

Acquisitional Evidence Bearing on the Account of Preverbal and Postverbal Subjects in Spanish

Julio Villa-García
University of Connecticut

1. Introduction

The main goal of this article is to shed new light on the controversial status of lexical/overt subjects in Spanish from the viewpoint of First Language Acquisition, hoping that evidence from the time-course of acquisition of preverbal and postverbal subjects in child Caribbean and European Spanish will be enlightening in this regard. The much debated analysis of overt subjects in null-subject languages like Spanish has raised—and tried to tackle—numerous questions in a body of research which spans several decades. One of the remaining questions of this enterprise is whether preposed (cf. (1a)) and postposed (cf. (1b)) subjects should be treated uniformly, or whether they should receive different analyses.

- (1) a. *Puerto Rico* me ha encantado [SV]
Puerto Rico cl. has enchanted
b. Me ha encantado *Puerto Rico* [VS]
cl. has enchanted Puerto Rico
Both: ‘Puerto Rico has captivated me.’

Additional questions posed by the existing works on subjecthood in Spanish-style null-subject languages to date include whether (some version of) Rizzi’s (1982) [Spec, IP/TP]/Extended Projection Principle (EPP) analysis of subjects assumed for languages like English can be maintained for Spanish, and whether overt preverbal subjects in Spanish have properties reminiscent of dislocated structures hosted in the CP area (Ordóñez 1997, *inter alia*). In parallel fashion, related questions involve the position(s) occupied by postverbal subjects, as well as the characterization of the syntactic, semantic, and pragmatic differences between preverbal and postverbal subjects. One important finding thus far is that the account of overt subjects in Spanish cannot rely solely on strictly syntactic factors: other aspects pertaining to the domains of the discourse-pragmatics interface (information structure) and the lexicon-syntax interface (type of verbal predicate involved) may also bear on the occurrence and distribution of subjects (Bel 2001, 2005; Barbosa 2009, among others). This state of affairs is further complicated by the fact that Spanish, contrary to other Romance *pro*-drop languages, frequently displays a third word order where the subject is sandwiched between the verb and the

* I would specially like to thank William Snyder for his guidance and help. This paper has benefited enormously from the comments of Eva Bar-Shalom, Jonathan David Bobaljik, Željko Bošković, Carole Boster, Eugenia Casielles, Mary Goodrich, John Grinstead, Diane Lillo-Martin, Mireia Llinàs-Grau, Krzysztof Migdalski, Antxon Olarrea, Iván Ortega-Santos, Ana de Prada, Yael Sharvit, Karen Zagona, and two anonymous reviewers. Similarly, I am grateful to the audiences at UUSLAW Smith College 2008, UUSLAW UConn 2009, GALA Lisbon 2009, HLS Puerto Rico 2009, BUCLD 34, LSRL-40 Seattle 2010, and UConn Language Fest 2010, especially José Camacho, Nina Hyams, Paula Kempchinsky, Letitia Naigles, Paco Ordóñez, Ana Teresa Pérez-Leroux, Tom Roeper, and Cristina Schmitt. I am also indebted to Luis Ortiz for his editorial assistance. This investigation owes much to the CLESS project at the University of Connecticut. This research has been supported in part by National Institute of Health grant #NIDCD 00183 to Diane Lillo-Martin and William Snyder. Likewise, this study has also been partially funded by the Full Graduate Fellowship “la Caixa,” awarded by the savings bank “la Caixa” (Barcelona, Spain), and by the Academic Excellence Covenant Insurance Co. Fellowship, conferred by the College of Liberal Arts and Sciences at the University of Connecticut. The usual disclaimers apply.

object/complement (i.e., VSO). Additionally, Caribbean Spanish adds to this puzzle inasmuch as it has been reported to be moving towards the negative setting of the Null Subject Parameter (Toribio 2000, Ticio 2004, among others).

In this paper I present results from five longitudinal case studies of children acquiring Caribbean and European Spanish in order to provide acquisitional evidence that helps tease apart competing analyses of the adult grammar. Despite the fact that the analyses in the literature do not straightforwardly predict preverbal and postverbal subjects to emerge concurrently in development in Spanish, I submit that the two constructions arise at the same time in child Caribbean and European Spanish. Overall, I propose that preverbal and postverbal subjects share a crucial prerequisite—the final “key ingredient”—that the child needs to acquire before making effective use of the constructions. Likewise, I observe a previously unacknowledged fact concerning the somewhat delayed emergence of the VSO order in child Spanish, in opposition to the results of the reports summarized in Pierce (1989), which I attribute to the extremely low frequency of occurrence of the construction.

The paper is organized into different sections, as follows: Section 2 offers an overview of the existing accounts of overt subjects in Spanish and their acquisitional predictions; Section 3 focuses on existing studies on the L1 acquisition of subjecthood in Spanish; Section 4 deals with the methods employed; Section 5 is dedicated to presenting and discussing the results of this study; Section 6 offers some concluding remarks and notes directions for future work.

2. The puzzle of preverbal and postverbal subjects. Dissenting views from syntax

As is well known, non-Caribbean Spanish is a prototypical null-subject language where overt subjects can appear in different positions with respect to the verb and the complements. The traditional account of *pro*-drop in Romance languages such as Catalan, Italian, and Spanish owes much to the ground-breaking work of the Italian linguist Luigi Rizzi (1982) in the framework of Government-Binding/Principles-and-Parameters (GB/P&P). An important aspect of Rizzi’s classical account is the assumption that the subject in languages like Spanish (be it overt or null) sits in the same position as the subject in languages like English, i.e., in [Spec, IP/TP], postverbal subjects being the result of a free subject inversion rule which optionally moves the subject to the right, adjoining it to the VP, possibly with an empty *pro* subject occupying the specifier of TP. A related view is defended in more contemporary proposals by authors including Zubizarreta (1999), who pursues the hypothesis that Spanish has a rightward specifier of TP, where postverbal nominatives are located. Other authors who have recently argued that preverbal subjects in Spanish occupy [Spec, TP] include Goodall (2001) and Ortega-Santos (2005 *et seq.*).

Nevertheless, research has shown that an account of preverbal subjects in languages like Spanish according to which subjects always move to [Spec, TP/AgrSP] to satisfy the EPP requirement may be inadequate, thus challenging (versions of) the traditional account sketched above. In fact, there has been a wealth of proposals to the effect that lexical subjects in Spanish are left-peripheral elements largely determined by the discourse-pragmatic context, and hence related to the CP field. This claim tends to go hand in hand with the contention that Spanish lacks the EPP, or alternatively that some parameterized version of the EPP is operative (e.g., the EPP can be satisfied by assuming head movement of the verb and its “rich” agreement morphemes to T°, à la Alexiadou & Anagnostopoulou 1998). On this view, lexical subjects in languages like Spanish do not necessarily have to be located in [Spec, TP]; this position may not be projected at all, or it may ultimately be occupied by *pro*. Proponents of this line of analysis whereby overt preverbal subjects in Spanish are hosted in a position above the inflectional layer include Contreras (1991), Olarrea (1996), Ordóñez (1997), Alexiadou & Anagnostopoulou (1998), Ordóñez & Treviño (1999), and Barbosa (2009), among many others. A logical question to pose is whether postverbal subjects are also left-peripheral elements under the view that (overt preverbal) subjects are topic-like constituents in the CP domain. For some authors, postverbal subjects are *in situ* (cf. Olarrea 1996, Ordóñez 1997, Barbosa 2009, *inter alia*). For instance, Barbosa (2009) argues that postverbal subjects can be either base-generated *in situ* in [Spec, vP] (cf. VSO), or *foci* (cf. VOS), in consonance with the empirical observation that postverbal subjects tend to be instances of new-information focus in languages like Spanish (Casielles 2001, *inter alia*).

The two major venues of research mentioned above, namely the [Spec, TP]/EPP-analysis and the CP-account, have polarized the spectrum of analyses of subjects in languages like Spanish. The

simplified labelled bracketings in (2a) and (2b) illustrate the derivation of a Spanish declarative sentence like *Pedro visitó San Juan* ‘Peter visited San Juan’ under each account.

- (2) a. [_{TP} Pedro [_T visitó] [_{VP} ... [_{VP} San Juan]]]
 b. [_{CP} Pedro [_C ∅] [_{TP} ... [_T visitó] [_{VP} ... [_{VP} San Juan]]]]

A third type of analysis in between the aforementioned accounts is the account according to which overt preverbal subjects can indeed occupy different syntactic positions (e.g., a CP specifier or [Spec, TP]), as suggested by Casielles (2001), Camacho (2006), López (2009), and Villa-García (in preparation), among others.

As regards postverbal subjects, there have been a number of accounts on a par with the right adjunction analysis of postverbal subjects advocated by the traditional account, and the rightward-specifier-of-TP analysis of Zubizarreta (1999) mentioned above. Thus, the VOS order can be derived by assuming that it results from scrambling of the object (Ordóñez 1997, Barbosa 2009), remnant movement of the TP over the subject in FocusP (Ordóñez 2000), remnant movement of the VP over the subject in the VP periphery/low IP area (Belletti 2004), and p(rosodic)-movement out of focus position past the subject (Zubizarreta 1998). Other proposals include LF movement of the subject to [Spec, TP] (assuming a weak-strong features/early minimalism framework), optional endowment of T° with an EPP feature (Chomsky 2000), and the Copy-Theory-of-Movement analysis advocated in the work of Stjepanović (1999) and Ortega-Santos (2006), according to which *all* movement is overt, the choice of copy to pronounce ultimately being a PF decision. A virtue of the latter analysis is that it is partly unified: as far as their syntactic treatment is concerned, preverbal and postverbal subjects are treated equally (i.e., in both cases the subject raises to [Spec, TP]), but different copies are pronounced, resulting in the different surface words orders.

It is important to mention that, along with notions pertaining to information structure (i.e., preverbal/topical-postverbal/focal), as argued by Casielles (2001), *inter alia*, properties of the lexicon-syntax interface also seem to influence the choice of subject position. By way of illustration, a sentence containing an unaccusative predicate such as *llegar* ‘to arrive,’ is more likely to exhibit a postverbal subject than, for instance, an out-of-the-blue sentence containing a transitive verb (*comprar* ‘to buy’), where the subject is not assumed to be the underlying object (Bel 2001, Barbosa 2009). Further, along with subjects of presentational unaccusatives, subjects of psych verbs and sentential subjects also display a tendency to appear postverbally (Contreras 1976).

In addition to the above, there are also a number of well-documented semantic asymmetries between preverbal and postverbal subjects in Spanish. For some differences regarding quantification, see Uribe-Etxebarria (1991). With regard to dissimilarities in interpretation regarding categorical judgments about given individuals (preverbal subjects) vs.thetic judgments conveying events (postverbal subjects), see, for instance, Uriagereka (2002) and Gallego (2005). Similarly, preposed and postposed subjects exhibit divergent agreement patterns, namely first-conjunct agreement with postverbal subjects (cf. Camacho 2003, Ortega-Santos 2005) and lack of number agreement with postverbal subjects (cf. Villa-García 2010b). Regarding Montalbetti’s (1984) Overt Pronoun Constraint (OPC), which posits that in languages that permit overt/null pronominal alternations, an overt pronominal must not have a quantified antecedent, an asymmetry is found in that postverbal—but not preverbal—overt pronouns can be bound by the quantified antecedent. In the last place, it is important to mention the well-documented ban on bare NPs in preverbal position, unless they are uncontroversially left-dislocated (Casielles 2001).

To make matters worse, according to a number of researchers (cf. Navarro Tomás 1948, Toribio 2000, Ticio 2004, *inter alia*), the Caribbean dialects of Spanish (Cuban, Dominican, and Puerto Rican Spanish) manifest a preference for overt preverbal subjects, although null and postverbal subjects are also attested in these varieties.

Even though the (necessarily over-simplified) discussion in the preceding paragraphs does not do justice to the amount and complexity of the analyses proposed in the literature, it should suffice to illustrate the lack of consensus in the field regarding the account of preverbal and postverbal subjects in Spanish, as well as the differences displayed by the two types of subject.

2.1. *Acquisitional predictions*

One key distinction between competing theories is the predictions each makes regarding the order of acquisition of preverbal and postverbal subjects. More concretely, a number of important research questions can be formulated, including whether preverbal and postverbal subjects are acquired at different times or simultaneously, whether preverbal subjects arise earlier than postverbal ones, or vice versa, and whether there are any differences between the onset of particular words orders (e.g., SVO, VSO, VOS). The accounts of preposed and postposed subjects in languages like Spanish summarized above make certain predictions about the timing of acquisition of the relevant constructions.

First, it is worthwhile mentioning the prediction of concurrent, simultaneous acquisition of overt subjects and topics made by the venue of research which places overt subjects in a CP specifier (see the studies conducted by Grinstead 1998 *et seq.* and Villa-García & Snyder, to appear). Similarly, if postverbal subjects are deemed to be *in situ* in [Spec, *v*P], then preverbal and postverbal subjects should be unrelated developmentally. The same holds for the accounts which place postverbal subjects in their base-generated position (i.e., [Spec, *v*P]), and preverbal subjects in [Spec, TP], since only the latter involve syntactic movement. Further, the rightward-specifier account of postverbal subjects defended in Zubizarreta (1999) predicts that the Spanish-acquiring child will need to determine that it is possible to project [Spec, TP] to the left (cf. SV) and to the right (cf. VS) in her target language. Thus, as noted by Jonathan Bobaljik (p.c.), the Spanish-acquiring child will have to acquire the relevant parameter underlying the direction of specifiers, in addition to the requirements needed to start using overt subjects. Note that this analysis predicts a different timing of acquisition of the VSO order, since this configuration cannot be derived by assuming a rightward specifier (i.e., the subject in VSO structures is not assumed to be in the same position as the subject in SVO and VOS configurations in this system). Moreover, the analyses which consider preverbal subjects to be topics and postverbal ones to be *foci* make the prediction that postverbal subjects will emerge earlier (cf. Casielles et al. 2006; see Section 3). Lastly, the partially unified account of preverbal and postverbal subjects defended by Stjepanović (1999) and Ortega-Santos (2006) *prima facie* predicts preverbal and postverbal subjects to be acquired concurrently. I return to this proposal in Section 5.

3. Previous studies on the acquisition of lexical subjects in Spanish

In influential work, Grinstead (1998, 2000) has suggested that preverbal and postverbal subjects emerge together in the spontaneous speech of Catalan- and Mexican-Spanish-speaking children. This contention is based on the proximity in time between the onset of each of the two constructions, as the author does not provide statistical support for his claim. If correct, this result has a number of non-trivial implications, as will be seen.

In her (2005) study on the acquisition of word order in Catalan and Spanish, Bel shows that, once acquired, preverbal subjects are more frequent than postverbal subjects, the latter having a higher rate of occurrence in the case of unaccusative verbs. Furthermore, Bel (2005) shows that children are sensitive to pragmatic notions pertaining to information structure from early on. Her results are indicative that, in actuality, the occurrence of overt subjects in Spanish is not reducible to a quintessentially grammatical parameter (e.g., the Null Subject Parameter), and that other factors such as the topic-comment and focus-presupposition articulations bear on the appearance of overt subjects in languages like Spanish, as noted earlier.

The Casielles et al. (2006) study advances the hypothesis that the syntactic and discourse features of preverbal subjects in Spanish may account for a likely emergence delay in comparison with null and postverbal subjects. Therefore, Casielles et al. (2006) predict that children learning Spanish will find it difficult to use preverbal subjects, since these nominals involve movement from [Spec, *v*P] to a higher position and encode a more complex discourse notion than postverbal or unpronounced subjects. According to Casielles et al. (2006), focal postverbal subjects do not involve movement inasmuch as stay in their *in-situ* position. For these authors, postverbal subjects are always focal and children are aware of new information from the earliest stages. In analogous fashion, it would be far from implausible to assume that it would be easier for the language acquirer to continue on a discourse-old topic and drop it rather than to start a new or contrastive topic which involves not only choosing the new topic but also syntactic movement (Eugenia Casielles, p.c.). In order to test this prediction, the acquisitional investigation carried out by Casielles et al. (2006) analyzed longitudinal data from a

bilingual child acquiring English and Spanish.¹ The age span considered in this study ranges from 2;4 to 3;5. In his Spanish, the child used null subjects 67.9% of the time, and overt subjects 32.1% of the time, postverbal subjects being used 20.7% of the time, as opposed to preverbal subjects, which occurred only 9.7% of the time. Thus, null and postverbal subjects constitute more than 88% of all subjects in the spontaneous speech of this bilingual child. Postverbal subjects started to appear at age 2;4, and preverbal subjects at age 2;5. Casielles et al. take these results to indicate that preverbal subjects are acquired later than null and postverbal subjects in Spanish, contra Grinstead (1998, 2000). However, in the absence of supporting statistics, the findings reported by Casielles et al. (2006) should be interpreted with caution. Consequently, these results cannot be regarded as conclusive evidence that postverbal subjects are acquired earlier than preverbal subjects, and further research is required to determine the time-course of acquisition of preposed and postposed subjects, an issue to which I turn in the following sections.

As far as Caribbean Spanish is concerned, Ticio (2002) has argued that Ana's linguistic development confirms that the Puerto Rican variety of Caribbean Spanish behaves differently from fully-*pro*-drop Spanish. Ticio's study draws on Grinstead's (1998) claims for early Catalan- and Mexican-Spanish that overt subjects, topics, and *wh*-questions emerge concurrently in development, as predicted by the venue of research which regards overt subjects as CP-related phenomena (cf. Section 2). Ticio found no occurrences of topics or *wh*-questions at a time when overt subjects were already attested in Ana's speech. The author takes this as an indication that subjects in Caribbean Spanish are not in the CP (as in non-Caribbean Spanish), but in [Spec, TP], much like in English, in accordance with her claim that Caribbean Spanish is no longer a null-subject language *per se*. The investigation conducted by Villa-García et al. (2010) provides acquisitional and statistical support for Ticio's speculations about Caribbean Spanish, though neither study tackled the issue of the time-course of acquisition of preverbal and postverbal subjects in Caribbean Spanish, a task which is undertaken in Section 5.

4. Research methods and data collection

The corpus data utilized in this study come from the CHILDES database (Child Language Data Exchange System (MacWhinney 2000)) and from the CLESS database (the University of Connecticut's Cross-Linguistic Early Syntax Study project). For Caribbean Spanish, I used longitudinal data from Ana, a middle-class girl raised in Mayagüez, Puerto Rico. Her corpus, retrieved from the CLESS database, includes transcripts and not-yet-transcribed recordings (see also Ticio 2002). The four European-Spanish-acquiring children of this study were raised in Spain.

Table 1: Longitudinal data

Variety of Spanish	Child	Corpus/ Database	Date of Retrieval	Age Span	# of Transcripts; Frequency
Caribbean	Ana (♀)	Ana (CLESS)	April 6, 2009	01;07-2;05,10	# = 41; weekly
European	Emilio (♂)	Vila (CHILDES)	March 7, 2009	0;11,09-2;11,24	# = 27; monthly/biweekly
	Inés (♀)	Inés (CLESS)	April 1, 2009	1;02,00-2;02,11	# = 36; (bi)weekly
	Irene (♀)	Llinàs-Grau/Ojea (CHILDES)	October 16, 2008	1;05,27-2;04,13	# = 31; biweekly
	Juan (♂)	Linaza (CHILDES)	March 10, 2009	1;07,02-3;09	# = 18; mostly monthly

¹ Note that if English had influenced the placement of subjects in the child's Spanish, then we would expect most subjects to be preverbal.

The data were counted manually, since at present there is no reliable computer-assisted method to find occurrences of lexical subjects in Spanish. In counting uses of the constructions of interest, imitations, repetitions, and formulaic expressions were discarded. All sentences containing overt subjects were tabulated. The relevant utterances were then coded for subject type (SV, VS), order of constituents (SVO, VSO, and VOS), sentence force (declarative, imperative, interrogative, or exclamative), and type of predicate involved (copula, unergative, unaccusative, psych, or (di)transitive). In order to ensure mastery of the pertinent constructions, the measure of acquisition was first clear use, followed soon after by additional, distinct uses (Stromswold 1996).

The statistical method used to check for concurrent acquisition was the Binomial Test, as outlined in Snyder (2007: Ch. 5). This method is an exact test of the statistical significance of deviations from a theoretically expected distribution of observations into two categories. The question which the Binomial Test aims to answer concerns the probability of a given outcome (e.g., several uses of postverbal subjects (construction A), before the first occurrence of preverbal subjects (construction B)), under the null hypothesis that construction B was available to the child as early as A, and had the same relative frequency of use as in later transcripts (e.g., the next ten transcripts after the appearance of the second construction). Put differently, the Binomial Test addresses the question of whether the presumed discrepancy between two given constructions is plausibly due to a lower frequency of use for the construction emerging later, or whether there is in fact a statistically significant difference between the two, as expected when the two constructions demand different prerequisites that the child needs to acquire prior to using the particular constructions successfully. A not significant, null result is consistent with chance, whereas a significant result refutes the null hypothesis of concurrent emergence of the two constructions, and in turn calls for a grammar-based explanation of the difference.

5. Results and discussion

This section is divided into two parts. The first subsection is devoted to exploring the timing of acquisition of preverbal and postverbal subjects in child Caribbean and European Spanish, the focus being placed on the SV(O) and VS/VOS word orders. Since VSO is shown to appear substantially later than the other word orders under consideration, Section 5.2 deals exclusively with the emergence of this word order.

5.1. On the time-course of acquisition of preverbal and postverbal subjects

The first clear uses of the constructions of interest are shown in Table 2:

Table 2: Emergence of preverbal and postverbal subjects

Variety of Spanish	Child	Onset Age of Preverbal Subjects (SV, SVO)	Onset Age of Postverbal Subjects (VS, VOS)
Caribbean	Ana	1;08,12	1;08,15
European	Emilio	1;11,12	1;09,19
	Inés	1;06,12	1;06,05
	Irene	1;07,22	1;07,05
	Juan	1;09,02	1;09,02

First of all, three of the children (namely Ana, Inés, and Irene) started using both preverbal and postverbal subjects roughly at the same time, the difference reducing to days. In Juan's spontaneous speech, preverbal and postverbal subjects emerged in the same transcript, while in the case of Emilio, postverbal subjects seem to have appeared substantially earlier than preverbal subjects. Note, however, that only two transcripts were available during this two-month period. Likewise, in the speech of three of the children (i.e., Emilio, Inés, and Irene), the first examples of postverbal subjects occurred before those of preverbal ones, whereas in the case of two children (i.e., Ana and Juan), preverbal subjects

were attested first.² These results alone may already be indicative that the two constructions emerge together in the acquisition of both Caribbean and European Spanish, as pointed out by Ana Teresa Pérez-Leroux (p.c.). It is important to bear in mind that no noteworthy differences were found regarding the time of emergence of SV vs. SVO and VS vs. VOS in any of the children under consideration. In this connection, note that all the children had already passed the “two-word” stage—during which children may produce the SV order but not the SVO order—at the point when overt subjects began to be used. An additional implication of the findings reported in Table 2 is that the time-course of acquisition of preverbal and postverbal subjects in child Caribbean Spanish proceeds in parallel with that of European Spanish, which is substantiated by the results of the statistical analysis reported below. Furthermore, the data from Ana clearly show that postverbal subjects do exist in her Puerto Rican variety of Spanish. The following are some examples of child utterances displaying the different word orders under consideration:

- | | |
|------------------------------|-------------------|
| (3) a. Peces no está | [SV] |
| fish not is | |
| ‘The fish are not (here).’ | (Emilio, 1;11,12) |
| b. Los busco yo | [VS] |
| cl. search I | |
| ‘I will look for them.’ | (Emilio, 2;04,17) |
| c. ¡Irene rompió la nariz! | [SVO] |
| Irene broke the nose | |
| ‘Irene has broken her nose.’ | (Irene, 1;11,13) |
| d. Pone teléfono mami | [VOS] |
| puts phone mom | |
| ‘Mom picks the phone.’ | (Inés, 2;01,22) |
| e. Salchicha cayó | [SV] |
| sausage fell | |
| ‘The sausage fell.’ | (Ana, 01;08,10) |
| f. Cayó bebé | [VS] |
| fell baby | |
| ‘The baby fell down.’ | (Ana, 01;08,15) |

At this point, it is imperative to consider the results of the Binomial Test.³ Recall that the goal is to determine whether the apparent discrepancy between the onset of preverbal and postverbal subjects in the speech of Ana, Emilio, Inés, and Irene can be due to chance, or whether the children needed to acquire additional prerequisites before making use of the construction appearing later (cf. Section 4). The Binomial Test takes into account the number of times that the construction emerging earlier occurs, as well as the relative frequency of the two constructions after the emergence of the second one, as shown in Table 3:⁴

² As shown in Table 3, even though Juan started using both constructions in the same session, three instances of preverbal subjects were found before the first clear use of a postverbal subject.

³ *Binomial Test* $p = (X / (X + Y))^Z$

X corresponds to the number of times construction A appears in the ten transcripts following the first clear use of B (cf. “Frequency” in Table 3); Y stands for the times B occurs in the ten transcripts after the first use of B (cf. “Frequency” in Table 3); Z refers to the uses of A before the first clear use of B (“# of earlier construction” in Table 3).

⁴ For the sake of clarity, I provide an example of the relevant calculations for Inés. The first clear uses of overt subjects in the spontaneous speech of this child were postverbal subjects (construction A; # = 4). Once the second construction (B), i.e., preverbal subjects, emerged, the relative frequency of use in the ten transcripts following the emergence of construction B was 41 for preverbal subjects and 38 for postverbal subjects. Hence, the desired probability can be calculated using the Binomial Test (cf. fn. 3): $p = (38 / (38 + 41))^4 = .054$.

Table 3: Results of the statistical analysis

Variety of Spanish	Child	# of Earlier Construction	Frequency		<i>p</i> -value
			Pre-	Post-	
Caribbean	Ana	1 (Pre-)	16	13	<i>p</i> = .552 not significant
European	Emilio	2 (Post-)	59	67	<i>p</i> = .283 not significant
	Inés	4 (Post-)	41	38	<i>p</i> = .054 marginally significant
	Irene	1 (Post-)	25	7	<i>p</i> = .781 not significant
	Juan	3 (Pre-)	47	38	<i>p</i> = .169 not significant + same transcript

As noted earlier, the Binomial Test tackles the crucial question of whether the differences reported in Table 2 are the result of chance, or whether there is a significant gap between the two constructions, in which case the prospect of concurrent acquisition would be refuted. In this sense, it is important to bear in mind that the fact that one construction appears chronologically earlier than the other may well be due to the lower frequency of the construction appearing later. Such a difference in frequency might occur because the child is copying the relative frequencies of use in the adult input, for instance, or because the child's immature processing system finds the second construction harder. After all, it is well known that chronological age, and differences in age, are extremely uninformative about the actual process of language acquisition. Concurrent acquisition does not require the unlikely—if not impossible—fact that the two constructions will appear for the first time in exactly the same utterance. The Binomial Test addresses this question, however, since the unit of measurement now becomes—not days or weeks—but rather occurrences of a particular grammatical construction. A significant difference by Binomial Test thus reveals that the age difference is “meaningful,” supporting a grammar-based hypothesis: the prerequisites that the child needs to acquire before making effective use of the relevant constructions are different for each construction.

Nevertheless, the *p*-values obtained after performing the Binomial Test are not significant, as shown in Table 3, and thus the evidence against the null hypothesis that preverbal and postverbal subjects were acquired simultaneously is weak.⁵ Null results reveal that the slightly earlier appearance of one construction in comparison with the other may be attributed merely to chance. In other words, the results reported in Table 3 are entirely consistent with the hypothesis that both preverbal and postverbal subjects emerged concurrently in the speech of the children of this study, in accordance with the proximity in time between the onset of preverbal and postverbal subjects in most cases in the transcripts examined. This conclusion is strengthened by the observation that all the children studied exhibit a robust pattern, regardless of their dialect of Spanish, i.e., European or Caribbean. Put another way, in terms of the time-course of acquisition of preverbal and postverbal subjects in Caribbean and European Spanish, no difference was found, which shows that Caribbean varieties do exhibit postverbal subjects, and therefore it cannot be concluded that they are null-subject languages *per se* (see Villa-García 2010a for evidence that Ana's early use of null subjects is akin to that of European-Spanish-acquiring children).

The discussion above provides novel acquisitional and statistical support for Grinstead's (1998, 2000) claim that both types of overt subject emerge at the same point in development in Spanish. Similarly, our results concur with those of Villa-García & Todorović (in preparation) for Serbo-Croatian, and contrast with those of Casielles et al. (2006) for Spanish. The findings above point in the direction that Spanish preverbal and postverbal subjects emerge together in acquisition, and that children seem to acquire the relevant linguistic requirements at the same time in development.⁶ Hence,

⁵ Note that the *p*-value obtained for Inés is marginally significant. It is important to keep in mind, however, that the *p*-values reported in this study have not been adjusted. If a Bonferroni correction was applied, the marginal significance of the *p*-value for Inés would vanish, hence becoming not significant.

⁶ A plausible question to ask at this juncture is whether the results of this paper may be accidental. Yet, in order to provide a positive answer to this question, it would be necessary to find a child who disproves the null hypothesis

the current results strongly suggest that preverbal and postverbal subjects should be analyzed uniformly. Nonetheless, the question which immediately arises is the extent to which the account of the two constructions can be unified. In this connection, the results of this paper are somewhat unexpected in light of the plethora of investigations which have shown that Spanish preverbal and postverbal subjects behave differently (cf. Section 2). An additional issue raised by Montrul (2004:268) is that if preverbal subjects are topics whose acquisition is concurrent with that of topicalizations and other CP-related phenomena (cf. Grinstead 1998 *et seq.*; Villa-García & Snyder, to appear), it is surprising that the emergence of preverbal and postverbal subjects proceeds in parallel. At first sight, the observation that both types of subject arise simultaneously challenges the conception that preverbal subjects are topics vs. postverbal subjects, which are *in-situ*, base-generated elements, or *foci*, under different theories.

Whatever the case may be, Table 3 provides a finding that any empirically adequate theory of the adult grammar of *pro*-drop languages like Spanish should capture. So far, out of all the theories proposed in the literature, no analysis straightforwardly predicts the SV and VS word orders to emerge together in child Spanish. The one account which treats preverbal and postverbal subjects uniformly from a syntactic point of view is the analysis based on the Copy Theory of Movement of Chomsky (1995) defended in recent work by Stjepanović (1999) and Ortega-Santos (2006), as has been seen. This analysis assumes that movement takes place overtly, the choice of copy to pronounce being contingent on PF factors such as the focus-last requirement of Spanish (cf. Zubizarreta 1998, among others). However, as noted by Jonathan Bobaljik (p.c.), it is not at all clear that this partially unified account makes the prediction of concurrent emergence of preverbal and postverbal subjects in child Spanish, since the PF considerations regulating the pronunciation of the relevant copies vary in each case.

One way of looking at the puzzle arising from the results reported above would be to pursue the hypothesis that preverbal and postverbal subjects have different prerequisites that the child needs to acquire before making use of each construction. Importantly, however, the last prerequisite that the child has to acquire in order to start using preverbal and postverbal subjects is shared by the two constructions. That preverbal and postverbal subjects are different has been demonstrated by the vast body of empirical evidence adduced in the existing literature, which argues against the prerequisites for the two constructions being identical (cf. Section 2). Yet, the last (shared, common) prerequisite that the child needs to effectively start using preposed and postposed subjects must be acquired at the same time for the two constructions, as indicated by the concurrent emergence of preverbal and postverbal subjects in child Spanish. In much the same way, this move would circumvent the problem acknowledged by Montrul (2004) mentioned above, since certain prerequisites (other than the last one) may be unshared by the two constructions. At this point, however, it is imperative to determine what the last “key” ingredient is, a matter which future research should care to address. A final note concerns the acquisition of the VSO order, whose emergence seems to be noticeably delayed with respect to the other word orders at issue. This is the focus of the next subsection.

5.2. *On the acquisition of the notorious VSO order*

In the count effected above, different word orders were attested, including SV, VS, SVO, and VOS. However, a question arises as to the emergence of the VSO order. It is well known that VSO has a narrower distribution than the other word orders at issue (Contreras 1991). Also, whereas Spanish exhibits the VSO order relatively frequently, the appearance of this word order in related languages like Italian is far more restricted (Belletti 2004).

The emergence of the VSO order is somewhat delayed in the speech of the children under consideration. By way of illustration, Irene produced her first VSO sentence (cf. (4)) at age 2;01,18, substantially later than the other word orders under discussion (see Table 2).

of concurrent emergence (i.e., a significant discrepancy by Binomial Test between the onset of preverbal and postverbal subjects, after any statistical adjustments/corrections have been applied).

- (4) Ahora voy a equibi yo un pipi [VSO]
 now gonna to write I a ??
 ‘Now I’m gonna write a XX.’ (Irene, 2;01,18)

Since the data available are sparse, it is worthwhile looking at a similar structure with copular verbs (i.e., a sentence displaying the VSC(omplement)/A(dverb) word order). Irene produced the A(dverb)VS word order eight times, while she used the order VAS only twice, with no occurrences of VSA at all. As for Emilio, he produced a total of two sentences showing the VSC word order, at ages 2;07,24 and 2;08,28. The rest of the children of this study did not produce one single instance of the VSO/VSC/VSA order in the sessions analyzed.

A logical question to pose at this point is whether this delay is due to the lower frequency of the VSO/C order, or whether the child needs to acquire certain extra prerequisites other than those necessary to use the SV(O), V(O)S orders. This question can again be addressed by the Binomial Test. In the case of Irene, the data available make it possible to apply this statistical method. Note that Irene’s use of the VSO order constitutes an isolated sentence, so it is not certain that the child has actually mastered the VSO construction, if we strictly keep to the criteria adopted in this study to determine first clear uses (cf. Section 4). In any case, before the first occurrence of the VSO word order in the transcripts under consideration, Irene produced 149 overt subjects displaying the other word orders under debate. After the first use of VSO, Irene’s examples of the remaining word orders amount to 197 in the following ten transcripts, with two (unclear) occurrences of the VSO order, at ages 2;04,28 and 2;05,13. Even if we consider the only two dubious examples of VSO in our statistical analysis—which implies biasing the result against the null hypothesis that the child does not need to acquire additional prerequisites to start using the VSO order, since we are maximizing the chances of obtaining a significant result—the Binomial Test still yields a not significant p -value ($p = .222$). This result is wholly consistent with concurrent emergence, though it should be taken as interim, since only one child could be studied for this hypothesis, and the data at hand are rather scarce. Be that as it may, if this preliminary result is taken at face value, the late emergence of the VSO order in comparison with the other word order combinations available in Spanish may be due to the low frequency of the particular construction, rather than to the construction demanding prerequisites different from those necessary to produce the other word orders under consideration. Put differently, the differences in frequency between VSO and the remaining word orders at issue explain the apparent discrepancy in the order of acquisition: in principle, our results are consistent with all the word orders under consideration becoming available to the child at the same time, less frequent tokens of the VSO order making it appear as if VSO was acquired significantly later.

The somewhat delayed emergence of the VSO order found in this study contrasts markedly with the findings of the diary studies summarized in Pierce (1989), which point to the conclusion that the VSO order is the first word order attested in child Spanish. Future research will show whether the provisional findings reported in the present study are on the right track.

6. Conclusions

The issue of subjecthood in genuine null-subject languages like Spanish has attracted much attention among linguists, as witnessed by the numerous existing analyses of preverbal and postverbal subjects proposed in the literature to date. This has led to a high degree of controversy as to the account of overt subjects, which is reinforced by the observation that Caribbean Spanish has been reported to exhibit a different pattern of behavior regarding subjecthood than other varieties such as European Spanish.

The present investigation was designed to shed light upon the on-again off-again debate over the analysis of lexical subjects in Spanish by means of presenting results from five longitudinal case studies disclosing information as to the manner in which the time-course of acquisition of preverbal and postverbal subjects proceeds. It has been seen that no existing theoretical proposal straightforwardly makes the prediction of concurrent, simultaneous acquisition of preposed and postposed subjects in languages like Spanish. In glaring contrast, however, the acquisitional and statistical evidence furnished in this paper strongly indicates that preverbal and postverbal subjects start to be used approximately at the same time in development in the acquisition of Caribbean and European Spanish, in line with Grinstead (1998, 2008) and Villa-García & Todorović (in preparation),

and contra Casielles et al. (2006). An additional conclusion of this paper is that child Caribbean and European Spanish behave similarly with respect to the timing of acquisition of overt subjects. Moreover, this paper has made an excursus into the timing of acquisition of the notorious VSO order, whose late emergence has been attributed to its considerably low frequency of occurrence.

Prima facie, a rather strong conclusion drawn from the results of this investigation would be that preverbal and postverbal subjects should be analyzed uniformly, since the findings reported in this paper suggest that preverbal and postverbal subjects are the *same*, in terms of the major linguistic prerequisites that a child must acquire to effectively start using the relevant constructions. However, syntactic research has shown that preverbal and postverbal subjects present various asymmetries ranging from informational load to agreement mismatches. Importantly, the results of this investigation do not necessarily require the acquisitional prerequisites for preposed and postposed subjects to be identical. Instead, the data at hand would also be compatible with a scenario in which any unshared prerequisites are acquired very early, and only the *last* prerequisite that the child must acquire is shared by the two constructions. It is therefore crucial to determine what that last “key ingredient” for preverbal and postverbal subjects is, an issue which I leave for future research. Correspondingly, further investigations should care to explore additional issues including the relevance of unaccusativity in the time-course of acquisition of overt subjects, and any likely differences in time between the emergence of weak pronominals and full DPs in no-longer-fully-*pro*-drop dialects like Caribbean Spanish (cf. Ordóñez & Olarrea 2006; Paco Ordóñez, p.c.).

Pending new advances in syntactic theory and future studies, I conclude the paper by stressing that the acquisitional findings reported in this article should be accounted for by any empirically adequate syntactic theory of the adult grammar of Spanish, a task which no analysis available in the literature has fully accomplished hitherto.

References

- Alexiadou, Artemis, and Elena Anagnostopoulou (1998). “Parameterizing AGR: Word Order, V-Movement and EPP-Checking.” *Natural Language and Linguistic Theory* 16, 491-539.
- Barbosa, Pilar (2009). “Two Kinds of Subject *pro*.” *Studia Linguistica* 63(1), 2-58.
- Bel, Aurora (2001). “Sujetos nulos y sujetos explícitos en las gramáticas iniciales del castellano y el catalán.” *Revista Española de Lingüística*, 31(2), 537-562.
- Bel, Aurora (2005). “Aspectos de la adquisición del orden de palabras: la posición del sujeto en Castellano y Catalán.” *RAEL* 4, 36-48.
- Belletti, Adriana (2004). “Aspects of the low IP area.” *The Structure of IP and CP. The Cartography of Syntactic Structures*. (vol.2.). Rizzi, L. (ed.). Oxford: Oxford University Press.
- Camacho, José (2003). *The Structure of Coordination: Conjunction and Agreement Phenomena in Spanish and Other Languages*. Dordrecht: Kluwer Academic Publishers.
- Camacho, José (2006). “Do Subjects have a Place in Spanish.” *New Perspectives in Romance Linguistics*. Montreuil, J-P. and C. Nishida (eds). Amsterdam-Philadelphia: John Benjamins.
- Casielles, Eugenia (2001). “The Syntax and Semantics of Preverbal Topical Phrases in Spanish.” *Current Issues in Spanish Syntax and Semantics*. Gutiérrez-Rexach, J., and L. Silva-Villar (eds.). Berlin/New York: Mouton DeGruyter, 65-82.
- Casielles, Eugenia, Jean Andruski, Sahyang Kim, Geoffrey Nathan, and Richard Work (2006). “Syntactic and Discourse Features of Subjects in Child Spanish: Evidence from Spanish/English Bilingual Acquisition.” *BUCLD* 30, 72-83.
- Chomsky, Noam (1995). *The Minimalist Program*. Cambridge, Mass.: MIT Press.
- Chomsky, Noam (2000). “Minimalist Inquiries: The Framework.” *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*. Martin, R., D. Michaels, and J. Uriagereka (eds.). Cambridge, Mass.: MIT Press, 89-156.
- Contreras, Heles (1976). *A theory of word order with special reference to Spanish*. Amsterdam: North Holland.
- Contreras, Heles (1991). “On the position of subjects.” *Syntax and Semantics* 25, 63-79.
- Gallego, Ángel (2005). “Phase Sliding.” Ms., Universitat Autònoma de Barcelona and University of Maryland, College Park.
- Goodall, Grant (2001). “The EPP in Spanish.” *Objects and Other Subjects: Grammatical Functions, Functional Categories and Configurationality*. Davies, W. D., and S. Dubinsky (eds.). Netherlands: Kluwer Academic Publishers, 193-223.
- Grinstead, John (1998). *Subjects, sentential negation and imperatives in child Spanish and Catalan*. Ph.D. Dissertation, UCLA.

- Grinstead, John (2000). "Tense, Number and Nominative Case Assignment in Child Catalan and Spanish." *Journal of Child Language* 27(1), 119-155.
- López, Luis (2009). *A derivational Syntax for Information Structure*. New York: Oxford University Press.
- MacWhinney, Brian (2000). *Guide to CHILDES data on the acquisition of Romance languages*. Mahwah, New Hampshire: Lawrence Erlbaum Associates.
- Montalbetti, Mario (1984). *After Binding: on the Interpretation of Pronouns*. Ph.D. Dissertation, MIT.
- Montrul, Silvina (2004). *The Acquisition of Spanish. Morphosyntactic Development in Monolingual and Bilingual L1 Acquisition and in Adult L2 Acquisition*. Amsterdam: John Benjamins.
- Navarro Tomás, Tomás (1948). *El español de Puerto Rico*. Río Piedras: Editorial Universitaria.
- Olarrea, Antxon (1996). *Pre and Postverbal Subjects in Spanish: A Minimalist Account*. Ph.D. Dissertation, University of Washington, Seattle.
- Ordóñez, Francisco (1997). *Word Order and Clausal Structure of Spanish and Other Romance Languages*. Ph.D. Dissertation, CUNY.
- Ordóñez, Francisco (2000). *The Clausal Structure of Spanish: A Comparative Perspective*. New York: Garland.
- Ordóñez, Francisco, and Antxon Olarrea (2006). "Microvariation in Caribbean/non Caribbean Spanish Interrogatives." *Probus* 18, 59-96.
- Ordóñez, Francisco, and Esthela Treviño (1999). "Left dislocated subjects and the pro-drop parameter: A case study of Spanish." *Lingua* 107, 39-68.
- Ortega-Santos, Iván (2005). "On Locative Inversion and the EPP in Spanish." *Actas del VIII Encuentro Internacional de Lingüística del Noroeste de la Universidad de Sonora*. Universidad de Sonora, México, 131-150.
- Ortega-Santos, Iván (2006). "On New Information Focus, Sentence Stress Assignment Conditions and the Copy Theory: A Spanish Conspiracy." *University of Maryland Working Papers in Linguistics* 14. Kazanina, N., U. Minai, P. Monahan, and H. Taylor (eds.). College Park, Maryland: UMWPL, 88-212.
- Pierce, Amy (1989). *On the emergence of syntax: A crosslinguistic study*. Ph.D. Dissertation, MIT.
- Rizzi, Luigi (1982). *Issues in Italian Syntax*. Dordrecht: Foris.
- Snyder, William (2007). *Child Language: The Parametric Approach*. Oxford: Oxford University Press.
- Stjepanović, Sandra (1999). *What Do Scrambling, Second Position Cliticization and Multiple WH-Fronting have in Common?* Ph.D. Dissertation, University of Connecticut, Storrs.
- Stromswold, Karin (1996). "Analyzing children's spontaneous speech." *Methods for assessing children's syntax*. McDaniel, D., C. McKee, and H. Smith-Cairns (eds.), Cambridge, Mass.: MIT Press, 22-53.
- Ticio, Emma (2002). "Dialectal Variation in the acquisition of the null subject parameter." *Proceedings of Generative Approaches to Language Acquisition 2001*. Costa, J., and M. J. Freitas (eds.). Lisbon: Associação Portuguesa de Linguística.
- Ticio, Emma (2004). "On the position of subjects in Puerto Rican Spanish." *The University of Connecticut Working Papers in Linguistics (UCONNWPL)* 12. Rodríguez-Mondoñedo, M., and E. Ticio (eds.).
- Toribio, Almeida Jacqueline (2000). "Setting Parametric Limits on Dialectal Variation in Spanish." *Lingua* 110, 315-341.
- Uriagereka, Juan (2002). "Evidential Contexts." Ms., University of Maryland, College Park.
- Uribe-Etxebarria, Myriam (1991). "On the Structural Positions of the Subjects in Spanish, their Nature and their Consequences for Quantification." Ms., University of Connecticut, Storrs.
- Villa-García, Julio (2010a). "A new look at the 'null subject' stage in early *pro*-drop and no-longer-*pro*-drop languages." Paper presented at the 40th Linguistic Symposium on Romance Languages (LSRL-40), University of Washington, Seattle.
- Villa-García, Julio (2010b). "To Agree or Not to Agree: Beyond Quintessentially Syntactic Agreement in Spanish." *Romance Linguistics 2009: Selected papers from the 39th Linguistic Symposium on Romance Languages (LSRL), Tucson, Arizona, March 2009*. Colina, S., A. Olarrea and A. M. Carvalho (eds.). Amsterdam/Philadelphia: John Benjamins, 249-266.
- Villa-García, Julio (in preparation). "Recomplementation and Locality of Movement in Spanish." Ms., University of Connecticut, Storrs.
- Villa-García, Julio, and Neda Todorović (in preparation). "What will the Serbo-Croatian-acquiring child use first, SV or VS?" Ms., University of Connecticut, Storrs.
- Villa-García, Julio, and William Snyder (to appear). "On the Acquisition of Overt Subjects, Topics and Wh-Questions in Spanish." *Proceedings of Generative Approaches to Language Acquisition 2009 (GALA 2009)*. Castro, A, J. Costa, M. Lobo, and F. Pratas (eds.). Cambridge: Cambridge Scholars Publishing.
- Villa-García, Julio, William Snyder, and José Riqueros-Morante (2010). "On the analysis of lexical subjects in Caribbean and Mainland Spanish: Evidence from L1 acquisition." *Proceedings of the 34th Annual Boston University Conference on Language Development (BUCLD 34)* (Vol. 2), Franich, K., L. Keil, and K. Iserman. Somerville, Mass.: Cascadilla Press, 433-444.
- Zubizarreta, María Luisa (1998). *Word Order, Prosody and Focus*. Cambridge, Mass.: MIT Press, 333-344.
- Zubizarreta, María Luisa (1999). "Word order in Spanish and the nature of nominative Case." *Beyond Principles and Parameters*. Johnson, K., and I. Roberts (eds.). Dordrecht: Kluwer, 223-250.

Selected Proceedings of the 13th Hispanic Linguistics Symposium

edited by Luis A. Ortiz-López

Cascadilla Proceedings Project Somerville, MA 2011

Copyright information

Selected Proceedings of the 13th Hispanic Linguistics Symposium
© 2011 Cascadilla Proceedings Project, Somerville, MA. All rights reserved

ISBN 978-1-57473-442-3 library binding

A copyright notice for each paper is located at the bottom of the first page of the paper.
Reprints for course packs can be authorized by Cascadilla Proceedings Project.

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Villa-García, Julio. 2011. Acquisitional Evidence Bearing on the Account of Preverbal and Postverbal Subjects in Spanish. In *Selected Proceedings of the 13th Hispanic Linguistics Symposium*, ed. Luis A. Ortiz-López, 178-189. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #2486.