Early null and overt subjects in the Spanish of simultaneous English-Spanish bilinguals and Crosslinguistic Influence

Julio Villa-García and Imanol Suárez-Palma
University of Manchester / University of Arizona

This study assesses the scope of the Crosslinguistic Influence (CLI) hypothesis’ predictions with regard to early bilingual acquisition. To this end, we analyze longitudinal corpus data from four bilinguals attesting the acquisition of subjecthood (null versus overt; preverbal versus postverbal) and the pragmatic adequacy of early null and overt subjects in a null-subject language (i.e., Spanish) in combination with a language differing in its pro-drop parameter setting (i.e., English). Our results indicate that CLI barely affects the development of subjects in the null-subject language at the initial stages, namely at the outset of null and overt subjects, and in turn support the Separate Development Hypothesis. Our bilingual cohort patterns with their Spanish-acquiring monolingual peer in that both groups display comparable proportions of null subjects as well as acquisitional trajectories of null and overt subjects at the early stages of acquisition. Much like monolinguals, bilinguals begin to produce preverbal and postverbal subjects concurrently. The bilingual children and the monolingual child of this study actually produce extremely high rates of pragmatically appropriate covert and overt subjects, which are for the most part target-like from the start, thus pointing to the absence of CLI effects. In light of monolingual and bilingual data, the paper also revisits the hotly debated issue of the ‘no overt subject’ stage of Grinstead (1998, et seq.), its existence in child Spanish being questionable.

Keywords: early bilingualism, null subjects, overt subjects, Crosslinguistic Influence, pragmatic appropriateness

1. Introduction

By analyzing corpus data from the domain of Bilingual First Language Acquisition (BFLA), we aim to achieve two objectives regarding the realms of acquisition and
syntax, with special reference to Spanish, a prototypical null-subject language that manifests non-overt subjects alongside preverbal and postverbal subjects.

The first goal is to further assess the Crosslinguistic Influence (CLI) hypothesis first laid out in Hulk and Müller (2000) in light of longitudinal data from bilingual children acquiring Spanish, a Romance pro-drop (i.e., null-subject) language, along with English, a non-pro-drop language. The CLI sets two major conditions that increase the likelihood of a phenomenon to be vulnerable to CLI effects: first, crosslinguistic interaction is likely to be found in linguistic phenomena pertaining to the syntax-pragmatics interface. Second, there must be a certain degree of superficial overlap between the two languages with respect to the phenomenon in question, although the underlying syntax of the construction in each language may be different. The focus of this study is null and overt subjects in Spanish, whose distribution is contingent on factors that include lexico-semantic and information-structure considerations. Such constructions are therefore related to the syntax-pragmatics interface, which plays an important role in their distribution, and are to a certain extent parallel and overlapping at the surface level in Spanish-style pro-drop and English-style non-pro-drop languages. Thus, null and overt subjects are deemed likely to be more susceptible to CLI (Hulk & Müller, 2000). In order to ensure that our assessment of the trajectory of bilingual acquisition of subjecthood is on the right track, we also analyze data from monolingual Spanish. The comparison of the two types of data will allow us to determine more precisely the degree of CLI.\(^1\) Since the CLI hypothesis is inextricably linked to the intersection between syntax and pragmatics, we investigate not only the proportion of null vs. overt subjects in early grammars, but also whether or not early null and overt subjects are pragmatically appropriate.\(^2\) While previous studies on bilingual English-Spanish acquisition, such as that of Paradis and Navarro (2003), focused on the pragmatic adequacy of overt subjects (with preposed and postposed subjects collapsed in one category), in this study we consider these overt subjects separately, concentrating on the adequacy of preverbal and postverbal subjects (see also Silva-Corvalán, 2014).

---

1. In future studies it would be relevant to contrast various language pairs with similar and different pro-drop settings in order to determine whether divergent outcomes between bilinguals and monolinguals are in fact a CLI effect or could be attributed to the nature of bilingualism itself (e.g., more reduced input in the case of bilinguals, whose linguistic experience is divided between two languages).

2. Note that the criterion here is felicisseness from the point of view of an adult speaker; a null-subject is likely to be used in non-emphatic first- and second-person contexts, as well as in cases where the referent of the covert subject is salient in the preceding discourse. We discuss the relevant constraints governing the distribution of overt subjects in due course.
The second goal of this paper is (ii) to shed light on the contentious analysis of preverbal (SV) and postverbal (VS) subjects in Spanish, an area of syntactic research that spans more than a quarter of a century. The analysis of pre- and postverbal subjects in Spanish remains unclear and continues to be the object of intensive research. By examining the time-course of acquisition of null and overt subjects in Spanish, we hope to also illuminate syntactic research. As previous studies on English-Catalan and English-Spanish bilinguals have not focused on the issue of SV/VS (see Juan-Garau & Pérez-Vidal, 2000; Liceras, Fernández-Fuertes, & Pérez-Tattam, 2008; Paradis & Navarro, 2003), with the exception of Silva-Corvalán (2014), the current project intends to investigate this aspect of acquisition in more detail. Despite Pierce’s (1992) claims that postverbal subjects have special status in child grammars, we believe that examining the trajectory of acquisition of preverbal and postverbal subjects in the evolving Spanish grammar of bilingual children acquiring Spanish alongside a non-pro-drop language like English is necessary. This is an area involving the syntax-pragmatics interface where CLI could occur (i.e., a greater incidence of the SV word order due to interference from English, where the order VS is not readily available). Thus, exploring this aspect through the lens of BFLA will also contribute to goal (i).

More generally, we hope to further our knowledge of and make a contribution to BFLA with particular attention to the acquisition of subjects in Spanish. This will be accomplished via supplementing existing studies in a novel way by looking at a larger number of corpora, by investigating the issue of the ‘null subject’ stage advocated by Grinstead (1998) for Catalan- and Spanish-acquiring children in light of bilingual acquisition, by devoting attention to the contexts in which early null and overt subjects occur, and by focusing on the timeline of the acquisition of preverbal and postverbal subjects in Spanish.

The paper is organized as follows: in Section 2 we discuss the CLI hypothesis in light of subjects; in Section 3, we offer a cursory look at the syntax of subjecthood in terms of the Null Subject Parameter; in Section 4, we review previous studies on the monolingual and bilingual acquisition of subjects in Spanish; in Section 5, we present the research questions and the methods used to address them; in Section 6, we report the results of this project; in Section 7, we further consider the results and offer some concluding remarks.

3. The diary studies summarized in Pierce (1989) lead to the conclusion that the VSO order is the first order attested in child Spanish. However, Villa-García (2011) shows that the VSO order, which is found in Spanish but not in other Romance languages like Italian, actually seems to appear later than the SVO and VOS orders, which the author attributes to the low frequency of the VSO construction in adult Spanish.
2. The Crosslinguistic Influence Hypothesis regarding subjecthood

In the realm of bilingualism, a host of studies has supported the view that in their bilingual mind, children have two differentiated grammatical systems that primarily develop separately. This is in accordance with the Separate Development Hypothesis (SDH) (De Houwer, 1990; Genesee, 1989; Meisel, 1989), but is in contrast to the Unitary System Language Hypothesis championed by Volterra and Taeschner (1978). This being said, the SDH does not exclude the possibility of interlinguistic influence from one language to the other. In fact, there is substantial evidence that there is cross-language interference in bilingualism (see, e.g., Serratrice & Hervé, 2015, and references therein).

The seminal work of Hulk and Müller (2000) and Müller and Hulk (2001) maintains that such cross-language interference is likely to be found in linguistic phenomena belonging to the intersection between syntax and pragmatics (i.e., the syntax-pragmatics interface). A second condition for CLI is that there must be a certain degree of superficial correspondence between the two languages with respect to the phenomenon in question, although the underlying syntax of the construction in each language might differ. To borrow Hulk and Muller’s (2000) terminology, Language A will influence Language B if the surface phenomena of Language B include the relevant surface phenomena of Language A.

Subjects constitute such a case: on the one hand, the distribution of null and overt subjects in pro-drop languages like Spanish is largely determined by discourse-pragmatic considerations such as old and new information, amongst others, and on the other, subjects in Spanish do show a certain degree of parallelism with their English counterparts at the surface level. For instance, the word order SVO is attested in both Spanish and English (e.g., Pedro produce muchos sujetos / Pedro produces many subjects), although it is well known that this is not the arrangement we always find in Spanish-style pro-drop languages, where clausal word order is not as rigid as in English. Consequently, it comes as no surprise that the CLI hypothesis has generated much research in the area of subjecthood in BFLA, particularly when it concerns pro-drop languages like Spanish and non-pro-drop languages like English.

With respect to longitudinal research, there have been studies investigating the degree of CLI in the acquisition of English and different pro-drop languages (Gutiérrez-Clellen, Simon-Cerejido, & Wagner, 2008; Hacohen & Schaeffer, 2007; Haznedar, 2010; Juan-Garau & Pérez-Vidal, 2000; Líceras et al., 2008; Líceras, Fernández-Fuentes, & Alba de la Fuente, 2012; Paradis & Navarro, 2003; Schmitz, Patuto, & Müller, 2012; Serratrice, Sorace, & Paoli, 2004; Silva-Corvalán, 2014; Zwanziger, Allen, & Genesee, 2005). As Serratrice and Hervé (2015) note, the studies conducted by Juan-Garau and Pérez-Vidal (2000), Líceras et al. (2008),
Silva-Corvalán (2014), and Zwanziger et al. (2005) did not find evidence in support of CLI based on the relative proportion of null and overt subjects in children’s naturalistic productions (see also Gutiérrez-Clellen et al., 2008; Liceras et al., 2012); their performance regarding the rate of non-overt and lexical subjects was reminiscent of that of monolinguals, challenging the CLI hypothesis. Nonetheless, Serratrice and Hervé (2015) emphasize the need to investigate the issue of whether subjects are pragmatically appropriate, since CLI is tightly connected to the interface between syntax and semantics/pragmatics, a consideration that was indeed taken into account by some studies. Of these studies, those that found CLI in the domain of subject realization (e.g., Hacohen & Schaeffer, 2007; Haznedar, 2010; Paradis & Navarro, 2003; Serratrice et al., 2004; Silva-Corvalán, 2014) note that the rate of pragmatically inappropriate subjects in the pro-drop language was higher in bilinguals than in monolinguals. Thus, it seems that CLI is more likely to be apparent once the syntax-pragmatics domain is taken into account. As will be seen below, Silva-Corvalán (2014) has recently drawn attention to the fact that the two bilinguals of her long-term longitudinal study began to manifest clear signs of CLI after age 4, once the non-pro-drop language (in this case, English) became more dominant (on the issue of language dominance underlying CLI, see Kupisch, 2007; Serratrice, Sorace, Filiaci, & Baldo, 2009; Yip & Matthews, 2009). We resume the discussion of these issues in due course. In the meantime, we turn to an overview of the major characteristics of null and overt subjects in contemporary Spanish.

3. On null and overt subjects in Spanish

The analysis of subjects has commanded much attention in Chomsky’s generative paradigm, particularly since the advent of the Principles and Parameters framework (Chomsky, 1981). For one thing, the requirement that languages like English have an overt preverbal subject – the traditional Extended Projection Principle (EPP) (e.g., *it rains / it rains) – remains the object of painstaking research. Similarly, Romance languages like Italian, Portuguese, and Spanish display tacitly implied, null (i.e., phonologically unrealized) subjects in finite clauses, a property that correlates with the availability of preverbal (SV) and postverbal (VS) subjects (Chomsky, 1981; inter alia). Clustering properties like these led to the postulation of the Null Subject Parameter (NSP) (aka Pro-drop Parameter), with languages like English exemplifying the negative setting, and languages like Spanish the positive one (Barbosa, 2009; Burzio, 1986; Chomsky, 1981; Jaeggli, 1984; Kenstowicz, 1989; Rizzi, 1982; inter alia). The existence of null referential subjects and the availability of subject-verb inversion (SV/VS), two of the major properties generally deemed to be subsumed under the scope of the NSP, are illustrated in (1) for

2nd proofs
Early null and overt subjects in the Spanish of simultaneous English-Spanish bilinguals

English and Spanish (the reader is referred to Camacho, 2013, for the view that the link between the properties typically associated with the NSP is not as direct as has traditionally been assumed). Note that (1b’) illustrates the NSP property of obligatorily null expletives in +NSP languages like Spanish.

(1)  
-NSP (e.g., English)        +NSP (e.g., Spanish)  
a. He walks / *Ø walks    a’. Él camina / Ø camina  
   he walks  
b. It rains / *Ø rains   b’. Llueva / *ello llueve  
rains it rains  
c. The king has died / *has died the king  c’. El rey ha muerto / ha muerto el rey  
   the king has died

Null subjects in Spanish occur in contexts where they are mandatory, such as with atmospheric predicates, as in (1b’), or when the referent of the subject is clear in the preceding discourse, as shown in (2). This example contains a null subject in the second sentence. Here, the subject is salient and carries the [-topic shift] pragmatic feature (i.e., it is a case of topic continuity); hence, a null subject (Ø) is favored.

(2) Bendicho y yo estuvimos en España. ¡Ø Nos encantó!
   Bendicho and I were in Spain. cl. charmed
   ‘Bendicho and I were in Spain. We loved it!’

Subject pronouns in English-style non-null subject languages perform a clearly grammatical function, namely that of being overt person and number markers. These functions are arguably performed by the morphological endings of the verb in pro-drop languages such as Spanish. Thus, pronominals in [+NSP] languages like Spanish are restricted to focal (i.e., emphatic and contrastive) contexts and to cases where they are needed to identify the referent, since their role is not to mark person and number, unlike their English counterparts (Silva-Corvalán, 2014). In this sense, therefore, pronouns in languages like Spanish are not merely optional but also serve a discursive function.4

---

4. The NSP has generated a great deal of research from different angles over three decades, trying to answer questions such as what licenses null subjects (i.e., the empty category pro), and how they are identified. The main accounts to date range from analyses in which pro is the empty category hosted in the canonical subject position (Spec, IP/Tc) and those in which it is in fact the agreement morphemes in the verb that function as the real subject, to those according to which null subjects are the result of an elided overt subject that has been deleted in P(phonological) F(form) (see Barbosa, 2009; Camacho, 2013, for an overview and recent developments).
In addition to analyzing the properties of null subjects in pro-drop languages like Spanish, much effort has been devoted to characterizing the behavior and distribution of overt preverbal and postverbal subjects. Preverbal subjects are generally featured in canonical (SVO) sentences and in sentences where the subject is topical, as illustrated in (3a) and (3b–e) respectively. In (3b), the object el periódico is the only constituent that represents new information (focus), with Juan compró as given or old information that is presupposed (a case of topic continuity, from the preceding sentence). In (3c), the nominal phrase las ondas gravitacionales constitutes given information; hence, it appears preverbally (note that in this example, the subject could have been null, much like in (2), since it would have been obvious by virtue of the context and the third-plural agreement on the verb; this alternation actually shows that null and overt subjects are not necessarily in complementary distribution). In (3d), el País Vasco and Asturias are preverbal subjects in their respective sentences, for they instantiate topic-shift subjects. In (3e), the subject precedes an uncontroversially CP-related element—a wh-item. Thus, such subjects are topic elements in the left periphery of the clause (see Villa-García, 2015, Ch. 3).

(3) a. **What happened?**
   A. Pedro se salió de cura
   Peter cl. left of priest
   ‘Peter left the priesthood.’

b. **What did John buy?**
   A. Juan compró el periódico
   John bought the newspaper
   ‘John bought the newspaper.’
   [Olarrea, 2012, p. 604]

b. **Qué ha venido pasando con las ondas gravitacionales?**
   what has come happening with the waves gravitational

b. Bueno, las ondas gravitacionales están relacionadas con la well the waves gravitational are related with the
   Teoría de la Relatividad…
   ‘Theory of the Relativity
   ‘A: What has been going on with gravitational waves? B: Well,
   gravitational waves are related to Relativity Theory…”
   [Spontaneous dialog. Canal 24 Horas, Spain, 22 February 2016]

d. Estuvimos viajando. El País Vasco tiene mucho que ofrecer.
   were traveling the country Basque has much that offer
   Y Asturias nos encantó
   and Asturias cl. charmed
   ‘We were traveling. The Basque Country has a lot to offer. And we really
   liked Asturias.’

e. **Juan, ¿cuándo viene?**
John when comes
‘As for John, when is he coming?’

Postverbal subjects, for their part, tend to occur if they constitute new information or corrective/contrastive focus (see (4a)); if they are subjects of unaccusative and psychological predicates ((4b) and (4c)); if they constitute subjects in locative inversion structures (4d); if they are subjects in thetic sentences expressing mere events (4e); if they constitute sentential subjects (4f); if they are subjects in focus/wh-questions (4g); and if they are part of the stylistic inversion pattern attested in direct speech (4h). The reader is referred to Ortega-Santos (2008, 2016) and to Silva-Corvalán (2014) for discussion and to Villa-García (2011) for an overview of the asymmetries found between preverbal and postverbal subjects in Spanish-style pro-drop languages.

(4) a. ¿Quién compró la carne? Me preguntó quién había comprado who bought the meat cl. asked who had bought la carne y le dije que la había comprado Juan the meat and cl. said that cl. had bought John ‘Who bought the meat? S/he asked me who had bought the meat and I told them that John bought it.’

b. Llegó un hombre a la puerta arrived a man to the door ‘A man got to my door.’

c. Me gusta Ottawa cl. pleases Ottawa ‘I like Ottawa.’

d. Aquí anidan gavilanes here nest sparrowhawks ‘Sparrowhawks nest here.’

e. Ha muerto Franco has died Franco ‘Franco has died.’

f. Es bueno que se estudie sintaxis is good that cl. studies syntax ‘It’s good that syntax is studied.’

g. ¿Qué compró Susan what bought Susan ‘What did Susan buy?’

h. “Nunca saldremos de la crisis”, aseguró el presidente never exit of the crisis assured the president ‘Never will we get out of the economic crisis,’ warned the president.”
Set against this background, the literature is full of different proposals intended to account for the syntax of preverbal and postverbal subjects. As regards the former, the major debate has revolved around whether preverbal lexical subjects are canonical subjects in the inflectional layer (i.e., SpecAgrSP/SpecTP), as is commonly assumed for English, whether they are left-peripheral (A-bar) constituents akin to topics in the CP layer, or whether they can occupy different preverbal slots (see, for instance, Villa-García, 2015, for recent discussion on the syntax of preverbal subjects). As far as postverbal subjects are concerned, there are several proposals that account for their syntax while trying to capture their being regulated by considerations that go beyond syntax proper. These range from analyses wherein postverbal subjects sit in a rightward specifier of TP, or where they are in-situ elements in vP, to accounts whereby subjects that appear after the verb are the result of pronouncing a low copy of the subject in a non-trivial (i.e., movement) chain (the reader is referred to Ortega-Santos, 2016, for recent discussion of the syntax of postverbal subjects and a thorough overview of the existing accounts). Whatever the right account of preverbal and postverbal subjects turns out to be, at present there is no unified account of the two types of subject (arguably, the account according to which subjects are the result of pronouncing different copies of a moved element depending on PF considerations is partially uniform, since the two types of subject would be treated exactly the same way in the syntax; however, there is a difference in the PF of the structures with preverbal subjects and those with postverbal ones, which constitutes a departure from a totally unified account). Of course this abundance of proposals is to be expected given the different behavior displayed by preverbal and postverbal subjects in Spanish and the various constraints that govern their distribution (see the different contexts for postverbal subjects illustrated in (4)).

The fact remains that the distribution of subjects in languages like Spanish is not dictated by purely grammatical constraints; it is largely determined by discourse-pragmatic considerations: the choice of a preverbal or a postverbal subject is made on the basis of criteria such as whether the subject is topical (SVX) or focal (VXS) in nature, for instance, and with respect to null subjects, their understanding is heavily dependent on the context (e.g., the referent of a non-overt third-person subject, for example, needs to be salient in the discourse). Thus, successful grammatical and pragmatic use of subjects in languages like Spanish in both children and adults implies that the speaker has access to the syntax-pragmatics interface, a hypothesis that we test acquisitionally in this study.
4. Previous studies on the early acquisition of subjects in Spanish in monolingual and bilingual contexts

The occurrence of null subjects that coexist with overt subjects in the speech of children exposed to non-pro-drop languages like English has attracted a great deal of attention in a body of research that has resulted in a plethora of proposals, ranging from grammatical (e.g., Hyams, 1986, *et seq.*; Hyams & Wexler, 1993; Rizzi, 1994; Valian, 1991) to processing (e.g., Bloom, 1990) explanations of the phenomenon in question.

As is known, null subjects are also attested in the early speech of monolingual children acquiring pro-drop languages like Catalan and Spanish (Grinstead, 1998, *et seq.*). In fact, Grinstead has argued that unlike children learning English, Catalan- and Spanish-acquiring children pass through a ‘null subject/no overt subject’ stage, during which all of the child’s subjects are null (see (5)). This phase lasts until approximately age 2 (see also Austin, Blume, Parkinson, Lust, & Núñez del Prado, 1997), which contrasts markedly with the observation that both null and overt subjects are found in the speech of English-acquiring children.

\[ (5) \text{ The 'null subject' stage in the acquisition of Spanish } \]

\[ \begin{array}{c}
100\% \text{ null subjects} \\
\downarrow \quad = \quad 2 \text{ yrs.} \\
\quad \text{null and overt subjects}
\end{array} \]

Several authors have raised a number of criticisms against Grinstead’s claim that children learning Catalan and Spanish do not use lexical subjects at an early stage in development. Most notably, Aguado-Orea and Pine (2002) argue against the ‘null subject’ stage by claiming that Grinstead’s (2000) data for Spanish are rather sparse and his conclusions are based on only one child. Similarly, the authors note that Grinstead’s (2000) hypothesis is difficult to test, in part because children’s early utterances occur during the ‘one word’ stage. In parallel fashion, Bel (2001, 2003) reports that Catalan and Spanish children produce overt subjects at a rate of approximately 33% from the earliest utterances, with no significant increments in the use of overt pronouns after that point. Thus, Bel concludes that English-, Catalan-, and Spanish-acquiring children do not differ from each other in this respect.

Nevertheless, the criticisms against Grinstead’s original claims are not fully justified. First, Grinstead (1998) and Grinstead and Spinner (2009) used data from more than one Spanish-acquiring child. Moreover, it is important to consider that Grinstead’s claim that the ‘null subject’ stage ends at around age 2 is just an.
approximation (see Brown’s 1973 ‘stages, not ages’). Grinstead and Spinner (2009) underscore that their claim is a grammatical, rather than a chronological claim. Therefore, Grinstead’s estimate cannot realistically be taken to mean chronological age. A given child may move to a period during which his/her speech starts to exhibit lexical subjects at age 1.6, whereas another child may do so at age 2.3 (Grinstead & Spinner, 2009). Along the same lines, note that the earliest transcript examined by Bel for a Catalan-acquiring child contains data recorded at age 1.6, with all remaining transcripts starting after age 1.7. Given the preceding discussion that the age of 2 is a mere approximation, it is entirely possible to find overt subjects in the earliest transcript analyzed, which is indeed confirmed by the data from one of the children studied by Bel, namely María (see the López-Ornat Corpus, CHILDES). In fact, as shown in the following sections, some of the children of the present study started using overt subjects before age 2. Consequently, it cannot be concluded on the basis of the issues raised by the aforementioned authors that children acquiring quintessentially null-subject languages like Spanish do not go through a ‘null subject’ stage, although the opposite cannot be proven either.

Grinstead and Spinner (2009) endeavor to show that Catalan- and Spanish-acquiring children do not show evidence of using overt subjects from the beginning of two-word speech. However, the authors focused on the Mean Lengths of Utterance (MLUs) at the time when overt subjects started to be used, not before. If there is actually a ‘no overt subject’ stage, one would expect to find only null subjects at an MLU where at least two slots could be filled (e.g., subject-verb). Similarly, the ‘null subject’ phase makes an important prediction: there should be pragmatic errors in the use of covert subjects, since the child would in principle be using null subjects in contexts that would require an overt subject in adult speech. To the best of our knowledge, this prediction has not been tested to date. Whatever the case may be, Villa-García (2013) has recently found a statistically significant difference between the onset of null subjects and that of overt subjects in the speech of two Spanish-acquiring monolingual children, with MLUs that suggest that the children were past the holographic (i.e., one-word) stage. This finding thus constitutes evidence in favor of Grinstead’s ‘null subject’ phase hypothesis. Hence, there seems to be evidence that children acquiring null-subject languages like Spanish do drop subjects substantially at the early stages. We revisit this far-from-settled question in the light of monolingual and especially bilingual data, a task that the existing works have not undertaken thus far.

Another claim arising from the work of Grinstead and colleagues is that children acquiring pro-drop languages like Spanish do not acquire lexical subjects until their left periphery (aka the CP layer) is active, since lexical subjects in Spanish are analyzed under certain accounts as being discourse-pragmatic constituents.
situated in the left-periphery (see Villa-García, 2015, for a recent overview of this longstanding debate). Therefore, the prediction made by this hypothesis is that children will start using overt subjects at the same time as uncontroversially left-peripheral phenomena such as topics/Clitic-Left Dislocations (CLLD) and foci/wh-questions. This proposal came to be known as the 'Interface Delay' hypothesis, support for which was in fact provided by the studies conducted by Grinstead (1998, 2004) and Grinstead and Spinner (2009), who found a correlation between the emergence of lexical subjects and CP-related phenomena in a study featuring Spanish- and Catalan-acquiring children (see also Villa-García & Snyder, 2010; however, see Villa-García, 2012, for a dissenting view). As noted by an anonymous reviewer, an alternative explanation for the alleged absence of focused and topicalized constructions could be provided based on syntactic deficits, characterized in terms of the nature and number of operations (i.e., syntactic complexity) involved the derivations that concern the left periphery. Similarly, note that although topicalizations are present at a low rate in child spontaneous production, children seem to display sensitivity to topic/comment structures with a marked topic as well as occurrences of wh-questions (along these lines, the reader is referred to Soares, 2003, who examines child European Portuguese, with children as young as 1;4). An additional question for the 'Interface Delay' proposal is what it has to say about languages like English, where subjects are typically assumed to occupy a position in the inflectional domain.

A noteworthy finding of Grinstead (1998, 2004) and Villa-García (2011) is that when lexical subjects begin to emerge in the spontaneous speech of Catalan-and Spanish-acquiring children, both preverbal and postverbal subjects arise concurrently. Villa-García (2011) takes this finding to argue for a unified account of preverbal and postverbal subjects in languages like Spanish, although currently there is no uniform theory that predicts this finding (see Section 3). In any case, it seems that the prerequisites for the child to start using preverbal and postverbal subjects successfully in Spanish may differ, but crucially the available evidence of simultaneous acquisition of the two types of subject suggests that at least the last prerequisite is shared by pre- and postverbal subjects.

Having discussed relevant aspects of the monolingual acquisition of subjects, we now turn to corpus studies investigating the BFLA of subjects in Spanish.5 The first study concerning English and Spanish bilinguals is that of Paradis and Navarro (2003). The authors studied Manuela, whose data we revisit in the present paper. Manuela’s mother is a speaker of British English, while her father is

---

5. The reader is referred to Camacho (2013, Ch. 9) and Silva-Corvalán (2014, Ch. 4–5) for thorough reviews of the existing accounts, including reference to relevant L2 and heritage-speaker studies.
a native speaker of Cuban Spanish. Paradis and Navarro (2003) is a pioneering study since the authors did not concern themselves only with subject proportions, but also with the contexts in which overt subjects occur as well as the potential role of the input in the process of language acquisition. The authors report that Manuela's proportion of overtly expressed subjects is higher than that of monolinguals, which at face value could be interpreted as CLI, but cautiously conclude that the Spanish input received by the child, Cuban Spanish, a variety of Caribbean Spanish with a higher rate of explicit subjects than its mainland Spanish counterparts, may have played a role in determining the higher quantity of lexical subjects used by the child. In fact, their analysis of the parental input confirms that the proportion of overt subjects in the adult speech is rather high. Paradis and Navarro (2003) go on to claim that Manuela employs lexical subjects inappropriately in low informativeness contexts, and conclude that crosslinguistic effects may occur in both the syntax and the pragmatics of the syntax-pragmatics interface. Silva-Corvalán (2014, p. 133) reviews this study and raises some valid criticisms. She argues against the use of a category devoted to low informativeness for lexical subjects and concludes that the evidence adduced in the Paradis and Navarro study is indicative that input plays a major role, rather than an "internal, psycholinguistic mechanism of transfer" (Silva-Corvalán, 2014, p. 134).

Using spontaneous production data from Leo and Simon, two twin brothers raised in Spain, Liceras et al. (2008) and Liceras et al. (2012) have shown that the children's developing grammars in English and Spanish follow separate routes of acquisition. Liceras et al. (2008) adopt the assumption that the directionality of CLI will be from the non-pro-drop to the pro-drop language, given that the mechanisms underlying a [-pro-drop] language are deemed to be simpler. First, the authors note that overt subjects in English are much more frequent from the earliest stages, attaining a stable 95% at the end, which mirrors the development of monolingual English-acquiring children. Their Spanish too is akin to that of monolinguals, a conclusion that we reinforce in our study of the two brothers, with an emphasis on the appropriateness of their early null and overt subjects. Liceras et al. (2008) claim that the twins acquired inflections very early, with a very low rate of omission of inflectional markers at stage I (identified according to MLU), namely 4.28%. The Spanish of these twins manifests a large percentage of null subjects that decreases only slightly as a rather stable ratio of null

---

6. In a similar vein, Sorace (2004) notes that crosslinguistic transfer is thought to occur from a more economical language A to a less economical language B, regardless of language dominance (see below on this issue). Spanish subjects may be more complex and "less economical" than English subjects, because of the syntactic, semantic, and pragmatic restrictions governing their distribution (see Section 3).
and lexical subjects across stages. The conclusion reached by Liceras et al. (2008, p. 131) is that there is “clear evidence that there are two separated systems in the bilingual mind.”

The longitudinal study conducted by Silva-Corvalán (2014), using a larger amount of data than any other study, considers the development of two siblings from age 1;6 to age 5;11, with differences in terms of language exposure that, over time, correlate with different patterns of lexical subject realization in Spanish (Serratrice & Hervé, 2015). From very early on, the siblings treat subjects differently in the two languages, the proportion of overt subjects in English being almost double than that in the Spanish data, revealing a developmental path similar to that of monolinguals. The use of subjects in the developing English of these two children mirrors that of monolinguals; their Spanish, however, shows that the use of subjects is affected by a reduction in their exposure to this language as time goes by. One of the siblings is comparable to monolinguals until age 4;0, at which time his pronoun use increases and pragmatically inappropriate uses occur (mainly a few cases of redundant pronouns). His brother, on the contrary, uses a higher proportion of overt subjects than expected from the beginning, with an evidently increased rate in the 3;0–3;11 period, and makes pragmatically inappropriate use of subjects at times. Silva-Corvalán (2014, p. 163) points to the possibility that the subject pronouns that are not pragmatically validated may be due to “the children’s copying of the overt subject requirement for their stronger language, English (i.e., external influence).” The [subject pronoun + verb] string frequently found in English could be active in the child’s mind, resulting in the child’s copying this structure into Spanish with the ensuing higher rate of pronominals. Silva-Corvalán’s (2014) overall conclusion is that no CLI from English was found in the existing studies (including her own study) before the age of 4.

As for the position of the subject, Silva-Corvalán (2014, p. 172) reports no influence from Spanish on the English canonical (SVO) word order from the outset, but with respect to Spanish, she notes that by the end of her study, at age 5;11, the siblings “did not appear to have reached complete productive mastery of the factors that govern the placement of overt subjects in Spanish.” First, Silva-Corvalán contends that pronominal subjects in Spanish are overwhelmingly preverbal across time for both children. For non-pronominal subjects, the author notes that preverbal subjects across verb types occur 62.27% of the time for one

7. In line with the study just reviewed, Silva-Corvalán (2014) predicts that the direction of interference will be from the non-null-subject language to the null-subject one, this vulnerability owing to the fact that subjects in pro-drop languages are syntactic phenomena constrained by lexico-semantics and discourse pragmatics, and thus their derivation involves a higher degree of complexity.
child and 44.4% of the time for the other from 1;7 to 1;11.30. The children also display sensitivity to the type of verb (e.g., a larger proportion of postverbal subjects with unaccusatives for both children) until the age of 4;0. However, from that point onwards, both children begin to manifest a stronger preference for the preverbal position, with the rate of preverbal subjects increasing to approximately 85%, including unaccusatives.8 In spite of the rise in the number of preverbal subjects, Silva-Corvalán encountered very few examples of clearly unjustified preverbal subjects. On the whole, Silva-Corvalán concludes that the English word order becomes more prevalent as English becomes more dominant and exposure to Spanish is reduced. An important conclusion drawn from the work of Silva-Corvalán (2014) is that no noticeable crosslinguistic interference, whose locus of vulnerability is deemed to lie at the syntax-pragmatics interface, is found before the age of 4;0 in bilingual children, a fact that the author attributes to English becoming the most dominant language of the two.9

Regarding other pro-drop/non-pro-drop combinations, Juan-Garau and Pérez-Vidal (2000, p. 189) have shown that a child acquiring Catalan and English simultaneously separates both systems from the outset with regard to subjects, pointing to "the absence of any major influence of one language on the other." Along the same lines, Zwanziger et al. (2005) found no evidence of interlinguistic interaction regarding subject realization in a study featuring Inuktitut-English bilinguals. Moreover, Serratrice et al. (2004) also report no evidence of CLI in the speech of an Italian-English balanced bilingual child, with some pragmatically inappropriate lexical subjects at later stages that should have been null from the perspective of an adult speaker of Italian. However, the authors also note that a monolingual control made similar pragmatic errors, which diminishes the hypothesis that there is interlinguistic interference in the bilingual child’s Italian development.

In sum, the monolingual studies reviewed above have focused on the putative 'no overt subject' stage in child Catalan and Spanish (Grinstead, 1998, et seq.), whose existence remains to be determined, and on the time-course of acquisition of overt subjects. As far as bilingual children are concerned, most existing studies are framed within the Crosslinguistic Influence (CLI) hypothesis. Paradis and Navarro’s (2003) seminal study indicates that Manuela, a bilingual child acquiring Cuban Spanish, uses a higher proportion of overt subjects than her monolingual


9. Thus, this is in line with the claim made by authors including Kupisch (2007), Serratrice et al. (2009), and Yip and Matthews (2009) that language dominance is another factor determining CLI.
peers, which could be ascribed to CLI, since cases of pragmatically inappropriate overt subjects were found in the speech of this child. However, their results are admittedly inconclusive since the input the child received in Spanish came from Cuban Spanish, a variety displaying high proportions of phonetically realized subjects, whose distribution does not seem to rely so much on discourse-pragmatic considerations, as is the case in other varieties of Spanish. The Liceras and colleagues (Liceras et al., 2008; Liceras et al., 2012) studies do not find a CLI effect regarding subject provision, and conclude that bilinguals pattern with monolinguals, although they do not take pragmatics into consideration. Finally, Silva-Corvalán (2014) does not find evidence for CLI at the early stages, but shows that as proficiency in Spanish decreases due to English dominance, CLI becomes more apparent with regard to subjects. The current study revisits the issue of CLI in the acquisition of subjects in bilingual English-Spanish children, with particular attention given to aspects that have not been the major focus of the existing BFLA studies, such as the ‘No Overt Subject’ phase, the timeline of acquisition of preverbal and postverbal subjects, and the appropriateness of early null and overt subjects.

5. Research questions and methods

Drawing on previous studies, we pose the following questions, the first one, about CLI, being more general, and the second and third being contingent on the first one (i.e., a positive answer to questions (ii)–(iii) would in principle be compatible with a scenario where CLI does not occur). We also include a number of attending predictions.

i. Does the bilingual acquisition of a pro-drop language (e.g., Spanish) in tandem with a non-pro-drop language (e.g., English) display CLI effects with respect to the grammatical development of null and overt subjects?
   - If English is exerting an influence on Spanish from the start, we should observe a higher proportion of overt subjects from the outset, rather than an overwhelming proportion of null subjects, as has been claimed for monolingual Spanish. The opposite would indicate that bilingual children’s Spanish grammatical development in terms of subjects is parallel to that of their monolingual peers.

ii. Does the pro-drop language of the bilingual child pass through a stage in which only non-overt subjects are attested, as has been reported for monolinguals (Austin et al., 1997; Grinstead, 1998, et seq.; Villa-García, 2013)? In that sense, are early null subjects pragmatically felicitous in the developing pro-drop language of bilingual children?
If it is true that children acquiring Spanish do display a ‘no overt subject’ stage, their bilingual peers should pass through the same phase if no CLI occurs. Similarly, if the ‘no overt subject’ stage exists, we should find instances of pragmatically inappropriate early null subjects in both monolingual and bilingual children’s speech. More specifically, regarding the pragmatic felicity of early null subjects, a number of predictions can be derived, as follows:

- If neither monolingual Spanish nor bilingual English-Spanish children produce infelicitous null subjects, then it could be the case that children do not go through a ‘null subject’ stage, contra Grinstead (1998, et seq.). However, it could also be the case that children may be producing null subjects because that is what Spanish pragmatics requires, behaving like monolingual Spanish children.

- If both bilinguals and monolinguals produce infelicitous null subjects in the same proportion, then it follows that monolingual and bilingual Spanish children do go through a ‘no overt subject’ stage. Consequently, in this scenario bilingual English-Spanish children would behave much like monolinguals, in support of the Separate Development Hypothesis.

- Finally, if monolingual Spanish children produce infelicitous null subjects but bilingual English-Spanish children do not, then the monolinguals go through a ‘null subject’ phase and bilinguals may be taken to show CLI from English.

iii. Do bilingual children acquire (overt) preverbal and postverbal subjects concurrently, in much the same way as their monolingual counterparts (Grinstead, 1998, et seq.; Villa-García, 2011)? Are preverbal and postverbal subjects used appropriately from the outset?

- If CLI from English takes place, then we should observe large proportions of preverbal subjects in Spanish, as demanded by the grammar of English, where postverbal subjects are highly constrained. The opposite result would argue that children in bilingual contexts behave like their monolingual counterparts.

- If bilingual children overuse overt subjects, then they will provide evidence for CLI from English, particularly in the case of bilingual Iberian-Spanish-acquiring children, whose parents do not overuse overt subjects, unlike their Caribbean (e.g., Cuban) counterparts.

10. These questions are still open for monolingual Spanish, and it is our hope that the current study can contribute to this ongoing debate.
In order to address these issues, we retrieved longitudinal data from the Child Language Data Exchange System (CHILDES) database (MacWhinney, 2000). A total number of five different corpora were considered, three of which correspond to four bilingual children acquiring languages differing in their pro-drop parameter setting (English-Spanish). The other corpus attests the acquisition of Spanish by a child in a monolingual setting, which serves as the control case study (i.e., the baseline), though see fn. 1.

The first bilingual corpus in this study is that of Carla (CHILDES, Pérez-Bazán, 2002). Carla is an American child living in Michigan, USA. She is an only child whose father is a US citizen and her mother is from Spain. The language they use at home is Spanish. A logical question to pose is whether Carla had enough exposure to English during the time period considered (2;0–3;3) to regard her as an early bilingual. The English sentences that appear in her Spanish transcripts actually indicate that this child was also acquiring English at the same time as Spanish. We also analyzed the data provided in the Deuchar corpus (CHILDES Deuchar; Deuchar & Quay, 2000), which registers the development of both the English and Spanish child grammars of Manuela, born in Brighton, England. She is the daughter of the investigator, also English, and her father is Cuban. Manuela was exposed to Cuban Spanish in the home and to English by caretakers and her maternal grandmother. The rest of the data dealing with the acquisition of the aforementioned language pairs comes from the CHILDES FerPuLice corpus (Liceras, Fernández Fuentes, Perales, Pérez-Tattam, & Spradlin, 2008), chronicling the linguistic evolution of Leo and Simon, twins. The two brothers were born in Salamanca, Spain to a US-born mother and a Spanish father. The twins were addressed in both languages from birth. The twins’ parents usually communicated in Spanish at home, except for the two months in the summer when they traveled to the USA. Finally, all the bilingual data were contrasted with those of Irene, from the CHILDES Llanás-Grau/Ojea corpus (Ojea, 1997). Irene is a monolingual Spanish child born in the Principality of Asturias, Spain. She was recorded by her mother, who was also the investigator. Both of her parents are L1 speakers of Spanish. Table 1 outlines relevant data such as the age ranges considered, and the number and frequency of the files analyzed.

---

11. Ideally, Manuela’s data should be compared to that of (at least) one monolingual child acquiring Cuban Spanish. The only data available would be that of a monolingual child acquiring Puerto Rican Spanish, namely Ana, from the UConn’s Cross-Linguistic Early Syntax Study (CLESS) project. However, Ana acquired Puerto Rican Spanish, and since Caribbean Spanish turns out not to be monolithic in the studies of adult overt subject use (Otheuy & Zentella, 2012; Otheugy, Zentella, & Livert, 2007), the comparison would still not be with a monolingual child acquiring the same dialect.
Table 1. Longitudinal data of the bilingual and monolingual children of this study

<table>
<thead>
<tr>
<th>Child</th>
<th>Language(s)</th>
<th>Corpus/Database</th>
<th>Date of Retrieval</th>
<th>Age Span</th>
<th>Transcripts Analyzed (N); Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Irene</td>
<td>(M) Spanish (Iberian)</td>
<td>Iliná-Grúa/Ojía (CHILDDES) (Ojía, 1997)</td>
<td>October 30, 2015</td>
<td>01:05:27–02:01:29</td>
<td>N = 26; biweekly</td>
</tr>
<tr>
<td>Carla</td>
<td>(B) English-Spanish (Iberian)</td>
<td>Pérez-Bazán (CHILDDES) (Pérez-Bazán, 2002)</td>
<td>February 25, 2016</td>
<td>02:00:00–03:03:00</td>
<td>N = 11; mostly monthly; every two months towards the end</td>
</tr>
<tr>
<td>Leo</td>
<td>(B) English-Spanish (Iberian)</td>
<td>Fer Fürce (CHILDDES) (Liceras et al., 2008)</td>
<td>October 12, 2015</td>
<td>01:10:22–02:10:21</td>
<td>N = 25; mostly biweekly</td>
</tr>
</tbody>
</table>

* For the twins, their Spanish naturalistic speech was recorded at intervals of two-three weeks until age 3:00 (with some interruptions during the summer holidays), and then monthly from that point onwards. Their English, on the other hand, was sometimes recorded more frequently, but the sessions are usually far shorter and recorded on consecutive days (database manual, CHILDDES).

The data were searched and coded manually, since at present there is no reliable computer-assisted method to find occurrences of finite verbs and of overt subjects. We examined all available Spanish and English transcripts in the corpora, from the beginning (aimed at ensuring that the first occurrences of subjects were recorded) to the moment of production of the first finite verb with covert subjects and until the first-of-repeated uses of overt subjects were found. We then compiled the data from ten (or more) transcripts in each language (if available) from that moment on, to facilitate the statistical analysis. In counting occurrences of the tokens of interest, imitations, repetitions, and lexicalized, formulaic expressions were discarded. In order to ensure that the children had indeed mastered the relevant constructions, we employed a conservative measure of acquisition, which was taken to be FRU (“first-of-repeated uses”), that is, first clear use, followed soon after by additional, distinct uses featuring different lexical items (see Stromswold, 1996; Snyder & Stromswold, 1997); accordingly, initial isolated uses of a construction were not taken into account. All relevant utterances were tabulated and then coded for subject type (overt, covert), order of constituents, sentence force (declarative, imperative, interrogative, or exclamative) and type of predicate involved (copula, raising, unaccusative, psych. or (di)transitive, etc.). Drawing on previous studies, we did not consider imperative sentences (with or without a subject) in our count. The context of every token containing a finite verb was analyzed...
manually, therefore allowing us to assess the felicity of each child’s production. In so doing, we examined the preceding and following lines, which also enabled us to discard direct imitations of the parents’ or caregivers’ utterances, as well as immediate repetitions of the same sentence by the child, which were not included in the count. Finally, the Mean Length of Utterance in Words (MLUw) for each file of each child was computed using the automated function “MLU” available from the CHILDES CLAN program (MacWhinney, 2000).

Regarding the analysis, the statistical method we used was the Binomial Test (Snyder, 2007), aimed at checking for concurrent, simultaneous acquisition. The Binomial Test, a non-distributional method for statistical hypothesis testing, addresses the question of whether the apparent onset interval between two given constructions is plausibly due to a lower frequency of use for the construction emerging later (i.e., consistent with chance), 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 have prior to using the particular constructions successfully. Once the pertinent figures have been collected, the desired probability can be calculated thus:

\[
(6) \quad \text{Binomial Test} \quad p = \frac{X}{(X + Y)}^Z
\]

In (6), X corresponds to the number of times construction A (e.g., null subjects) appears in the ten transcripts following the first-of-repeated uses of B (e.g., overt subjects); Y stands for the times B occurs in the transcripts (as a minimum, ten) after the first use of B; and Z corresponds to the uses of A before the first clear use of B.12

12. In interpreting the results of the Binomial Test, we adopt the .05 significance level assumed in the Social Sciences. Significant results (i.e., \( p < .05 \)) are thus taken to refute the null hypothesis that constructions A and B emerged simultaneously. If, on the contrary, the \( p \)-value is higher than the .05 significance level, that is, if the result is null or not significant, the evidence against the null hypothesis that constructions A and B appeared at the same time is weak. A note of caution is in order, though: a null result is not strong evidence for concurrent acquisition, but can still be considered consistent with chance (i.e., it may well be the case that one structure could have been used before the other just due to chance). Put differently, a null result, albeit not conclusive, is compatible with the hypothesis that both structures arose at the same time. In short, the Binomial Test allows us to say that a result is statistically significant or not significant, based on the relative frequency of A and B in the ten transcripts following the first clear use of B.
6. Results and discussion

In this section, we focus on the results of this study, intercalating some discussion as we proceed. We begin by exploring the issue of null subjects, focusing on their rate of occurrence and timing of acquisition with respect to overt subjects. We first concentrate on the monolingual child and then on the bilingual children of the study. This leads us to reconsider the existence of a ‘no overt subject’ stage in the acquisition of Spanish. In so doing, we also pay attention to early null subjects and their feliciteness, which will allow us to assess not only whether children produce them successfully from a grammatical point of view, but also pragmatically. Beyond production, looking into the contexts of null-subject use will enable us to determine whether children also comprehend subjects (null and overt) correctly in the context of a conversation with adults. We then turn to the time-course of acquisition of SV/VS in subject use in mono- and bilingual Spanish, as well as their appropriateness from the standpoint of the adult language.13

6.1 Null vs. overt subjects

Table 2 provides the onset age of null and overt subjects for each child, including the relevant MLUs in words, as well as the results of the statistical analysis used to check for the concurrent emergence of null and overt subjects, employing the Binomial Test outlined in the previous section.

<table>
<thead>
<tr>
<th>Child</th>
<th>Bilingual/monolingual</th>
<th>Variety of Spanish</th>
<th>Onset Age of Null Subjects</th>
<th>Onset Age of Overt Subjects</th>
<th>Binomial Test (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Irene</td>
<td>Monolingual</td>
<td>Iberian Spanish</td>
<td>01:07,05 (MLUw 1.72)</td>
<td>01:07,22 (MLUw 2.05)</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td>Carla</td>
<td>Bilingual</td>
<td>Iberian Spanish</td>
<td>02:00,00 (MLUw 1.71)</td>
<td>02:03,00 (MLUw 3.33)</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td>Manuela</td>
<td>Bilingual</td>
<td>Cuban Spanish</td>
<td>01:09,05 (MLUw 1.25)</td>
<td>01:11,07 (MLUw 1.34)</td>
<td>p = 0.0735</td>
</tr>
</tbody>
</table>

13. We are aware of the need to supplement longitudinal studies that utilize corpus data, drawn mainly from spontaneous speech, with additional data gathered from additional sources. In this sense, it could be the case that children may not use a specific construction—which may have mastery of—simply because the contextual requirements in a specific conversation or speech sample do not enable its use. However, this is not to say that naturalistic data are not informative at all, although it is true that the results obtained from corpus studies should be mirrored by the results of studies employing different methodologies. Converging results from different studies constitute the strongest sources of evidence.
Early null and overt subjects in the Spanish of simultaneous English-Spanish bilinguals

<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
<th>Language</th>
<th>MLUw (years)</th>
<th>MLUw (weeks)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leo</td>
<td>Bilingual</td>
<td>Iberian Spanish</td>
<td>01;10,22</td>
<td>02;04,09</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Simon</td>
<td>Bilingual</td>
<td>Iberian Spanish</td>
<td>02;00,16</td>
<td>02;01,28</td>
<td>0.120</td>
</tr>
</tbody>
</table>

All the children considered began using non-overt subjects before they started to use overt ones. First, the monolingual child’s Spanish started displaying null subjects at age 01;07,05, followed shortly after by overt subjects, at age 01;07,22, with no transcripts in between these two, since Irene was recorded fortnightly. Despite this proximity in time, the significant difference after performing the Binomial Test indicates that the two constructions emerged in Irene’s speech at separate times. Over the whole period considered, Irene produced null subjects at a rate of 71.78%, with the remaining 28.22% being cases of overt subjects, as shown in Table 3 (see also Figure 1).

Starting with the bilingual female child whose data belong to the Pérez-Bazán corpus, the gap found between Carla’s null subjects and her overt ones turns out to be statistically significant, in much the same way as in the case of Irene. Chronologically, the time span between the two types of subject is three months (equalling two transcripts), with MLUw compatible with the child being on her way out of the ‘one-word’ stage. During the entire period, Carla produced null subjects at a rate of 54.76% and overt subjects at a rate of 45.24%. Her results are, therefore, parallel to those of the monolingual Spanish-acquiring child of this study –Irene.

Regarding the bilingual girl from the Deuchar corpus, Manuela’s first-of-repeated uses of null subjects with inflected verbs in the first- and third-person in Spanish occurred in the first Spanish transcript available, at age 1;09,05 (the first transcript actually contains Manuela’s first English one-word utterances at age 1;07,12). Thus, it is likely that Manuela started producing null subjects before the age of one year and nine months. In any case, we find two uses of overt subjects in the same transcript (likely memorized chunks), although it is not until the following transcript that we find a larger number of instances of overt subjects (i.e., first-of-repeated uses), which suggests that their emergence took place slightly later (1;11,04) (again, there is only one transcript between those two, at age 1;09,09, concerned with the child’s English). The result of the Binomial Test aimed at checking for concurrent acquisition of null and overt subjects is only marginally significant, which is compatible with the hypothesis of simultaneous acquisition of null and overt subjects in the speech of this bilingual child. Note, however, that since the result is not significant, it cannot be concluded either that the two constructions were acquired at the same time; failure to find a significant difference does not automatically lead to the conclusion that the two constructions were acquired concurrently, as shown in the previous section. Nonetheless, given the few

2nd proofs
transcripts available and the lack of enough null subjects before the occurrence of overt subjects, we cannot reach any firm conclusions about the time-course of null and overt subjects in this child’s linguistic development. Recall also from Paradis and Navarro (2003) and Silva-Corvalán (2014) that this child was mainly exposed to Cuban Spanish, which is a variety where the rate of overt subjects is higher than in other varieties. From the beginning of the transcripts until age 2;05.05, we find that the proportion of null vs. overt subjects was 55.77% (null) vs. 44.23% (overt). Despite the fact that this child was exposed to both English and Cuban Spanish, the rate of null vs. overt subjects in the child’s early speech is wholly consistent with the results found for monolingual children acquiring Iberian Spanish. Even though Manuela acquired Cuban Spanish, a variety reported to manifest a larger proportion of overtly expressed subjects, her overall percentage of null subjects is higher than that of Carla, who acquired Iberian Spanish, although still lower than the rest of the Iberian-Spanish acquiring children (see Table 3/Figure 1).

As for Leo, from the FerFuLice corpus, his first instances of non-overt subjects started at age 1;10.22, which is actually in the first transcript available (suggesting that covert subjects may have appeared in the child’s spontaneous speech even earlier) and the first productive instance of an overt subject occurred at age 2;04.09. There is a total of eight transcripts between the emergence of null subjects and overt subjects.14 The difference between the two constructions is significant by Binomial Test, which points out that the prerequisites for null subjects that the child needed to acquire are different from those of overt subjects. It is of note that the MLU_w at the moment the very first non-overt subject occurs is 1.39, which indicates that at this point the child already has the ability to produce more than one word in his speech (i.e., it would not be impossible for an overt subject to co-occur with the verb in the same sentence); as indicated by fn. 14, Leo’s MLU_w during this period seem to allow for more than one ‘slot,’ which amounts to saying that the occurrence of a SV or VS combination would in principle be possible in terms of MLU (although still low to consider that the child is already in the two-word stage:

14. For Leo, the relevant transcripts containing only null-subject sentences until the emergence of overt subjects concern the following ages and MLU_w:

<table>
<thead>
<tr>
<th>Age</th>
<th>MLU_w</th>
</tr>
</thead>
<tbody>
<tr>
<td>01:11:20</td>
<td>1.422</td>
</tr>
<tr>
<td>02:00:16</td>
<td>1.609</td>
</tr>
<tr>
<td>02:01:01</td>
<td>1.465</td>
</tr>
<tr>
<td>02:01:28</td>
<td>1.358</td>
</tr>
<tr>
<td>02:01:29</td>
<td>1.333</td>
</tr>
<tr>
<td>02:02:07</td>
<td>N/A</td>
</tr>
<tr>
<td>02:02:21</td>
<td>N/A</td>
</tr>
<tr>
<td>02:03:25</td>
<td>1.712</td>
</tr>
</tbody>
</table>
see below). Finally, the proportion of null vs. overt subjects in Leo’s early speech is 75.51% (null) and 24.49% (overt).

As far as Leo’s twin brother, Simon, is concerned, we find his first uses of null subjects at age 02;00,16, and the first instances of overt subjects at age 2;01,28, with only one transcript, at age 02;01,01 (MLUw; 1.07), in between. Unlike in the case of Leo, we did not find a significant difference between the emergence of null and overt subjects. However, the percentage of null subjects (63.16%) for this child is still much higher than that of overt subjects (36.84%) during the entire period considered, fully consistent with the pattern found for monolingual Spanish.

Table 3. Percentage of null and overt subjects in early child Spanish\textsuperscript{15}

<table>
<thead>
<tr>
<th>Child</th>
<th>Variety of Spanish</th>
<th>% of null subjects</th>
<th>% of overt subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Irene (M)</td>
<td>Iberian Spanish</td>
<td>71.78 (N = 295)</td>
<td>28.22 (N = 116)</td>
</tr>
<tr>
<td>Carla (B)</td>
<td>Iberian Spanish</td>
<td>54.76 (N = 92)</td>
<td>45.24 (N = 76)</td>
</tr>
<tr>
<td>Manuela (B)</td>
<td>Cuban Spanish</td>
<td>55.77 (N = 29)</td>
<td>44.23 (N = 23)</td>
</tr>
<tr>
<td>Leo (B)</td>
<td>Iberian Spanish</td>
<td>75.51 (N = 37)</td>
<td>24.49 (N = 12)</td>
</tr>
<tr>
<td>Simon (B)</td>
<td>Iberian Spanish</td>
<td>63.16 (N = 24)</td>
<td>36.84 (N = 14)</td>
</tr>
</tbody>
</table>

\textsuperscript{15} It should be noted that these percentages reflect the total number of occurrences found in the transcripts studied (from the beginning of the corpus until the tenth transcript following the first-of-repeated uses of the last construction emerging; in the case of Irene, 18 transcripts following the first overt subjects were considered, since the relevant data had already been coded for a previous study), without considering potential intermediate stages and thus perhaps masking developmental stages. The goal is just to note the early proportions of null and overt subjects in the children studied. Furthermore, as the careful reader will note, the raw numbers furnished in brackets suggest that the overall numbers are not very strong for some of the children analyzed, hence our caution when interpreting the findings.
The sentences in (7) are examples of the early null subjects produced by the children of this study. (7a) constitutes an example from Irene, the monolingual Spanish-acquiring girl; sentences (7b–e), for their part, correspond to the Spanish of the four English-Spanish bilingual children of this study.

(7) a. ya ababó  
   'It’s already over.'
   [Irene, 01/07/05]

b. cantan
   'They are singing.'
   [Carla, 02/00/00]

c. abrió
   'S/he opened.'
   [Manuela, 01/09/05]

d. cae
   'It fell.'
   [Leo, 01/10/22]

e. ahí está
   'Here it is.'
   [Simon, 02/00/16]

The evidence available so far points to the conclusion that with respect to the timeline of acquisition of null and overt subjects in bilingual English-Spanish children, no CLI effects from English are observed at the initial stages when subjects emerge in the naturalistic speech of children. Of course this does not fully answer the
question of whether there is CLI in the acquisition of subjects, a task that the rest
of the paper intends to undertake in more detail. The next section reviews the
evidence adduced up to this point in light of the much debated ‘null subject’ stage
in the acquisition of pro-drop languages like Spanish, an issue which has not been
investigated in BFLA to date.

6.1.1 A ‘no overt subject’ stage? Statistics, chronology, pragmatic
appropriateness of child null subjects, and CLI

As noted, Grinstead (1998, et seq.) has claimed that monolingual children acquiring
paradigmatic null-subject languages like Catalan and Spanish pass through a ‘null subject’ stage, which lasts approximately until age 2. Given that a theory
that provides an explanation for early stages must also include an explanation for
changes in development, the question arises as to what can be relevant for children
in order to change their grammar from one that allows only null subjects. Another
question is how the prospect of a ‘no overt subject’ stage can be reconciled with
the fact that children seem to attain an adult-like distribution of null and overt
subjects immediately after the completion of the ‘no overt subject’ stage. If this is
the case, then there would be a sudden, almost instantaneous change, and not a
gradual and protracted process (see Villa-García, 2013, for the claim that children
are rather conservative in that they do not start using overt subjects until they have
identified the prerequisites needed to produce them correctly, in line with Snyder’s
(2007) hypothesis of Grammatical Conservatism). The goal of this section is not to
provide a definite answer to this debate, since this requires additional developmental research examining the trajectory of subject acquisition more closely. Instead,
we intend to evaluate the available bilingual and monolingual data to determine
if bilinguals and monolinguals are comparable regarding early subjects, and to see
if this can shed new light on the contentious ‘null subject’ stage. In this regard,
it is important to note that the ‘no overt subject’ stage has not been examined in
BFLA, the existing studies have not checked for statistical differences between null
and overt subjects (with the exception of Villa-García, 2013), and, crucially, prior
research has not explored the question of whether or not null subjects are pragmatically appropriate both before overt subjects emerge and once the two types of
subjects are part of the monolingual and bilingual children’s speech (recall that the
‘no overt subject’ phase predicts pragmatic errors in the use of null subjects, that is,
we should in principle find covert subjects in contexts that would require an overt
subject from the point of view of an adult speaker of Spanish, since all subjects are
expected to be null during this early period).

Starting with what the statistical analysis reveals, for three out of five children,
our monolingual child and two of the bilingual children, we find a statistically
significantly result by Binomial Test (see Table 2 above). These findings indicate
that for these children, the fact that null subjects occurred before overt ones is not due to chance, but to different prerequisites (be they grammatical and/or semantic-pragmatic) that the child needed to attain before making use of the second construction (i.e., overt subjects) in a successful manner. This result is wholly consistent with Grinstead’s hypothesis: if children start using only null subjects, we expect there to be a significant difference between null and overt subjects. As noted by two anonymous reviewers, however, simply because two bilingual children and one monolingual child produce null subjects significantly before overt subjects does not automatically lead to the conclusion that Spanish-speaking children go through a ‘null subject’ stage. This difference could be due to other factors. For instance, those children may not have had the opportunity to produce overt subject sentences in some transcriptions because the discourse pragmatics of the context never required a lexically realized subject. It is conceivable that in some cases, the parents initiated most of the conversations (by means, for instance, of an overt subject topic), and their children followed up with a null subject (whose referent was the topic introduced by the parents).

For the other two children, however, the results did not reach significance (only marginally for Manuela; note that for Simon, the statistical difference was not significant, unlike in the case of his twin brother Leo). Non-significant results cannot be taken to imply that the two constructions’ difference is necessarily due to chance, but they are compatible with such an interpretation.

Although the results of the statistical analysis are coherent with the prospect of a ‘no overt subject’ stage, we believe that further criteria should be taken into account, such as the duration of such a period, whether the MLU_w’s allowed for an overt subject to occur with a verb, as well as whether we find errors in the use of early null subjects from a pragmatic point of view, as predicted by the ‘no overt subject’ stage hypothesis.

As far as Irene’s putative ‘null subject’ stage is concerned, it is important to note that we considered her transcripts from the beginning of the corpus at age 1;05,27. The first-of-repeated uses of null subjects with fully inflected verbs occurs at 1;07,05 (MLU_w: 1.72), with one transcript between this one and the one reporting Irene’s naturalistic speech at age 1;07,22 (MLU_w: 2.05), the time when we find her first-of-repeated uses of overt subjects. This means that at least in the case of Irene, the ‘no overt subject’ stage—if real—is rather short-lived, spanning over two weeks. Note that Irene’s MLU’s suggest that she was on the way out from the one-word stage, which is compatible with two-word utterances.

Regarding the appropriateness of Irene’s early null-subject utterances, it is of note that the numerous imperative sentences uttered by Irene, albeit not counted for purposes of the analyses run above, show 100% of accuracy, that is, they are pragmatically appropriate commands directed to the addressee. We first look at

2nd proofs
the adequacy of Irene’s null-subject sentences before the occurrence of overt subjects (i.e., during the presumed ‘no overt subject’ stage) and then once both types of subject are in place. During the period lacking overt subjects, Irene produced five examples of first-person (speaker) and second-person (hearer) null subjects, which are anchored to the particular communicative situation and thus felicitous. Moreover, she uttered a sentence featuring an atmospheric predicate, hence containing an obligatorily covert subject. Similarly, she uttered three third-person null-subject sentences where the subject was salient in the discourse, and thus fully appropriate. She also produced a sentence where the null subject referred back to the object of the preceding utterance, thus realizing the [–topic shift] feature associated with null subjects. At this early stage, only two cases can potentially be considered errors. At age 1;07,05, for instance, Irene uttered the example within the dialog in (8). In this particular example, the mother and the child are talking about a piece of jewelry, and suddenly Irene utters the null-subject sentence no, no oye (lit. ‘no, not hears’). The mother is confused as a result, since the reference is not clear (the other example illustrates exactly the same situation, but featuring the verb caer ‘to fall’ in the preterit instead).

(8) Context: the child and the mother are talking about a bracelet

CHI (Irene): no, no oye
   no, not hears
   ‘S/he doesn’t hear’

MOT: ¿Quién?
   who
   ‘Who (doesn’t hear)?’

Although this sentence can be considered pragmatically inadequate due to the fact that the referent of the covert subject is not salient in the context (as witnessed by the mother’s confusion), it may well be the case that it is in fact salient, but the mother was, for instance, not paying attention. In any case, what the data reviewed suggest is that Irene actually made very few pragmatic errors — if any — given the potential pitfalls she had. This result is not predicted by the ‘no overt subject’ stage hypothesis, although the findings should be interpreted cautiously, since none of Irene’s early utterances actually required an overt subject, except, perhaps, the two sentences alluded to above and exemplified by the utterance in (8).

Once Irene’s overt subjects emerged in her spontaneous speech (and until the end of the transcripts considered), her null-subject production can be deemed to be virtually errorless, with numerous instances of no overt subject sentences in which the referent is salient in the discourse. The only potential error (cayó ‘it fell’), which does not have a distinct referent in the discourse at the beginning, actually turns out to be a case in which the child is referring to herself in the third
person. Thus, this error is actually not a pragmatic error per se. Overall, before the appearance of overt subjects, Irene used null subjects in a pragmatically appropriate way (the examples available did not seem to require an overt subject), and after the emergence of overt subjects, Irene continued to use null subjects successfully, with all instances being [–topic shift] occurrences with a salient referent in the communicative discourse.

In the case of Carla, the bilingual child from Michigan, MLUw's are also indicative of the availability of two slots (null subjects at age 2:00:00 — MLUw: 1.714; overt subjects at age 2:03:00 — MLUw: 3.333). Note that for Carla, the first transcript in the corpus already contains 14 instances of null subjects, which suggests that the child may have started using such constructions earlier. The first-of-repeated uses of overt subjects were identified three months later, which is consistent with Grinstead's proposal. Recall that Carla is one of the two bilingual children for whom a statistically significant result between null and overt subjects was found.

Regarding the pragmatic adequacy of covert subjects in the spontaneous speech of this child, before the rise of overt subjects in her speech, Carla produced 19 null-subject sentences, all of which are pragmatically appropriate (1st- and 2nd-person singular forms, instances of non-overt subject sentences where the referent is salient in the preceding discourse, and cases where the null subject is a topic continuation of an immediately preceding nominal). Again, it is possible that none of Carla’s early sentences required an overt subject, since they were context-bound and therefore the referent was salient in the discourse at all times. Once overt subjects are in place, and until the age of three years and two months, the sample contains 56 null-subject utterances. Much like the null subjects uttered before the appearance of overt subjects, most null subjects were cases where the referent was salient in the discourse, or cases of first-person singular/plural and second-person singular sentences. Carla's performance is also error-free from the stance of pragmatics during this later period, much like the monolingual Spanish-acquiring child —Irene.

Concerning the early null subjects uttered by Manuela, we find extremely few instances in the sample available. More specifically, we only find four null-subject sentences before the first-of-repeated uses of overt subjects (not surprisingly, the statistical difference for this child is not significant). However, it is of note that Manuela’s first transcript at age 1;09:05 already contains examples of null-subject utterances, which suggests that perhaps she started using those before that moment. Similarly, Manuela’s MLUw was 1.250 at age 1;09:05, which suggests that she was not completely past the holographic stage. Thus, no conclusion can be drawn regarding Manuela’s putative ‘no overt subject’ stage. Her early null subjects are pragmatically appropriate, however. The four instances preceding overt subjects are all cases whose referent is salient in the preceding discourse. From age 1;11.07
Early null and overt subjects in the Spanish of simultaneous English-Spanish bilinguals

until 2;05.05, all of Manuela’s null subjects, which amount to 39 occurrences, are adult-like, including instances where the null subject is first- and second-person singular or its referent is salient in the specific communicative situation.

With respect to the twins, Leo produced 14 null-subject sentences before he began to use overt subjects productively, according to the files available. As in the case of Carla, the result of the Binomial Test checking for concurrent acquisition of null and overt subjects is significant for this child. All of his early null subjects concern either salient referents in the discourse or first-person singular cases, none of which can be deemed infelicitous, despite his MLU’s during this period, which suggest that he is not totally out of the one-word stage (see fn. 14). After the emergence of overt subjects in his speech, we find 22 covert subject sentences in the transcripts analyzed until age 2;10.21. Once again, these sentences contain null subjects in the first- and second-person singular and first-person plural (and two examples with third-person morphology in which Leo is actually referring to himself), together with covert subjects with a clearly identifiable referent in the preceding discourse.

As for Simon, we only find 4 null-subject sentences before the first-of-repeated uses of overt subjects, with MLU’s in the first four transcripts between ages 1;10.22 and 2;01.28 which indicate that Simon had not yet left the one-word stage (1.634, 1.208, 1.296, and 1.070). Nevertheless, once again all of his early null subjects are pragmatically appropriate (salient referent), and the same applies to the null subjects he produced after overt subjects started. Although just 14 in total, Simon’s null-subject sentences until the age of 2 years and 10 months point out that his use of null subjects is errorless, with instances involving salient references along with first-person singular forms.

Overall, we find that both the monolingual child and bilingual children of the present study use null subjects in a target-like fashion from the moment these constructions emerge in their naturalistic speech. With respect to the ‘null subject’ stage, its existence in the children of this study is not apparent. The statistically significant results for two bilinguals and one monolingual indicate that the difference between null and overt subjects for such children was not due to chance (rather the results point to deeper, grammatical and/or pragmatic reasons for this difference). However, the short-lived nature of this putative phase for some of the children (notably, Irene and Simon) casts doubt on the existence of a firm, clearly delimited null-subject stage. If this were the case, we would need to account for the abrupt, almost immediate appearance of overt subjects in place. Furthermore, the children’s virtually errorless production in terms of null subjects from a pragmatic point of view is also not predicted by the ‘no overt subject’ stage hypothesis, in spite of the fact that the covert-subject sentences spontaneously uttered by the children of this study seemed to allow for a null subject from the stance of an adult
native speaker of Spanish. Future research investigating much larger sample sizes will delve into this issue further.

As far as the issue of CLI is concerned, recall that if the CLI hypothesis is to be taken seriously, attention should also be devoted to the pragmatic felicitousness of children’s early subjects (Paradis & Navarro, 2003; Serratrice & Hervé, 2015). The early performance of the Spanish-acquiring children in bilingual contexts regarding null subjects shows no apparent CLI effects. In principle, CLI would predict an effect in terms of the use of null subjects in bilinguals for whom one of the two languages (i.e., English) dictates that null subjects are extremely restricted (i.e., diary-drop) and overt preverbal subjects are mandatory, in compliance with the well-known requirement that sentences have phonetically realized (preverbal) subjects in English. However, the children of this study seem to abide by the pragmatic constraints regulating the occurrence of null subjects in Spanish from the first utterances containing subjects, a result also found for the monolingual child.

A caveat is in order, however. Based on Silva-Corvalán’s (2014) findings that CLI becomes more apparent after age 4 in the naturalistic speech of two children acquiring English and Spanish simultaneously, it may well be the case that the children of this study did manifest CLI at later stages of development, particularly if at some point once of their two languages becomes stronger. Future research considering longer periods of development and larger data will be able to tackle this issue, but the fact remains that when the provision of subjects begins in the speech of bilingual children, the hypothesis of CLI is not supported by the spontaneous data in this study, which further substantiates Liceras et al.’s (2008) and Silva-Corvalán’s (2014) conclusions.

6.2 Preverbal and postverbal subjects

We have noted that most existing studies have concentrated on overt subjects without distinguishing between preverbal and postverbal subjects (apart from Silva-Corvalán, 2014). It is our belief that the time-course of acquisition of preverbal and postverbal subjects is an important aspect of children’s linguistic development in pro-drop languages like Spanish, in particular when evaluating the hypothesis of CLI: if English were determining the word order in the Spanish of English-Spanish bilinguals, then SVO would be the preferred and prominent word order, and the different word order possibilities (e.g., SV and VS) would not be determined by factors that go beyond purely grammatical considerations (e.g., information structure and lexico-semantics). Table 4 provides information related to the onset age of preverbal and postverbal subjects (with first-of-repeated uses being once more the criterion for determining the age of acquisition) as well as the pertinent MLUs. The results of the Binomial Test are also furnished, with the aim
of checking for simultaneous acquisition of the two relevant types of subject when the first instances of these constructions did not occur in the same file. The reader should note that the frequencies obtained in this regard, in particular for three of the bilingual children, are not very high, and although the results are consistent across children, as we will see, caution should be taken when drawing conclusions based on limited amounts of data.

Table 4. Onset age of preverbal and postverbal subjects in Spanish

<table>
<thead>
<tr>
<th>Child</th>
<th>Variety of Spanish</th>
<th>Onset Age of Preverbal Subjects (SV, SVO)</th>
<th>Onset Age of Postverbal Subjects (VS, VOS)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Irene (M)</td>
<td>Iberian Spanish</td>
<td>1:07,22 (MLUw 2.05)</td>
<td>1:07,22 (MLUw 2.05)</td>
<td>(same transcript)</td>
</tr>
<tr>
<td>Carla (B)</td>
<td>Iberian Spanish</td>
<td>02:03,00 (MLUw 3.333)</td>
<td>02:03,00 (MLUw 3.333)</td>
<td>(same transcript)</td>
</tr>
<tr>
<td>Manuela (B)</td>
<td>Cuban Spanish</td>
<td>01:11,07 (MLUw 1.335)</td>
<td>01:11,07 (MLUw 1.335)</td>
<td>(same transcript)</td>
</tr>
<tr>
<td>Leo (B)</td>
<td>Iberian Spanish</td>
<td>02:05,00 (MLUw 1.53)</td>
<td>02:04,09 (MLUw 1.55)</td>
<td>p = 0.127</td>
</tr>
<tr>
<td>Simon (B)</td>
<td>Iberian Spanish</td>
<td>02:05,00 (MLUw 1.34)</td>
<td>02:01,28 (MLUw 1.27)</td>
<td>p = 0.250</td>
</tr>
</tbody>
</table>

Starting with Irene, the monolingual child of this study, the first-of-repeated uses of overt subjects occurred at age 1:07,22 (MLUw 2.05). Importantly, the same transcript contains instances of both preverbal and postverbal subjects, which indicates that the child began using both constructions concurrently, a result that corroborates Grinstead’s (1998) results for child Catalan and Spanish (see also Villa-García, 2011, for similar findings with child Iberian and Puerto Rican Spanish).

Turning now to the bilingual children of this study, the picture that emerges is almost identical to that of the monolingual child. In no case did we find a statistically significant difference between the appearance of preverbal subjects and that of postverbal subjects, which is consistent with the hypothesis of concurrent acquisition of the two kinds of subject. Actually, much like in the case of Irene, two bilingual children’s first clear uses (i.e., first-of-repeated uses) of preposed and postposed subjects were found in the same transcript. Furthermore, two of the children, namely the twin brothers, started using postverbal subjects (slightly) earlier than preverbal ones, an unexpected result if the children’s English were regulating the word order of their Spanish.
Table 5 (see also Figure 2) shows the percentages of preverbal and postverbal subjects during the entire period for each child. Even though for some children the frequencies are very low, it seems safe to conclude that all the children of this study employ preverbal and postverbal subjects in their spontaneous speech.

Table 5. Percentage of preverbal and postverbal subjects in early child Spanish

<table>
<thead>
<tr>
<th>Child</th>
<th>Variety</th>
<th>% of Preverbal Subjects (SV, SVO)</th>
<th>% of Postverbal Subjects (VS, VOS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Irene (M)</td>
<td>Iberian Spanish</td>
<td>46.6 (N = 54)</td>
<td>53.4 (N = 62)</td>
</tr>
<tr>
<td>Carla (B)</td>
<td>Iberian Spanish</td>
<td>69.7 (N = 53)</td>
<td>30.3 (N = 23)</td>
</tr>
<tr>
<td>Manuela (B)</td>
<td>Cuban Spanish</td>
<td>69.6 (N = 16)</td>
<td>30.4 (N = 7)</td>
</tr>
<tr>
<td>Leo (B)</td>
<td>Iberian Spanish</td>
<td>50 (N = 6)</td>
<td>50 (N = 6)</td>
</tr>
<tr>
<td>Simon (B)</td>
<td>Iberian Spanish</td>
<td>57.1 (N = 8)</td>
<td>42.9 (N = 6)</td>
</tr>
</tbody>
</table>

Figure 2. Percentage of preverbal and postverbal subjects in early child Spanish

The percentages reveal an apparent asymmetry between the monolingual child and her bilingual peers: Irene seems to be the only child for whom postverbal subjects are more numerous, a result that concurs with what Bates (1976) found for child Italian and Gavarró and Cabré-Sans (2009) for early Catalan. For the bilingual children, Leo appears to use preverbal subjects half of the time, and the rest of the children use preverbal subjects more often. If taken at face value, this finding could be interpreted as suggesting that there is a CIJ effect from English: bilinguals are using preverbal subjects more than postverbal ones, the reason being that the canonical word order enforced in English is determining their Spanish word order. However, the question arises as to whether this conclusion can be
extended to the population at large. Data from previous studies suggests that this cannot be the case. First, the results reported by Silva-Corvalán (2014) reviewed in Section 4 suggest that one of the bilingual children in her study patterns with our monolingual child in showing a higher percentage of postverbal subjects while the other child patterns with three of our bilingual children in showing a higher percentage of preverbal subjects during similar periods. Bel’s (2003, p. 17) study of Catalan-Spanish bilingual children revealed 67.2% of preverbal subjects in their Spanish and 61.3% in their Catalan. These percentages are comparable to the ones found for two of the English-Spanish bilingual children of the present study, the crucial difference being that Bel’s study was concerned with two null-subject languages—Catalan and Spanish. The data from Villa-Garcia (2009) for three monolingual children acquiring Spanish in Spain actually show patterns akin to those found for the bilingual children of this study (i.e., a higher percentage of preverbal subjects from the outset), despite the fact that the three monolingual children of that study were not in contact with a non-pro-drop language.16 Paradis and Navarro (2003) have posed the question of whether such proportions (i.e., an apparently higher rate of preverbal subjects for bilinguals) are due to CLI or to the frequencies in the input the children are exposed to. The data reviewed so far seem to suggest that the second hypothesis is better motivated, since the explanation for such high overall percentages of preverbal subjects for children not exposed to a non-pro-drop language cannot be attributed to the influence of the non-pro-drop language in the pair. Silva-Corvalán (2012, p. 797) has stressed the importance of determining whether it is the non-pro-drop language that motivates the high

<table>
<thead>
<tr>
<th>Child</th>
<th>Corpus</th>
<th>Age span considered</th>
<th>% of Preverbal Subjects</th>
<th>% of Postverbal Subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emilio (cf)</td>
<td>Vila (CHILDES) (Vila, 1990)</td>
<td>0:11.09–02:11.24</td>
<td>53.42 (N = 86)</td>
<td>46.58 (N = 75)</td>
</tr>
<tr>
<td>Inés (cp)</td>
<td>Inés (UC Conn-CLESS)</td>
<td>01:02.00–02:02.11</td>
<td>70.88 (N = 185)</td>
<td>29.12 (N = 76)</td>
</tr>
<tr>
<td>Juan (cf)</td>
<td>Linaza (CHILDES) (Linaza, Sebastián, &amp; Barrio, 1981)</td>
<td>01:07.02–03:09</td>
<td>52.32 (N = 79)</td>
<td>47.68 (N = 72)</td>
</tr>
</tbody>
</table>

Thus, the data show that even in the case of children not exposed to a non-null-subject language like English, the percentage of preverbal subjects may still be higher than that of postverbal ones, hence refuting the idea that a larger proportion of preverbal subjects is (necessarily) owing to the influence of English. The case of Emilio from the Vila corpus is particularly interesting, since even though he is listed as a monolingual Spanish speaker in CHILDES, he was exposed to Catalan from birth and in the files analyzed the adults occasionally use Catalan with him. Therefore, this is one of those cases of a bilingual child acquiring two pro-drop languages (see fn. 1). Note in passing that Villa-Garcia (2013) found a statistically significant gap between the onset of null subjects and the onset of overt subjects for Emilio.

16 The relevant data from Villa-Garcia (2009) are as follows:
number of preverbal subjects in child BFLA, or bilingualism *per se*, that is, “the demands posed by having to learn and use to languages.” The evidence gathered in this paper suggests that the higher rates of preverbal subjects found in some of the children studied in this and previous studies may be due to reasons other than the presence of a non-pro-drop language or the demands of having to cope with two linguistic systems. It may actually be the case that the frequencies in the input or, as we will see in the following subsection, the type of subject (SV/VS) required in each particular context, may play a role in the frequencies for each type of subject. Further research taking into account a larger sample of children will determine whether this conclusion is corroborated.

The sentences in (9) constitute examples of constructions containing preverbal and postverbal subjects produced by the children in this study:

(9) a. La India machó
the Indian left
“The Indian girl left.”

b. Me lo prestó Jessica
cl. cl. lent Jessica
“Jessica lent it to me.”

c. Este es de mamá
this is of mom
“This one belongs to mom.”

d. Sí cabe nejo (= conejo)
yes fits rabbit
“The rabbit does fit.”

e. A@l yo miro
there I look
‘I look there.’

In the next section we deal with the appropriateness of early preverbal and postverbal subjects in the naturalistic speech of the children of this study.

6.2.1 On the appropriateness of preverbal and postverbal subjects
In order to determine if overt subjects—namely preverbal and postverbal subjects—are also adequate from the first-of-repeated occurrences, we manually analyzed each child utterance containing overt subjects, which were then classified into one of the categories reviewed in Section 3. Note that the criteria employed are different from those of Paradis and Navarro (2003), since the authors were looking at the pragmatic appropriateness of overt subjects as opposed to null subjects (that is, whether a lexically realized subject was informative and required in the sentence). Our objective here, however, is to determine whether preverbal and postverbal subjects are
used in a target-like fashion, which will in turn allow us to evaluate their pragmatic appropriateness, since the distribution of preverbal and postverbal subjects largely depends on pragmatic (information structural) as well as lexico-semantic factors.

Irene, our monolingual control, used preverbal and postverbal subjects in a variety of contexts. We identified the following categories, with raw numbers in brackets:

- **Preverbal subjects:**
  - Topic shift and topic-continuity contexts (see (3d) and (3b,c)) (N = 33)
  - Topic subject + wh-item (see (3e)) (N = 10)
  - Emphatic or contrastive subject with the pronoun yo (‘I’) or tú (‘you’) (N = 9)\(^\text{17}\)
  - Focused subject (focal or a subject that is itself a wh-item) (N = 2)

- **Postverbal subjects:**
  - Contrastive/new information focus (see (4a)) (N = 17)
  - Unaccusatives (see (4b)) (N = 20)
  - Locative inversion (see (4d)) (N = 8)
  - Obligatory postverbal subjects in wh-questions (see (4g)) (N = 16)
  - Stylistic inversion (reports of direct speech) (see (4h)) (N = 1)

From the above, coupled with the fact that we did not find any “odd sounding” sentences in terms of word order in the speech of Irene between the ages of 1;07,22 and 2;04,13, we can conclude that Irene’s spontaneous production of overt subjects from the outset was, much like that of null subjects, error-free. Although we have only analyzed data from one monolingual child, these results are compatible with the conclusion that children acquiring Spanish have mastery of the factors governing the distribution of subjects from the moment they begin to use them productively in their spontaneous speech. Needless to say, further research analyzing the appropriate use of preverbal and postverbal subjects of other monolingual children is required to determine if this conclusion holds.

As for the bilingual children of this study, the reader should keep in mind that the data for the four bilinguals are not as abundant as those of Irene, both in terms of the utterances provided by the child and the availability of files in their corpora, which may explain why we failed to find examples in all the categories identified for Irene’s overt subjects above. Starting with Carla, the distribution of her early preverbal and postverbal subjects in Spanish from 2;03,00 until 3;03,00 was as follows:

17. Arguably, this category could be part of the topic shift one as well, since even though pronominal subjects are emphatic/contrastive, they can likewise serve to mark topic-shift at the same time.
- Preverbal subjects:
  - Topic-shift and topic-continuity contexts (see (3d) and (3b,c)) (N = 30)
  - Emphatic or contrastive subject with the pronoun yo ('I') or tú ('you') (N = 22)
  - Focused subject (focal or a subject that is itself a wh-item) (N = 1)
- Postverbal subjects:
  - Contrastive/new information focus (see (4a)) (N = 5)
  - Unaccusatives (see (4b)) (N = 7)
  - Psychological predicates (see (4c)) (N = 1)
  - Locative inversion (see (4d)) (N = 7)
  - Obligatory postverbal subjects in wh-questions (see (4g)) (N = 3)

As far as Manuela is concerned, the following represents the distribution of her early overt subjects (from 01;11,07 to 2;05,05): 18

- Preverbal subjects:
  - Topic-shift and topic-continuity contexts (see (3d) and (3b,c)) (N = 12)
  - Emphatic or contrastive subject with the pronoun yo ('I') or tú ('you') (N = 4)
- Postverbal subjects:
  - Unaccusatives (see (4b)) (N = 3)
  - Locative inversion (see (4d)) (N = 2)
  - Obligatory postverbal subjects in wh-questions (see (4g)) (N = 2)

For Leo, the following overt subject distribution was found between the ages of 2:04,09 and 2:10,21:

- Preverbal subjects:
  - Topic-shift and topic-continuity contexts (see (3d) and (3b,c)) (N = 3)
  - Emphatic or contrastive subject with the pronoun yo ('I') or tú ('you') (N = 2)
  - Focused subject (focal or a subject that is itself a wh-item) (N = 1)
- Postverbal subjects:
  - Contrastive/new information focus (see (4a)) (N = 3)
  - Unaccusatives (see (4b)) (N = 3)

18. For Manuela, a relatively high percentage of pronominal preverbal subjects was found, which could be due to her using them more often than in other varieties. Still, the cases found appear to imply contrast with the hearer (yo vs. tú), which makes it difficult to conclude that they are instances of what in other varieties of Spanish would be redundant overtly realized pronouns.
As for Simon, between the age of 2;01.28 and 2;10.21, we found the following types of subjects:

- **Preverbal subjects:**
  - Topic-shift and topic-continuity contexts (see (3d) and (3b,c)) (N = 3)
  - Emphatic or contrastive subject with the pronoun yo (T) or tú (‘you’) (N = 4)
  - Focused subject (focal or a subject that is itself a wh-item) (N = 1)

- **Postverbal subjects:**
  - Unaccusatives (see (4b)) (N = 4)
  - Locative inversion (see (4d)) (N = 1)
  - Obligatory postverbal subjects in wh-questions (see (4g)) (N = 1)

Although the data for the twins in this regard are rather sparse, it seems that their results are comparable to those of the other children considered here in that they all employ word order in their Spanish in an adult-like, target-like manner from the outset. These findings are thus consonant with the view that there is no CLI with respect to overt subjects in the BFLA of Spanish coupled with English; even though their productions are not as abundant as that of the monolingual child, the English-Spanish bilingual children of this study pattern with the monolingual in terms of early overt subject use.

7. **Further discussion and conclusions**

This paper took as its point of departure the hypothesis of Crosslinguistic Influence (Hulk & Müller, 2000) in the domain of the acquisition of subjects in pro-drop languages paired with non-pro-drop languages in BFLA, since subjects in null-subject linguistic varieties constitute phenomena whose behavior and distribution is conditional upon syntactic, lexico-semantic, and pragmatic considerations. After reviewing previous studies on the acquisition of subjects in Spanish-style languages, we formulated a number of questions to further test CLI in light of data from four English-Spanish bilinguals, who were compared with a monolingual.

The first question (see (i) in Section 5) was whether the bilingual acquisition of a non-pro-drop language (e.g., English) influences the acquisition of a pro-drop language (e.g., Spanish) with respect to the development of null and overt subjects. The child data indicate that the bilingual children in this study behave like their monolingual peers in terms of the timeline of the acquisition of null vs. overt subjects, which points to the absence of CLI effects from English in this regard. The second question (see (ii) in Section 5) raised was whether the pro-drop language of the bilingual child passes through a stage in which only non-overt subjects are
attested, as has been (controversially) claimed for monolinguals. To the best of our knowledge, this hypothesis had not previously been tested for Spanish in the context of BFLA. We initially identified first-of-repeated uses of null and overt subjects and then ran the statistical analysis. Our assumption for monolinguals, following Grinstead’s work, was that initially, children acquiring pro-drop languages like Spanish pass through a phase during which only null subjects are attested. We showed that the evidence gathered from our monolingual child and the bilingual children of this study is not conclusive. First, the percentage of null subjects is rather high for all children (much higher than the rate of null subjects reported for child English in the literature). Second, we found a significant difference between null and overt subjects for Irene, a monolingual child, but also for Carla and Leo, two bilingual children. This result indicates that null and overt subjects are acquired at separate times, and thus is consistent with the ‘null subject’ stage proposed for monolingual pro-drop as well as with the claim that the two constructions demand different prerequisites that the child needs to attain before using the constructions successfully.

Nevertheless, this conclusion is not irrefutable, since it may be the case that the interactions in which the young children engaged did not require overt subjects. Similarly, we observed that children’s chronological gap between null and overt subjects seems rather short in some cases (especially in the case of Irene, the monolingual child). As has been shown, children use null and overt subjects in an adult-like manner from the beginning, which raises the non-trivial question of whether the transition from the putative ‘null subject’ stage to the next stage in acquisition is an abrupt, almost immediate process. Furthermore, since the ‘no overt subject’ stage hypothesis predicts numerous pragmatic errors in children’s early null-subject sentences and CLI is intimately associated with the syntax-pragmatics interface, we also endeavored to address the question of whether early null subjects are pragmatically felicitous in the developing pro-drop language of bilingual children. For all the children of this study, we found that their performance was target-like both before the emergence of overt subjects and afterwards, with most of their null subjects adhering to the pragmatic rules governing them (suggesting as well that children’s comprehension of subjects is in place from early on). However, little can be concluded in terms of the ‘no overt subject’ phase proposal based on the felicity of the no overt subject sentences analyzed, since the contexts in which the null subjects occurred in the transcripts analyzed did not seem to require an overt subject in the place of a covert one. With respect to CLI, the results reported here are suggestive that no CLI effects are apparent in the early null-subject sentences analyzed.

Overall, with respect to the emergence of subjects, the conclusion drawn here is that the grammatical development of null and overt subjects in the bilingual
children's null-subject language mirrors that of their monolingual peers, in line with previous studies (e.g., Liceras et al., 2008).

We then moved on to preverbal and postverbal subjects and asked whether bilingual children acquire them concomitantly, much like their monolingual counterparts, and in keeping with prior research (see (iii) in Section 5). The answer to this question turned out to be positive, according to both chronological and statistical evidence, with some of the children producing their first instances of SV and VS in the very same transcript. Again, this outcome points to the conclusion that the development of overt subjects in the pro-drop language of bilinguals runs in parallel with that of their monolingual peers, suggesting the absence of CLI in this area of acquisition. This in turn argues that bilingual children indeed construct two linguistic systems which are for the most part independent, since the children's pro-drop language exhibits the word order found in the adult language, not the English word order (i.e., SVO). Furthermore, such results provide an argument for a unified account of preverbal and postverbal subjects in languages like Spanish, although such an analysis does not seem easy to accomplish in light of the well-known differences between preposed and postposed subjects. Still, such results are compatible with a scenario in which the necessary prerequisites for each type of subject are different, but crucially the last prerequisite acquired by the child is the same for SV and VS (Villa-García, 2011).

In order to provide a more complete answer with respect to CLI, we analyzed the contexts in which the children's SV and VS utterances occurred, and classified each overt-subject sentence into one of the major categories identified for preverbal and postverbal subjects. The findings suggest that, in much the same way as the null subjects, the monolingual and bilingual children's spontaneous speech contains adult-like preverbal and postverbal subjects from the start, with instances of SV(X) and V(X)S utterances that adhere to the patterns observed in the adult language. Again, despite the relatively small sample size, this result is compatible with the conclusion reached above that no CLI effects are found with respect to the acquisition of subjects in Spanish in combination with English. Put another way, the fact that the earliest utterances containing null and overt subjects in the pro-drop language of the bilingual children of this study were appropriate from the perspective of pragmatics goes against CLI. The evidence adduced in this paper, when combined with the findings of previous studies, is consistent with the contention that whether a child exhibits a higher percentage of preverbal or postverbal subjects is not necessarily attributable to the influence of English (i.e., CLI) or to bilingualism itself, but rather to the frequencies found in the input or to the type of subject (i.e., SV/VS) required in particular contexts. A note of caution is in order when drawing conclusions about CLI, however, since Silva-Corvalán (2014) has shown that early speech does not show evidence of CLI in general, but soon
after the age of 4 and if the non-pro-drop language becomes more dominant, CLI becomes more apparent in the naturalistic speech of bilingual children. Similarly, it should be noted that for Carla, Spanish was the dominant language, though it is far from clear whether this was the case for Leo and Simon (Liceras et al., 2008). Further studies should examine whether language dominance is responsible for the remarkably good performance of these children with regard to null and overt subjects in their Spanish.

It is of note that it is not only the pragmatic felicitousness of subjects that was target-like, but also the grammatical competence of the children studied. Indeed, in analyzing the transcripts considered in this study, we found no noteworthy errors of co-occurrence in the naturalistic speech of the children of this study (most of their errors were morphological in nature or simply errors of omission), which provides an argument for Snyder’s (2007, et seq.) Grammatical Conservatism proposal. This hypothesis contends that children do not make productive use of a particular construction until they have acquired the necessary requirements to generate the structure in a successful fashion. It seems that the prospect of Grammatical Conservatism can be extended to bilingual children as well.

In general, our results add to the plethora of studies in the domain of BFLA arguing for the Separate Development Hypothesis. Further research employing different methodologies (e.g., experiments), larger data samples, other language combinations, and more children will ascertain whether the results of this study are on the right track. Such studies will bring us closer to unveiling the mechanisms underlying the emergence, licensing, distribution, and use of subjects in natural language.

Acknowledgements

We are grateful to Ludovica Serratrice for initial guidance and suggestions, as well as Antxon Olarrea for comments on an early draft. We are also indebted to the audience at the regular seminar series of the School of Psychology and Clinical Language Sciences of the University of Reading, especially Jorge González Alonso and Jason Rothman, and to the audience at the Department of Spanish and Portuguese Linguistics Colloquium at the University of Arizona. We are thankful to the anonymous abstract reviewers and the audiences of The Romance Turn VIII (UAB, Barcelona) and UIC Bilingualism Forum (Chicago) for valuable feedback. This paper owes much to the comments and criticisms of three anonymous expert reviewers. We would also like to express our gratitude to Rachel Klassen and Juana Liceras for supporting this project and for their thorough comments on the last version. All errors are our own.
Early null and overt subjects in the Spanish of simultaneous English-Spanish bilinguals

References


doi:10.4159/harvard.9780674732469


doi:10.1017/CBO9781139524407


2nd proofs


Abstract

Este artículo evalúa las predicciones de la Hipótesis de la Influencia Interlingüística en el ámbito de la adquisición bilingüe temprana. Con este fin, se analizan datos de corpus longitudinales de cuatro niños bilingües con atención a la adquisición de los sujetos (nulos versus explícitos; preverbales versus postverbales), así como la adecuación pragmática de los mismos en una lengua pro-drop (el español), en conjunción con una lengua que muestra el valor opuesto del parámetro del sujeto nulo (el inglés). Los resultados obtenidos sugieren que los efectos de la Influencia Interlingüística son mínimos en el desarrollo de los sujetos en español en los estudios iniciales, a saber, en el momento en que aparecen los primeros ejemplos de sujetos omitidos y explícitos. En consecuencia, este estudio apoya la Hipótesis del Desarrollo Independiente. El grupo de niños bilingües muestra un comportamiento similar al de la niña monolingüe de este estudio: en las
etapas iniciales, ambos grupos exhiben proporciones de sujetos elididos así como trayectorias de adquisición de sujetos nulos y léxicos comparables. Al igual que los niños monolingües, los bilingües comienzan a producir sujetos preverbales y postverbales de forma simultánea. Ambos grupos utilizan los sujetos nulos y explícitos de forma adecuada desde el punto de vista pragmático, lo que sugiere una vez más la ausencia de un efecto de interferencia del inglés. A la luz de los datos monolingües y bilingües disponibles, este artículo cuestiona la existencia de la Fase del Sujeto Nulo propuesta por Grinstead (1998 en adelante) para la adquisición del español.

**Palabras clave:** bilingüismo temprano, sujetos nulos, sujetos explícitos, Influencia Interlingüística, adecuación pragmática

**Authors’ addresses**

Dr. Julio Villa-García  Imanol Suárez-Palma
Linguistics and English Language  Department of Spanish & Portuguese
The University of Manchester  The University of Arizona
Oxford Road  1423 E. University Blvd. Rm. 545
Manchester  Modern Languages Building
M13 9PL  P.O. Box 210067
UK  Tucson, Arizona 85721
UK  USA
julio.villa-garcia@manchester.ac.uk  ispalma@email.arizona.edu

**About the authors**

Dr. Julio Villa-García is a Lecturer in Spanish Linguistics and Syntax in the Division of Linguistics and English Language at the University of Manchester, in the UK. His Ph.D. from the University of Connecticut (2012) concentrated on syntax and first language acquisition, with special attention to Spanish. He is the author of several peer-reviewed articles and book chapters, and has recently published *The syntax of multiple-que sentences in Spanish: Along the left periphery* (2015, John Benjamins).

Imanol Suárez-Palma is a doctoral student in Hispanic Linguistics at the University of Arizona, where he works under the supervision of Antxon Olarra. He holds M.A. degrees from the University of Oviedo in Spain and the University of Kent in the UK. His focus is on morphosyntax and language development.