RESLA
Revista Española de Lingüística Aplicada

https://matrix.aesla.org.es/RESLA
ISSN 0213-2028 | e-ISSN 2254-6774

Crosslinguistic influence and language similarity: the oral production of sentential subjects in child L2 English

Sonja Mujcinovic * & Raquel Fernández Fuertes

Universidad de Valladolid, Spain

Abstract: This study seeks to examine crosslinguistic influence in the L2 English of different child participant groups (L1 Spanish, L1 Bosnian and L1 Danish) when compared to L1 English controls. The focus is placed on the L2 English sentential subjects orally produced by these children in a semi-guided individual interview. The subjects were classified following three criteria: type (overt/null), grammaticality (grammatical/ungrammatical) and adequacy (adequate/non-adequate). Two formal proposals are put to the test taking into account the difference between [+/-null subject] languages: typological similarity between the L1 and the L2 (in the case of the L1 Danish group) and lexical transparency where the properties of the superset L1 influences those of the subset L2 and thus, facilitates L2 acquisition (in the case of the L1 Spanish and L1 Bosnian groups). Our results show that typological similarity can to a certain point explain the differences between the [+null subject] and the [-null subject] language groups, but it does not justify the high production of both grammatical and adequate subjects. These high rates can be interpreted in terms of lexical specialization in that there is facilitative crosslinguistic influence effects from the superset language (i.e., Spanish and Bosnian) to the subset language (i.e., English).

Keywords: *language similarity, sentential subjects, crosslinguistic influence, L2 English, lexical specialization.*

Influencia interlingüística y similitud lingüística: la producción oral de sujetos oracionales de niños bilingües de inglés como L2

Resumen: Este estudio examina la influencia interlingüística en el inglés como L2 de diferentes grupos de niños participantes (español L1, bosnio L1 y danés L1) en comparación con el grupo de control de inglés L1. La atención se centra en los sujetos oracionales en el inglés L2 producidos oralmente por estos niños en una entrevista individual semi-guiada. Los sujetos se clasificaron siguiendo tres criterios: tipo (explícito/nulo), gramaticalidad (gramatical/agramatical) y adecuación (adecuado/no-adecuado). Se ponen a prueba dos propuestas formales teniendo en cuenta la diferencia entre las lenguas [+/sujeto nulo]: la similitud tipológica entre la L1 y la L2 (para el grupo danés L1) y la especialización léxica en la que las propiedades de la L1 *superset* influyen en las de la L2 *subset* y, por tanto, facilitan la adquisición de la L2 (para los grupos español L1 y bosnio L1). Los resultados muestran que la similitud tipológica puede explicar hasta cierto punto las diferencias entre los grupos lingüísticos [+sujeto nulo] y [-sujeto nulo], pero no justifica la elevada producción de sujetos gramaticales y adecuados, algo que puede interpretarse en términos de especialización léxica: existe influencia interlingüística facilitadora desde la lengua *superset* (el español y el bosnio) hacia la lengua *subset* (el inglés).

Palabras clave: *similitud lingüística, sujetos oracionales, influencia interlingüística, L2 inglés, especialización léxica.*

How to cite: Mujcinovic, S., & Fernández Fuertes, R. (2026). Crosslinguistic influence and language similarity: the oral production of sentential subjects in child L2 English. *Revista Española de Lingüística Aplicada, 39*(2), 68-94. https://doi.org/10.58859/resla.1544

**Corresponding author:* *sonja.mujcinovic@uva.es*

AESLA
Asociación Española de Lingüística Aplicada

Available under the
CC BY 4.0 license. ©

1. Introduction

Within the framework of generative grammar, sentential subjects have been extensively discussed from the point of view of formal linguistics as well as in acquisition studies. In the case of the former, the null subject parameter was proposed and used to address the differences that emerge across languages in terms of the availability of null subjects (e.g., Perlmutter, 1971; Chomsky & Lasnik, 1977; Chomsky, 1981; Rizzi, 1982; Phinney, 1987; Platzack, 1987; Liceras, 1988, 1989; Jaeggli & Safir, 1989; Belletti, 2001, 2004; Sheehan, 2006; Camacho, 2006, 2013, 2016; Holmberg, 2010).

In the case of the latter, analyses have been conducted on monolingual acquisition (e.g., Bloom, 1990; Hyams & Wexler, 1993; Hyams, 2011; Hoff, 2017; Orfitelli & Hyams, 2008, 2012), and on bilingual acquisition, including both simultaneous bilinguals (e.g., Montrul, 2004; Liceras et al., 2008; Cuza, 2013; Cuza & Camacho, 2017; Liceras & Fernández Fuertes, 2019) and sequential bilinguals (e.g., Lozano, 2002; Park, 2004; Montrul & Rodríguez-Louro, 2006; Rothman, 2008, 2009; Montrul et al., 2009; Pladevall-Ballester, 2012, 2016; Cuza et al., 2013; Quesada, 2014; Mitkovska & Bužarovska, 2018; Pladevall-Ballester et al., 2022). When considering bilinguals, most of these studies mainly compare bilinguals whose two languages exhibit a different configuration with respect to the null subject parameter. Much less frequently have previous studies dealt with the analysis of bilinguals whose two languages share the same value of the null subject parameter. Few studies have been conducted on the contact between two [+null subject] languages (e.g., Bini, 1993; Margaza & Bel, 2006; Sorace & Filiaci, 2006; Bel et al., 2016; Lozano, 2018) and even fewer on that of two [-null subject] languages (e.g., White, 1985; Liceras, 1989; Liceras & Alba de la Fuente, 2015; Mujcinovic, 2015, 2020). In all cases the focus of attention has been placed in analyzing potential crosslinguistic influence effects between the two languages of the bilingual.

The present study takes as a point of departure previous formal and empirical works on the null subject parameter to offer a characterization of the English as a second language (L2) of three groups of child participants that differ in their first language (L1): Spanish, Bosnian and Danish. While previous works on L2 acquisition have mainly addressed adult grammars, we focus on the oral semi-spontaneous production of L2 English children which we analyze in terms of grammaticality and pragmatic adequacy in the case of three subject forms (i.e., full DPs, overt personal pronominal subjects, and null subjects; see examples 1-3 and discussion below). Furthermore, the language combination that we include allows us to offer a comparison between languages that have been more studied (i.e., Spanish) and others that have been less so (i.e., Bosnian and Danish) and to address the role played by language similarity. Given that L1 crosslinguistic influence characterizes and shapes (at least) the initial stages of L2 acquisition (e.g., Hawkins 2001, Herschensohn, 2000; Montrul, 2016), attention has been placed on how and in which respect the L1 and the L2 differ. Being crosslinguistic influence an overt manifestation of the L1 (Gass, 1996, p. 385), the similarities between the languages in contact have some bearing on the type of crosslinguistic influence that might appear. In the case of typological similarity,

the focus of this investigation, it has often been put together with typological proximity. However, these are two different concepts and, while typological proximity has centered the attention of L2 acquisition research, typological similarity has received much less attention (e.g., Rothman, 2010, 2011; Rothman & Cabrelli Amaro, 2010; Montrul et al., 2011; Liceras & Alba de la Fuente, 2015; Westergaard et al., 2017; Cuza et al., 2018). Typological proximity applies to languages that belong to the same family and share the same origin; an example is Spanish and French, both derived from Latin and considered Romance languages. However, typological similarity refers to languages that share the same option at a micro-parametric level so that "a typological or formal universal is equally realized in these two typologically close languages" (Liceras & Alba de la Fuente, 2015, p. 333). In other words, Spanish and Bosnian are not typologically proximate as they do not belong to the same family since Spanish is a Romance language while Bosnian is a Slavic language. However, they are typologically similar languages with respect to subject realization as both are [+null subject] languages. Danish and English are both typologically proximate (i.e., Germanic languages) and typologically similar (i.e., [-null subject] languages). In the same vein, Spanish and Bosnian are neither typologically similar nor typologically proximate to English.

Therefore, the analysis of the contact between English, on the one hand, and Danish, Spanish, or Bosnian, on the other, allows us to explore more in depth how language typological similarity works in child L2 acquisition.

This paper is organized as follows. In section 2, we offer a formal description of sentential subjects with a focus on the four languages under consideration in this study (i.e., English as the target language and as the L2 of our participants and Danish, Spanish and Bosnian as their L1s). In section 3, we provide an overview of different acquisition works that have analyzed sentential subjects and the crosslinguistic influence effects that have been reported in this case. In section 4, we present the research questions that we have formulated. Section 5 provides information about the participants that took part in this study, the task used to elicit the data and the data collection and classification procedures. We present the data analyses and the discussion in section 6 and finish with the conclusions that we have obtained and that appear in section 7.

2. On the nature of sentential subjects: typological similarities and differences

The availability of null subjects across languages divides languages into two groups: [+null subject] languages and [-null subject] languages (e.g., Perlmutter, 1971; Chomsky & Lasnik, 1977; Chomsky, 1981; Rizzi, 1982; Phinney, 1987; Platzack, 1987; Liceras, 1988, 1989; Jaeggli & Safir, 1989; Belletti, 2001, 2004; Sheehan, 2006; Camacho, 2006, 2013, 2016; Holmberg, 2010). The examples below illustrate these two groups by referring to three possible subject forms: full Determiner Phrases (full DPs) in 1, overt personal pronouns in 2 and null subject pronouns in 3.

(1) full DP subject

a.	*the man*	ate	two bananas	[English]
b.	*manden*	spiste	to bananer	[Danish]
	"man-the	ate	two bananas"	
c.	*el hombre*	comió	dos bananas	[Spanish]
	"the man	ate	two bananas	
d.	*čovjek*	je[1] pojeo	dvije banana	[Bosnian]
	"man	was ate	two bananas"	

(2) overt personal pronominal subject

a.	*he*	ate	two bananas	[English]
b.	*han*	spiste	to bananer	[Danish]
	"he	ate	two bananas"	
c.	*él*	comió	dos bananas	[Spanish]
	"he	ate	two bananas"	
d.	*on*	je pojeo	dvije banane	[Bosnian]
	"he	was ate	two bananas"	

(3) null subject

a.	**pro*	ate	two bananas	[English]
	"(he)	ate	two bananas"	
b.	**pro*	*spiste*	*to bananer*	[Danish]
	"(he)	ate	two bananas"	
c.	*pro*	comió	dos bananas	[Spanish]
	"(he)	ate	two bananas"	
d.	*pro*	pojeo je	dvije banane	[Bosnian]
	"(he)	ate was	two bananas"	

All four languages allow for full DPs (1) and overt pronouns (2) to occupy subject position. They differ, however, in the possibility of having null subjects (3). In this respect, the fact that both Spanish and Bosnian allow their subjects to be null (3c and 3d) makes them

[1] *je* is the 3rd person singular auxiliary of the verb "to be."

[+null subject] languages. On the contrary, since English and Danish only permit overt subjects (1a and b and 2a and 2b) and do not generally allow their subjects to be null (3a and 3b), they are classified as [-null subject] languages.

Different theories have been put forward to explain the availability of null subjects in a language. Derived from the Principles and Parameters Theory within Universal Grammar, the Extended Projection Principle (EPP) dictates that all sentences must have a syntactic subject (Chomsky, 1981). In the case of [+null subject] languages (i.e., Spanish and Bosnian), since these languages have a rich verbal agreement morphology that licenses and identifies null subjects, the EPP is satisfied (Rizzi, 1986). In the case of [-null subject] languages (i.e., English and Danish), the EPP is satisfied via overt pronouns, because these languages have a poor verbal agreement morphology that lacks pronominal features and, therefore, cannot satisfy the EPP (Camacho, 2013).

In the last decades, reformulations of the null subject parameter in terms of phonological overtness have been proposed (e.g., Holmberg, 2005, 2010; Sheehan, 2006; Martínez-Sanz, 2011 and Liceras & Fernández Fuertes, 2019). Overt pronouns in both [+null subject] languages and [-null subject] languages are said to occupy the same position (i.e., SpecIP) and are, therefore, equally interpreted and have the same value. Since overt pronouns are spelled out, they have a phonological form. Null pronouns, on the contrary, are not spelled out (i.e., they have no phonological form). That is, [-null subject] languages have one subject type available (i.e., overt), whereas [+null subject] languages have two subject types (i.e., overt and null). This makes the [+null subject] the superset option and the [-null subject] the subset one (e.g., White, 1989; Ayoun, 2003; Holmberg, 2005; Sheehan, 2006; and Liceras & Fernández Fuertes, 2019). A [+null subject] language is considered the superset language when compared to a [-null subject] language, because [+null subject] languages have two possible realizations of the subject (i.e., the phonologically realized option and the phonologically null option) whereas [-null subject] languages have one (i.e., the phonologically realized option) (Holmberg, 2005; Sheehan, 2006).

3. On the acquisition of sentential subjects

In the case of L1 acquisition, different studies have shown that all children go through a stage where they omit subjects regardless of whether the child's L1 is a [+null subject] or a [-null subject] language. In the case of the acquisition of [+null subject] languages, children start producing both overt and null subjects at a very early stage, both subject types generally conforming to adult grammars (e.g., Bloom, 1990; Grinstead, 1998; Hyams & Wexler, 1993; Juan-Garau & Pérez-Vidal, 2000; Liceras et al., 2006, 2008; Montrul, 2004; Serratrice, 2002, 2005; Paradis & Navarro, 2003; Valian, 1991; Valian & Eisenberg, 1996). Whereas, for children acquiring [-null subject] languages, these initial null subjects do not conform to adult grammar requirements and gradually disappear.

In the case of bilingual grammars in general and of L2 bilingual grammars in particular, the main issue is whether differences across the two languages of the bilingual (i.e., the L1 and the L2) may lead to crosslinguistic influence effects (Cook, 2016).

Crosslinguistic influence appears when L2 bilinguals seek to find an equivalence between the patterns in their L1 and those in their L2. Depending on the degree of similarity between the languages in contact, two types of crosslinguistic influence can be identified: facilitative (or positive) and interfering (or non-facilitative or negative). Facilitative crosslinguistic influence is said to occur when equivalent structures are available and have the same distribution in both languages resulting in a successful acquisition of L2 properties. Interfering crosslinguistic influence is the result of overlapping structures when these are wrongly perceived as formally and functionally similar in the L1 and the L2, resulting in ungrammatical or inadequate structures. Crosslinguistic influence has been said to be shaped by two main issues: the degree of L2 proficiency that the speakers have and the amount of similarity the L2 bilinguals can identify between their two languages (Cenoz, 2001, 2003 and Gass & Selinker, 2008). Thus, the more proficient the L2 speakers are the less interference appears (Gathercole, 2016:123). Therefore, it is plausible to believe that crosslinguistic influence is more likely to occur in the beginning and in the intermediate stages of L2 acquisition and is expected to diminish but not necessarily disappear in more advanced stages. This is said to be true for the syntactic or morphological domain, but previous studies have shown that the syntax-pragmatics interface continues to be vulnerable even in very proficient stages (e.g., Park, 2004; Tsimpli & Sorace, 2006; Rothman, 2009; Slabakova & Ivanov, 2011; Pladevall-Ballester, 2012, 2016; Lozano, 2018; Mitkovska & Bužarovska, 2018).

In the case of L2 acquisition, numerous studies have analyzed how differences between the L1 and the L2 shape the participants' L2 production (for example, subject omission and overproduction of overt subjects), based on typological proximity and similarity and crosslinguistic influence.

Under the Facilitation Hypothesis (Gundel & Tarone, 1992), when languages are typologically similar, the L1 can facilitate the acquisition of the L2. In the same line, if linguistic differences are found between the L1 and the L2, a potential effect of crosslinguistic influence would be a lower L2 learnability given the interference effect of the L1 in the learning of the L2 (Schepens et al., 2016). Therefore, relatedness between the L1 and the L2 seems to play a crucial role so that, if languages in contact are similar, acquisition would be facilitated, while if languages in contact are different or distant, this difference or distance will interfere with acquisition. For instance, in the case of subjects, when the speakers' L1 is a [-null subject] language and the L2 a [+null subject] language, different studies have reported an overproduction of overt subjects in the data of speakers with different proficiency levels (e.g., Rothman, 2008, 2009; Lozano, 2002; Montrul & Rodríguez-Louro, 2006; Montrul et al., 2009; Quesada, 2014). However, studies conducted on speakers of a [+null subject] L1 learning a [-null subject] L2 have reported that a non-native-like rate of subject omission is not that frequent in the L2 (e.g., Park, 2004; Pladevall-Ballester, 2012, 2016; Mitkovska & Bužarovska, 2018). These different patterns may suggest that, overall, learning the superset language is more costly than learning the subset.

In the case of two [-null subject] languages in contact, the scarce amount of research conducted to date shows that there is a facilitation effect in that results obtained are native-like (White, 1985; Liceras, 1989; Liceras & Alba de la Fuente, 2015 and Mujcinovic, 2015, 2020). It seems felicitous to assume that typological similarity seems to play a crucial role.

In the case of two [+null subject] languages in contact, the amount of research is also scarce and most of the studies deal with adult speakers (e.g., Bini, 1993; Margaza & Bel, 2006; Sorace & Filiaci, 2006; Bel et al., 2016; Lozano, 2018). The results from these studies show an overproduction of overt subjects which is interpreted as related to interface vulnerability and not to typological proximity or typological similarity.

In the case of 2L1 acquisition research (i.e., simultaneous bilingualism), a proposal has been put forward to account for the presence as well as for the directionality and effect of crosslinguistic influence that we would like to adapt to L2 contexts: the lexical specialization approach (Liceras & Fernández Fuertes, 2019). In language contact situations between a superset and a subset language, these authors argue that crosslinguistic influence will go from the superset language to the subset language (and not the other way around) and it will have a facilitation effect. Liceras & Fernández Fuertes (2019) analyzed the sentential subjects that appear in the spontaneous production of English-Spanish simultaneous bilingual children. In this case and given that Spanish is the superset language, no interference from English into Spanish would occur and, in fact, no overproduction of overt pronouns appears in the Spanish data. Rather, the superset language is the one causing crosslinguistic influence (i.e., interference from Spanish into English) and with a specific effect: acceleration in the development of the overt subject requirement in 2L1 bilingual English when compared to L1 monolingual English. These authors explain these results in terms of lexical specialization as follows: the availability of two subject types in Spanish (i.e., null and overt pronouns) makes this language lexically specialized when it comes to person features. These are always reflected in the verbal agreement markers but can also appear realized as overt pronouns (e.g., *(yo) com-**o***, 'I eat'). This type of lexical specialization does not appear in the case of English as overt pronouns are the only available option and do not alternate with null subject pronouns. As a result, Spanish is the superset language because it incorporates both subject types, whereas English is the subset language because it requires its subjects to be overt.

4. Research questions

The aim of our study is to account for possible crosslinguistic effects in the oral production of sentential subjects by child L2 bilinguals. To do so and to analyze the data produced, two different approaches are considered: typological similarity and lexical specialization. Thus, to provide a more fine-grained approach to crosslinguistic influence, the following two research questions were formulated:

1. What role, if any, does typological similarity play in the oral production of sentential subjects by L2 English bilingual children?

CC BY

2. What role, if any, does lexical specialization play in the oral production of sentential subjects by L2 English bilingual children?

Previous studies point towards the saliency of language typology. For instance, Liceras & Alba de la Fuente (2015) claim that the closer the languages in contact are typologically speaking, the lower rates of ungrammaticality (i.e., interfering crosslinguistic influence) are expected. If typological similarity plays a role, then the results from our study should show that the L1 Danish group outperforms the L1 Spanish and L1 Bosnian groups, because of the facilitative crosslinguistic influence effects that derive from both English (L2) and Danish (L1) being typologically similar. If languages in contact are not typologically similar and differences are found between the L1 (Spanish and Bosnian, in this case) and the L2 (English), lower L2 learnability can arise (i.e., more crosslinguistic influence which can delay the learning of the L2) (Schepens et al., 2016). If this is so, we expect to find interfering crosslinguistic influence in both the L1 Spanish and the L1 Bosnian groups. In other words, the less typologically similar the languages are, the more crosslinguistic influence is expected. This would be reflected in a higher rate of ungrammatical subjects.

On a different note, if Liceras & Fernández Fuertes's (2019) proposal on lexical specialization is adapted to the L2 acquisition of sentential subjects, the following prediction could be made in the case of L2 English in contact with L1 Spanish, L1 Bosnian and L1 Danish: crosslinguistic influence is expected from the L1 if the L1 is a superset language (i.e., Spanish and Bosnian) into the L2 (i.e., English) and the effect should not be interfering, because the L2 bilinguals are expected to rely on one of the options available in Spanish or Bosnian (i.e., the overt subject) as a reinforcement for the only available option in their L2 English. Thus, these bilinguals' output is expected to be mostly grammatical and no crosslinguistic influence from the [+null subject] L1 into L2 English would take place (e.g., Park, 2004; Pladevall-Ballester, 2012, 2016 and Pladevall-Ballester et al., 2022). In the case of the L1 Danish group and given that both Danish and English are subset languages, no such learnability rule could be established, although a low error rate is to be expected if language similarity has an effect.

5. Methodology

The data that we analyze were extracted from the soraUVALAL corpus (Mujcinovic & Fernández Fuertes 2021) available in the SLABank of the TalkBank system (MacWhinney, 2019). General information about the participants and tasks is described below.

5.1 The participants

A total of 91 L2 bilingual children participated in the study: 26 L1 Spanish children from Spain, 26 L1 Bosnian children from Bosnia & Hercegovina, 26 L1 Danish children from Denmark and 13 L1 English children from Canada as controls (mean age=11; SD age = 0.71). All the L2 participants learned English in an institutional context where on average they received L2 English input for two sessions per week over a period of 3 years.

The parents' of the participants completed a Language Background Questionnaires (LBQ) in their L1 where the children's language history (e.g., how and when they have acquired their L1, whether languages other than the L1 are or were part of their language background), self-reported language proficiency (i.e., how good they believe their command of each of the languages is), and language attitudes (e.g., whether they like learning English) were detailed. The LBQ functions as a subjective measure of linguistic behavior that assesses linguistic experience and status.

In order to take part in the study, the participants had to meet the following inclusion criteria: (i) both the parents and the child had to share the same L1 (i.e., Spanish, Bosnian or Danish); (ii) English had to be the participants' only L2 and, in this respect, the participants that were learning an L3 were excluded; (iv) instruction could only take place in institutional settings in the school that they were attending so that participants who received extracurricular lessons in English were excluded; (v) participants who had spent longer periods of time abroad were also excluded; and (vi) only the participants that had been exposed to L2 English for a period of 2 to 4 years at their primary school were included in this study. These criteria ensure that all the children in the three experimental groups have a rather similar linguistic profile.

The L2 English participants' age ranged from 9 to 12. The difference in age is due to the time when the participants were first exposed to L2 English. In the case of the L1 Spanish participants, who were the youngest, they were first exposed to English at the age of 7, whereas for the L1 Danish and the L1 Bosnian participants the first exposure to English took place two years later (age 9). To compensate for the age gap, the time of exposure was controlled for (2 to 4 years; mean: 3 years). The Spanish participants attended grades 3 to 5 and the Bosnian and Danish participants attended grades 5 to 7.[2]

Regarding the amount of L2 English input received in the institutional setting, all the participants have received on average 2 lessons per week. In the case of the Spanish participants, since they were taking part in a CLIL (Content and Language Integrated Learning) program, they were exposed to more input because two other subjects (Natural and Social Sciences and Arts) were taught in English. More exposure can lead to better results. Nonetheless, even though these participants have received more L2 English input, their overall production is similar to that of the other L2 English groups, as table 1 below shows.

Implementing the above-mentioned criteria resulted in rather homogenous participant groups across the three countries.

In the case of the control group, just the participants who have only been exposed to English at home were included in this study. These participants also belonged to a monolingual English education system and their age ranged from 10 to 11 years to match the age of the L2 participants (9-12).

[2] The information relative to the children's academic profile comes from the information sheet we gathered from each school regarding the general grade system as well as the L2 English instruction curricula that is followed in each school.

Proficiency could not be measured via proficiency tests due to restrictions imposed by schools. We have used the mean length of utterance value measured in words (MLUw) as a production measure.[3] The different MLUw values per participant group appear in Table 1.

Table 1. Descriptive statistics for English MLUw values per participant group.

Group	Spanish	Bosnian	Danish	English
words	10,539	10,676	18,086	5,691
utterances4	2,504	3,515	2,836	1,010
MEAN	4.189	3.468	5.109	5.689
SD	1.079	0.905	1.658	0.950

The MLUw values were calculated in terms of the number of words over utterances. The information in Table 1 shows that there is a difference in terms of word production when comparing across the experimental groups with the L1 Danish participants producing more words. However, this is balanced out when we take into account the number of utterances produced by each L2 English group. When we consider the control group, the amount of production is the lowest. This is not surprising given that it is the group that contains less participants (n=13) when compared to the experimental groups (n=26). The mean of words over utterances (i.e., the MLUw) shows that participants produce utterances that, on average, contain between 3 and 5 words. We use these MLUw values as a production measure and have performed the statistical analyses that are reported next.

To determine whether there are any significant differences between the groups, a one-way ANOVA was conducted after confirming homogeneity of variance of the data. A significant effect of language group was confirmed with a large effect size ($F_{(3,87)}$=13.05, p<.001, η^2=0.31). The results show that there is significant difference between the L1 Bosnian (p<.001) and L1 Spanish (p<.001) groups when compared to the L1 English group. The same difference is found between L1 Bosnian (p<.001) and L1 Spanish (p=.038) when compared to the L1 Danish group. But no such significant difference is found between the L1 Danish and L1 English group (p=.501) nor between the L1 Spanish and L1 Bosnian (p=.151). These results show that the [+null subject] language groups group alike and that the [-null subject] language groups group alike; and that, in turn, there is a difference between participants with a [+null subject] L1 and participants with a [-null subject] L1.

[3] The MLU has been used as an indicator of L1 children's language complexity and language development (Brown, 1973). L2 studies have also resorted to the MLU as an indicator of relative proficiency (e.g., Håkansson, 2001; Unsworth, 2008; Unsworth & Blom, 2010; Hawkins & Filipovic, 2012; Lundell & Lindqvist, 2012, 2014; Faitaki & Murphy, 2023).

[4] In the transcription system that is used in CHILDES, CHAT (Codes for the Human Analysis of Transcripts), an utterance is a speech turn which does not necessarily contain a clause. Thus, an utterance could be an interjection, a DP, or a clause (MacWhinney, 2000).

5.2 The oral experimental task

The oral data obtained for this study were elicited via a semi-guided oral interview. Following a protocol designed to ensure uniformity in data collection and naturalistic speech, the participants were interviewed individually, and the interviews were audio-recorded. The topics treated in the interview related to family, hobbies, school, music, friends, etc. The participants were always allowed to ask for vocabulary which was provided to them in a non-inflected form. The interviews were conducted in a small room at the schools which the participants attended.

Prior to data collection, warming-up sessions were conducted (e.g., Hedge, 2000; Ellis & Yuan, 2004; Kuiken & Vedder, 2011; Ellis, 2014). The purpose of these sessions was to familiarize the participants with the interviewer and to make sure that they understood and were comfortable with the task. These sessions were presented as a class activity, so that the participants would not feel like they were being examined and would perform the task in a more natural and relaxed environment. They were reminded that they were not going to be graded for part-taking and that the activity was not part of the evaluations that their own school followed.

5.3 Coding procedure

To analyze the data obtained, sentential subjects were extracted both from matrix and embedded clauses. These subjects were then classified in terms of form, grammaticality (i.e., syntax) and adequacy (i.e., syntax-pragmatics).

In terms of form, five categories were distinguished: full DPs and proper names (as in 4), overt personal pronouns (as in 5), grammatical null pronouns (as in 6), and ungrammatical null pronouns (as in 7).

(4) *CHI: *my mother* is a teacher [full DPs]
(L1 Spanish group)

(5) *CHI: I have a hamster [overt personal pronouns]
(L1 Bosnian group)

(6) *CHI: you *pla*y online and *pro* try to win [grammatical null pronouns]
(L1 Danish group)

(7) *CHI: bec*ause pro* is working [ungrammatical null pronouns]
(L1 Bosnian group)

All subjects were classified in terms of their being grammatical or ungrammatical. The subjects were classified as ungrammatical if a null subject was used in any other than coordinated structures. If the ungrammaticality was attributed to issues that were not related to the subject and the subject was correct, the instance was considered as grammatical (for example: incorrect case assignment, as in 8, or incorrect object omission, as in 9).

(8) *CHI: *him* have a job [incorrect case assignment]
(L1 Bosnian group)

(9) *CHI: yes, I love Ø very much [incorrect object omission]
(L1 Bosnian group)

In terms of adequacy, each subject form was classified in relation to its referent. Full DPs are adequate if used for referent introduction, referent reintroduction and disambiguation. Pronouns are adequate if used for reference maintenance and null pronouns are adequate if used for reference maintenance in coordinated structures. Any other uses than the ones indicated were considered non-adequate (e.g., DPs used for reference maintenance (RM), as in 10, pronouns used for reference introduction (RI), as in 11, or pronouns used for reference reintroduction (RRI), as in 12).

(10) DPs used for reference maintenance

Previous context:

*CHI: and then after football my mum come

Example:

*CHI: and &eh my mum drive me to taekwondo (L1 Danish group)

(11) pronoun used for reference introduction

Previous context: *there is no previous context as this is the first sentence produced by the participant*

Example:

*CHI: and then after football *he* come (L1 Bosnian group)

(12) pronoun used for reference reintroduction

Previous context:

*CHI: the last we have done a bit of Guy Forks when he was a catholic terrorist.

*CHI: they tried to explode the whole the House of Parliament.

*CHI: while the king was inside.

*CHI: he meant that the Bibel@s:dan should be read in English and not only Latin.

Example:

*CHI: so *he* decided to try to kill him (L1 Danish group)

The following were excluded from the study as it was not possible to analyze the subject present in the sentences or it was unclear how it could be classified for the purposes of our study: incomplete sentences, codeswitching involving subjects, *wh*-pronouns in subject position and fixed expressions. Expletive subjects and imperatives were also excluded as they are non-referential, in the first case, and constitute an instance of overlap in the four languages, in the second case.

6. Results and discussion

The sentential subjects produced by the experimental groups (i.e., L1 Spanish, L1 Danish and L1 Bosnian) and by the control group (i.e., L1 English) are presented next. The overall amount and distribution of subjects produced by all the participant groups appears in Table 2 and will be discussed in the subsequent paragraphs.

Table 2. Overall distribution of subjects produced.

	grammatical				
L1	**adequate**	**non-adequate**	**total**	**ungrammatical**	**Total**
Spanish	83.79% [853]	1.96% [20]	85.76% [873]	14.24% [145]	100% [1018]
Bosnian	91.18% [982]	3.80% [41]	94.98% [1023]	5.02% [54]	100% [1077]
Danish	96.23% [1838]	1.05% [20]	97.28% [1858]	2.72% [52]	100% [1910]
English	99.01% [663]	0% [0]	99.01% [663]	0.09% [6]	100% [669]
total	92.77% [4336]	1.73% [81]	94.50% [4417]	5.50% [257]	100% [4674]

To run the statistical analyses, two General Linear Mixed Models (GLMMs) were fitted: one that focuses on the grammaticality of the subjects produced (i.e., grammatical vs. ungrammatical subjects) and one that is concerned with the adequacy of the grammatical subjects (i.e., adequate vs. non-adequate subjects). A negative binomial distribution model was created to fit the data, due to overdispersion in the data (dispersion ratio=1.306, p<.005). With DHARMa package overdispersion, residuals distribution and outlier tests were checked. The optimal model included language (Spanish, Danish, Bosnian and English), type of subject (grammatical vs ungrammatical and adequate *vs*. non-adequate) and the interaction as fixed effects parameters. The random effect structure includes random intercept by participant (R^2 = 0.677). The model includes the log of total number of subjects produced by each participant as offset term which allows to use negative binomial distribution to model the rate of occurrences based on the total number of subjects. A model including the MLUw as a variable was also run but no main effect was found (F(1,3)=.054, p=.941). The AIC score is lower in the model run with the MLUw (R^2 = 0.425) when compared to the model run without the MLUw (R^2 = 0.633). Therefore, we report below the analysis corresponding to the latter model.

As for the grammaticality of the sentential subjects produced, the results in terms of subject type (full DPs, overt pronouns, grammatical null subjects and ungrammatical null subjects) are summarized in Table 3.

Table 3. Distribution of subjects per language group: subject type.

L1		full DPs	overt pronouns	gram null	ungram null	total
[+null subject] languages	Spanish	17.88% [182]	66.50% [677]	1.38% [14]	14.24% [145]	100% [1018]
	Bosnian	12.81% [138]	80.78% [870]	1.39% [15]	5.01% [54]	100% [1077]
[-null subject] languages	Danish	11.57% [221]	81.26% [1552]	4.45% [85]	2.72% [52]	100% [1910]
	English [control]	14.50% [97]	75.64% [506]	8.97% [60]	0.90% [6]	100% [669]

To conduct the statistical analysis in terms of grammaticality, Table 4 shows the sets of pairwise comparisons with Tukey adjustments using the emmeans package that were conducted. The model includes the data set of grammatical and ungrammatical subjects produced with two predictors: participants' L1 and grammaticality (i.e., grammatical *vs*. ungrammatical subject). The results in Table 4 show the comparisons obtained, which are discussed below.

Table 4. Summary of the GLMM fixed effects for grammaticality.

fixed effect	*estimate*	*lower CI*	*higher CI*	*z value*	*p value*
(intercept)	0.995	0.908	1.090	-0.102	0.918
L1 Spanish	0.864	0.766	0.975	-2.366	0.017
L1 Bosnian	0.958	0.853	1.077	-0.705	0.480
L1 Danish	0.980	0.881	1.089	-0.369	0.712
ungram subjects: total	0.008	0.003	0.022	-9.776	<0.001
ungram subjects: L1 Spanish	18.258	6.843	48.711	5.801	<0.001
ungram subjects: L1 Bosnian	5.757	2.089	15.864	3.385	<0.001
ungram subjects: L1 Danish	3.164	1.150	8.705	2.232	0.025

These results show that independently of the L1, the preferred subject type for all the participants is the overt pronoun. This preference is significant both for the L2 English groups as well as for the L1 English group (p<.001). All the language groups follow the same pattern: overt pronouns > full DPs > null subjects.

An across group comparison shows that the L1 Danish and the L1 English groups produce more grammatical null subjects in contrast to the L1 Spanish and L1 Bosnian groups who produce more ungrammatical null subjects.

In Table 5 the overall distribution of grammatical subjects appears as one variable (grouping full DPs, overt pronouns and grammatical null subjects) and ungrammatical subjects as the other.

Table 5. Distribution of subjects per language group: grammaticality.

L1		grammatical		ungrammatical		total % [# of cases]
		% [# of cases]	SD	% [# of cases]	SD	
[+null subject] languages	Spanish	85.76 [873]	14.99	14.24 [145]	14.99	100 [1018]
	Bosnian	94.99 [1023]	10.17	5.01 [54]	10.17	100 [1077]
[-null subject] languages	Danish	97.28 [1858]	4.43	2.72 [52]	4.43	100 [1910]
	English [control]	99.10 [663]	2.15	0.90 [6]	2.15	100 [669]

Most of the sentential subjects produced by the three L2 groups, as shown in Table 5, are grammatical and this is significantly so both for the L2 groups as well as for the control group (all p<.001). When focusing on ungrammaticality rates, a between group comparison shows that the L1 Spanish group (p<.001) and the L1 Bosnian group (p=.001) significantly differ from the L1 English group. However, no significant difference is found between the L1 Danish and the L1 English groups (p>.05). Hence, the [-null subject] language groups behave similarly, and they differ from the [+null subject] language groups. In the case of the production of grammatical subjects, the analysis shows no statistically significant difference between neither of the language groups (all p>.05).

To further explore the production of subjects, Table 6 shows the distribution of adequate subjects.

Table 6. Distribution of subjects per language group: adequacy.

L1		adequate		non-adequate		total % [# of cases]
		% [# of cases]	SD	% [# of cases]	SD	
[+null subject] languages	Spanish	97.71 [853]	3.63	2.29 [20]	3.45	100 [873]
	Bosnian	95.97 [982]	11.28	4.03 [41]	10.51	100 [1023]
[-null subject] languages	Danish	98.92 [1836]	4.53	1.07 [20]	4.47	100 [1856]
	English [control]	100 [663]	0.00	0 [0]	0.00	100 [663]

To conduct the statistical analysis in terms of adequacy, zero-inflation condition was checked before fitting the model. The results in Table 7 show the comparisons obtained, which are discussed below.

Table 7. Summary of the GLMM fixed effects for adequacy (zero-inflation model coefficients).

fixed effect	***estimate***	***z value***	***p value***
(intercept)	18.004	0.010	0.992
L1 Spanish	-20.870	-0.012	0.991
L1 Bosnian	-19.063	-0.011	0.991
L1 Danish	-19.516	-0.011	0.991
non-adequate subjects: pronoun_RI	-0.056	0.000	1.000
non-adequate subjects: pronoun_RI	-4.075	-0.003	0.998
non-adequate subjects: L1 Spanish (pronoun_RI)	19.256	0.010	0.992
non-adequate subjects: L1 Spanish (pronoun_RRI)	19.259	0.012	0.990
non-adequate subjects: L1 Bosnian (pronoun_RI)	3.178	0.002	0.999
non-adequate subjects: L1 Bosnian (pronoun_RRI)	5.154	0.003	0.997
non-adequate subjects: L1 Danish (pronoun_RI)	0.183	0.000	1.000
non-adequate subjects: L1Danish (pronoun_RRI)	-4.765	-0.003	0.998

RI = reference introduction and RRI = reference re-introduction

In this case, since all the subjects produced by the control group are adequate and thus, there is no variability, they have been excluded from the statistical analysis and only the comparison within the experimental groups is carried out. Also, within these experimental groups, adequate subjects significantly outnumber non-adequate ones (all p<.005).

In the case of non-adequate subjects, the L1 Bosnian group produces the highest rate (4.03%) followed by the L1 Spanish group (2.29%) and finally the L1 Danish group (1.07%). Thus, the three L2 groups pattern alike in this respect.

To further explore the non-adequate subjects and bearing in mind the very low rates seen in Table 6, Table 8 shows that the majority of these are produced when the participants have used a full DP for reference maintenance where a pronoun is expected. This is especially relevant in the case of the L1 Spanish and L1 Bosnian groups. It seems that these [+null subject] language groups prefer to repeat the full DP.

Table 8. Distribution of subjects per language group and adequacy type: non-adequate subjects.

L1	DP_RM	PRON_RI	PRON_RRI	total
Spanish	100% [20]	0% [0]	0% [0]	100% [20]
Bosnian	80.49% [33]	12.19% [5]	7.31% [3]	100% [41]
Danish	55% [11]	35% [7]	10% [2]	100% [20]
English [control]	0% [0]	0% [0]	0% [0]	100% [0]

RM = reference maintenance; RI = reference introduction and RRI = reference re-introduction

In spite of the very low rate of non-adequate subjects, the results in Table 8 point to so-called over-explicitation effects (Hendriks, 2003). In particular, overt DPs are used in contexts of reference maintenance where a pronoun is more pragmatically adequate. This phenomenon has been found to be true independently of the participants' L1 or the L2 being learnt (e.g., Ahrenholz, 1998; Chini, 1998; Muñoz, 2000; Hendriks, 2003). In the case of L2 studies, over-explicitation has been attributed to the learner's concern for communicative success and to a search for hyper clarity linked to the leaners' awareness of their deficiencies in the L2 (e.g., Véronique et al., 2000; Hendriks, 2003).[5]

To further explore the results presented above, we focus first on typological similarity and then on lexical specialization approach with the aim of providing an answer to both research questions.

Previous studies on crosslinguistic influence and language typology have argued that crosslinguistic influence is affected by typological similarity so that when typologically similar languages are in contact, the L1 can facilitate the acquisition of the L2, as per the Facilitation Hypothesis (Gundel & Tarone, 1992). If typological similarity plays a role, then the L1 Spanish and L1 the Bosnian groups (i.e., both with [+null subject] L1s), on the one hand, and the L1 Danish and the L1 English groups (i.e., both with [-null subject] L1s), on the other, should pattern alike. What regards the production of ungrammatical subjects, our results show that the L1 Danish and the L1 English groups pattern alike. Therefore, it can be argued that in the case of [-null subject] languages (i.e., L1 Danish), the L1 functions as a facilitator. Our results also show that in the case of [+null subject] languages (i.e., L1 Spanish and L1 Bosnian), they also pattern alike since they both produce more ungrammatical subjects. Thus, in the case of [+null subject] languages, the L1 may be responsible for these participants' higher rate of ungrammatical subjects when compared to the L1 Danish and the L1 English groups. However, despite this difference between

[5] Similar results are also found in the case of L1 speakers' interpretation data. Lapidus Shin & Smith Cairns (2012) analyzing L1 Spanish children (and adults) found that there was a tendency to accept DPs in contexts where pronouns are more adequate. They claim that children's tolerance for redundancy (overacceptance of DPs) and ambiguous structures (overacceptance of pronouns) is related to difficulties with perspective and referent accessibility.

the two typological groups, data show a very high rate of grammaticality for the three experimental groups, scores ranging from 85% to 99% and the difference between the groups not being statistically significant. In other words, the production of grammatical subjects is similar across the three experimental groups. Therefore, it can be argued that, when it comes to the production of grammatical subjects, typological similarity does not seem to be a very salient factor as it does not predict (nor account for) differences that could be attributed to the typological similarity or difference between participants' L1 and the L2.

The production of adequate subjects is not only contingent on the acquisition of the syntactic properties that characterize sentential subjects in each language. It also involves the combination and mastery of other linguistic domains which, therefore, places the production of adequate subjects at the interface level. Different L2 studies have suggested that the interfaces are one of the salient problems L2 bilinguals have when mastering sentential subjects (e.g., Sorace, 2005; Tsimpli & Sorace, 2006; Sorace & Filiaci, 2006). Thus, an analysis of both grammaticality and adequacy allows for a more comprehensive interpretation of the production of sentential subjects by L2 bilinguals. By separating between syntactic requirements issues and the interface conditions, we can see whether the interface at stake (i.e., syntax-pragmatics) is a conditioning factor in the production of speakers with [+null subject] L1s as well as in that of speakers with [-null subject] L1s. Or rather, whether only the purely syntactic constraints without the influence of the pragmatic factors can account for the L2 bilinguals' production.

The analysis of sentential subjects attending to adequacy yields even more homogeneous results across the four participant groups (experimental and control). So, once more, typological similarity does not seem to explain the results we have obtained as both participants with [+null subject] L1s and those with a [-null subject] L1 behave very similarly showing very high correctness rates. Analyzing the non-adequate subjects, only the L1 Bosnian group seems to perform differently from the rest of the groups. Being the L1 Bosnian group the one that has the lowest MLUw values (Table 1), it could be argued that their production rate may be an indicator of their being less engaged in the task and, possibly so, of their having a different proficiency level. However, these differences across the three L2 English groups never reach statistical significance. In fact, the production for the three experimental groups show similar patterns in their use of over-explicitation, which suggests that L2 English learners, regardless of their L1s, resort to similar strategies. Thus, typological similarity does not seem to be the influencing factor as it does not allow us to predict a different behavior between the L1 Spanish and the L1 Bosnian groups on the one hand, and the L1 Danish group on the other. Quite the opposite: the participants' performance does not seem to change as a function of whether their L1 is typologically similar or typologically different from the L2.

Hence, what our results show is that, for all the languages under study, production of both grammatical and adequate subjects is at ceiling and, consequently, the production of ungrammatical and non-adequate subjects is low. These results seem to suggest that all

the groups adhere to the overt subject rule in English (i.e., as a one–subject–type language), making the production of null ungrammatical subjects rather marginal. The availability of two subject types (i.e., null and overt) in the superset language (i.e., Spanish and Bosnian in this case) seems to have a facilitative effect in the acquisition of a subset language (English in this case, which only has the overt subject type available).

These results are in line with the lexical specialization approach regarding how the existence and directionality of crosslinguistic influence is shaped when superset and subset languages are in contact, as Liceras & Fernández Fuertes (2019) propose for the L1. That is, crosslinguistic influence takes place from the L1 superset language (i.e., Spanish and Bosnian) into the L2 (i.e., English) and it facilitates the acquisition of sentential subjects, as the results show.

The lexical specialization proposal lies on the availability of two subject types (i.e., overt and null) and it predicts that the language that is lexically specialized (i.e., Spanish and Bosnian in this case as they allow both subject types) would help reinforce the only available option in the language that is not lexically specialized (i.e., English in this case). That is, the prediction is that crosslinguistic influence from the lexically specialized language into the non-lexically specialized language would not have a negative effect making the learner produce (more) illicit null subjects in English. Rather, crosslinguistic influence would have a positive effect by reinforcing the only available option in the non-lexically specialized language (i.e., the overt subject).

In the case of the L1 Danish participants, crosslinguistic influence from their L1 would also yield to a high production of overt subjects given that both the L1 and L2 share the same value of the parameter (i.e., they are both non-lexically specialized languages). That is, the fact that the overtness rates of these two groups of participants is at ceiling is expected. To learn a property that is already available in your L1, as is the overtness of the subject, has been systematically proven to be easier than to learn a property that differs from that in your L1. So, in the case of the L1 Danish group, either facilitative crosslinguistic influence or fast learning of this obligatoriness yields the same results.

Going back to our two research questions, we believe our data lend support to lexical specialization over typological similarity in accounting for the subject production of these three L2 English child groups. It can be argued that signs of crosslinguistic influence interfering effects can be seen in the L1 Bosnian and L1 Spanish groups when compared to the L1 Danish group, because the first two groups pattern alike while the L1 Danish group does so with the L1 English control group. However, while this is true, these differences cannot hide the fact that (i) grammaticality rates are very high for all three groups and that (ii) accurate rates are at ceiling for all three groups. Therefore, as argued above, the availability of two subject types in Spanish and Bosnian (i.e., null and overt pronouns) makes Spanish and Bosnian lexically specialized languages when it comes to grammatical person information because, even if person features are always realized via verbal agreement markers, two subject types are indeed possible. This lexical specialization is what is behind the superset/subset language difference. And it is not only

a matter of having more options, but crucially so that, out of the two options that are available in Bosnian and Spanish (i.e., null and overt), only one of these is available in Danish and English (i.e., overt). This superset relation is what, we believe, is behind the high grammaticality and adequacy rates that these L2 bilingual children have in their L2 English production. And it is also what explains their lack of high null subject rates as a sign of negative crosslinguistic influence from their L1s.

7. Conclusion

Taking formal proposals on sentential subjects as a point of departure, this study has offered an analysis of crosslinguistic influence effects in the case of two L2 contact situations: typologically similar languages (L2 English-L1 Danish) and typologically different languages (L2 English-L1 Spanish and L2 English-L1 Bosnian). Assuming L2 English children rely on their L1 when producing sentential subjects, the focus has been placed on characterizing what has more weight: typological similarity between the L1 and the L2 or lexical specialization whereby the properties of the L1 may be seen as a superset of those of the L2 and this facilitates L2 acquisition. While typological similarity can explain how the L1 Bosnian and the L1 Spanish children pattern similarly and somehow differently from the L1 Danish and the L1 English, it fails to account for the high grammaticality and accuracy rates that all three experimental groups show. These, however, can be accounted for if we claim that lexical specialization in the L2 child bilinguals' L1s has a facilitating effect. That is, as in the case of 2L1 acquisition studies, that there is facilitative crosslinguistic influence effects from the superset language (i.e., Spanish and Bosnian) to the subset language (i.e., English).

The conclusions we reach in the present investigation can be further explored by considering other issues that can help explore the interaction between typological similarity and lexical specialization. These include the analysis of the written production of sentential subjects to see if modality plays a role and of other variables such as time of exposure and proficiency in the L2 bilingual acquisition. This will offer further insight into the representational nature of subjects in child L2 bilinguals.

As previous L2 studies have shown (e.g., Ferrero, 2020; Gómez Garzarán, 2023), empirical research on different grammatical properties can contribute to more felicitous teaching policies and practices. Although this is not the objective of the present investigation, the conclusions obtained in this study could be used as a point of departure to conduct pedagogical research where the design of the curricula and the materials for L2 instruction targets in a more accurate way potential crosslinguistic influence effects.

Together with these pedagogical implications this study leaves the door open to explore more in-depth issues that are fundamental in L2 instruction contexts such as age and language modality. The combination of data elicited via different methodologies (e.g., judgment data, processing data) could also allow us to gain further insight into the representational nature of sentential subjects and put into perspective the results obtained in the present investigation.

Acknowledgments
This work was carried out as part of the UVALAL (University of Valladolid Language Acquisition Lab) research activities. We would like to acknowledge the collaboration of Iban Mañas Navarrete in the statistical analysis.

CRediT Author contribution / *Contribución de los autores*
Conceptualization / *Conceptualización:* Sonja Mujcinovic & Raquel Fernández Fuertes.
Formal Analysis / *Análisis formal:* Sonja Mujcinovic & Raquel Fernández Fuertes.
Methodology / *Metodología:* Sonja Mujcinovic & Raquel Fernández Fuertes.
Writing / *Redacción:* Sonja Mujcinovic & Raquel Fernández Fuertes.
Research dataset / *Datos de investigación:* Sonja Mujcinovic.

Funding, data availability, and copyright / *Financiación, disponibilidad de datos y derechos de autoría*
Funding / *Financiación:* Ministry of Science, Innovation and Universities of Spain and the European Regional Development Fund (ERDF) [PGC2018097693-B-I00] & [PID2025-168895NB-I00]. Ministry of Science and Innovation of Spain [RED2022-134920-T].
Image use consent / *Consentimiento para el uso de imágenes:* Not aplicable / No procede.
Conflict of interest / *Conflicto de intereses:* The authors declare no conflict of interest / Los autores declaran no tener ningún conflicto de intereses.
Data availability statement/ *Declaración de disponibilidad de datos:* The data are available in open access through the University of Valladolid online repository (UVaDoc): https://uvadoc.uva.es/handle/10324/53753 & https://uvadoc.uva.es/handle/10324/53750 / Los datos están disponibles en acceso abierto a través del repositorio en línea de la Universidad de Valladolid (UVaDoc): https://uvadoc.uva.es/handle/10324/53753 & https://uvadoc.uva.es/handle/10324/53750
License/ *Licencia:* This article is published under the CC BY 4.0 License / *Este artículo se publica bajo la Licencia CC BY 4.0.*

Editorial history / *Fechas del proceso editorial*
Received / *Recibido:* 21/12/2023
Accepted / *Aceptado:* 09/04/2024
Published / Publicado: 15/11/2026

References

Ahrenholz, B. (1998). *Modalität und Diskurs: Instruktionen auf deutsch und italienisch*. Stauffenburg Verlag.

Ayoun, D. (2003). *Parameter setting in language acquisition*. Cambridge, UK: Continuum Press.

Bel, A., García, E., & Rosado, E. (2016). Reference comprehension and production in bilingual Spanish: the view from null subject languages. In A. Alba de la Fuente, E. Valenzuela, & C. Sanz-Martínez (Eds.), *Language Acquisition Beyond Parameters: Studies in Honour of Juana Liceras* (pp. 27–70). John Benjamins. https://doi.org/10.1075/sibil.51.03bel

Belletti, A. (2001). Inversion as focalization. In A. Hulk & J-Y. Pollock (Eds.), *Subject Inversion in Romance and the Theory of Universal Grammar* (pp. 60–91). Oxford University Press. https://doi.org/10.1093/oso/9780195142693.003.0003

Belletti, A. (2004). Answering with a “cleft”: The role of the null subject parameter and the VP periphery. In L. Brugé, G. Giusti, N. Munaro, W. Schweikert, & G. Turano (Eds.), *Proceedings of the XXX Incontro di Grammatica Generativa* (pp. 63–82). Cafoscarina.

Bini, M. F. (1993). La adquisición del italiano: más allá de las propiedades sintácticas del parámetro pro-drop. In J. M. Liceras (Ed.), *La Lingüística y el Análisis de los Sistemas no Nativos* (pp. 126–140). Doverhouse Editions.

Bloom, P. (1990). Subjectless sentences in child language. *Linguistic Inquiry, 21*, 491–504.

Brown, R. (1973). *A first language: The early stages*. Harvard University Press.

Camacho, J. (2006). Do Subjects have a place in Spanish? In J. P. Montreuil & C. Nishida (Eds.), *New Perspectives in Romance Linguistics* (pp. 51–66). John Benjamins. https://doi.org/10.1075/cilt.275.06cam

Camacho, J. (2013). *Null subjects*. Cambridge University Press. https://doi.org/10.1017/CBO9781139524407

Camacho, J. (2016). The null subject parameter revisited. In M. A. Kato & F. Ordoñez (Eds.), *The Morphosyntax of Portuguese and Spanish in Latin America* (pp. 27–48). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780190465889.003.0002

Cenoz, J. (2001). The effect of linguistic distance, L2 status and age on cross-linguistic influence in third language acquisition. In J. Cenoz, B. Hufeisen, & U. Jessner (Eds.), *Crosslinguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives* (pp. 8–20). Multilingual Matters. https://doi.org/10.21832/9781853595509-002

Cenoz, J. (2003). The additive effect of bilingualism on third language acquisition: A review. *International Journal of Bilingualism, 7*, 71–87. https://doi.org/10.1177/13670069030070010501

Chomsky, N. (1981). *Lectures on Government and Binding*. Foris. https://doi.org/10.1515/9783110884166

Chomsky, N., & Lasnik, H. (1977). Filters and control. *Linguistic Inquiry, 8(3)*, 425–504. https://doi.org/10.1007/978-94-015-6859-3_4

Chini, M. (1998, September). *L'emploi de la subordination chez des apprenants germanophones d'italien: entre agrégation et intégration*. Paper presented at the Euroconference: The structure of learner language. From pragmatics to syntax: Organizational principles of second language acquisition. Maratea, Italy.

Cook, V. (2016). *Second language learning and language teaching*. Routledge. https://doi.org/10.4324/9781315883113

Cuza, A. (2013). Crosslinguistic influence at the syntax proper: Interrogative subject-verb inversion in heritage Spanish. *International Journal of Bilingualism, 17(1)*, 71–96. https://doi.org/10.1177/1367006911432619

Cuza, A., & Camacho, J. (2017). Pronominal subject expression with inanimate reference in heritage speakers of Cuban Spanish. In A. Cuza (Ed.), *Cuban Spanish Dialectology: Variation, Contact and Change* (pp. 229–246). Georgetown University Press. https://doi.org/10.1353/book62540

Cuza, A., Guijarro-Fuentes, P., Pires, A., & Rothman, J. (2013). The syntax-semantics of bare and definite plural subjects in the L2 Spanish of English natives. *The International Journal of Bilingualism, 17(5)*, 632–652. https://doi.org/10.1177/1367006911435594

Cuza, A., Jiao, J., & López-Otero, J. (2018). Does typological proximity really matter? Evidence from Mandarin and Brazilian Portuguese-speaking learners of Spanish. *Languages, 3(2)*, 13. https://doi.org/10.3390/languages3020013

Ellis, R. (2014). Principles of instructed second language learning. In M. Celce-Murcia, D. M. Brinton, & A. S. Marguerite (Eds.), *Teaching English as a Second of Foreign Language* (pp. 31–45). Heinle.

Ellis, R., & Yuan, F. (2004). The effects of planning on fluency, complexity, and accuracy in second language narrative writing. *Studies in Second Language Acquisition, 26(1)*, 59–84. https://doi.org/10.1017/S0272263104026130

Faitaki, F., & Murphy, V. A. (2024). Subject realization in Greek preschool learners of English. *Second Language Research, 40*(4), 911-933. https://doi.org/10.1177/02676583231195308

Ferrero, M. (2020). Can research contribute to improve educational practice? *The Spanish Journal of Psychology, 23*, 1–6. https://doi.org/10.1017/SJP.2020.24

Gass, S. M. (1996). Second language acquisition and linguistic theory: The role of language transfer. In W. C. Ritchie & T. K. Bhatia (Eds.), *The Handbook of Second Language Acquisition* (pp. 317–345). Academic Press. https://doi.org/10.1163/9789004653450_014

Gass, S., & Selinker, L. (2008). *Second language acquisition: An introductory course*. Routledge. https://doi.org/10.4324/9780203932841

Gathercole, V. C. M. (2016). Factors moderating proficiency in bilingual speakers. In E. Nicoladis & S. Montanari (Eds.), *Language and the Human Lifespan Series. Bilingualism across the Lifespan: Factors Moderating Language Proficiency* (pp. 123–140). American Psychological Association; De Gruyter Mouton. https://doi.org/10.1515/9783110341249-009

Grinstead, J. (1998). *Subjects, sentential negation and imperatives in child Catalan and Spanish* [Unpublished doctoral dissertation]. University of California.

Gundel, J. K., & Tarone, E. (1992). Language transfer and the acquisition of pronouns. In S. M. Gass & L. Selinker (Eds.), *Language Transfer in Language Learning* (pp. 87–100). John Benjamins. https://doi.org/10.1075/lald.5.08gun

Gómez Garzarán, E. (2023). *Instructed Second Language Acquisition & Cross-Linguistic Influence: Word Order in Nominal Structures* [Unpublished doctoral dissertation]. University of Valladolid. https://doi.org/10.35376/10324/62568

Håkansson, G. (2001). Tense morphology and verb-second in Swedish L1 children, L2 children and children with SLI. *Bilingualism: Language and Cognition, 4*(1), 85-99. https://doi.org/10.1017/S1366728901000141

Hawkins, R. (2001). *Second Language Syntax: A Generative Introduction*. Blackwell.

Hawkins, J. A., & Filipović, L. (2012). *Criterial features in L2 English: Specifying the reference levels of the Common European Framework*. Cambridge: Cambridge University Press.

Hedge, T. (2000). *Teaching and learning in the language classroom*. Oxford University Press.

Hendriks, H. (2003). Using nouns for reference maintenance: A seeming contradiction in L2 discourse. In A. Giacalone Ramat (Ed.), *Typology and second language acquisition* (pp. 291–326). Mouton de Gruyter. https://doi.org/10.1515/9783110891249.291

Herschensohn, J. (2000). *The Second Time Around-Minimalism and L2 Acquisition*. John Benjamins. https://doi.org/10.1075/lald.21

Hoff, E. (2017). How bilingual development is the same as and different from monolingual development. *Bilingual from Birth: Process, Pedagogy and Policy, 8*, 3-16. https://doi.org/ 10.18192/olbiwp.v8i0.2114

Holmberg, A. (2005). Is there a little pro? Evidence from Finnish. *Linguistic Inquiry, 36(4)*, 533–564. https://doi.org/10.1162/002438905774464322

Holmberg, A. (2010). Null subject parameters. *Parametric variation: Null Subjects in Minimalist Theory*, 88–124. Cambridge University Press. https://doi.org/10.1017/CBO9780511770784.003

Hyams, N. (2011). Missing subjects in early child language. In J. de Villiers & T. Roeper (Eds.), *Handbook of Generative Approaches to Language Acquisition* (pp. 13–52). Springer. https://doi.org/10.1007/978-94-007-1688-9_2

Hyams, N., & Wexler, K. (1993). On the grammatical basis of null subjects in child languages. *Linguistic Inquiry, 24(1)*, 421–459.

Jaeggli, O., & Safir, K. J. (1989). The null subject parameter and parametric theory. In in O. Jaeggli & K. J. Safir (Eds.), *The Null Subject Parameter* (pp. 1–44). Springer. https://doi.org/10.1007/978-94-009-2540-3_1

Juan-Garau, M., & Pérez-Vidal, C. (2000). Subject realization in the development of a bilingual child. *Bilingualism: Language and Cognition, 3(3)*, 173–192. https://doi.org/10.1017/S1366728900000328

Kuiken, F., & Vedder, I. (2011). Task complexity and linguistic performance in L2 writing and speaking. *Second Language Task Complexity: Researching the Cognition Hypothesis of Language Learning and Performance, 2*, 91–104. https://doi.org/10.1075/tblt.2.09ch4

Lapidus Shin, N., & Smith Cairns, H. (2012). The Development of NP Selection in School-Age Children: Reference and Spanish Subject Pronouns. *Language Acquisition, 19(1)*, 3–38. https://doi.org/10.1080/10489223.2012.633846

Liceras, J. M. (1988). Syntax and stylistics: more on the pro-drop parameter. *Learnability and second languages*, 71-93.

Liceras, J. M. (1989). On some properties of the "pro-drop" parameter: Looking for missing subjects in non-native Spanish. In S. M. Gass & J. Schachter (Eds.), *Linguistic Perspectives on Second Language Acquisition* (pp. 109–133). Cambridge University Press. https://doi.org/10.1017/CBO9781139524544.009

Liceras, J. M., & Alba de la Fuente, A. (2015). Typological proximity in L2 acquisition. In T. Judy & S. Perpiñán (Eds.), *The Acquisition of Spanish in Understudied Language Pairings* (pp. 329–358). John Benjamins. https://doi.org/10.1075/ihll.3.13lic

Liceras, J. M., & Fernández Fuertes, R. (2019). Subject omission/production in child bilingual English and child bilingual Spanish: The view from linguistic theory. *Probus, 31(2)*, 245–278. https://doi.org/10.1515/probus-2016-0012

Liceras, J. M., Bel, A., & Perales, S. (2006). 'Living with optionality': Root infinitives, bare forms and inflected forms in child null subject languages. In N. Sagarra & A. J. Toribio (Eds.), *Subject and copula omission. Selected proceedings of the 9th Hispanic Linguistics Symposium* (pp. 203–216). Cascadilla Proceedings Project.

Liceras, J. M., Fernández Fuertes, R., & Pérez-Tattam, R. (2008). Null and overt subjects in the developing grammars (L1 English/L1 Spanish) of two bilingual twins. *A Portrait of the Young in the New Multilingual Spain, 9*, 111–134. https://doi.org/10.21832/9781847690241-009

Lozano, C. (2002). Knowledge of expletive and pronominal subjects by learners of Spanish. *ITL-International Journal of Applied Linguistics, 135(1)*, 37–60. https://doi.org/10.1075/itl.135-136.02loz

Lozano, C. (2018). The development of anaphora resolution at the syntax-discourse interface: pronominal subjects in Greek learners of Spanish. *Journal of Psycholinguistic Research, 47(2)*, 411–430. https://doi.org/10.1007/s10936-017-9541-8

Lundell, F. F., & Lindqvist, C. (2012). Vocabulary aspects of advanced L2 French: Do lexical formulaic sequences and lexical richness develop at the same rate? *Language, Interaction and Acquisition, 3*(1), 73-92. https://doi.org/10.1075/lia.3.1.05for

Lundell, F. F., & Lindqvist, C. (2014). Lexical aspects of very advanced L2 French. *Canadian Modern Language Review, 70*(1), 28-49. https://doi.org/10.3138/cmlr.1598

MacWhinney, B. (2000). *The CHILDES Project: Tools for Analyzing Talk (3rd ed.)*. Lawrence Erlbaum Associates.

MacWhinney, B. (2019). TalkBank and SLA. In N. Tracy-Ventura & M. Paquot (eds.), *The Handbook of SLA and Corpora*. New York: Routledge.

Margaza, P., & Bel, A. (2006). Null subjects at the syntax-pragmatics interface: Evidence from Spanish interlanguage of Greek speakers. Proceedings of GASLA, 88–97.

Martínez-Sanz, C. (2011). *Null and overt subjects in a variable system: The case of Dominican Spanish* [Unpublished doctoral dissertation]. University of Ottawa. https://doi.org/10.20381/ruor-5066

Mitkovska, L., & Bužarovska, E. (2018). Subject pronoun (non) realization in the English learner language of Macedonian speakers. *Second Language Research, 34(4)*, 463–485. https://doi.org/10.1177/0267658317747925

Montrul, S. (2004). Subject and object expression in Spanish heritage speakers: A case of morpho-syntactic convergence. *Bilingualism, Language and Cognition, 7*, 125–142. https://doi.org/10.1017/S1366728904001464

Montrul, S. (2016). *The acquisition of heritage languages*. Cambridge University Press. https://doi.org/10.1017/CBO9781139030502

Montrul, S., & Rodríguez-Louro, C. (2006). Beyond the syntax of the null subject parameter. In V. Torrens & L. Escobar (Eds.), *The Acquisition of Syntax in Romance Languages* (pp. 401–418). John Benjamins. https://doi.org/10.1075/lald.41.19mon

Montrul, S., Dias, R., & Santos, H. (2011). Clitics and object expression in the L3 acquisition of Brazilian Portuguese: Structural similarity matters for transfer. *Second Language Research, 27(1)*, 21–58. https://doi.org/10.1177/0267658310386649

Montrul, S., Dias, R., & Thomé-Williams, A. (2009). Subject expression in the non-native acquisition of Brazilian Portuguese. In A. Pires & J. Rothman (Eds.), *Minimalist Inquiries into Child and Adult Language Acquisition: Case Studies across Portuguese* (pp. 301–325). De Gruyter Mouton. https://doi.org/10.1515/9783110215359.2.301

Mujcinovic, S. (2015). The analysis of subjects in the oral and written production of L2 English learners: transfer and language typology. *Procedia-Social and Behavioral Sciences, 198*, 344–349. https://doi.org/10.1016/j.sbspro.2015.07.453

Mujcinovic, S. (2020). *English subjects in the linguistic production of L1 Spanish, L1 Bosnian and L1 Danish speakers: comparative grammar, typological similarity and transfer* [Unpublished doctoral dissertation]. University of Valladolid. https://doi.org/10.35376/10324/43635

Mujcinovic, S., & Fernández Fuertes, R. (2021). *soraUVALAL corpus*. SLABank-TalkBank. https://doi.org/10.21415/cksf-ch67

Muñoz, C. (2000, October). The over-explicitation of personal reference in L2 narratives. Proceedings of the Euroconference on Information Structure, Linguistic Structure and the Dynamics of Acquisition, San Feliu de Guixols, Spain.

Orfitelli, R., & Hyams, N. (2008). An experimental study of children's comprehension of null subjects: Implications for grammatical/performance accounts. In H. Chan, H. Jacob, & E. Kapia (Eds.), *BUCLD 32 Proceedings* (pp. 335–346). Cascadilla Press.

Orfitelli, R., & Hyams, N. (2012). Children's grammar of null subjects: Evidence from comprehension. *Linguistic Inquiry, 43(4)*, 563–590. https://doi.org/10.1162/ling_a_00106

Paradis, J., & Navarro, S. (2003). Subject realization and crosslinguistic interference in the bilingual acquisition of Spanish and English. *Journal of Child Language, 30*, 1–23. https://doi.org/10.1017/S0305000903005609

Park, H. (2004). A minimalist approach to null subjects and objects in second language acquisition. *Second Language Research, 20*(1), 1-32. https://doi.org/10.1191/0267658304sr228oa

Perlmutter, D. M. (1971). A typological difference among languages. In D. M. Perlmutter (Ed.), *Deep and Surface Structure Constraints in Syntax* (pp. 99–122). Holt, Rinehart & Winston.

Phinney, M. (1987). The pro-drop parameter in second language acquisition. In T. Roeper & E. Williams (Eds.), *Parameter Setting* (pp. 221–238). Reidel. https://doi.org/10.1007/978-94-009-3727-7_10

Pladevall-Ballester, E. (2012). Child L2 English acquisition of subject properties in an immersion bilingual context. *Second Language Research, 28(2)*, 217–241. https://doi.org/10.1177/0267658312438534

Pladevall-Ballester, E. (2016). Developmental asynchrony in the acquisition of subject properties in Child L2 English and Spanish. *Language Acquisition, 23(4)*, 387–406. https://doi.org/10.1080/10489223.2016.1187610

Pladevall-Ballester, E., Puig-Mayenco, E., Tubau, S., & Capdevila, M. (2022). Asymmetries in child foreign language acquisition: production and interpretation of L2 English subjects. *Revista Electrónica de Lingüística Aplicada, 21(1)*, 40–57. https://doi.org/10.58859/rael.v21i1.496

Platzack, C. (1987). The Scandinavian languages and the null-subject parameter. *Natural Language & Linguistic Theory, 5(3)*, 377–401. https://doi.org/10.1007/BF00134554

Quesada, M. L. (2014). Subject pronouns in second language Spanish. In K. L. Geeslin (Ed.), *The Handbook of Spanish Second Language Acquisition* (pp. 253–269). Wiley-Blackwell. https://doi.org/10.1002/9781118584347.ch15

R Core Team (2017). *R: A language and environment for statistical computing. R Foundation for Statistical Computing*, Vienna, Austria. https://www.R-project.org/

Rizzi, L. (1982). *Issues in Italian syntax*. Foris. https://doi.org/10.1515/9783110883718

Rizzi, L. (1986). Null objects in Italian and the theory of pro. *Linguistic Inquiry, 17*(3), 501-557.

Rothman, J. (2008). How pragmatically odd! Interface delays and pronominal subject distribution in L2 Spanish. *Studies in Hispanic and Lusophone Linguistics, 1(2)*, 317–340. https://doi.org/10.1515/shll-2008-1022

Rothman, J. (2009). Pragmatic deficits with syntactic consequences? L2 pronominal subjects and the syntax-pragmatics interface. *Journal of Pragmatics, 41(5)*, 951–973. https://doi.org/10.1016/j.pragma.2008.07.007

Rothman, J. (2010). On the typological economy of syntactic transfer: Word order and relative clause high/low attachment preference in L3 Brazilian Portuguese. *IRAL-International Review of Applied Linguistics in Language Teaching, 48(2–3)*, 245–273. https://doi.org/10.1515/iral.2010.011

Rothman, J. (2011). L3 syntactic transfer selectivity and typological determinacy: The typological primacy model. *Second Language Research, 27(1)*, 107–127. https://doi.org/10.1177/0267658310386439

Rothman, J., & Cabrelli Amaro, J. (2010). What variables condition syntactic transfer? A look at the L3 initial state. *Second Language Research, 26(2)*, 189–218. https://doi.org/10.1177/0267658309349410

Schepens, J. J., van der Slik, F., & Van Hout, R. (2016). L1 and L2 distance effects in learning L3 Dutch. *Language Learning, 66(1)*, 224–256. https://doi.org/10.1111/lang.12150

Serratrice, L. (2002). Overt subjects in English: evidence from the marking of person in an English-Italian bilingual child. *Journal of Child Language, 29(2)*, 327–356. https://doi.org/10.1017/S0305000902005068

Serratrice, L. (2005). The role of discourse pragmatics in the acquisition of subjects in Italian. *Applied Psycholinguistics, 26(3)*, 437–462. https://doi.org/10.1017/S0142716405050241

Sheehan, M. (2006). *The EPP and null subjects in Romance* [Unpublished doctoral dissertation]. Newcastle University.

Slabakova, R., & Ivanov, I. (2011). A more careful look at the syntax-discourse interface. *Lingua, 121(4)*, 637–651. https://doi.org/10.1016/j.lingua.2010.05.003

Sorace, A. (2005). Syntactic optionality at interfaces. In Cornips, L., & Corrigan, K. (eds.), *Syntax and Variation: Reconciling the Biological and The Social*, 46-111. Amsterdam: John Benjamins.

Sorace, A., & Filiaci, F. (2006). Anaphora resolution in near-native speakers of Italian. *Second Language Research, 22(3)*, 339–368. https://doi.org/10.1191/0267658306sr271oa

Tsimpli, I., & Sorace, A. (2006). Differentiating interfaces: L2 performance in syntax-semantics and syntax-discourse phenomena. In D. Bamman, T. Magnistkaia, & C. Zaller (Eds.), *Proceedings of the 30th Annual Boston University Conference on Language Development (BUCLD 30)* (pp. 653–664). Cascadilla Press.

Unsworth, S. (2008). Age and input in the acquisition of grammatical gender in Dutch. *Second Language Research, 24*(3), 365-395. https://doi.org/10.1177/0267658308090185

Unsworth, S., & Blom, E. (2010). Comparing L1 children, L2 children and L2 adults. *Experimental Methods in Language Acquisition Research, 27*, 201-222.

Valian, V. (1991). Syntactic subjects in the early speech of American and Italian children. *Cognition, 40*, 21–81. https://doi.org/10.1016/0010-0277(91)90046-7

Valian, V., & Eisenberg, Z. (1996). The development of syntactic subjects in Portuguese speaking children. *Journal of Child Language, 23*, 103–128. https://doi.org/10.1017/S0305000900010114

Véronique, D., Carroll, M., & von Stutterheim, C. (2000, October). Anaphoric linkage. Proceedings of the Euroconference on Information Structure, Linguistic Structure and the Dynamics of Acquisition, San Feliu de Guixols, Spain.

Westergaard, M., Mitrofanova, N., Mykhaylyk, R., & Rodina, Y. (2017). Crosslinguistic influence in the acquisition of a third language: The Linguistic Proximity Model. *International Journal of Bilingualism, 21(6)*, 666–682. https://doi.org/10.1177/1367006916648859

White, L. (1985). The "pro-drop" parameter in adult second language acquisition. *Language Learning, 35(1)*, 47–61. https://doi.org/10.1111/j.1467-1770.1985.tb01014.x

White, L. (1989). Linguistic universals, markedness and learnability: Comparing two different approaches. *Interlanguage studies bulletin (Utrecht), 5(2)*, 127–140. https://doi.org/10.1177/02676583890050020 https://doi.org/10.1177/026765838900500202