

# On the Position of Preverbal Subjects in Spanish. A View from the Discourse Perspective

José Camacho, jcamacho@rutgers.edu

November 1, 2011

*Abstract.* This article advances a proposal about the position of preverbal subjects (PS) in null subject languages based on the discourse co-construal possibilities of PS and CLLD as antecedents to pronominal subjects (overt and null). Based on an antecedent choice preference task conducted with 68 native speakers of Spanish, we conclude that PS have mixed structural properties: they can either be topics or bound topic variables. By contrast to other mixed-structural approaches to PS, it is argued that the two possible PS positions do not correlate with the strong/weak partition initially proposed by Cardinaletti and Starke (1999).

## 1 Introduction

The status of subjects in null subject languages (NSLs) like Spanish has been actively debated over the past decades (cf. Rizzi, 1982; Rigau, 1988; Contreras, 1991; Koopman and Sportiche, 1991; Solà, 1992; Zubizarreta, 1994; Barbosa, 1995; Barbosa et al., 2005; Cardinaletti, 1997, 2004; Alexiadou and Anagnostopoulou, 1998; Ordóñez and Treviño, 1999; Goodall, 1999; Suñer, 2003; Camacho, 2006, 2011; Frascarelli, 2007; Sheehan, 2007, among others). For Spanish, one central divide among proposals is whether PS occupy a canonical, A-bar, Spec, TP position (cf. Goodall, 1999 and Suñer, 2003, for example), or whether they are dislocated and peripheral (cf. Contreras, 1991; Alexiadou and Anagnostopoulou, 1998; Ordóñez and Treviño, 1999 among others).<sup>1</sup>

---

<sup>1</sup>I would like to thank Christina Hunt for invaluable help in codification of the data, and Liliana Sánchez for discussion on the methodology and content of the paper.

A second important issue relates to whether all PS are created equal. In this sense, Cardinaletti (1997, 2004) proposes that although the position of subjects is the same in NSLs and non-NSLs like English, weak pronominal PS occupy a lower structural position from strong pronominal and DP PS.

In this paper, I approach this debate from a slightly different perspective. If PS are left-dislocated, then one would expect them to equally serve as potential discourse antecedents for other subjects. Likewise, when pronominal subjects are co-construed with an antecedent in a preceding clause, one should not find differences between weak, overt pronominals and weak null pronominals. In order to test these assumptions, we analyze experimental evidence both for discourse antecedent choice and for acceptability of null vs. overt pronouns when the subject is co-construed with another subject, and find that CLLD were selected significantly more frequently than PS as antecedents, and that null subjects co-construed with a preceding subject are significantly more acceptable than overt, weak ones. We conclude from the first result that PS subjects have two possible positions, one similar to CLLD items, another one in Spec, TP, and from the second result, that these two positions do not correlate with the strong/weak position, and that overt, weak pronominals can appear in either position.

The evidence also argues in favor of Chomsky's (1981) **Avoid Pronoun Principle**, which states that overt pronouns should be dispreferred whenever possible (i.e. whenever a null counterpart is available). Our analysis suggests that this principle can be restated in the following terms: null pronouns are preferred because they correspond to an unambiguous structure, whereas overt ones are structurally ambiguous.

## 2 The position of preverbal subjects

### 2.1 An invariant position for PS

Cardinaletti (1997, 2004) argues that NSLs and non-NSLs have identical structural positions for subjects. In particular, she denies that the existence of null subjects somehow allows for the dislocation of overt preverbal subjects, as some prominent alternatives suggest (see below). In her proposal, PS can indeed occupy distinct positions, but this option cuts across the null subject status of the language, and relates to whether the subject is a DP/strong pronoun or a weak pronoun. DPs/strong pronouns occupy a higher position than weak ones (including pro), as seen in (1).

- (1) [SubjP DPs/strong pronouns [AgrSP weak pronouns/pro]]

The evidence for Cardinaletti's uniform analysis of PS positions across languages comes from a number of language-internal and crosslinguistic facts. For example, she observes that in most languages, weak pronouns must be adjacent to the verb in I, as indicated by the fact that they cannot be separated by a parenthetical. This is illustrated in (2), where (2a) has a strong DP/pronominal separated from the verb in I by a parenthetical, whereas in Dominican Spanish, the expletive *ello* 'it' cannot (cf. (2b) and Suñer, 2003). This suggests that in Dominican Spanish, the preverbal field looks like (2c). The same argument can be reproduced for Italian and French weak pronouns.

- (2) a. Juan/Él, a mi parecer, es muy simpático.  
 Juan/he, according to me is very nice  
 'Juan/He, according to me, is very nice.'
- b. Ello (\*a mi parecer) no sería malo estudiar.  
 it according to me, not would be bad to study  
 'It (according to me) wouldn't hurt to study.'
- c. [TP {Juan/él} [XP parenth. [TP {ello/pro} V]]] (Dominican Spanish)

In order to show that *pro* is indeed a weak pronoun in AgrSP, Cardinaletti (following Burzio, 1986 and Rizzi, 1987) first argues that *pro* is always preverbal, using the distribution of floated quantifiers and agreement patterns in copular sentences with nominal predicates. She then notes the existence of a parallelism between *pro* and overt expletives in non-NSLs. Thus, only weak pronouns, not strong ones can appear in contexts where the subject is not referential in French (cf. (3)). Likewise, in Spanish, only *pro* is possible in those contexts, as seen in (4).<sup>2</sup>

- (3) a. Il/\*lui est arrive trois filles.  
 it/he is arrived three girls  
 'Three girls arrived.'
- b. Il/\*lui pleut.  
 it rains  
 'It is raining'
- c. Ils/\*Eux m'on vendu un livre.  
 they CL-have sold a book  
 'They have sold me a book.' (French)

<sup>2</sup>(4c) is not possible in the impersonal interpretation.

- (4) a. *Pro/\*ello* llegaron tres niñas.  
*pro/it/he* arrived three girls  
 ‘Three girls arrived.’  
 b. *Pro/\*ello* llueve.  
*pro/it* rains  
 ‘It is raining’  
 c. *Pro/#ellos* me vendieron un libro.  
*pro/they* CL sold a book  
 ‘They sold me a book.’ (Spanish)

It is worth noting that Cardinaletti’s proposal of two distinct structural positions for subjects does not necessarily correlate those positions with interpretive differences, in particular in terms of topichood. Rather, Cardinaletti’s two positions follow from prosodic, semantic and structural differences between weak and strong pronouns, not from discourse-interpretive distinctions, by contrast to Frascarelli’s (2007) proposal to be discussed below.

## 2.2 The CLLD analysis

Contreras (1991); Barbosa (1995); Alexiadou and Anagnostopoulou (1998) and Ordóñez and Treviño (1999), among others, argue that PS are clitic left-dislocated constituents akin to left-peripheral direct and indirect objects, illustrated in (5). In the stronger version of this proposal, PS are obligatorily dislocated in NSLs, whereas in non-NSLs PS occupy the “canonical” subject position, let us assume for the sake of argument that it is Spec, TP, as schematically represented in (6).

- (5) a. A Pedro, le regalaron un libro.  
 to Pedro, CL gave a book  
 ‘Pedro, they gave him a book.’  
 b. A la hermana de Susana, la vi el otro día.  
 to the sister of Susana, CL saw the other day  
 ‘Susana’s sister, I saw her the other day.’ (Spanish)

- (6)  $\text{Subj}_i$  [<sub>TP</sub> *pro*<sub>*i*</sub>/CL<sub>*i*</sub> [<sub>I</sub> V [<sub>VP</sub> . . . ]]]

This analysis builds on the availability of null subjects in NSLs. The lexical DP is dislocated because there is a null resumptive category in the canonical subject position. This category is either a null *pro*, as in (6), or the inflection morpheme, as proposed by Ordóñez and Treviño (1999), following Taraldsen (1992) and Baker (1995).

Proponents of the PS-as-CLLD analysis point out to the parallelism between both types of constituents, among them relative scope patterns (cf. Barbosa, 1995; Alexiadou and Anagnostopoulou, 1998, henceforth A & A, and Ordóñez and Treviño, 1999). For example, A & A (p. 504) note that quantificational and indefinite PS have unambiguous scope, as CLLD elements do. Thus, in (7a), the indefinite subject can only have wide scope in Greek, as it does in Spanish.<sup>3</sup>

- (7) a. Kapios fititis stihiothetise kathe arthro.  
 some student filed every article  
 ‘Some student filed every article.’ (from A & A, 505, ex. 20a, Greek)
- b. Algún/un estudiante archivó todos los artículos.  
 some/a student filed every article  
 ‘Some student filed every article.’ ( $\exists > \forall$ )
- c. Algún/un artículo lo archivaron todos los estudiantes.  
 some/an article CL filed all the students  
 ‘An article, all the students filed it.’ ( $\exists > \forall$ )

Suñer (2003, 344, ex. 5), however, points out that PS can have narrow scope, and in fact she argues that this interpretation (illustrated in (8b)) is the preferred one for the example in (8a). As Sheehan (2007, 57-58) notes, in addition to Spanish, European Portuguese, Romanian and Italian can have PS with narrow scope.

- (8) a. En la biblioteca departamental, algún estudiante sacó prestado  
 in the library departmental, some student took loaned  
 cada libro.  
 each book.
- b. ‘In the departmental library, each book was taken out by some student.’ ( $\exists > \forall$ )
- c. ‘In the departmental library, some (specific) student took out each book.’ ( $\forall > \exists$ , dispreferred in Spanish)

Finally, if a PS indefinite is in an A-bar position and cannot reconstruct (like

---

<sup>3</sup>Notice that the scope in these examples corresponds to the surface order of the quantifiers. If the lower quantifier remains in situ, in (7b), one would predict that the higher quantifier would always have scope over it, even if it reconstructed to Spec, VP. In (7c), on the other hand, if the CLLD constituent is base-generated in its peripheral position, as A & A assume, it follows that it will not reconstruct. So it is possible that the similar scope pattern shown by CLLD elements and PS stems from independent principles.

CLLD items), it should not have scope over a modal or negation, assuming that modals do not raise covertly. This is illustrated in (9), after A & A's examples, where the only possible interpretation is *many* > *not*.<sup>4</sup>

- (9) Muchos hombres no se enamoraron de María.  
 Many men not CL fell in love of Mary  
 'Many men didn't fall in love with Mary.' (Spanish)

It should be noted that the strongest version of this analysis would argue that preverbal subjects have exactly the same properties of CLLD items. In particular, if one assumes that CLLD items occupy a dedicated structural position with certain interpretive properties (most notably, being a topic), then preverbal subjects fall within that category and interpretation.

Just as PS and CLLD pattern in similar ways, they also diverge in a number of other ways. For example, Goodall (1999) observes that PS cannot be bare (cf. (10a)), whereas topics can (cf. (10b)).<sup>5</sup>

- (10) a. \*Niños jugaban en el parque.  
 children were playing in the park  
 b. Marta, niños no cuida.  
 Marta, children not looks after  
 'Marta doesn't look after children.' (Spanish)

Zubizarreta (1994) and Suñer (2003) also point out that PS and CLLD are interpreted in different base positions. Thus, according to Suñer (2003), PS do not reconstruct to a lower position in VP, whereas CLLD do, as seen in her examples in (11). In (11b), the DP inside the PS (*Hugo*) can be coindexed with the clitic, whereas that same coindexing is impossible if the subject is postverbal (cf. (11a)).

- (11) a. \*Lo<sub>i</sub> defendió [una amiga de Hugo<sub>i</sub>]  
 CL defended a friend of Hugo  
 b. [Una amiga de Hugo<sub>i</sub>] lo<sub>i</sub> defendió.  
 a friend of Hugo CL defended

<sup>4</sup>Data on modals are less clear-cut than with negation.

<sup>5</sup>Notice, however, that this property must be independent of whether the subject is a topic or not, since in (i) a clearly dislocated subject with a topic interpretation cannot be bare.

(i) \*Niños, con Marta, no juegan nunca en el parque.  
 children, with Marta, not play never in the park

‘A friend of Hugo’s defended him.’ (from Suñer, 2003, 349,  
Spanish)

The pattern with CLLD constituents is the opposite, as shown in (12), adapted from one of Suñer’s examples. In that case, if there is a postverbal subject, coindexing of it with the DP inside the CLLD constituent is possible (cf. (12a)), but if the subject is null (by her assumption, a preverbal *pro*), coindexing with a CLLD item is impossible (cf. (12b)).<sup>6</sup>

- (12) a. [La primera página de las novelas de Cortázar<sub>i</sub>] siempre la  
the first page of the novels of Cortazar always CL  
corregía él<sub>i</sub> en ayunas.  
proofread he before breakfast  
‘The first page of Cortazar’s novels, he always proofread it before  
breakfast.’
- b. ??[La primera página de las novelas de Cortázar<sub>i</sub>] *pro*<sub>i</sub> siempre la  
the first page of the novels of Cortazar always CL  
corregía en ayunas.  
proofread before breakfast  
‘The first page of Cortazar’s novels, he always proofread it before  
breakfast.’ (Spanish)

These contrasts follow, according to Suñer, from the fact that the PS occupies an A-position, and leaves no trace when moving, hence it cannot reconstruct. For this reason, in (12a) there will no Principle C violation, whereas in (12b), the DP inside the subject is lower than the coindexed clitic, hence Principle C is violated.

To sum up, the CLLD-as-PS view relies on several ideas. For our purposes, the most important two are a) the notion that a NS is potentially connected to an overt subject and b) the idea that overt PS and CLLD are both topics. Together, these two assumptions predict that CLLD constituents and PS should pattern the same when they are antecedents of a null subject, a prediction we tested in an experiment described in section 5.

### 2.3 Frascarelli (2007)

Frascarelli (2007) presents a proposal that shares some of the properties of the

---

<sup>6</sup>Not all speakers share these judgements, in particular, for some speakers (12a) is ungrammatical under the intended reading. My own judgements agree with Suñer’s.

PS-as-CLLD and of Cardinaletti’s account. For her, weak subjects (null and intonationally weak pronouns) belong to the T-domain and more importantly for our purpose, are bound by an “aboutness-shift Topic.” This topic introduces new discourse information or shifts from the preceding topic. By contrast, intonationally strong pronouns (and full DPs), are not bound by a Topic, rather, they are in the C-domain and can themselves be Topics or Foci. The resulting representations for weak, strong and null subjects are given in (13).<sup>7</sup>

- (13) a. Weak pronouns: [<sub>CP</sub> [<sub>TopP</sub> Top<sub>i</sub> [<sub>TP</sub> *pro*<sub>i</sub>/*she*<sub>i</sub> . . . ]]]  
 b. Strong pronouns [<sub>CP</sub> [<sub>TopP</sub> SHE [<sub>TP</sub> . . . ]]]

Frascarelli assumes that overt PS are in an A’ position, particularly when they serve as antecedents to a NS (p. 714) and also that CLLD elements can be either an Aboutness-Shift Topic (i.e introduce a new topic), or a continuity Topic (Familiar Topic in Frascarelli’s terms). This part of Frascarelli’s analysis is fairly close to the PS-as-CLLD analysis.

She suggests that the discourse division of labor between weak and strong subjects is a consequence of certain general linguistic principles. In particular, she argues that Chomsky’s Avoid Pronoun principle in (14) and Economy requirements “lead to the expectation that in a NS language (like Italian) overt subjects are only realized for emphasis/contrast or to ensure recoverability in discourse” (Frascarelli (2007, 694)). Similar conceptions have been adopted by Rizzi (1997) and by Cardinaletti (2004).

- (14) **Avoid Pronoun** (cf. Chomsky, 1981 formulated by Frascarelli, 2007, 694)  
 Avoid overt pronoun, whenever possible.

## 2.4 Summary

To summarize the preceding sections, two competing analyses have been proposed for PS in NSLs. On the one hand, they have been argued to be CLLD constituents in the default case. On the other, Cardinaletti has argued that subjects occupy different structural positions: Strong pronominal subjects and overt DPs are located higher than weak pronouns and *pro*, but neither of those positions are peripheral, A’-positions. Finally, Frascarelli explicitly correlates the position of strong PS

<sup>7</sup>Although Frascarelli accepts Cardinaletti and Starke’s (1999) division between weak and strong pronouns, she does not formalize in structural terms, but rather in terms of prosodic strength.

with topics, and weak subjects with topic-bound categories.

### 3 Predictions

The strong version of the PS-as-CLLD analysis equates subjects and CLLD items, and more specifically, assumes that they are topics, hence it predicts that both should act in similar fashion as potential antecedents in contexts such as (15).

- (15) a.  $[_{CP1} [_{TopP} \text{Subj}_i \dots]] [_{CP2} [_{TP} \text{pro}_i/\text{she}_i \dots]]$   
 b.  $[_{CP1} [_{TopP} \text{CLLD}_i \dots]] [_{CP2} [_{TP} \text{pro}_i/\text{she}_i \dots]]$

Second, since weak pronouns (null and overt unstressed ones) are topic bound (particularly in Frascarelli's view), they should maintain the topic of the preceding clause as has been repeatedly proposed for null subjects. To see why this is the case, consider the structure in (16). The first clause has a possibly overt topic, followed by the second clause with a null topic that binds the null subject. One can safely assume that the default situation would favor topic continuity in this context.

- (16)  $[_{CP1} [_{TopP} \text{DP}_i \dots]] [_{CP2} [_{TopP} \emptyset_i [_{TP} \text{pro}_i \dots]]]$

Third, if Chomsky's Avoid Pronoun principle holds, one should find a clear contrast between overt and null pronouns in co-construal, topic continuity contexts. What is less clear is whether this contrast also holds between strong and overt unstressed pronouns.

In sum, the different analyses make the predictions presented in ((17)).

- (17) a. CLLD items and subjects should both pattern alike as subject antecedents.  
 b. Weak pronouns (overt unstressed and null) should pattern alike with respect to the antecedents they take.  
 c. Both weak overt and null pronominal (weak) subjects should maintain the topic from the preceding clause.  
 d. Chomsky's Avoid Pronoun principle predicts preference for null pronouns under topic continuity.

In order to test these predictions, we conducted an experiment to investigate speakers' discourse choice under different conditions. Before describing it in detail, I will review previous studies of discourse choice in Italian and Spanish.

## 4 Antecedents of null subjects. Previous studies

Several researchers have looked at antecedent choice for null subjects, in cross-clausal contexts (cf. Carminati, 2002; Alonso-Ovalle et al., 2002; Sorace and Filiaci, 2006; Jegerski et al., 2011). Alonso-Ovalle et al. (2002) replicating Carminati's results for Spanish, conclude that speakers preferred for *pro*'s antecedent to be in Spec, IP (regardless of the nature of that antecedent), in the sense of a preverbal subject. However, this study did not take into account the discussion in the preceding section, therefore did not explicitly include CLLD constituents as potential antecedents.

Alonso Ovalle et al. added another variable to their experiment, namely the distinction between stressed/unstressed pronouns. By doing this, they wanted to clarify whether the relevant feature that determines antecedent preference was structural position or some type of lexical content. Thus, they tested coreference patterns in examples such as (18), which included an unstressed pre- or postverbal pronoun and a stressed, contrastive postverbal one.

- (18) a. Pedro piensa que está cansado él.  
'Peter thinks that he is tired.'
- b. Pedro piensa que él está cansado.  
'Peter thinks that he is tired.'
- c. Pedro piensa que está cansado ÉL.  
'Peter thinks that HE is tired.' (from Alonso-Ovalle et al., 2002, 23, ex. 13, Spanish)

Their results suggest that coreference is highly associated with the preverbal subject position, regardless of whether the pronoun is stressed or not.

Both Carminati and Alonso Ovalle et al.'s results are partially compatible with Frascarelli's analysis. If overt DP subjects are Topics (as Frascarelli assumes), and NSs are topic-bound, then it follows that NSs should be easily co-construed with a subject.

## 5 The experiments

Our experiments were designed to test two specific items: Whether strong version of the PS-as-CLLD hypothesis holds, and whether the strong and weak distinction correlates with distinct structural positions. Specifically, Frascarelli's analysis proposes that null/weak subjects are topic-bound, whereas stressed, strong ones

are topics or foci. For null/weak subjects, this entails topic continuity (i.e. co-construal with a topic in the preceding clause, typically the subject), whereas for strong, overt subjects it does not.

In order to test these two consequences of the analysis, we conducted two experiments. The first one asked participants to select a pronominal subject's antecedent in a preceding clause, focusing in particular on whether the antecedent was an overt subject or a CLLD item, as schematized in (19).

(19) [Clause1 CLLD<sub>i</sub> . . . S<sub>j</sub> . . .] [Clause2 *pro*<sub>i?</sub>/<sub>j?</sub> / *he*<sub>i?</sub>/<sub>j?</sub> / HE<sub>i?</sub>/<sub>j?</sub> . . .]

The second one asked participants to rate items with two clauses, one with an overt DP subject, one with a pronominal subject, interpreted as co-construed, as in (20).

(20) [Clause1 S<sub>j</sub> . . .] [Clause2 *pro*<sub>i</sub> / *he*<sub>i</sub> . . .]

## 5.1 Experiment 1: Antecedent choice for pronominal subjects

### 5.1.1 Methods and participants

In the first experiment, participants listened to two clauses such as the ones in (21a). After listening to them, they had to answer a question on the screen about the potential antecedent of the second clause, as in (21b). The possible answers in (21c) were given in a multiple choice format, and participants could choose more than one of those responses. Each example had three variants, one with an overt unstressed pronoun, one with a stressed pronoun, and one with a null pronoun. All of the examples were presented online orally, randomized, and participants could listen to them as many times as they wanted.

- (21) a. A Juan lo llevó Pedro. Ahora *pro* / *el* / *ÉL* está esperando.  
to Juan CL took Pedro. Now / he / HE is waiting. Now (he) / HE is waiting.  
'Juan, Pedro took him. Now, he is waiting.'
- b. ¿Quién está esperando?  
'Who is at home?'
- c. (i) Juan  
(ii) Pedro  
(iii) Nadie 'no one'  
(iv) Otra persona 'another person'

Each of the initial clauses contained two possible animate antecedents. Participants heard a total of 21 items with S/CLLD potential antecedents, equally divided between null, overt unstressed and stressed subjects. Additionally, 12 distracters were also included. Of the original 85 participants, 68 completed most of the questions in this experiment.

### 5.1.2 General results for antecedent choice and subject status

The table in (22) presents general results for antecedent choice (S or CLLD) across all subject types. As seen, CLLD antecedents were selected much more frequently than subject antecedents (65% vs. 31%),

(22) Antecedent choice (S/CLLD) for all subjects

	Selected	%	N
CLLD antecedent	922	65	1428
Subject antecedent	442	31	1428
Other antecedent	64	4	1428

If we look at more specific results, the table in (23) represents antecedent choice (S or CLLD) depending on whether the subject was overt or null. Once again, CLLD antecedents were more frequently selected. For overt pronouns, CLLD antecedents were substantially and significantly more frequent ( $M = .41, SE = .013$ ) than subject antecedents ( $M = .22, SE = .011$ ),  $t(1427) = -9.86, p < .001$ ). For null pronouns, CLLD antecedents were even more frequent ( $M = .72, SE = .021$ ) than subject antecedents ( $M = .25, SE = .020$ ),  $t(475) = -13.03, p < .001$ )

(23) Antecedent choice (S v. CLLD) for overt and null pronominal subjects

Antecedent	Overt pronoun (N=952)		Null pronoun (N=476)	
	CLLD	581	61%	341
Subject	320	34%	121	25%
Other	51	5%	14	3%
Total	952	100%	476	100%

In the tables in (24) we can see antecedent-choice results for stressed and unstressed pronominal subjects (including both overt unstressed and null). When

the subject was a stressed (strong) pronominal subject, a CLLD antecedent was selected more frequently ( $M = .57, SE = .023$ ) than a subject antecedent ( $M = .37, SE = .022, t(475) = -5.11, p < .001$ ). This difference was more widespread when the subject was unstressed: CLLD antecedents were selected even more frequently ( $M = .68, SE = .015$ ) than subject antecedents ( $M = .28, SE = .015$ ),  $t(951) = -15.57, p < .001$ ).

(24) Antecedent choice (S/CLLD) for stressed vs. unstressed pronominal subjects

Antecedent	Stressed pronoun (N=476)		Unstressed pronoun (N=952)	
	CLLD	271	57%	651
Subject	176	37%	265	28%
Other	29	6%	36	4%
Total	476	100%	952	100%

Finally, the tables in (25) present results for antecedent choice (S or CLLD) for overt unstressed pronominal subjects and null pronominal subjects. Once again, when the subject was overt, unstressed, CLLD antecedents were more chosen significantly more frequently ( $M = .57, SE = .023$ ) than subject antecedents ( $M = .37, SE = .022$ ),  $t(475) = -5.11, p < .001$ ). As mentioned earlier, for null pronominal subjects, the CLLD antecedent was also statistically more frequent than the subject antecedent (see above).

(25) Antecedent choice (S vs. CLLD) for overt unstressed vs. null subjects

Antecedent	Overt unstressed pronoun (N=476)		Null pronoun (N=476)	
	CLLD	310	65%	341
Subject	144	30%	121	25%
Other	22	5%	14	3%
Total	476	100%	476	100%

These results are summarized in figure (25).

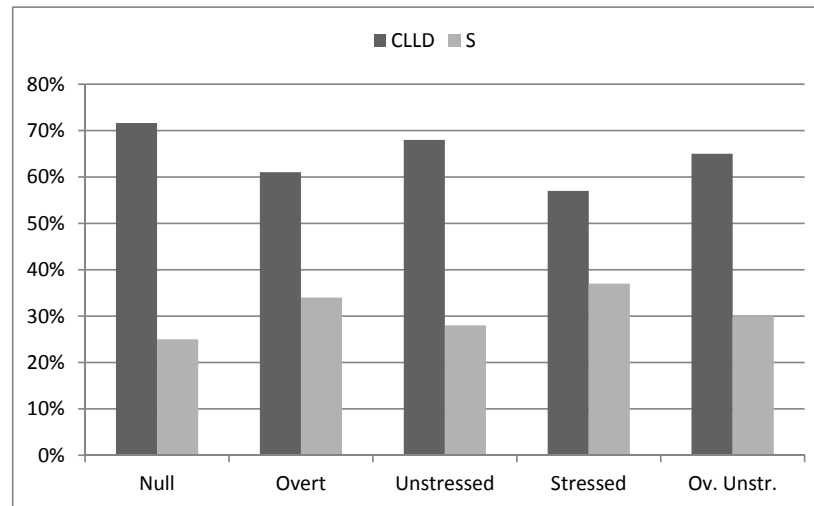


Figure 1: Percentage of antecedent choice (S vs. CLLD) by type of pronominal subject

### 5.1.3 Antecedent choice and subject status. Discussion

The results presented in the preceding section argue against the predictions in (17) above, repeated below in (26). Starting with prediction (26a), the fact that antecedent CLLDs were much more frequent antecedent subjects both in general and for every subject type suggests that CLLD and subjects do not have the same antecedent status and hence do not always occupy the same position.

- (26)
- a. CLLD items and subjects should both pattern alike as subject antecedents.
  - b. Weak pronouns (overt unstressed and null) should pattern alike with respect to the antecedents they take.
  - c. Both weak overt and null pronominal (weak) subjects should maintain the topic from the preceding clause.
  - d. Chomsky's Avoid Pronoun principle predicts preference for null pronouns under topic continuity.

One natural alternative hypothesis to make is that antecedent CLLDs are always topics, as several authors have suggested, whereas preverbal subjects are not.<sup>8</sup> The proposed structure for CLLD antecedents is presented in (27). In the first sentence (CP1), DP<sub>CLLD</sub> is presented as a topic, therefore in the second clause, it will be the natural topic antecedent. If, as Frascarelli (2007) suggests, *pro* is topic-bound, we predict that *pro* will overwhelmingly chose a CLLD as antecedent when available.

(27) [CP1 [TopIP DP<sub>CLLD-i</sub> . . . ]] [CP2 [Top2P  $\emptyset_i$  [TP . . . *pro\_i*/ella<sub>i</sub> . . . ]]]

In the case of an antecedent subject, the situation is more complicated. In general, if Frascarelli is correct, the subject can occupy two positions, one as a new topic, one as a bound topic. As a new topic in the first clause (cf. (28a)) the antecedent S will pattern like a CLLD antecedent. However, there is an alternative possibility, illustrated in (28b), where the subject does not introduce a new topic, but sits in Spec, TP. In this structure, the subject will not be an easy antecedent in the subsequent clause.

(28) a. [CP1 [TopIP DP<sub>S-i</sub> . . . ]] [CP2 [Top2P  $\emptyset_i$  [TP . . . *pro\_i*/ella<sub>i</sub> . . . ]]]  
 b. [CP1 [TopIP [TP1 DP<sub>S-i</sub> ]]] [CP2 [Top2P  $\emptyset_j$  [TP . . . *pro\_j*/ella<sub>j</sub> . . . ]]]

Note, however, that unlike Frascarelli, we do not assume that the new topic/bound topic-variable distinction correlates with the strong/weak pronominal distinctions, for reasons that will become clear below. Rather, we assume that the Spec, TP position is a non-topic position, by contrast to TopP.

If we turn now to the implications of the weak or strong nature of the pronominal for antecedent choice, recall that Cardinaletti (1997, 2004) and Frascarelli (2007) consider both null and overt, unstressed pronouns weak, as opposed to stressed or phonologically strong ones. For Cardinaletti, weak pronouns occupy a structural position closer to TP, whereas the strong pronouns are higher. For Frascarelli, on the other hand, strength is a matter of prosodic content. From the point of view of discourse interpretation, Frascarelli proposes that weak pronouns are bound by a topic and should therefore pattern the same with respect to the type of antecedent they select. This leads to the predictions in (26b)-(26c).

The results we have just presented do not seem to lend support to the discourse correlates of weak/strong pronouns (in Frascarelli's analysis). First, weak pronouns do not pattern like with respect to the antecedents they take (prediction (26b)). As seen, CLLD antecedents were more frequent with null pronouns than

<sup>8</sup>With the right stress, they can be contrastive topics as well, but those fall beyond scope of this paper.

with overt unstressed pronouns (72% vs. 65%), whereas for S antecedents, the tendency is the opposite (25% with null pronouns vs. 30% with overt unstressed ones). This, on the other hand, is consistent with the analysis proposed above, where PS are ambiguous but CLLD are always topics.

Nevertheless, these numbers cannot be statistically compared, so they are not conclusive. In the following section, I report the results of the second experiment, where participants were asked to rate sentences when subjects were coreferent. This experiment specifically addresses topic continuity.

## 5.2 Experiment 2: Acceptability of coreferential subjects

In the second experiment, participants were asked to rate the acceptability of a paragraph that contained two sentences with coreferential subjects. About half of those sentences included an overt, unstressed pronominal, and half a null one, as illustrated in (29). These minimal pairs were not presented sequentially. After each item, participants were asked a question such as: “if Maria is also the person who returned last night, this example sounds. . .” and then they were given a scale of 1-4: ‘Totally acceptable’ (1), ‘somewhat acceptable’ (2), somewhat unacceptable (3) and ‘totally unacceptable’ (4).

- (29) a. María salió. Ella volvió anoche.  
 Maria left she returned last night  
 ‘Maria left. She returned last night’  
 b. María salió. Volvió anoche.  
 Maria left returned last night  
 ‘Maria left. She returned last night’

A total of 41 items were selected, 20 with an overt subject, 21 with a null subject. The participants were the same as for the preceding experiment. Overall results and results for null and overt pronominal subjects are presented in the table in (30).

(30) Acceptability of coreferential subjects

	Overall		Overt pronominal		Null pronominal	
	N=2870	SD	N=1400	SD	N=1470	SD
Acceptability	1.68	.92	1.92	1	1.46	.75

As this table shows, the overall acceptability rate is 1.68, somewhere between to-

tally acceptable (1) and somewhat acceptable (2). However, acceptability of coreferential subjects is statistically significantly lower for overt pronominal subjects (1.92) than for null pronominal subjects (1.42) ( $F(1, 2868) = 190.3, p < .001$ ).

We wanted to test whether discourse cohesiveness affected co-construal acceptability, so some of the examples included a temporal adverb, as illustrated in (31), whereas others (like (29) above) did not, and (30) show results when that distinction is taken into account.<sup>9</sup>

- (31) a. Julio se cayó. Después lloró unos minutos.  
 Julio CL fell. after cried some minutes  
 ‘Julio fell. After, he cried for some minutes’  
 b. Julio se cayó. Después él lloró unos minutos.  
 Julio CL fell. after he cried some minutes  
 ‘Julio fell. After, he cried for some minutes’
- (32) Acceptability of coreferential subjects depending on presence of temporal adverbs

	With temporal adverb			With out temporal adverb		
	Total	Overt prn	Null prn	Total	Overt prn.	Null prn.
N	700	420	280	2170	980	1190
Accept. (SD)	1.85 (1)	2.09 (1.1)	1.49 (.78)	1.63 (.88)	1.84 (.97)	1.45 (.74)

As seen in this table, participants rated co-construed subjects less acceptably without a temporal adverb (1.85) than without one (1.63), although as mentioned, these two ratings are not statistically comparable. For items with a temporal adverb, participants significantly graded subject co-construal much higher with null pronouns (1.49) than with overt ones (2.09,  $F(1, 698) = 70.3, p < .001$ ). Likewise, for items without a temporal adverb, subject co-construal was judged significantly more acceptable with null pronominals (1.45) than with overt ones (1.84,  $F(1, 2168) = 110.6, p < .001$ ). The average acceptability judgements are presented in figure (32). Note that while acceptability for null-pronominal co-construed subjects remains fairly constant across the board (1.45 without temporal adverbs, 1.49 with temporal adverbs and 1.46 overall), for overt pronouns,

<sup>9</sup>These were not minimal pairs, that is, not all examples included a version with an adverb and one without an adverb, the statistics presented below reflect contrasts only for cases with or without an adverb, but these two groups are not directly comparable. Additionally, in some of the examples, the adverb was initial, in others, final.

acceptability varies (1.84 without temporal adverbs, 2.09 with temporal adverbs, 1.92 overall).

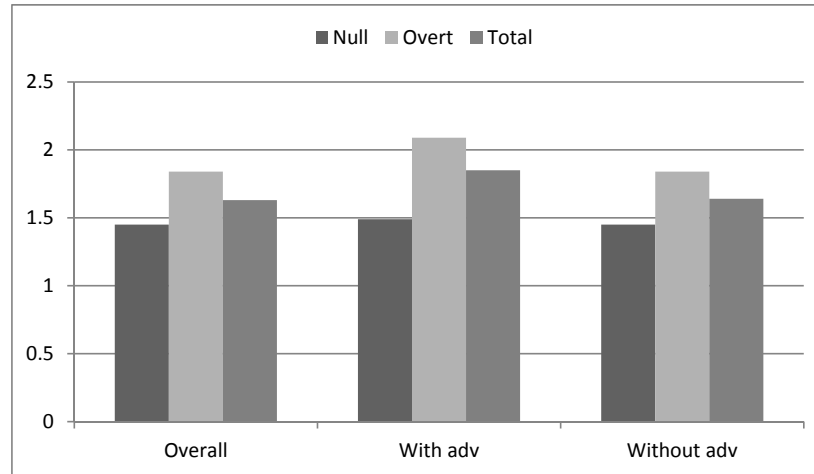


Figure 2: Average acceptability for coreferential subjects

### 5.3 Acceptability of coreferential subjects. Discussion

The results from the preceding section seem to argue against analyzing all weak pronouns as topic-bound and located in the same structural position. If that were the case, one would expect them to behave in similar ways with respect to acceptability under co-construal (cf. predictions (26b)-(26c) above), but we find significantly different degrees of acceptability. On the one hand, it seems that *pro* is the default option to maintain the topic from a preceding clause, as a number of authors have argued. This observation is consistent with Frascarelli’s treatment of *pro* as a topic-bound variable, as already suggested in (28a), repeated below without the overt weak pronominal option.

$$(33) \quad [_{CP1} [_{Top1P} DP_{S-i} \dots [_{CP2} [_{Top2P} \emptyset_i [_{TP} \dots pro_i \dots ]]]]]$$

On the other hand, overt weak pronouns do not consistently maintain the topic. If, as we have argued earlier, PS can occupy two distinct positions regardless of their strong/weak status, this result would follow. In one of those positions, shared with *pro*, the pronoun is also topic-bound as in (34a); in the other one, the subject is a topic itself, as in (34b). Presumably, these two structural options result in

speakers' lower acceptability for overt, weak pronouns: a structural analysis like (34a) will be more consistent with co-construal than (34b).

- (34) a.  $[_{CP1} [_{Top1P} DP_{S-i} \dots]] [_{CP2} [_{Top2P} Top_i [_{TP} \dots ella_i \dots]]]$   
 b.  $[_{CP1} [_{Top1P} DP_{S-i} \dots]] [_{CP2} [_{Top2P} ella_i [_{TP} \dots ]]]]$

The proposed analysis can also provide an explanation for the acceptability differences between presence and absence of the temporal adverb, given some additional assumptions. Specifically, if the temporal adverb provides the paragraph with higher discourse cohesiveness, those paragraphs where the adverb joins two sentences will be more prone to co-construal than those without the adverb. It follows that those co-construal contexts will penalize structures where the overt element is a topic (as in (34a)) as opposed to a topic-bound variable (as in (34b)).

Under this view, higher unacceptability rates for overt pronouns reflect a structural ambiguity. This means that speakers potentially entertain both structures, and depending on which one is assigned as a parse for the stimulus, they will accept it at higher or lower rates.

Finally, this analysis suggests a new perspective on Chomsky's Avoid Pronoun Principle presented in (14), which states that overt pronouns should be avoided whenever possible (i.e. whenever there is a null counterpart). Given the analysis presented here, overt weak pronominals in Spanish are structurally ambiguous between TopP and Spec, TP, whereas null pronominals are unambiguous: They are always in Spec, TP. Thus, Avoid Pronoun can be seen as a way of avoiding structural ambiguity by selecting the structure that is unambiguous vs. the one that can correspond to two possible structural parsings. In this sense, Avoid Pronoun can be rephrased in terms of Avoid structural ambiguity.

## 6 Summary and conclusions

This paper has tested two related proposals on the status of subjects. First, the hypothesis that preverbal subjects in null subject languages of the Spanish-type have the same status as CLLD items. Assuming the strongest version of this analysis (which completely equates PS and CLLD items), we analyzed antecedent choices for null, overt unstressed and overt stressed pronominal subjects, comparing preference for CLLD and S as potential antecedents. The underlying idea is that if PS = CLLD items, they should be selected as antecedents for pronominal subjects at roughly similar rates. Our results indicate that CLLD items are chosen more frequently than subjects, which led to the proposal that PS have two distinct struc-

tural positions, one as topics, another one as topic-bound variables, much along the lines of Frascarelli's analysis. Because all CLLD items are topics but PS are ambiguous, we account for the higher preference for CLLD items as antecedents.

Second, we tested whether these two structural positions correspond to the strong/weak partition by examining acceptability rates for co-construal of subjects in two clauses, one of which is null or overt unstressed. If null and overt unstressed pronominals occupy the same position (as both Cardinaletti and Frascarelli assume), then there should be no difference in acceptability in the context of co-construal, particularly if all types of weak pronominals occupy the same position. However, our results indicate that null and overt unstressed pronouns do not behave in the same way: in contexts of co-construal, null pronouns are accepted at a higher rate than overt, unstressed ones. This result follows as well if null pronouns are always topic-variables, whereas overt, unstressed ones are structurally ambiguous.

Third, we tested the notion that discourse-paragraph cohesiveness (through a temporal adverb) may affect co-construal acceptability rates, and our results indicate that more cohesive a paragraph is (i.e. cases with a temporal adverb), the less acceptable an overt pronoun is. We speculate that this is the result of clashing expectations: a more cohesive discourse should favor topic continuity, but an overt, unstressed pronoun is ambiguous between a topic-continuity reading (when it is in Spec, TP) and a new Topic reading. Given this ambiguity, speakers will accept the overt version less.

Finally, we have argued that Chomsky's Avoid Pronoun Principle reflects the structural ambiguity overt, weak pronouns present vs. null pronominals, which are structurally unambiguous. Thus, the principle can be restated as Avoid structural ambiguity.

## References

- Alexiadou, A. and Anagnostopoulou, E. (1998). Parametrizing AGR: Word order, v-movement and EPP-checking. Natural Language and Linguistic Theory, 16:491–539.
- Alonso-Ovalle, L., Clifton, C., Lyn, F., and Fernández-Solera, S. (2002). Null vs. overt pronouns and the topic-focus articulation in Spanish. Journal of Italian Linguistics, 14(2):151–169.
- Baker, M. (1995). The Polysynthesis Parameter. Oxford University Press, Oxford.

- Barbosa, P. (1995). Null Subjects. PhD thesis, MIT.
- Barbosa, P., Kato, M., and Duarte, M. E. (2005). Null subjects in European and Brazilian Portuguese. Journal of Portuguese Linguistics, 4:11–52.
- Burzio, L. (1986). Italian Syntax. Riedel, Dordrecht.
- Camacho, J. (2006). Do subjects have a place in Spanish. In Montreuil, J.-P. and Nishida, C., editors, New Perspectives in Romance Linguistics, pages 51–66. John Benjamins, Amsterdam/Philadelphia.
- Camacho, J. (2011). On null subjects: Towards a unified analysis of the null subject variation. Rutgers University.
- Cardinaletti, A. (1997). Subjects and clause structure. In Haegeman, L., editor, The New Comparative Syntax, pages 33–63. Longman, London/New York.
- Cardinaletti, A. (2004). Towards a cartography of subject positions. In Rizzi, L., editor, The Cartography of Syntactic Structures, Volume 2, pages 115–165. Oxford University Press, Oxford/New York.
- Cardinaletti, A. and Starke, M. (1999). The typology of structural deficiency. In Riemsdijk, H. V., editor, Clitics in the Languages of Europe (Empirical Approaches to Language Typology, 20; Eurotyp, 5), pages 145–233. Mouton, Berlin.
- Carminati, M. N. (2002). The processing of Italian subject pronouns. PhD thesis, University of Massachusetts, Amherst.
- Chomsky, N. (1981). Lectures on Government and Binding. Foris, Dordrecht.
- Contreras, H. (1991). On the position of subjects. In Rothstein, S., editor, Perspectives on Phrase Structure: Heads and Licensing. Syntax and Semantics 25., pages 63–79. Academic Press, San Diego.
- Frascarelli, M. (2007). Subjects, topics and the interpretation of referential pro. an interface approach to the linking of (null) pronouns. Natural Language and Linguistic Theory, 25:691–734.
- Goodall, G. (1999). On preverbal subjects in Spanish. In Satterfield, T., Christina, T., and Cresti, D., editors, Current Issues in Romance Languages, pages 95–110. John Benjamins, Amsterdam.

- Jegerski, J., VanPatten, B., and Keating, G. D. (2011). Cross-linguistic variation and the acquisition of pronominal reference in 12 Spanish. Second Language Research, 27(4):481–507.
- Koopman, H. and Sportiche, D. (1991). The position of subjects. Lingua, 85:211–258. Special issue edited by J. McCloskey.
- Ordóñez, F. and Treviño, E. (1999). Left dislocated subjects and the pro-drop parameter: A case study of Spanish. Lingua, 107:39–68.
- Rigau, G. (1988). Strong pronouns. Linguistic Inquiry, 19:503–511.
- Rizzi, L. (1982). Issues in Italian Syntax. Foris, Dordrecht.
- Rizzi, L. (1987). Three remarks on null subjects. Talk given at the Workshop on Dialectology, GLOW X, University of Venice.
- Rizzi, L. (1997). A parametric approach to comparative syntax properties of the pronominal system. In Haegeman, L., editor, The new comparative syntax, page 268–285. Longman, London.
- Sheehan, M. (2007). The EPP and Null Subjects in Romance. PhD thesis, Newcastle University.
- Solà, J. (1992). Agreement and subjects. PhD thesis, Universitat Autònoma de Barcelona, Barcelona, Spain.
- Sorace, A. and Filiaci, F. (2006). Anaphora resolution in near-native speakers of Italian. Second Language Research, 22(3):339–368.
- Suñer, M. (2003). The lexical preverbal subject in a Romance null subject language: Where are thou? In Núñez-Cedeño, R., López, L., and Cameron, R., editors, A Romance Perspective on Language Knowledge and Use. Selected Papers from the 31st Linguistic Symposium on Romance Languages (LSRL), pages 341–58. John Benjamins, Amsterdam.
- Taraldsen, T. (1992). Agreement as pronoun incorporation. Paper presented at the 1992 GLOW Colloquium.
- Zubizarreta, M. L. (1994). The grammatical representation of topic and focus: Implications for the structure of the clause. University of Venice Working Papers in Linguistics, 4:97–126.