element (a trace, *pro* or PRO). Still, the contrast between (41a) and (41b) remains a mystery.

4.2 Background assumptions

Barbosa 1995, Pollock 1997, Alexiadou and Anagnostopoulou 1998, Kato 1999, among others, proposed that, in consistent NSLs of the rich agreement type, Agr is invariably “(pro)nominal” in the sense that it has a D/N feature capable of checking the EPP in T. The particular implementations of this proposal vary, but the basic idea is that the set of *phi*-features in T (Agr) is [+N] and valued, thus behaving like a pronominal clitic for all purposes. As a consequence of this, the EPP is not checked by Merge/Move XP; in other words, there is no A-movement of subjects to pre-verbal position and the thematic position for the subject is the “in-situ” position inside the v/VP.

When the subject is a fully specified nominal, this yields a post-verbal subject construction given that v/V raises to T (so-called “free inversion”). Thus, the Portuguese example (45a) is analysed as in (45b):

- (45) a. Telefonou a Maria
 called the Maria
 b. [[T telefonou] [vP a Maria ~~telefonou~~]]

When the subject is silent, the configuration is the following, where *pro* has the same status as empty categories associated with pronominal clitics in general:

- (46) a. Telefonaram
 b. [[T telefonar-**am**_i] [vP *pro*_i]]

In this approach, apparent SVO constructions are instantiations of independently attested mechanisms of placing an argument in the front of the clause. One of them is Clitic Left Dislocation (henceforth CLLD). Again, I assume that Clitic Left Dislocated Topics are base-generated in a position of adjunction to the clausal projection that is predicated of them (see the arguments in Barbosa 2009). Thus (example (47a) is analysed as in (47b):

- (47) a. A Maria telefonou.
 b. [[A Maria]_i [TP telefonou [*pro*_i]]

In (47b) the DP *a Maria* is licensed by “rules of predication” (Chomsky 1977). In this case, the “open” position inside the complex predicate is supplied by *pro*.

Another mechanism of argument fronting found in the Romance NSLs is

direct movement to an A-bar position in front of the clause — known as Focus movement (see Martins 1994, Uriagereka 1995, Rizzi 1997) or Emphatic movement (see Raposo 1994, Zubizarreta 1998):

(48) [*DP* ... [_T V [~~*DP*~~ ...]]]

In Barbosa (1995, 2001, 2009) I argued that the non-referential QPs that cannot be left dislocated are fronted by A-bar movement whenever they appear in pre-verbal position (regardless of their status as subjects or objects). Thus, the following example is analysed as in (49b) (where FP stands for the neutral term “Functional Projection”):

(49) a. *Alguém telefonou.* ‘Someone called’
 b. [FP *alguém* ... [_T [*telefonou*] [VP ~~*alguém*~~ ...]]

Given that, in this view, the thematic position for the subject is post-verbal, the prediction is that, if the language allows left dislocation with doubling by an explicit pronominal subject, the pronominal doubler should be a post-verbal pronoun rather than a pre-verbal one. In effect, as seen in the previous section, such “double subject” constructions do exist. They are the structures with emphatic pronouns. As observed, these are incompatible with non-referential QPs, which suggests that they are instances of left dislocation. Doubling by an overt pre-verbal pronoun, by contrast, is extremely rare in the Romance NSLs even though it is very productive in French (cf. 33a) or Brazilian Portuguese (see Barbosa 2009 and the references cited there for discussion).

One other fact that is immediately captured under this analysis is the following. In all of the infinitival complements under discussion, the explicit subject must occur in post-verbal position. This observation is made in Torrego (1996) and Belletti (2005), who mention examples (51) and (52), respectively (in (50) I quote an example in EP):

- (50) *Não sabemos como (*nós os dois) assinar (nós os dois) a carta.*
 not know.1stPI how we the two to-sign we the two the letter
 ‘We don’t know how the two of us will sign the letter’.
- (51) **No sabemos si (*algunos de nosotros) asistir (algunos de nosotros)*
 not know.1stPL if some of us to-attend some of us
 ‘We don’t know if some of us should attend’
- (52) *I ragazzi risultarono [(*loro) aver (loro) riposto alla question]*
 the boys turned out they to have they answered to-the question

According to the analysis proposed here, pre-verbal subjects require a configuration of CLLD (cf. (47b), where the relation between the dislocated topic and the clausal projection that is predicated of it is mediated by silent *pro* licensed by (clitic-like) verbal agreement morphology. Since (non-inflected) infinitival clauses lack person and number agreement morphology, *pro* is not licensed in these environments. Hence, the configuration of CLLD is not available and so pre-verbal subjects are rightly predicted not to be possible in these contexts⁴.

Even though the theory sketched captures the absence of pre-verbal subjects in (non-inflected) infinitival clauses, it doesn't have an answer for why, in this particular kind of language, a subject may occur at all in these contexts, in contrast to what happens in other languages such as English, French or German. This issue will be taken up in the next sub-section.

4.3 Proposal

Belletti (2005) as well as Torrego (1996) propose to deal with the issue of the overt subjects of infinitivals in terms of the configuration of Clitic Doubling. This is also our take on the matter, even though the details of the analysis are substantially different.

Recall that, according to the theory of the Null Subject Property presented in the previous section, Agr has clitic-like properties, having a nominal feature capable of checking the EPP. Above, we claimed that this is the reason why the lexical subject merged in argument position inside the v/VP doesn't raise to Spec-TP, but we haven't been very clear about the relation that is established between the agreement affix and the lexical subject. Alexiadou and Anagnostopoulou (2001) argue that VS order corresponds to a Clitic Doubling configuration in Greek and Spanish though not in Italian. Here I would like to suggest that the data examined in this paper indicate that the Clitic Doubling approach is more generalized and should be extended to (at least) all of the Romance NSLs. In particular, I suggest that VS structures have properties in common with Clitic Doubling of datives in languages such as Spanish, as exemplified below:

- (53) No le ablé a nadie.
 not to him talked to noone 'I talked to noone'

As we can see, this kind of clitic doubling can occur with any type of DP

⁴ In Hungarian, the subject of the infinitival clause may appear in pre-verbal position, but this is a pre-verbal Focus position, which can also be occupied by focused objects (cf. Szabolcsi (2009)).

including non-referential quantifiers. Unlike Torrego (1995) and Belletti (2005), however, I do not assume a “big” DP approach, for the reasons sketched in section 2. In the VS case, the configuration in question is the following:

(54) [_{TP} [_T [_V [Agr]] [_{V/VP} *subject* ...]]
| *Agree* |

I assume that the agreement affix in merged in T with its *phi*-features valued, and values T's uninterpretable features. Since the agreement affix is itself uninterpretable in LF (it is not an argument), it acts as a probe, finds an active goal — the lexical subject — and Agree matches the features of the subject with those of the agreement affix. The basic idea is that, in the NSLs of the consistent type, the relation that is established between the argumental subject (be it null or overt) and T is not direct and is invariably mediated by the agreement affix. This idea will play a key role in our account of the occurrence of explicit subjects in obligatory control complements. In the next section, we will see how this proposal works for the raising and control complements⁵.

⁵ Torrego (1996) observes that the agreement relation established between the DP in post-verbal position and verbal agreement is, in some cases, semantic and not morphological, as witnessed by the following examples:

- (i) *Spanish*:
- a. Firmamos **los linguistas** / **muchos (de los linguistas)** la carta.
 signed.1st.Pl the linguists / many (of the linguists) the letter
 - b. Firmamos **la clase entera** la carta.
 signed the whole class the letter
- EP doesn't have the counterparts to (ia), but allows for (iia,b):
- (ii) a. Fomos embora **a turma inteira**.
 went.1st.pl away the whole class
- b. Odiámos ter de falar **o grupo todo com ele**.
 Hated to-have to talk the whole class with him
- Interestingly, the patterns of agreement found with subjects are replicated in the case of clitic doubling of objects:
- (iii) *Spanish (Torrego 1996)*
- a. **Nos** vio **a muchos de los linguistas**.
 us saw to many of the linguists
 - b. **Nos** vio **a los linguistas**.
 us saw to the linguists
 - c. **Nos** vio **a la clase entera**.
 us saw to the whole class

4.4 Applying the proposal to raising and control infinitives

Our approach to infinitival constructions is based on the nominal properties of the infinitival affix. Let us assume that infinitival Agr is [+N/D] in a NSL. The only difference between finite and infinitival clauses is that infinitival Agr is anaphoric, as originally proposed in Borer (1989).

4.4.1 Raising complements

Consider the following example:

- (55) Acabaram por receber **só elas / só as crianças** bons papéis.
ended up to receive only they / only the children good roles
'Only the children ended up receiving good roles'

I propose that the representation of (55) is as in (56):

- (56) [TP [T acabaram] [por [TP [T receber] [_{VP} [[só elas] / [só as c.] [bons ...]]]]]
- └────────── Agree ─────────┘

In this representation, the verbal agreement affix checks the EPP in the embedded clause as well as in the matrix; there is no expletive *pro* and the agreement affix in the matrix establishes a long distance Agree relation with the “in situ” subject of the embedded clause. I assume that raising complements are not strong phases.

Now recall that, according to A. Szabolcsi, the NSLs, Finnish and Russian pattern alike in this case. This fact can be explained if, in all of these languages, subjects may stay “in situ” (or in a Focus position inside the embedded clause) and establish a long distance Agree relation with matrix T. As matter of fact, Holmberg (2005) shows that subjects may stay “in situ” in

EP only has clitic doubling of objects with pronouns, QPs containing a pronoun and a restricted kind of morphologically singular collective DPs. With the latter, we find exactly the same patterns of agreement found with subjects:

- (iv) **Viu-nos à turma inteira / ao grupo** todo a copiar.
saw-us to-the class whole / to-the group all to cheat
'He saw the whole class / group cheating'

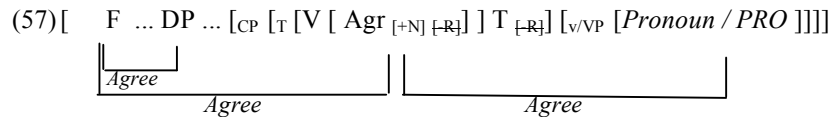
This parallelism between post-verbal subject constructions and clitic doubling of objects suggests that we are dealing with the same kind of structure.

Finnish and Bailyn (2004) makes a similar observation for Russian.

4.4.2 Obligatory Control

In obligatory control complements, if the subject of the infinitival is explicit, it must be anaphoric. Since, as shown above, the data are problematic for a *backward control* analysis, I adopt an Agree-based approach. In particular, I essentially follow Landau (2000, 2004), who adopts Borer's (1989) original suggestion that abstract Agr in infinitival T is anaphoric. In Landau's (2004) implementation of this idea, I/T contains a [-R] feature. In addition, Landau assumes that PRO is an anaphor (that is, it is [-R]) and that only a [-R] element can check the [-R] feature of I/T. The feature [-R] makes PRO active to enter an Agree relation with the matrix functional head that licenses the controller DP (T or *v*) and acts as an instruction to co-index the features of the antecedent with those of PRO.

Associating these ideas with our suggestion that infinitival Agr in a NSL is nominal and anaphoric, this means that, besides containing the [+N/D] feature, infinitival Agr is also [-R]. Thus, it can check off the [-R] feature in I/T as well as the EPP. In addition, it is active to establish an Agree relation with Agr or *v* in the matrix, thus co-indexing its features with those of the antecedent (I assume that this process suffices to delete the [-R] feature on the agreement affix). Since the [-R] feature in I/T is checked off by infinitival Agr, the argumental subject is free to be either [-R] or [+R], that is, it may be PRO or an overt pronoun. Agree between the affix and the subject argument (PRO or a lexical pronoun) guarantees coindexation between the antecedent in the matrix and PRO or the lexical pronoun:



In the languages where Agr is not nominal, however, the [-R] feature in T can only be checked by a [-R] subject argument and this is why only PRO is admitted. In other words, what enables the presence of the explicit pronoun in OC infinitives is the fact that it doesn't directly check any uninterpretable feature in T.

One obvious question raised by the above representation is why (57) is not a violation of the Phase Impenetrability Condition. Here I tentatively suggest that infinitival clauses are not strong phases, but I leave this issue for future work.

Recall that, in these cases, the subject of the infinitival clause cannot be

an R-expression, a restriction that we attributed to a violation of Condition C. I assume that coindexation is translated as variable binding so the pronoun is obligatorily interpreted as a bound variable.

5. Conclusions

In this paper, I have discussed evidence that control and raising infinitives have overt subjects in European Portuguese as well as in the other Romance NSLs. This evidence is in conformity with Szabolcsi's (2009) cross-linguistic study where the following generalization is put forward:

- (58) The overt subjects of control complements can only be pronouns. The overt subjects of raising complements can be pronouns or lexical DPs.

I have argued that the evidence underlying (58) constitutes a strong case for a non-raising approach to Control. Relying on the observation that all of the Romance Null Subject Languages allow for explicit subjects in raising and control complements (in compliance with (58)), I have developed an Agree-based theory of (58) that aims to capture the association between this phenomenon and the Null Subject Property. Drawing on previous proposals that verbal agreement morphology in the NSLs is "nominal", (cf. Barbosa 1995, Kato 1999 Alexiadou and Anagnostopoulou 1998), I suggested that the relation between a lexical subject "in situ" and T is invariably mediated by verbal agreement in a configuration of Clitic Doubling. This assumption, coupled with an Agree based theory of control such as Landau's (2000, 2004) captures the occurrence of explicit subjects of raising and control complements in these languages.

References

- Alexiadou, Artemis & Elena Anagnostopoulou (1998) Parametrizing AGR: Word Order, V-Movement and EPP-Checking, *NLLT* 16, pp. 491-539.
- Alexiadou, Alexiadou & Anagnostopoulou, Elena (2001) The subject *in situ* generalization, and the role of Case in driving computations, *Linguistic Inquiry*, 32, 193-231.
- Alexiadou, Alexiadou, Anagnostopoulou, Elena, Iordachioaia, G. & Marchis, M. (2008) 'A stronger argument for backward control', talk presented at NELS 39, Cornell University, November 2008.
- Anagnostopoulou, Elena. (1997) Clitic Left Dislocation and Contrastive Left Dislocation. In *Materials on Left Dislocation*, Anagnostopoulou, E., Henk Van Riemsdijk and Frans Zwarts (orgs.), pp.151:192. Amsterdam: John Benjamins

- Bailyn, J. F. (2004) Generalized Inversion. *Natural Language and Linguistic Theory* 22, 1-49.
- Barbosa, Pilar (1995) *Null Subjects*, Ph.D. dissertation, MIT, MITWPL, Cambridge, Mass.
- Barbosa, Pilar (2001) On inversion in *wh*-questions in Romance. *Romance Inversion*, ed. A. Hulk e J.-Y. Pollock, 20-90. New York: Oxford Press.
- Barbosa, Pilar (2009) Two kinds of subject *pro*. *Studia Linguistica* 63. Blackwell Publishing.
- Belletti, Adriana (2005) Extended doubling on the VP periphery. *Probus* 17-1: 1-35.
- Borer, Hagit (1989) Anaphoric AGR, in Osvaldo Jaeggli and Kenneth J. Safir (eds.), *The Null Subject Parameter*, Kluwer Academic Publishers, pp. 69-109.
- Burzio, Luigi (1986) *Italian Syntax*. Dordrecht: Reidel.
- Cardinaletti, Anna (1998) A second thought on *emarginazioni*: destressing vs. 'right dislocation.' *University of Venice Working Papers in Linguistics* 8.2., 1-28.
- Chomsky, Noam (1977) On *wh*-Movement. In *Formal Syntax* (P. Culicover, T. Wasos and A. Akmajian, editors), pp. 71-132. New-York: Academic Press.
- Cecchetto, Carlo (2000). Doubling structures and reconstruction. *Probus* 12: 93-126.
- Cinque, Guglielmo (1993) A null theory of phrase and compound stress. *Linguistic Inquiry* 24:239-267.
- Costa, João (1998) *Word order variation. A constraint-based approach*. The Hague: Holland Academic Graphics.
- Costa, João. (2004) Subjects in Spec,vP: locality and agree. In A. Castro, A., M. Ferreira, V. Hacquard & A. Salanova (orgs.) *Romance Syntax*, MIT Working Papers in Linguistics 46.
- Dermidache, Hamida (1992) *Resumptive Chains in Restrictive Relative Chains , Appositives and Dislocation Structures*. MIT Ph.D. Dissertation, MITWPL.
- Holmberg, Anders (2005) Is there a little *pro*? Evidence from Finnish. *Linguistic Inquiry* 36: 533-564.
- Hornstein, Norbert (1999) Movement and control. *Linguistic Inquiry* 30. 69-96.
- Landau, Idan (2000) *Elements of control: structure and meaning in infinitival constructions*. Dordrecht: Kluwer Academic Publishers.
- Landau, Idan (2004). The scale of finiteness and the calculus of control. *Natural Language and Linguistic Theory* 22: 811-877.
- Kato, Mary (1999) Strong and weak pronominals in the Null Subject Parameter. *Probus* 11, pp. 1-37, The Netherlands, Walter de Gruyter.
- Martins, Ana Maria (1994) *Clíticos na História do Português*. Ph.D. dissertation, University of Lisbon.
- Ordóñez, Francisco (1998) Post-verbal asymmetries in Spanish. *NLLT* 16:313-346.
- Piera, Carlo (1987) Sobre la estructura de las cláusulas de infinitivo. In V. Demonte & M. Lagunilla (orgs.) *Sintaxis de las Lenguas Románicas*. Madrid: Ediciones El Arquero: 148-163.
- Polinsky, Maria & Eric Potsdam (2002) Backward control. *Linguistic Inquiry* 33: 245-282.
- Pollock, Jean Yves (1997) *Langage et Cognition: Introduction au Programme Minimaliste de la Grammaire Générative*. Paris: Presses Universitaires de France.

- Raposo, Eduardo (1994) Affective operators and clausal structure in European Portuguese and European Spanish. Unpublished ms., University of California at Santa Barbara, California.
- Rigau, Gemma (1987) Sobre el carácter quantificador de los pronombres tónicos en Catalán. In Violeta Demonte and Marina Fernández Lagunilla (eds.), *Sintaxis de las lenguas Románicas*. Madrid: Textos Universitarios.
- Rizzi, Luigi (1978) A Restructuring Rule in Italian Syntax. In S. J. Keyser (ed.), *Recent Transformational Studies in European Languages*. Cambridge: MIT Press.
- Rodrigues, Cilene (2004) *Impoverished Morphology and A-movement out of Case-domains*. Ph.D Dissertation, University of Maryland.
- Sola, Jaume (1992) *Agreement and Subjects*. Ph.D. dissertation, Universitat Autònoma de Barcelona.
- Szabolcsi, Anna (2009) Overt nominative subjects in infinitival complements cross-linguistically: Data, diagnostics, and preliminary analyses. NYU Working Papers in Linguistics, Vol. 2. Papers in Syntax, Spring 2009.
- Torrego, Esther (1995) On the nature of clitic doubling. In *Evolution and Revolution in Linguistic Theory*, Hector Campos and Paula Kempchinsky (orgs.), 251-275. Washington D.C.: Georgetown University Press.
- Torrego, Esther (1996) On quantifier float in control clauses. *Linguistic Inquiry* 27-1: 111-126.
- Torrego, Esther (1998) *The Dependencies of Objects*. Canbridge, Mass: The MIT Press
- Uriagereka, Juan (1995) Aspects of the Syntax of Clitic Placement in Western Romance. *Linguistic Inquiry* 26: 79-123.