Running headline: **Null subjects in bilingualism and SLI**

Title: **Null subjects in Monolingual and Bilingual, Typical and Atypical Development**

Acknowledgements: We are very grateful to all the children and professors in the kindergartens and schools where this experiment has been conducted.
Abstract
In this paper, we investigate the production of Null Subjects by three groups of children: monolinguals acquiring Capeverdean; bilingual speakers of Capeverdean and European Portuguese, typical language development; bilingual speakers of Capeverdean and European Portuguese with SLI features. Capeverdean, a Portuguese-based Creole, does not allow null referential subjects in matrix clauses. European Portuguese, on the other hand, is a prototypical pro-drop language.

The results obtained in an elicited production task involving a picture stimulus and the answer to a question of the type ‘What is X doing to Y?’ have revealed that the production of null subjects is not a feature that enables a distinction between bilingual children with and children without language impairment.

As for the contrast between the early abandon of subject drop by Capeverdean monolinguals and the maintenance of Null Subjects among the bilinguals in both languages, the latter confirms that this is an area of interference.
Introduction

Monolingual children, bilingual children and children with specific language impairment have similarities and differences in their linguistic behaviors (Armon-Lotem, 2009; Armon-Lotem, Danon & Walters, 2010; Paradis, 2007; Roeper, 2009). The identification of the linguistic profile of different populations is crucial for a proper characterization of each group, and for defining strategies for intervention whenever relevant.

One topic of main interest in this area of linguistic studies concerns the acquisition of subjects. The syntax of subjects is known to develop with age (Hyams, 1986; Borer & Wexler, 1987; Bloom, 1990; Valian, 1991; Wexler, 1998; Rizzi, 1992, 2000, 2005a,b; a.o.), and to be delayed in children with Specific Language Impairment (Paradis, 2007; Roeper, 2009; a.o.).

In languages with no null subjects, like English, French, Dutch, German or Danish, it is known that typically developing children drop subjects in their early productions. This subject drop is limited to the sentence-initial position, as illustrated in the examples in (1), taken from Guasti (2002:151):

(1) a. Tickles me. (Adam, 3;6) [English]
   b. Se, blomster har. (Jens, 2;2) [Danish]
      look flowers have/has
      ‘Look (I/you/she/we) have/has flowers.’
   c. Mange du pain. (Grégoire, 2;1) [French]
      eat.3SG some bread

Interestingly, this root subject drop disappears quite early. 3 year old children display very low rates of sentences with null subjects (Hyams & Wexler, 1992), and there is quite robust evidence to say that, even at the stage in which they produce utterances like (1), they know that their languages do not allow null subjects.

These observations have given further support to the idea that different stages of cognitive development and different levels of exposure condition the expression of UG-constrained systems. The controversy arises when scholars take some particular position as to whether this early Null Subjects phenomenon is due to the children’s competence or to some limitations in their performance.
In order to explore his own proposal regarding the systematic discrepancies of early linguistic productions, Rizzi (2005a,b) lists the three main previous approaches in the literature:

A. Performance account: the underlying competence/grammar of a two-year old is basically adult-like, but there is a performance filter which makes child’s productions not faithful to the internalized grammar. As for the early subject drop in non pro-drop languages, this means that the child makes use of certain grammar-independent strategies of structural reduction (Valian, 1991; Bloom, 1990).

B. Grammar-based accounts, which establish a link with properties of the particular grammar: an early grammar may differ from the target adult system in the use of some legitimate grammatical options; in other words, certain parameters may be initially set on non-target values, undergoing later resetting. This is what has been proposed in Hyams (1986) as for the subject drop phenomena in early English: children start assuming that the Null Subject Parameter has a positive value (this means that a null pronominal subject is possible), which is expressed in their early productions. The evidence available to the English learner later leads to the resetting of the Null Subject Parameter on the negative value.

C. Grammar-based accounts, which establish a link with properties of Universal Grammar: it may be the case that UG is not operative in its full-fledged form initially, as certain principles or computational devices may be subject to maturation (Borer & Wexler, 1987; Wexler, 1998). One example of this approach as to the early subject drop is Rizzi’s (1992, 2000) proposal that, since it only occurs when the subject is sentence-initial (Root Subject Drop), it may be reduced to the option the child has of ‘truncating’ the external layers of the structure, and that this option could result from the non operativity in early systems of the particular UG principle that requires root clauses to be full CP’s.

Crucially, the observation that certain adult languages have a subject drop phenomenon, distinct from the Null Subject Parameter and with the same structural characteristics as early subject drop (this is for instance the case described for adult grammar in Haegeman (1995) regarding diary-style English), has led Rizzi to propose a different explanation to early subject drop. In Rizzi (2005b), the author argues for a fourth type of account that combines A. and B.: ‘discrepancies between child and target systems are grammar-based and performance-driven. Certain parameters are initially set
on values which facilitate the task of an immature production system (although here – as opposed to the account in C. – it is not assumed that particular aspects of UG mature).’ (Rizzi, 2005b:3). In other words, the existence of that other subject drop phenomenon, distinct from the Null Subject Parameter and with the same structural characteristics as early subject drop, suggests ‘that the relevant kind of subject drop appears to be a UG parameter, rather than a manifestation of a somehow ‘immature’ UG.’

According to this approach, ‘the child uses grammatically determined strategies consisting in the initial adoption of certain parametric values.’ One known effect of these values is that they alleviate the task of production, through, for instance, the use of grammatically licit processes of ellipsis (meaning here the omission of pronounced material). ‘After the production system has grown more efficient, the ‘facilitating’ values are kept if supported by experience, and abandoned otherwise.’

With this background in mind, let us turn to the goals of the present paper. We intend to investigate whether the production/omission of pronominal subjects is a distinctive factor for three populations: monolingual children acquiring Capeverdean, a Portuguese-based Creole language, bilingual children acquiring Capeverdean and European Portuguese, and bilingual children acquiring Capeverdean and European Portuguese at risk for Specific Language Impairment.

Pronominal subjects are particularly interesting for studying Capeverdean and European Portuguese, because the two languages have different parametric options. European Portuguese allows referential null subjects, whereas Capeverdean forbids them in main clauses, as shown in Table 1:

(insert table 1)

European Portuguese is a prototypical pro-drop language, allowing root null subjects, expletive null subjects also referential null subjects in certain embedded contexts, as (2) illustrates:

(2) a. root null subject:

\[
\text{Vou à praia.} \\
\text{go:1SG to:the beach}
\]
‘I’m going to the beach.’

b. *expletive null subject:

Chove.

rain

‘It rains.’

c. *embedded null subject:

Ele disse que Ø vai à praia.

3sg said that go to-the beach

‘He said that he is going to the beach.’

Capeverdean creole, unlike English, among other languages, is not a consistent non null subject language, qualifying instead as a partial pro-drop language (using Holmberg’s 2005 terminology). In fact, it forbids referential null subjects in root contexts (3a), but it allows—has obligatory expletive null subjects (3b), and also obligatory referential null subjects in embedded contexts whenever they are co-referential with a matrix quantified subject (3c) (Costa & Pratas, in press):

(3) a. No root referential null subject:

*(N) ta kanta.

1SG TMA sing

‘I sing.’

b. Expletive null subject:

Sata txobi.

PROG rain

‘It is raining.’

c. Embedded null subject if co-referent with matrix quantified subject:

Tudu, fla ma Ørj kanta na festa.

everyone said that sing in party

‘Everyone said that they sang in the party.’

On the basis of this description, it should now be obvious what is at stake for different children and what can be a research question for different populations:
- Monolingual children acquiring European Portuguese have to figure out that the language is consistent pro-drop. It has been demonstrated that they do so from very early on by Gonçalves (2004).

- Monolingual children acquiring Capeverdean creole have to figure out that the language lacks referential null subjects in root contexts. However, these children have to find out that the language permits a subset of null subjects in a very specific context (embedded subject positions when they are co-referent with a matrix quantified subject).

- Bilingual children have to master two languages with conflicting parametric options. In this sense, these two languages provide a good test case to detect areas of interference. It may also be the case that this is an area in which monolingual and bilingual children behave differently.

- Since it is known that SLI children may display delayed acquisition of obligatory subjects, their performance on this area may be a differential factor for SLI children and bilingual children.

On the basis of this list of factors, we ask the following questions:

A. Do monolingual Capeverdean children know that the language lacks referential null subjects in root declarative sentences from early on, as was found for other non null subject languages?

B. Do bilingual children behave like monolinguals or is the performance on null/overt subjects a criterion for differentiating bilingual and monolingual typically developing children?

C. Do bilingual children at risk for SLI perform differently from the other two groups? In other words, is the performance on null/overt subjects a marker for the identification of SLI in bilingual children acquiring Capeverdean and European Portuguese?

The paper is structured as follows. In the next section, we present the method used for eliciting sentences with the relevant type of subjects. In the following sections present the results and discussion, respectively, emphasizing the consequences for establishing criteria for differentiating the linguistic profile of the groups under consideration. Finally, we present some conclusions.
Method

In order to assess the production of subjects, we used an elicitation procedure, modeled after Jakubowicz, Nash, Rigaut & Gérard (1998), in which children were asked a question about a picture with the format *What is X doing?*. For instance, for a picture like the one in Picture 1, we used the question in (4):

(4) Kusé ki gatu sata fazi?
   what Q cat PROG do
   “What is the cat doing?”

The advantage of this type of question for elicitation is that the subject is salient, making it very likely to be null in an answer (in a null subject language) (cf. Samek-Lodovici, 1998, among many others). In a non pro-drop language, the expectation is that a weak subject pronoun is used in this context. This is indeed confirmed for adult Capeverdean and adult European Portuguese:

Capeverdean:
(5) A: Kusé ki gatu sata fazi?
   What Q cat PROG do
   “What is the cat doing?”
   B: a. *Ø sata morde katxor
      PROG bite dog
   b. E sata morde katxor
      3SG PROG bite dog
      ‘It is biting the dog.’

European Portuguese:
(6) A: O que é que o gato está a fazer?
   What Q the cat is doing
‘What is the cat doing?’

B: a. Ø está a morder o cão.
   BE PREP bite the dog
   ‘It is biting the dog.’

b. #Ele está a morder o cão.
   3SG BE PREP bite the dog

We used ten different pictures, for a total of ten items. All items included transitive verbs, and had temporal and aspectual markers that were correspondent in the two languages. The test was run on three groups of participants:

   a) Typically developing monolingual children acquiring Capeverdean creole aged from 2;4 to 3;6.
   b) Typically developing bilingual children acquiring Capeverdean creole and European Portuguese aged between 8 and 9.
   c) Bilingual children at risk for SLI acquiring Capeverdean creole and European Portuguese aged between 8 and 9.

Although there are no standardized tests for Capeverdean creole, and these children were not diagnosed for SLI, they had features that made them strongly resemble SLI cases: they all had a low score on language tests run in Portuguese (percentile 10 or below), tested with the language development skills test for Portuguese TALC (Kay and Tavares 2008), normal non-verbal IQ (according to Raven test), and absence of any type of articulatory, hearing, neurological or mental impairment.

The monolingual group was assessed in Capeverdean only. The bilingual groups were assessed in the two languages, on separate days, by independent experimenters, native speakers of the languages being tested. The order of application of the tests for the bilingual groups was: Capeverdean first, Portuguese after. This was done this way in order to prevent the influence of Portuguese on Capeverdean.

All participants were tested individually in a quiet room by two experimenters. One of the experimenters asked the questions, and the other transcribed the child’s response (the same role was played by the same experimenter for all children). No time limit was imposed, and no stimulus or correction was given depending on the type of response,
besides general encouragement to pursue the task and a final reward after completion of the task (for the children only). The tests were recorded using a digital voice recorder. Children’s productions were fully transcribed during and after the session by the experimenters. One transcription was made by one of the experimenters during the session, right after the child’s utterance. A second transcription was made on the basis of the audio recording. In case of false starts or reformulations, everything was transcribed, but only the last utterance of the child was counted and coded. The second transcriptions were double-checked and compared with the original transcriptions for reliability. No case of disagreement between the original and the final transcript occurred.

The data were coded for overt or null subject.

The three groups of participants are as follows:

**Group A** included 13 monolingual children acquiring Capeverdean creole, living in Santiago Island, Cape Verde. These children (6 boys, 7 girls) were aged between 2;4 and 3;6, and had no reported malfunction in language or any other developmental, cognitive or physical impairment. This group was further subdivided into two subgroups (I and II), according to age. Group I includes 7 children aged between 2;4 and 2;11. Group II includes 6 children aged between 3;0 and 3;6. This division was made because the literature reports that age 3 is critical for the emergence of lexical subjects in languages in which null subjects are ruled out.

**Group B** includes 6 sequential bilingual children aged between 8;3 and 9;6 (XX boys, XX girls). These children first acquired Capeverdean and then Portuguese. They were all born in Portugal, are spoken to in Capeverdean at home and in their neighborhood and have schooling in Portuguese.

**Group C** includes 7 sequential bilingual children aged between 6;9 and 9;3 (two boys, five girls) with SLI features. Although there are no standardized tests for Capeverdean creole, and these children were not diagnosed for SLI, they had features that made them strongly resemble SLI cases: they all had a low score on language tests run in Portuguese (percentile 10 or below), normal non verbal IQ (according to Raven test), and absence of any type of articulatory, hearing, neurological or mental impairment.

**Results**
The experiment yielded the following results for the different groups:

**Group A: Monolingual children**
Production of sentences with null subjects
Group I: 37.14% (26/70) Ages 2;4-2;11
Group II: 16.66% (10/60) Ages 3;0-3;6

**Group B: Typically developing bilingual children**
Production of sentences with null subjects
In Portuguese: 98.5% (69/70)
In Capeverdean: 88.5% (62/70)

**Group C: Language impaired bilingual children**
Production of sentences with null subjects
In Portuguese: 100% (77/77)
In Capeverdean: 100% (77/77)

The results are extremely clear. First, there is an obvious asymmetry between monolingual and bilingual children, since only the former do not prefer null subjects, whereas bilingual children opt for producing sentences with null subjects.

The second clear result is that there is a developmental effect in first language acquisition similar to the one reported for other languages: the youngest subgroup of monolinguals drop more subjects than the children who are older than 3.

As for bilinguals, it is very clear that there was a massive production of sentences with null subjects independently of the language they were tested on, and independently of there being features of Specific Language Impairment.

The following list summarizes the main results of the experiment:

a) Only monolingual children avoided the production of null subjects in Capeverdean creole;

b) There was a developmental effect on monolingual children similar to what has been reported for other languages in which null subjects are ruled out, since children mastered the avoidance of null subjects in the relevant context at around age 3;
c) Bilingual children produced null subjects both in Capeverdean creole and in European Portuguese;

d) Typically developing and impaired bilingual children behaved alike.

Discussion

The results presented in the previous section are quite revealing. On the basis of them, we are able to draw a couple of conclusions we would like to elaborate on in this section.

First of all, we think that the results of the experiment provide evidence in favor of Wexler’s (1998) idea that the value of parameters is set very early. In fact, young monolingual children acquiring Capeverdean creole produce a very low rate of null subjects: Group I, ages 2;4-2;11: 37.14% of subject drop; Group II, ages 3;0-3;6: 16.66% of subject drop. This contrasts with what has been reported for pro-drop languages. Gonçalves (2005) and Grinstead (1998) report that monolingual children acquiring European Portuguese and Spanish, respectively, produced very high rates of null subjects in the relevant contexts. Altogether, this indicates that children know whether their first language is a null subject language or not. In other words, young children are setting the relevant parameters related to null subjects very early.

These rates of subject drop, even though they are low, call for an explanation. In Pratas (2009) it has been proposed that these results are consistent with the hypothesis of a truncated structure (Rizzi, 1992, 2000). This truncation is said to occur, in this case, above VP (in the children’s elicited productions there was always some preverbal morphology marking the progressive; hence no form similar to a Root Infinitive), presumably at the TP level.

Given what we have described in the present paper, however, we have now evidence to consider, as an adaptation of the main proposal in Rizzi (2005b), that these Capeverdean children are using this type of grammatically determined strategy, by adopting a certain parametric value that alleviates the task of production. Rizzi (2005b) proposes that the specific properties of the production system around the age of two lead to the adoption, by the Language Acquisition Device, of a formal strategy which
has the functional role of simplifying the performance task. This strategy can be stated as follows:

(7) Adopt parametric values which reduce computational load on the production system and are not contradicted by positive evidence.

Rizzi further makes it clear that ‘it looks plausible that any grammatically licit ellipsis will reduce the burden of the production system, by allowing it not to pronounce chunks of the linguistic representation’. This is why there is such a contrast in the acquisition of parameters that have grammatically licit ellipsis, on the one hand (there is an attested parametric discontinuity in the following aspects: root subject drop; determiner drop; ellipsis of copulas and auxiliaries; root infinitives), and word order phenomena parameters, on the other hand (there is no attested parametric discontinuity in the following parameters: head-complement; Vo to I; generalized V2; Null Subject). The point is that the latter may be largely irrelevant in increasing or decreasing the burden of the production system.

The relevant adaptation of Rizzi’s proposal, given the properties of Capeverdean, is that, not only do these very young monolingual children not have any positive evidence to contradict this other subject drop parameter value (theirs is one of those adult languages that allow for a type of diary-style), as their language provides some kind of positive evidence that may be taken to confirm it: the obligatorily null expletive subjects and also the obligatorily null embedded subjects whenever they are co-referent with a matrix quantified subject (Costa & Pratas, in press).

In later stages of development, Capeverdean monolingual children abandon this strategy and learn to distinguish between contexts. This idea is supported by the decrease in the rates of subject from Group A.I to Group A.II.

The second result worth discussing is the difference between monolinguals and bilinguals. Crucially, and as mentioned in the previous section, a big difference was found between both bilingual groups (typically developing and language impaired), on the one hand, and monolinguals, on the other hand. Whereas monolinguals avoided sentences with null subjects, bilinguals uttered them most of the time, and did not exhibit a differential behavior in Capeverdean and in European Portuguese.
Since the children in Groups B and C are older (around 8 years old), we must take into account the previously mentioned observation of Rizzi concerning the use of a facilitating strategy that involves ellipsis: ‘After the production system has grown more efficient, the ‘facilitating’ values are kept if supported by experience, and abandoned otherwise.’ Therefore, if we want to stick with this approach to subject drop in its more general terms, we must now discuss whether, in the case examined here, there is this kind of ‘support by experience’ to keep the facilitating values or not.

But, before we develop our arguments to answer this question, we must handle the fact that these children are sequential bilingual. Under Muller & Hulk’s (2001) analysis of areas of crosslinguistic interference in bilingualism, our pattern of responses can be readily accounted for in the following terms. It is well known that typically developing bilingual children separate the systems they are acquiring from very early on. This is known as the dual system hypothesis (cf. Meisel, 2004, for a comprehensive review). Accordingly, one might expect that our bilingual participants would produce low rates of null subjects when they were tested in Capeverdean, which did not happen. Müller & Hulk (2001) suggest that, in spite of the fact that the two languages are acquired separately, there are clear areas of interference. According to these authors, interference is expected in the areas of interface. Now, if we think of the syntax of null subjects in Capeverdean, there is a clear area of interface. First, the children must figure out that the language does not have null subjects. Moreover, they must figure out that this conclusion is only partial, because the language has obligatory expletive null subjects and, more importantly, under specific discourse settings (if there is a matrix co-referential quantified antecedent) the referential subject must also be null in embedded contexts.

Following Müller & Hulk’s (2001) suggestion, one may hypothesize that bilingual children overuse null subjects in Capeverdean because this is an area in which an accurate mastery of the interface requirements on discourse, binding by eventual matrix antecedents, and the choice between lexical, weak and null pronouns converge.

One obvious question that comes up, however, is why the sense of interference is dictating that there are more null subjects than lexical subjects and not the other way round. We think there are two clear reasons for this. First, as shown in the introduction, Capeverdean does have null subjects in some very specific contexts. As such, the interference is actually the result of choosing one option independently made available
in their system. Second, there is independent evidence from the acquisition of complement pronouns and null objects in European Portuguese showing that, whenever children are confused for some reason, they opt for the weakest possible form – in this case, the null option. In other words, and following Rizzi’s terms for younger children’s productions, they opt for the strategy that alleviates the production system.

Now, we can finally go back to the question that we have left behind: is there this kind of ‘support by experience’ available to these bilingual children, allowing them to keep any facilitating values ‘after the production system has grown more efficient’, as opposed to what seems to happen with the monolinguals (this was hinted by the decrease of subject drops from Group A.I to Group A.II)? We suppose that the answer to this question is clear by now. Yes, there is. Considering that this is a clear area of interface and, thus, of interference (Müller & Hulk, 2001), for these older children the ‘facilitating’ values are supported by experience, through the evidence of another language, which allows their maintenance.

A final aspect needs to be discussed regarding the results: there was no difference between typically developing children and children with SLI features. Both groups B and C had very high rates of sentences with null subjects. In this sense, it should be evident that subject omission is not a clear marker for differentiating bilingualism and language impairment, since the two groups behave alike. Note that we are not ruling out the possibility that subject omission is a feature of SLI. However, if we would like to find a structure or construction that clearly differentiates the two groups, null subjects are not a good candidate, since we find no evidence for a differential performance.

Conclusion

In this paper, we provided evidence in favor of the following claims on the basis of the investigation of children acquiring Capeverdean creole and European Portuguese:

a) Monolingual Capeverdean children set the value of parameters for null subjects from very early on, ruling out root referential null subjects;

b) The production of null subjects is a feature for the identification of bilingualism, since all bilingual speakers of Capeverdean creole (unlike monolinguals) produced very high rates of null subjects;
c) The production of null subjects by bilinguals with language impairment is not a feature that enables a distinction between bilingual children with or without language impairment, since there was no difference between the typically developing group and the group at risk for SLI;
d) The interference found in bilinguals can be attributed to the fact that the mastery of null subjects in Capeverdean involves a mastery of interface issues. As such, null subjects are an area vulnerable to interference, according to Müller and Hulk (2001).

References


<table>
<thead>
<tr>
<th></th>
<th>European Portuguese</th>
<th>Capeverdean Creole</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg</td>
<td>(eu) canto.</td>
<td>*(N) kanta.</td>
<td>‘I sing.’</td>
</tr>
<tr>
<td>2sg</td>
<td>(tu) cantas.</td>
<td>*(bu) ta kanta.</td>
<td>‘You sing.’</td>
</tr>
<tr>
<td>3sg</td>
<td>(ele) canta.</td>
<td>*(E) ta kanta.</td>
<td>‘He sings.’</td>
</tr>
<tr>
<td>1pl</td>
<td>(nós) cantamos.</td>
<td>*(Nu) ta kanta.</td>
<td>‘We sing.’</td>
</tr>
<tr>
<td>2pl</td>
<td>(vocês) cantam</td>
<td>*(Nhos) ta kanta.</td>
<td>‘You sing.’</td>
</tr>
<tr>
<td>3pl</td>
<td>(eles) cantam</td>
<td>*(Es) ta kanta.</td>
<td>‘They sing.’</td>
</tr>
</tbody>
</table>

Table 1: Null referential subjects in simple clauses: licit in European Portuguese, illicit in Capeverdean
Picture 1

(courtesy Naama Friedmann)