



---

# Universidad de Valladolid

FACULTAD de FILOSOFÍA Y LETRAS  
DEPARTAMENTO de FILOLOGÍA INGLESA  
Grado en Estudios Ingleses

## TRABAJO DE FIN DE GRADO

The analysis of English sentential subjects in the spontaneous production of English/Spanish simultaneous bilingual children

Andrea San José Prieto

Tutora: Raquel Fernández Fuertes

2018-2019



## **ABSTRACT**

This dissertation presents a study on the analysis of the production of English sentential subjects on the part of simultaneous English/Spanish bilingual children and on the examination of the children's development on the realization of the subject element in English from the age of 2;11 to 5;11. The data collected are spontaneous and have been selected from CHILDES. Data analysis focuses on three variables: referentiality, overtness and grammatical correctness; and the objective is to test the children's linguistic production of sentential subjects, both overall and by developmental stages. The results obtained are in line with what previous studies have found: the children reveal an early consciousness on how subjects are used in English; and the children present some patterns in their subject production which could be related to crosslinguistic influence from Spanish into English.

**KEY WORDS:** Null Subject Parameter, language acquisition, crosslinguistic influence, English, sentential subjects, bilingualism.

## **RESUMEN**

Este trabajo presenta un estudio de los sujetos oracionales del inglés en la producción espontánea de niños bilingües simultáneos inglés/español además de un análisis del desarrollo lingüístico en la producción del sujeto desde los 2;11 años de edad hasta los 5;11. Los datos que se han estudiado son espontáneos y han sido extraídos del proyecto CHILDES. El análisis de este estudio se centra en tres variables: la referencialidad del sujeto, la presencia de sujetos nulos o explícitos y la corrección gramatical del sujeto. El objetivo se centra en examinar la producción lingüística del sujeto oracional del inglés de los niños tanto en la producción global como a lo largo de las distintas fases del desarrollo. Los resultados obtenidos coinciden con los de estudios previos: los niños muestran una conciencia temprana sobre cómo se usan los sujetos en inglés; y los niños presentan algunos patrones en su producción de sujetos oracionales que podrían estar relacionados con la influencia interlingüística del español al inglés.

**PALABRAS CLAVE:** parámetro del sujeto nulo, adquisición del lenguaje, influencia interlingüística, inglés, sujetos oracionales, bilingüismo.

## Table of contents

<b>1. Introduction</b> .....	1
<b>2. Theoretical Framework</b> .....	2
<i>2.1. The Grammatical Properties of Subjects</i> .....	2
<i>2.2. Principles and Parameters</i> .....	3
<i>a. The Extended Projection Principle</i> .....	3
<i>b. The Null Subject Parameter</i> .....	4
<b>3. Previous Works</b> .....	6
<b>4. Methodology</b> .....	9
<i>4.1. Research Objectives and Hypotheses</i> .....	9
<i>4.2. Data Selection</i> .....	10
<i>4.3. Participants</i> .....	10
<i>4.4. Data Classification Criteria</i> .....	13
<b>5. Data Analysis and Discussion</b> .....	16
<i>5.1. Referentiality of Sentential Subjects</i> .....	17
<i>5.2. Subject Overttness</i> .....	21
<i>5.3. Grammatical Correctness</i> .....	25
<b>6. Conclusion</b> .....	28
<b>7. Bibliography</b> .....	31

## **1. Introduction**

Lyons (2008) defines language acquisition as the process whereby children achieve a fluent control of their native language and considers that the study of language acquisition has been strongly influenced by Chomsky's (1968) theory of generative grammar, which suggests that children are born with a knowledge of universal formal principles that determine the grammatical structure of language. Moreover, during the developmental process of language acquisition there are some factors that intervene: the input children are exposed to and the time they are exposed to language which will enable them to acquire the different properties of a specific language, or the problems children might deal with when acquiring a language.

This dissertation is focused not only on the analysis of the production of sentential subjects on the part of simultaneous English/Spanish bilingual children; but also, on the examination of the children's development regarding the realization of the subject element from the age of 2;11 to 5;11 - which is the age in which children are expected to have acquired the basic properties of the language they are exposed to. Hence, the present dissertation takes previous studies on bilingual language acquisition together with some theoretical implications about the subject element so as to establish the basis of this study.

This dissertation is divided into eight sections starting with the introduction. The second section deals with the theoretical framework which presents a discussion on sentential subjects and the different properties that define them, with a special emphasis on *The Extended Projection Principle* and *The Null Subject Parameter*. Then, in section three, an overview of the previous studies on bilingual language acquisition is provided. Section four presents the objectives and the different hypotheses of the study together with the methodology followed in order to perform the analysis. Next, section five presents the results gathered from the analysis of the data and the discussion of those results. The following section (section six) deals with the conclusion reached once the results have been analysed and discussed. And, finally, section seven provides all the bibliographical sources referred to throughout this dissertation. In addition, the CD

annexed includes the Excel database where data selection and data classification have been compiled.

## **2. Theoretical Framework**

This section presents a discussion on sentential subjects and the properties that define them in agreement with different scholars. Furthermore, a concise explanation about each of the principles that characterize the subject element will be presented together with the divergences that could emerge when this grammatical category is analysed in two different languages which, in this case, are English and Spanish. Moreover, a specification of the characteristics and the different situations where the *Null Subject Parameter* occurs will be included due to the fact that the major issue dealt with in this paper revolves around the subject element.

### **2.1. The Grammatical Properties of Subjects**

Andrews (1985) explains that there are 5 different properties which are usually shared by subjects. The first property states that subjects normally take the form of a DP (1a), but also, they can adopt the form of a clause (1b) or a PP (1c).

- (1) a. *The child* is playing
- b. *Reading a book* is relaxing
- c. *Between 1 or 2* is enough

Furthermore, subjects are the external argument of a verb and, therefore, they are usually marked with nominative case because of verbal inflection.

The third property deals with verbal-agreement since subjects agree in person and number with the verb that follows them. For example, in English verbs are marked with an -s for the third person singular in the present tense (2a); and in Spanish, verbal-agreement appears in all verb forms since, in this language, verbal inflection almost uniquely identifies which person performs the action (2b).

- (2) a. She plays the piano
- b. Yo como macarrones, pero mis hermanos prefieren sopa.

The fourth property is related with the preverbal position subjects typically occupy. That is, subjects are normally recognized as the leftmost constituent within a sentence, as in (3); however, sometimes other elements whose function are pragmatic, or emphatic can also appear in this initial position typically associated with subjects, as in (4).

- (3) *Peter* likes playing football
- (4) *Last week*, I went to the cinema

Finally, the last property deals with the fact that subjects assign person, number and gender to the co-referential reflexive pronoun that appears within the sentence. This property can be seen in example (5), where the pronoun *she* assigns person, number and gender features to the reflexive pronoun *herself* (i.e. third person singular and feminine) which functions as the direct object of the sentence.

- (5) *She* kicked *herself*

## **2.2. Principles and Parameters**

The properties aforementioned typify subjects across languages although they are focused on English. Therefore, a more in-depth discussion of the abstraction of subjects in Universal Grammar (UG) is worthily mentioned. Accordingly, UG was proposed by Chomsky (1965) as a way to capture that human beings not only have innate capacities connected with language acquisition, but also, that all human languages share a set of systems and categories which, together with a series of principles and parameters, constitute UG. Consequently, due to the fact that this dissertation is focused on the analysis of subjects, the principle that is

going to be taken into consideration is the *Extended Projection Principle* and in the case of parameters the parameter at stake is the *Null Subject Parameter*.

### **a. The Extended Projection Principle**

The *Extended Projection Principle* was first developed by Chomsky in 1986. This principle states that “a lexical structure must be represented categorically at every syntactic level” (Chomsky, 1982:8); that is, any verb demands a particular number of arguments.

Following Chomsky (1995), and according to the formulation of the *Extended Projection Principle*, the [Spec, IP] position (specifier of the inflectional phrase) - which is the subject of IP - is a compulsory slot in any sentence. This may be so as a morphological property of inflection (I) or due to the predicational character of the VP (verbal phrase). To put it in a different way, “the EPP (*Extended Projection Principle*) demands that a clause have a subject at every syntactic level” (Chomsky, 1995:65); in other words, all clauses must have a subject. Due to the fact that principles are universal, they must appear and be obeyed in all languages. Nevertheless, even though all clauses in any language require a subject, the nature of the subject could change depending on the language at issue. This is identified as subject variability across languages which is condensed in the *Null Subject Parameter*.

### **b. The Null Subject Parameter**

As previously stated, the *Null Subject Parameter* and the *Extended Projection Principle* are the focus of this study. The *Null Subject Parameter* is also identified as the ‘pro-drop’ parameter. According to Hyams (1989), this parameter explains how in languages like Spanish or Italian the production of null subjects in tensed sentences is allowed (Hyams, 1989:215); in other words, this parameter assesses whether a language can omit the subject element or not. The *Null Subject Parameter* divides languages into two types: [+ null subject] languages, and [- null subject] languages. The fact that a language is classified as a [+/- null subject] language is determined by the compliance with a certain



number of properties, but the remarkable one is related with the possibility of null subjects with inflected verbs.

Due to the fact that this dissertation is concentrated on the analysis of English and Spanish, it is important to take into account that English is set as a [- null subject] language while Spanish is a [+ null subject] language; therefore, the explanation of the *Null Subject Parameter* that will be presented hereafter will show a divergence between the two languages.

While the projection of the subject position is obligatory (as per the *Extended Projection Principle*), there are some languages that permit its omission, as Spanish in (6), where the pronominal subject *tú* is omitted. In other words, this means that in [+ null subject] languages like Spanish the lexical subject is completely optional because verbal inflection (-s in *vas*) indicates the person and number of the subject; thus, there is no necessity of having an explicit subject. In (7), the inflection of the verb -n indicates third-person plural and, consequently, the explicit pronominal subject can be omitted. Subject omission is indicated by the null category *pro*.

On the other hand, English is a [- null subject] language as it has weak verbal inflection. Therefore, the inflection of the verb does not provide enough information to grammatically identify the person and number of the subject. Consequently, the subject has to be overtly expressed, as in (8).

(6) *pro* vas a comprar chocolate

(7) *pro* cantan

(8) *they* sing

In addition, the subject element has some referential properties that give way to two subject types: referential and non-referential. Referential subjects are subjects which have a referent. In Spanish, referential subjects present two forms: null and overt. Thus, Spanish shows not only null referential subjects as it

can be observed in (9a), but also, overt referential subjects through the use of DPs and pronouns, as in (9b) and (9c) respectively.

- (9) a. *pro* me mandó un mensaje esta mañana  
b. *Laura* me mandó un mensaje esta mañana  
c. *Ella* me mandó un mensaje esta mañana

On the contrary, non-referential subjects deal with subjects that do not have a referent. In Spanish, non-referential subjects are always null, as in (10).

- (10) *pro* ha nevado.

English, as opposed to Spanish, requires its subjects to always be overt whether they are referential and non-referential subjects because it is a [- null subject] language. In this case, referential subjects are presented through pronouns (11a) and DPs (11b), two of the three options Spanish has.

- (11) a. *She* is dancing.  
b. *My sister* is dancing.

English non-referential subjects need to be overt, too, and thus expletives are used. Expletives do not have lexical meaning but are used due to syntactic reasons. This can be observed in (12).

- (12) *It* is snowing.

Summarizing, the *Null Subject Parameter* presents two different settings and while English is a [- null subject] language in which the subject cannot be omitted; Spanish is a [+ null subject] language since it allows the presence of null elements in subject position. With regards to the referentiality of subjects, Spanish allows the presence of null elements both in the case of referential subjects (where null elements are an option) and in the case of non-referential

subjects (where null elements are obligatory); while in the case of English, both referential and non-referential subjects must be overt.

### **3. Previous Works**

Sentential subjects have long been studied in the case of acquisition works. Thus, due to the fact that this dissertation is focused on the analysis of the acquisition and production of sentential subjects on the part of simultaneous English/Spanish bilingual children, a selection of previous works regarding the study of subject realization on simultaneous bilingual children have been used in order to establish the basis of the present study (i.e. Paradis & Navarro, 2003; Juan-Garau & Pérez-Vidal, 2000; Serratrice et al. 2004). Consequently, the results of these works will show some important perceptions and implications which will serve to put in context the issue under analysis in this dissertation. Moreover, as the children whose English production is analysed are bilingual, and given the differences between English and Spanish discussed in the previous section, the question is how children acquire these distinct properties. That is, if each of the languages of the bilingual has a different setting of the *Null Subject Parameter*, will bilinguals' English production be affected by their other language, that is, Spanish?

Paradis and Navarro (2003) centre their study on several hypotheses. The first one is related to whether crosslinguistic influence occurs in the domain of subject realization in Spanish in a bilingual acquisition context; and, the second hypothesis explores whether the source of the interference is due to child-internal crosslinguistic contact between English and Spanish, or due to the nature of the language input the child receives in a bilingual family.

Paradis and Navarro (2003) examine the use of overt subjects in Spanish by two Spanish monolingual children, one Spanish-English bilingual child and their parental interlocutors. They calculate the rates of overt versus null subjects as well as the discourse pragmatic contexts where overt subjects are used by the children so as to find out both bilingual and monolingual differences in the distributional properties and the functional determinants of subject realization. Moreover, within their study, they

investigate the potential influence of the adult's input on the children's output. The results obtained suggest that the bilingual child showed patterns in her subject realizations in Spanish which could be interpreted as a consequence of crosslinguistic effects from English (that is, the child uses more overt subjects in Spanish than her monolingual counterparts because there is an influence from the obligatory nature of overt subjects in English); however, the authors also found evidence that this result may have its origin in the input the child is exposed to, instead of resulting from internal crosslanguage contact (that is, the Spanish variety this child is exposed to - Caribbean Spanish - makes her more exposed to overt subjects in Spanish than the Spanish variety the monolingual children are exposed to - Peninsular Spanish).

Juan-Garau and Pérez-Vidal (2000) report on a case study of bilingual first language acquisition in Catalan, which is a [+null subject] language, and English, which is a [-null subject] language. They focus their study on the subject realization in the two contrasting languages the child is acquiring simultaneously. The main aim of Juan-Garau and Pérez-Vidal is to research on language separation in the early stages of acquisition and to prove whether there is evidence of crosslinguistic influence on the child's development; for example, as the authors explain, they want to ascertain that, if the child maintains subject optionality in English for an extended period of time, it could be interpreted as a result of Catalan interference since Catalan would be the language in which the child develops faster. On the contrary, if the child acquires subject obligatoriness in English, it would provide evidence that the child's different grammatical systems evolve separately from an early stage.

Juan-Garau and Pérez-Vidal (2000) analyse one bilingual participant that has been exposed to the two languages, Catalan and English, from birth. After organizing the data, the results they obtain suggest a connection between the optional subject phenomenon and the development of the inflection component in the child's grammatical system for Catalan, but not so clearly for English. In essence and according to their conclusions, the child seems to have developed an early awareness that subjects are required in English but not so in Spanish and he appears to follow separate routes in his syntactic development in each language.

Finally, Serratrice et al. (2004) focus their study on testing the interaction between syntax and pragmatics, and on the comparison of the distribution of subjects and objects in the Italian and English of a bilingual child, so as to demonstrate the existence of crosslinguistic influence. Their main hypotheses are connected not only with the fact of finding any evidence of crosslinguistic influence in the distribution of subject and object arguments in an English-Italian bilingual child, but also with the use of overt pronominal subjects in Italian by the bilingual child in contexts where monolinguals would use a null subject. The results they obtain confirm their predictions and support the argument that crosslinguistic influence may occur in bilingual first language acquisition in specific contexts in which syntax and pragmatics interact. However, they only find evidence of crosslinguistic influence in the distribution of overt pronominal subjects in Italian and not from Italian into English.

All in all, these studies on bilingual acquisition and subjects show that some kind of interaction may occur between the two grammars the bilingual is acquiring. Taking these studies as the point of departure, the next section deals with the methodology carried out to select the data so as to do the analysis that will lead to address objectives of the present study and to confirm or reject the different hypotheses put forward.

## **4. Methodology**

This section of the dissertation discusses the methodological process followed in order to perform the analysis of the data as well as the research objectives and hypotheses that guide the analysis. This section is divided into different subsections in which the research objectives and hypotheses, the data selection procedure, the description of the participants' profile, and the data classification process will be described.

### **4.1. Research Objectives and Hypotheses**

Bearing in mind the preceding discussion of comparative analyses, both empirical (section 3) and theoretical (section 2) on subject production, the present research has two aims: to provide an analysis of the production of English sentential subjects on the part of simultaneous bilingual children; and to

test the children's development on the realization of the subject element from the age of 2;11 to 5;11. For that purpose, the children's subject production has been analysed both overall as well as by stages. In this last case, it has been assessed every six months until the age of 5;11, the age at which it is considered that children have fully acquired the basic properties of the language they are exposed to.

Additionally, the hypotheses that will be tested against the data gathered are as follows and the three of them will be addressed considering the overall production of the children as well as their production by developmental stages:

1. Bilingual English/Spanish children would produce more referential subjects than non-referential ones in English.
2. Bilingual English/Spanish children would produce more null subjects than overt subjects in English during their early stages of acquisition.
3. With regard to ungrammaticality in English, bilingual English/Spanish children would produce more commission cases than omission cases during their process of acquisition.

With regards to hypothesis 1, children are expected to produce more referential subjects (13) than non-referential subjects (14), because referential subjects are directly tied to a referent while non-referential subjects are grammatical categories that contribute no meaning.

(13)            *My brother* is eating an ice-cream

(14)            Yesterday, *it* was raining

As for hypothesis 2, the focus is to examine whether the over-production of null subjects in English during the early stages of acquisition is due to the contact with the Spanish language; in other words, whether in the simultaneous acquisition of English and Spanish, there is crosslinguistic influence on the realization of sentential subjects from Spanish into English. This could be so

because null subjects are frequently used in Spanish and they are a grammatical option.

As for hypothesis 3, children are intended to produce more commission cases (15), which are those errors related with the selection of the correct pronoun that functions as the subject; than omission cases (16), which refer to the omission of the subject element, during their acquisition process.

(15) because *it* (i)s all my toys

(16) *pro* don't see it

As it can be observed in (15), the children have produced a commission case by using the pronoun *it* so as to refer to more than one toy (*they*); while, in (16), the child has produced an omission case because he has completely omitted the subject (*I*).

## 4.2. Data Selection

With the aim of selecting the data for this dissertation, the corpora accessible in CHILDES (MacWhinney, 2000) have been used. It is important to mention that CHILDES is a database where a collection of corpora containing the spontaneous production of children and the adults that interact with them is included; in other words, it contains audio/video recordings and the corresponding transcriptions of the original oral production. Furthermore, each corpus incorporates information of the children's linguistic environment and of how the data were collected.

Due to the fact that the core issue of this study is the analysis of the subject production in English/Spanish simultaneous bilingual children, the selected corpus was the FerFuLice corpus. The FerFuLice corpus contains a longitudinal study of the spontaneous production of a set of English/Spanish bilingual identical twins. Therefore, the transcriptions are sorted in different folders depending on the language of the recordings: English recordings and Spanish

recordings. Due to the fact that the objective of the present dissertation is to analyse sentential subjects in the children's English production, the English recording folder is the one from which data have been selected.

### **4.3. Participants**

The participants in the FerFuLice corpus are Simon and Leo. Simon and Leo are two identical English/Spanish bilingual twins from a middle-class family of Salamanca (Spain). Their father is a native speaker of Peninsular Spanish while their mother is a native speaker of American English. The parents use the one parent-one-language principle when they communicate with their children. The parents mostly use the Spanish language when communicating with each other except when there is a monolingual English speaker present, or during summer when they go to the United States. In addition, during the first year of the children, the mother was the children's main caretaker while the father was present all-day during weekends and less on weekdays. When the twins were 1;10, they experienced some more contact with Spanish since they went to day care. Moreover, the children had additional contact with English not only during sporadic visits by their maternal grandparents, but also during their residence in the United States each summer.

Once the selection of the specific files containing the participants' production has been carried out, the MLU (Mean Length of Utterances) of both children in the different stages of language production was calculated through the use of the MLU program available in the CLAN software. Thus, the MLU value illustrates the linguistic productivity of the children and their linguistic development.

*Table 1* presents a general view of the data selected divided into the 7 developmental stages to be considered.



**Table 1: Participants Selection**

<b>Stage</b>	<b>File selected</b>	<b>Age</b>	<b>Simon MLU-English</b>	<b>Leo MLU-English</b>
Stage 1	“021105”	2;11.05	2.592	2.274
Stage 2	“030404”	3;04.04	3.905	3.278
Stage 3	“031000”	3;10.00	3.915	3.914
Stage 4	“040400”	4;04.00	4.846	4.356
Stage 5	“041013”	4;10.13	4.527	3.992
Stage 6	“050426”	5;04.26	4.275	3.041
Stage 7	“051123”	5;11.23	4.745	5.456

As shown in *Table 1*, 7 files have been selected each corresponding to a developmental stage as indicated by the different ages and MLU values each child has in each of these stages.

#### **4.4. Data Classification Criteria**

The classification of the data selected for the present dissertation has been compiled in an Excel database adjoined to this dissertation in electronic format. Hereafter, the classification sheet that has been designed and the different variables used to classify the data will be explained.

Three blocks of information have been considered: a first block containing general contextual information; a second block containing the specific subject instance produced; and a third block with the syntactic classification of the subject instance. The information in the first block includes the name of the participant (Simon and Leo), the age, the developmental stage, the MLU and the file. The third block includes the three syntactic variables considered for the analysis of the subject element: referentiality, overtiness, and grammatical correctness.

In relation to the variable of referentiality, the data are classified into *referential* and *non-referential* depending on whether the subject has a referent or not, respectively. Therefore, sentences like (17a) and (17b) produced by Leo; and (18a) and (18b) in the case of Simon are codified as in *Table 2*.

- (17) a. *you* turn it off  
 b. *there is* some more xxx
- (18) a. *our soap* is n(o)t clean  
 b. *there (i)s* a poopie

**Table 2. Referentiality**

<b>Participant</b>	<b>Example</b>	<b>Referential</b>	<b>Non-referential</b>
Leo	<i>you</i> turn it off	1	0
Leo	<i>there is</i> some more xxx	0	1
Simon	<i>our soap</i> is n(o)t clean	1	0
Simon	<i>there (i)s</i> a poopie	0	1

Regarding the variable of overttness, the data are classified into *overt* subjects and *null* subjects. Sentences with overt subjects as (19a) and (19b) in the case of Leo; and sentences with null subjects like (20a) and (20b) in the case of Simon were classified as in *Table 3*.

- (19) a. *she* did n(o)t come back  
 b. *pro* pick one!
- (20) a. *the(re) a(re)* exactly fifty two lette(r)s  
 b. *pro* don't ask me

**Table 3. Overttness**

<b>Participant</b>	<b>Example</b>	<b>Overt subject</b>	<b>Null subject</b>
Leo	<i>she</i> did n(o)t come back	1	0
Leo	<i>pro</i> pick one!	0	1
Simon	<i>the(re) a(re)</i> exactly fifty two lette(r)s	1	0
Simon	<i>pro</i> don't ask me!	0	1

Concerning grammatical correctness, this third variable is linked to the second one (overttness) as it focused on null subjects (omission cases) and on wrong used of subjects (commission cases). In this case, data are classified into grammatical null subjects and ungrammatical null subjects. Consequently, sentences like (21a) and (21b) produced by Leo; and (22a) and (22b) produced by Simon are classified as in *Table 4*.

- (21) a. *pro* come downstairs!  
 b. here \**pro (i)s* a window
- (22) a. *pro* close your eyes!  
 b. \**pro* can't believe he got a match

**Table 4. Grammatical Correctness**

<b>Participant</b>	<b>Example</b>	<b>Null subject</b>	<b>*Null subject</b>
Leo	<i>pro</i> come downstairs!	1	0
Leo	here * <i>pro (i)s</i> a window	0	1
Simon	<i>pro</i> close your eyes!	1	0
Simon	* <i>pro</i> can't believe he got a match	0	1

Finally, during the data classification process, some instances were excluded. These correspond to fixed expressions and unclear language productions. As excluded from the classification they were not taken into account in the analysis either.

With regard to fixed expressions, examples like (23a) or (23b) were not taken under consideration as they did not reveal productive language use.

- (23) a. come on please  
b. see?

Moreover, unclear instances produced by the children were only included in the analysis when they were clarified by the researchers in the transcription. For instance, examples of unclear productions like (24a) and (24b) were included in the study as the information added in square brackets by the transcribers was enough to clarify the child's production. On the contrary, if those problematic instances were not solved, not easy to understand, or appeared with the "xxx" symbol used to mark incomprehensible material, they were not useful for the analysis and thus, they had been discarded. This is the case of example (25).

- (24) a. *ya* [: you] *folf* [: wolf] or *bea(r)* ?  
b. *dey* [: they] *don('t)* get *sca(r)e*  
(25) *xxx* wanna play

## 5. Data Analysis & Discussion

In this section, the results obtained from the analysis of the data will be discussed. Furthermore, as the results serve to confirm or reject the different hypotheses stated in the methodology section, this section will follow the same organization as that in the research hypotheses. Consequently, the children's results related with the referentiality of sentential subjects would be firstly argued; next, the issue of subject overtness will be explained; and, finally, the children's results connected with grammatical correctness in English will be discussed.

### 5.1. Referentiality of Sentential Subjects

As for the referentiality of subjects in which children are supposed to produce more referential subjects than non-referential subjects (hypothesis 1), the relevant information appears in *Table 5a* and *5b*.

*Table 5a. Referential and Non-referential Subjects by stages and overall (Simon)*

Stage	Referential SUs	Non-referential SUs
1 (2;11)	85 (96.6%)	3 (3.4%)
2 (3;04)	11 (100%)	0
3 (3;10)	80 (94.1%)	5 (5.9%)
4 (4;04)	249 (98%)	5 (2%)
5 (4;10)	128 (97.7%)	3 (2.3%)
6 (5;04)	67 (100%)	0
7 (5;11)	140 (98%)	3 (2%)
<b>Total Simon</b>	<b>760 (97.6%)</b>	<b>19 (2.4%)</b>

**Table 5b. Referential and Non-referential Subjects by stages and overall (Leo)**

<b>Stage</b>	<b>Referential SUs</b>	<b>Non-referential SUs</b>
1 (2;11)	32 (91.4%)	3 (8.6%)
2 (3;04)	15 (93.7%)	1 (6.3%)
3 (3;10)	54 (98.1%)	1 (1.9%)
4 (4;04)	142 (98%)	3 (2%)
5 (4;10)	68 (97.1%)	2 (2.9%)
6 (5;04)	20 (95.2%)	1 (4.8%)
7 (5;11)	151 (97.4%)	4 (2.6%)
<b>Total Leo</b>	<b>482 (97%)</b>	<b>15 (3%)</b>

Table 5a and 5b illustrate the classification of subjects produced by both children on each of the 7 developmental stages and overall in terms of referentiality. As it can be observed, both children produce more referential subjects than non-referential subjects; in other words, Simon produces a total of 779 subjects where 760 (97.6%) of them are referential - that is, the subjects have a referent, as in (26); and where only 19 (2.4%) subjects are non-referential - those subjects do not have a referent, as in (27).

(26) *we* won't be able to see

(27) *there (i)s* (a) letter here

Leo follows the same pattern, as he produces a total of 497 subjects, 482 (97%) are referential subjects and 15 (3%) are non-referential subjects.

Additionally, it can be said that, overall, Simon produces more subjects (781) than Leo (497); but, Leo produces more non-referential subjects (3%) than

Simon (2.4%), although the difference in both cases is not very significant. Moreover, if the results are observed through each of the 7 developmental stages, the children also produce more referential subjects than non-referential ones on each of the stages; for example, on stage 4 for both children, Simon produces 250 (98%) referential subjects and 5 (2%) non-referential subjects; and Leo produces 142 (98%) referential subjects and 3 (2%) non-referential subjects. Thus, bearing in mind the results obtained, hypothesis 1 in which children were expected to produce more referential subjects than non-referential subjects has been confirmed.

Furthermore, the subject realization of Simon and Leo in terms of referentiality has been classified in concordance with the adult grammar as in *Table 6a* and *6b*.

**Table 6a. Referentiality in compliance with the adult grammar (Simon)**

Stage	Referential SUs		Non-referential SUs	
	Adult-like SUs	Non-adult like SUs	Adult-like SUs	Non-adult like SUs
1 (2;11)	66 (77.7%)	19 (22.3%)	3 (100%)	0
2 (3;04)	11 (100%)	0	0	1 (100%)
3 (3;10)	75 (93.8%)	5 (6.2%)	5 (100%)	0
4 (4;04)	233 (93.6%)	16 (6.4%)	5 (100%)	0
5 (4;10)	122 (95.3%)	6 (4.7%)	3 (100%)	0
6 (5;04)	66 (98.6%)	1 (1.4%)	0	0
7 (5;11)	137 (97.9%)	3 (2.1%)	2 (100%)	0
<b>Total Simon</b>	<b>710 (93.4%)</b>	<b>50 (6.6%)</b>	<b>18 (94.8%)</b>	<b>1 (5.2%)</b>

*Table 6b. Referentiality in compliance with the adult grammar (Leo)*

Stage	Referential SUs		Non-referential SUs	
	Adult-like SUs	Non-adult like SUs	Adult-like SUs	Non-adult like SUs
1 (2;11)	25 (78.1%)	7 (21.9%)	3 (100%)	0
2 (3;04)	15 (100%)	0	1 (100%)	0
3 (3;10)	49 (90.8%)	5 (9.2%)	1 (100%)	0
4 (4;04)	135 (95%)	7 (5%)	1 (33.3%)	2 (66.7%)
5 (4;10)	63 (92.7%)	5 (7.3%)	1 (50%)	1 (50%)
6 (5;04)	15 (75%)	5 (25%)	1 (100%)	0
7 (5;11)	142 (94%)	9 (5,9%)	4 (100%)	0
<b>Total Leo</b>	<b>444 (92.1%)</b>	<b>38 (7.9%)</b>	<b>12 (80%)</b>	<b>3 (20%)</b>

*Table 6a* and *6b* show the children's subject realization by stages and overall according to the adult grammar; that is, in this case, the data have been classified into 'adult-like' and 'non-adult like'. On the one hand, Simon produces a total of 779 referential subjects and 19 non-referential subjects, as previously explained; if the data are observed through each of the 7 developmental stages, it can be said that in both cases (referential and non-referential subjects) Simon produces more adult-like subjects than non-adult like ones; as an example, in stage 1 - which is the stage in which children are supposed to produce more incorrect grammatical instances - Simon produces more adult-like referential (77.7%) and non-referential subjects (100%) respectively than non-adult like ones - 22.3% in the case of referential subjects and 0% in the case of non-referential subjects. Furthermore, in an overall view, Simon produces more adult-like referential subjects (93.4%) and non-referential subjects (94.8%) than non-adult like ones (6.6% in the case of referential subjects and 5.2% in case of non-referential subjects).



On the other hand, Leo - who produces a total of 482 referential subjects and 15 non-referential subjects - also produces more adult like referential and non-referential subjects on each of the developmental stages; except on stage 5 where Leo produces a 50% of adult-like non-referential subjects and 50% of non-adult like non-referential subjects; nevertheless, in an overall view, Leo follows the same pattern as Simon since he produces more adult-like referential subjects (92.1%) and non-referential subjects (80%) than non-adult like ones.

In essence, bearing in mind the results obtained, it can be said that in terms of subjects' referentiality, both children have acquired the adult grammar since the early stages of acquisition.

## 5.2. Subject Overttness

As for the overttness of subjects in which children are supposed to produce more null subjects than overt subjects in English during their early stages of acquisition (hypothesis 2), the relevant results appear in *Table 7a* and *7b*.

*Table 7a. Overttness in English by Stages and Overall (Simon)*

Stage	Overt SUs	Null SUs
1 (2;11)	63 (71.6%)	25 (28.4%)
2 (3;04)	11 (91.7%)	1 (8.3%)
3 (3;10)	73 (85.9%)	12 (14.1%)
4 (4;04)	235 (92.5%)	19 (7.5%)
5 (4;10)	124 (94.7%)	7 (5.3%)
6 (5;04)	63 (94%)	4 (6%)
7 (5;11)	120 (84.5%)	22 (15.5%)
<b>Total Simon</b>	<b>689 (88.4%)</b>	<b>90 (11.6%)</b>

*Table 7b. Overtness in English by Stages and Overall (Leo)*

Stage	Overt Sus	Null SUs
1 (2;11)	27 (77.1%)	8 (22.9%)
2 (3;04)	16 (100%)	0
3 (3;10)	49 (89%)	6 (11%)
4 (4;04)	134 (92.4%)	11 (7.6%)
5 (4;10)	59 (84.3%)	11 (15.7%)
6 (5;04)	14 (66.7%)	7 (33.3%)
7 (5;11)	127 (82%)	28 (18%)
<b>Total Leo</b>	<b>426 (85.7%)</b>	<b>71 (14.3%)</b>

*Table 7a* and *7b* show the children's subject production in terms of overtness. In the case of Simon, out of the 779 subjects that he produces, 88.4% are overt subjects and 11.6% are null subjects. In addition, Leo produces a total of 497 subjects, from which 85.7% are overt subjects and 14.3% are null subjects. These results illustrate that, in an overall view, Simon's rate of overt subjects (88.4%) is higher than that of null subjects (11.6%); and, in the case of Leo, he also produces a higher rate of overt subjects (85.7%) than that of null subjects (14.3%). Moreover, when comparing the results of both children, it can be observed that Leo produces a higher rate of null subjects (14.3%) than Simon (11.6%); however, the difference between both children is not very significant. Additionally, if the results are observed on each of the 7 developmental stages, Simon and Leo produce more overt subjects than null subjects even in the early stages of acquisition; for example, in stage 1, which is the stage in which children are expected to produce more errors according to the adult grammar, Simon produces more overt subjects (77.1%) than null subjects (22.9%); and in the case of Leo, he also produces more overt subjects (77.1%) than null subjects (22.9%). Therefore, the rate of null subjects produced by Simon and Leo in each of the stages and overall is lower than the rate of overt subjects due to the fact that English is a [- null subject] language in which overt subjects are required

and null subjects are heavily restricted. Consequently, hypothesis 2, which states that the children would produce more null subjects than overt subjects during the early stages of acquisition due to crosslinguistic influence from the Spanish language, has not been confirmed. That is, these children do not show any sign of crosslinguistic influence in their English production in the case of sentential subjects.

Moreover, the children's null/overt subject production has also been classified according to its compliance with the adult grammar, as in *Table 8a* and *8b*.

*Table 8a. Overtness in compliance with the adult grammar (Simon)*

	<b>Overt SUs (Total S: 689)</b>		<b>Null SUs (Total S:90)</b>	
<b>Stage</b>	<b>Adult-like SUs</b>	<b>Non-adult like SUs</b>	<b>Adult-like SUs</b>	<b>Non-adult like SUs</b>
1 (2;11)	62 (98.4%)	1 (1.6%)	7 (28%)	18 (72%)
2 (3;04)	11 (100%)	0	0	1 (100%)
3 (3;10)	73 (100%)	0	7 (58.3%)	5 (41.7%)
4 (4;04)	235 (100%)	0	3 (15.8%)	16 (84.2%)
5 (4;10)	124 (100%)	0	1 (14.3%)	6 (85.7%)
6 (5;04)	63 (100%)	0	0	4 (100%)
7 (5;11)	120 (100%)	0	19 (86.4%)	3 (13.6%)
<b>Total Simon</b>	<b>688 (99.9%)</b>	<b>1 (0.1%)</b>	<b>37 (41.1%)</b>	<b>53 (58.9%)</b>

*Table 8b. Overtness in compliance with the adult grammar (Leo)*

	<b>Overt SUs (Total L: 426)</b>		<b>Null SUs (Total L: 71)</b>	
<b>Stage</b>	<b>Adult-like SUs</b>	<b>Non-adult like SUs</b>	<b>Adult-like SUs</b>	<b>Non-adult like SUs</b>
1 (2;11)	26 (96.3%)	1 (3.7%)	2 (25%)	6 (75%)
2 (3;04)	16 (100%)	0	0	0
3 (3;10)	49 (100%)	0	1 (16.7%)	5 (83.3%)
4 (4;04)	132 (98.5%)	2 (1.5%)	4 (36.4%)	7 (63.6%)
5 (4;10)	59 (100%)	0	5 (45.5%)	6 (54.5%)
6 (5;04)	14 (100%)	0	2 (28.6%)	5 (71.4%)
7(5;11)	127 (100%)	0	19 (67.9%)	9 (32.1%)
<b>Total Leo</b>	<b>423 (99.3%)</b>	<b>3 (0.7%)</b>	<b>33 (46.5%)</b>	<b>38 (53.5%)</b>

*Table 8a* and *8b* show the children's realization of overt and null subjects according to the adult grammar. In the case of Simon, out of the 779 subjects he produces, 11.6% are null subjects; and, out of the 90 null subjects he produces, 53 (58.9%) are non-adult like null subjects - which is the equivalent to 6,8% of the total number of subjects Simon produces. In case of Leo, out of the 497 subjects he produces, 14.3% are null subjects; and, out of the 71 null subjects he produces, 38 (53.5%) are non-adult like null subjects - which is the equivalent to 7.6% of the total number of subjects Leo produces. Hence, most of these incorrect null subject rates are produced during the process of the simultaneous acquisition of both languages (English and Spanish); thus, this production of non-adult like null subjects in English could be attributed to the contact with the Spanish language; in other words, there are some signs of crosslinguistic influence from Spanish into English on the realization of sentential subjects because null subjects are a grammatical option in Spanish. However, given that omission is a property of early grammars that affects subjects, verbal inflection,

copula verbs, among other categories, these low percentages of non-adult-like subject omission in English could also be the result of a developing grammar (as in Liceras and Fernández Fuertes 2018).

As it can be seen in example (28a) in the case of Simon and (28b) in the case of Leo, both children produce non-adult like subjects at some point of their acquisition process.

- (28) a. \**pro* now go to nursery school  
b. \**pro* don't know

However, in essence, the rate of null subjects in both children is lower than that of overt subjects; consequently, it can be inferred that, in an overall view, Simon and Leo have definitely acquired the English adult grammar during their acquisition process.

### 5.3. Grammatical Correctness

When dealing with grammaticality, two error types could be detected: omission cases (as discussed above in the case of non-adult-like null subjects) and commission cases. As for omission cases, example (28) illustrates this error type; in this case, the subject of the sentence has been omitted and, therefore, it is ungrammatical since English is a [- null subject] language in which subjects are required to be overt. On the contrary, a commission case is identified when the subject of the sentence is not the correct one (as discussed in the methodological section), as seen in (29a) and (29b) where the child has used an incorrect subject. In (29a) the child has used the non-nominative subject pronoun *me* instead of the nominative subject pronoun *I* and in (29b) the child has used the third person singular form of the verb to be (*is*) instead of using the pronoun *it*.

- (29) a. \**me* know  
b. \**is* goes like this

The children are expected to produce more commission cases than omission cases in English during their process of acquisition (hypothesis 3). The classification of results taking this into account appear in *Table 9a* and *9b*.

*Table 9a. Types of Ungrammaticality (Simon)*

<b>Stage</b>	<b>Omission cases</b>	<b>Commission cases</b>
1 (2;11)	18	1
2 (3;04)	1	0
3 (3;10)	5	0
4 (4;04)	16	0
5 (4;10)	6	0
6 (5;04)	4	0
7 (5;11)	3	0
<b>Total Simon</b>	<b>53 (6.8%)</b>	<b>1 (0.1%)</b>

*Table 9b. Types of Ungrammaticality (Leo)*

<b>Stage</b>	<b>Omission cases</b>	<b>Commission cases</b>
1 (2;11)	6	1
2 (3;04)	0	0
3 (3;10)	5	0
4 (4;04)	7	2
5 (4;10)	5	1
6 (5;04)	5	0
7 (5;11)	9	0
<b>Total Leo</b>	<b>37 (7.4%)</b>	<b>4 (0.8%)</b>

*Table 9a* and *9b* show the two types of ungrammaticality (omission cases and commission cases) produced by the children during the process of acquisition of the subject element. In the case of Simon, out of the 779 subjects he produces, 6.8% are omission cases (as in (28) above); and 0.1% are commission cases (the pronoun which functions as the subject is not correct as previously seen in (29). Additionally, out of the 497 Leo produces, 7.4% are omission cases and 0.8% are commission cases. These results suggest that Simon's omission rate (6.8%) is higher than that of commission (0.1%); moreover, results show that Leo follows the same pattern since he also produces more omission cases (7.4%) than commission cases (0.8%). However, it is also true that when comparing the results of both children, Leo produces more commission cases (0.8%) than Simon (0.1%); this could occur because Leo's overall subject production is lower (497) than that of Simon (779). All in all, the rate of omission cases produced by Simon and Leo in each of the 7 developmental stages and overall is higher than that of commission cases; consequently, hypothesis 3 in which children were expected to produce more commission cases than omission cases during their acquisition process has not been confirmed. This is could be linked, as in the case of the analysis in section 5.2. above, to omission being a defining property of early grammars.

## **6. Conclusion**

This section presents the final conclusions achieved in this paper. The present dissertation has presented a study not only on the analysis of the production of English sentential subjects on the part of simultaneous English/Spanish bilingual children, but also on the examination of the children's development on the realization of the subject element from the age of 2;11 to 5;11. Once the grammatical properties of subjects have been explained with a particular focus on *The Extended Projection Principle* and *The Null Subject Parameter*, this dissertation has illustrated some relevant studies on language acquisition used for establishing the basis of the present study. Furthermore, despite the fact that this dissertation might help to expand our knowledge on the acquisition of sentential subjects on the part of simultaneous bilingual children, there is still a need for further research. In addition, section 4 of this paper presents the

methodological process followed to perform the analysis; in this section, the objectives and research hypotheses have been presented. As aforementioned, the objectives of this dissertation dealt with the analysis of the subject production from simultaneous English/Spanish bilingual children and their development on the realization of the subject element from the age of 2;11 to 5;11. So as to carry out the analysis, the data were extracted from CHILDES and collected and classified into different variables in an excel database.

The analysis of data, which was explained in section 5 of this paper, showed the different variables used to test the children's realization of sentential subjects: referentiality, overtness and grammatical correctness. With regard to referentiality, both children were expected to produce more referential subjects than non-referential subjects as stated in hypothesis 1. The results obtained confirmed that hypothesis because Simon and Leo produced more referential subjects than non-referential subjects in each of the 7 developmental stages as well as overall. Also, the children's production of referential and non-referential subjects was assessed according to the adult grammar, and the results showed that neither of the two children produce high percentages of non-adult like subjects in terms of referentiality.

Moreover, as for the variable of overtness, the children were expected to produce more null subjects than overt subjects during the early stages of acquisition due to crosslinguistic influence of the Spanish language, which is a [+ null subject] language and their other L1. However, the results did not confirm this second hypothesis because both children produce more overt subjects than null subjects on each of the stages as well as overall. Hence, these children do not show any sign of crosslinguistic influence in their English production in the case of sentential subjects. This may happen because English is a [- null subject] language in which an overt subject is required. Nevertheless, the presence or absence of crosslinguistic influence on the children's linguistic production of sentential subjects would need further research since the data collected in this study are not enough to determine whether there is crosslinguistic influence or not due to the fact that English monolingual speakers have not been included in the analysis. As in the case of referentiality, the children's production of null subjects was



also tested according to the adult grammar and, once again, the results showed that both children produced more adult-like null subjects than non-adult like ones.

Finally, regarding the variable of grammatical correctness, the children were expected to produce more commission cases than omission cases as stated in hypothesis 3. The results did not confirm the hypothesis since the children produce more omission cases than commission cases on each of the 7 developmental stages as well as in an overall view of their production. These results are in line with some of the conclusions reached on previous studies: the children show some patterns in their subject production which could be understood as a consequence of crosslinguistic influence (Paradis & Navarro, 2003); and, the children seemed to have developed an early awareness on how subjects are used in English (Garau & Pérez-Vidal). However, the present study includes a small part of all the reachable topics that could be under analysis in this area; consequently, there are some grammatical instances which would need further research in order to provide more linguistic details connected with the realization of sentential subjects on the part of simultaneous bilingual children.

In essence, once the analysis of the subject production of simultaneous bilingual children has been done, it can be said that the children produced some misstatements during the process of acquisition. But both children showed an early development in terms of subject realization in English because they have mostly produced adult-like subjects since the beginning of their acquisition process.

## 7. Bibliography

Andrews, A. 1985 «The major functions of the noun phrase.» *Language typology and syntactic description*, 62-154.

Bloom, P. 1990 «Subjectless sentences in child language.» *Linguistic inquiry* 21 (4), 491-504.

Chomsky, N. 1965 *Aspects of the theory of syntax*. MIT.

Crystal, D. 2003 *A dictionary of linguistics and phonetics*. Blackwell.

Freidin, R. (ed.) 1991 *Principles and parameters in comparative grammar*. MIT.

Haegeman, L. and J. Guéron 2002 *English grammar: a generative perspective*. Blackwell.

Hyams, N., and K. Wexler 1993 «On the grammatical basis of null subjects in child language.» *Linguistic inquiry* 24, 421-459.

Huddleston, R. and G.K. Pullum 2002 *The Cambridge grammar of the English language*. CUP.

Jaeggli, O. and K. Safir 1989 *The null subject parameter*. Kluwer.

Juan-Garau, M. and C. Pérez-Vidal 2000 «Subject realization in the development of a bilingual child. » *Bilingualism: language and cognition* 3 (3), 173-92.

Liceras, Juana M., and R. Fernández Fuertes 2016 «Subject omission/production in child bilingual English and child bilingual Spanish: The view from linguistic theory.» *Probus*. <https://uvadoc.uva.es/handle/10324/32985>

MacWhinney, B. 2000. *The CHILDES project: Tools for analyzing talk*. 3rd ed. Lawrence Erlbaum.

Paradis, J. and S. Navarro 2003 «Subject realization and crosslinguistic interference in the bilingual acquisition of Spanish and English: what is the role of the input? » *Journal of child language* 30, 1-23.

Serratrice, L., et al. 2004 «Crosslinguistic influence at the syntax–pragmatics interface: subjects and objects in English–Italian bilingual and monolingual acquisition.» *Bilingualism: language and cognition* 7 (3), 183–205.