

Estela García-Alcaraz* and Aurora Bel

Does empirical data from bilingual and native Spanish corpora meet linguistic theory? The role of discourse context in variation of subject expression

<https://doi.org/10.1515/applirev-2017-0101>

Abstract: The goal of this study is to shed light on how empirical data on the discourse constraints of null and overt third person subject pronouns in L1 and bilingual Spanish meet linguistic theory. A (semi)spontaneous production task was administered to 34 Moroccan Arabic (MA)/Spanish early sequential bilinguals and 30 L1 Spanish controls. All 3rd person subject positions were coded: (1) morphosyntactic form (null pronoun vs. overt pronoun); (2) discourse function ([-Topic Shift] vs. [+Topic Shift]); (3) sentence relation (intrasentential vs. intersentential); (4) clause order within intrasentential contexts (main-subordinate vs. subordinate-main); and (5) access to the antecedent (clear vs. ambiguous antecedent). The results reveal general patterns of use in both L1 and bilingual Spanish: null pronouns express topic maintenance both in inter- and intrasentential contexts (both clause orders) and overt pronouns, especially in intersentential contexts, are generally used for topic change. However, additional analyses provide evidence that null pronouns in L1, but not in bilingual Spanish, are often used in change of reference contexts where the antecedent is not ambiguous. This reveals patterns that have gone unreported by most previous descriptive and theoretical studies. Finally, a higher use of ambiguous null pronouns is attested among bilingual speakers, which suggests a lower control of the mechanisms by which reference is established in discourse and supports, to some extent, the predictions derived from the Interface Hypothesis.

This paper derives from the first author's dissertation (directed by the second author), which was defended at Pompeu Fabra University in 2015 under the title *Comprensión y producción de los pronombres nulos y explícitos de tercera persona en posición de sujeto en la adquisición temprana del español L2*.

***Corresponding author: Estela García-Alcaraz**, Modern Languages and Literatures, University of Ottawa, Ottawa, Canada, E-mail: egarcia@uottawa.ca

Aurora Bel, Translation and Language Sciences, Pompeu Fabra University, Barcelona, Spain, E-mail: aurora.bel@upf.edu

Keywords: subject pronouns, anaphora resolution, production data, early sequential bilingual acquisition

1 Introduction

1.1 The topichood dimension in discourse reference tracking through subject pronouns

The success of communicative acts depends on how speakers encode information and how receivers decode it. In this process, reference to entities plays a key role and entails the selection of the appropriate expression among several referential forms differing in their degree of explicitness. Ariel's (1990) Accessibility Theory predicts a correlation between referring expressions (REs) and the cognitive processes involved in the accessibility of discourse referents: whereas full content REs (e.g. Determiner Phrases, DPs) refer to antecedents whose accessibility is low, pronouns, as less informative forms, pick up highly accessible antecedents in discourse. Pronouns, being referentially empty, must recover referential content from other discourse elements and, therefore, by performing co-reference, reference is solved and redundancy can be avoided. This requirement applies specifically to third person pronouns, since first and second pronouns have a constant referential denotation (see the classical approach from Benveniste 1966). In this paper, we focus on the process of reference production and expression alternation in reference tracking through third person subject pronouns across clauses.

Null subject languages like Spanish or Moroccan Arabic (MA) allow the omission of subjects and have two types of pronouns: null pronouns (NPs) and overt pronouns (OPs). However, the choice between them is not arbitrary as it obeys discourse constraints. Focusing on Spanish, the distribution of NPs and OPs has been largely discussed in descriptive studies (Luján 1985; Fernández Soriano 1999) and in corpus studies from variationist (de Prada Pérez 2009; Carvalho et al. 2015), generative (Domínguez 2013; Lozano 2016) and cognitive perspectives (Blackwell and Quesada 2012). According to these studies, NPs are mainly used in topic continuity contexts (thus having a [-Topic Shift] feature, henceforth [-TS]) and OPs express a change of referent (thus having a [+Topic Shift] feature, henceforth [+TS]). These discourse specializations seem to be consistent across null subject languages, as previous studies have also found similar patterns for Italian (Sorace et al. 2009) and Catalan (Bel et al. 2010), among others. MA is a pro-drop language with null subjects,

expletives and overt subject pronouns that convey particular pragmatic roles, such as contrast or emphasis. In embedded contexts, similar biases towards subject or object antecedents for NPs and OPs have been experimentally attested (Bel and García-Alcaraz 2015) signaling (non)switch of reference in similar ways to those of Spanish. Moreover, in MA, as is the case in Spanish, NPs are the unmarked option (except for present tense copular constructions, where the OP is compulsory). Based on this description, it seems that MA does not differ greatly from the typical pro-drop languages.

While most studies agree in explaining the alternation of NPs and OPs as dependent of [\pm TS] feature, some diverge in what might be understood as topic. For instance, under the topic maintenance label factors such as previous mention of the antecedent, distance from the antecedent or the syntactic function of the antecedent, to mention a few, are included (see de Prada Pérez 2009 for a review). Whatever the case, reference continuity, broadly understood, favors the use of NPs over OPs. In the present study, we adopt the proposal by Zubizarreta (1999), who defines topic as what the sentence is about (see Section 3.3).

By means of experimental data, Carminati's (2002) seminal Position of Antecedent Hypothesis (PAH) established that NPs tend to solve co-reference with subject antecedents and OPs with object antecedents in Italian. Although Carminati's stimuli were short and controlled, her results may be reinterpreted from the angle of the [\pm TS] discourse feature since the subject always appeared in first sentence position and, then, matched the topic. Accordingly, NPs occur in [$-$ TS] contexts whereas OPs occur in [$+$ TS] in null subject languages. Since its postulation, different studies have experimentally tested the PAH (Mayol and Clark 2010 in Catalan, Bel and García-Alcaraz 2015 in MA; Alonso-Ovalle et al. 2002 and Filiaci 2011 in Spanish) and agree in showing the overwhelming NP preferences for recovering antecedents in subject position (thus articulating [$-$ TS]). However, a complementary bias towards antecedents in object position by OPs (thus expressing [$+$ TS]) seems to be less pervasive since variation is attested among null subject languages.

Previous studies focusing on production data reported an overextension of OPs in L2 Spanish, which means that non-native speakers used OPs in [$-$ TS] contexts at a higher rate than monolinguals (Blackwell and Quesada 2012; Domínguez 2013; Lozano 2016); these findings can be accommodated within Tsimplici's (2011) learner-default's proposal (i.e. the OP), in contrast with the idea of linguistic-default (i.e. the NP). Different from previous corpus studies and in line with the detailed work by Lozano (2016), in our study specific features that may affect the distribution and choice of NPs and OPs are scrutinized: type of sentence linking (intersentential and intrasentential), clause order (main-subordinate vs.

subordinate-main within intrasentential contexts) and felicitous conditions of usage (ambiguity and redundancy).

Theoretical linguistic approaches have put features and interfaces at the heart of second language acquisition and bilingualism studies (Sorace 2011), while complementary descriptive approaches in L1 data have shown that NPs and OPs are not in free variation but mediated by discourse function and have identified some discourse properties that regulate their distribution (Luján 1985; Zubizarreta 1999). These constraints have also been observed in experimental data (see Quesada 2015 for an overview). In the current study, we aim to contribute to a line of research that offers a more fine-grained exploration of pronoun discourse constraints in extensive texts produced by bilinguals by using empirical data that allows us to make reliable predictions for acquisition. Furthermore, comparing native and bilingual preferences and strategies in pronominal anaphora production can be useful for both descriptive and formal approaches since closely examining empirical data may reveal previously unobserved patterns of usage under particular conditions or in specific contexts and thus contribute to a better understanding of the target language. In carrying out this comparison, it is also our goal to build connections between formal and acquisition research.

1.2 The acquisition of pronominal anaphora in bilingual and L2 speakers

According to previous literature (Licerias 1989; Montrul and Rodríguez-Louro 2006), the syntax of pronouns does not represent a great challenge for acquisition. However, anaphora resolution, that is, assigning a referential meaning to an anaphoric expression by identifying an appropriate and meaningful antecedent, involves the activation of different sources of information from different linguistic levels, which has been said to cause difficulties in learning development, as Sorace's (2011) Interface Hypothesis (IH) assumes. Previous studies have shown that OPs are the most problematic RE for acquisition and that even in native-like levels of proficiency residual optionality, displayed by a redundant use of OPs, is attested (Tsimpli et al. 2004; Sorace 2011). This optionality consists of an overuse of OPs in [-TS] continuity contexts in bilingual and L2 comprehension and production (Paradis and Navarro 2003 for simultaneous English-Spanish bilingualism, Serratrice et al. 2004 for simultaneous English-Italian bilingualism, Tsimpli et al. 2004 for adult English-Greek bilingualism, Lozano 2009 for adult English-Spanish bilingualism). Given that numerous studies have analyzed the acquisition of subject pronouns in

languages differing in the setting of the null subject parameter, the L2 overuse of OPs for [-TS] has been attributed to cross-linguistic influence from the L1 obligatory subject pronoun (Pérez-Leroux and Glass 1999). However, the cross-linguistic explanation is ruled out as the only explicative factor when, in studies involving bilingualism and two null subject languages, a redundant use of OPs is also attested (Margaza and Bel 2006 for adult Spanish-Greek bilingualism, Sorace et al. 2009 for simultaneous Spanish-Italian bilingualism). To explain this “universal” redundant use of OPs, Sorace and collaborators suggested that OPs are used for [-TS] as a strategy to reduce the processing load the online computation of simultaneous syntactic and pragmatic information imposes on individuals that manage more than one language (Tsimplici 2011). On the other hand, bilingual and L2 speakers at very advanced levels of proficiency appear to show NPs at native-like usage (Lozano 2009; García-Alcaraz and Bel 2011; Bel et al. 2016; but see Montrul and Rodríguez-Louro 2006; Domínguez 2013 for divergent results at lower levels of proficiency). As Lardiere (2011) points out, the IH “is about what is left after almost everything else has been acquired”.

To date, little attention has been paid to the production of REs and reference tracking among speakers of two null subject languages either from the experimental perspective (Kras 2008 for Croatian-Italian, Sorace et al. 2009 for Spanish-Italian, Margaza and Bel 2006 and Lozano 2009 for Greek-Spanish), or from corpus studies (García-Alcaraz and Bel 2011 for MA-Spanish). Moreover, it is important to pay attention to a relatively neglected bilingual population: sequential bilinguals. In this paper, we address the unexplored combination of MA-Spanish (two null subject languages from different language families and with attested similar pronoun biases) and analyze NPs and OPs discourse functions in narrative production of early sequential bilinguals. As a null subject language like Spanish, MA has two different forms of pronouns and NPs realize topic continuity and OPs tend to convey topic change (Bel and García-Alcaraz 2015). To the best of our knowledge, this is the first study that focuses on the role of topichood in modulating pronominal referencing in narrative production, and offers the possibility to test patterns of pronoun choice and resolution in different syntactic configurations and variable degrees of ambiguity that, no doubt, will inform current theories of pronominal anaphora.

2 The study: research questions and hypothesis

Based on the previous findings, we formulate two research questions and the subsequent hypotheses. We also address the issue of how linguistic theory and

empirical data from bilingual production can benefit from each other, which is the main topic of this Special Issue.

RQ 1: Do NPs and OPs have complementary discourse functions in Spanish narrative discourse as has been shown in experimental and theoretical approaches? Do discourse configuration features such as sentence linking (intersentential vs. intrasentential) and clause order (main-subordinate vs. subordinate-main within intrasentential contexts) play a role?

Hypothesis 1: Following previous studies (mainly Bel and García-Alcaraz 2015), NPs are expected to be used mainly for [-TS] and OPs for [+TS] in different discourse configurations, being stronger in intrasentential contexts (chiefly in subordinate-main clause order) since the local discourse link between an anaphor and its antecedent in embedded subordinate clauses is much narrower than in independent sentences.

RQ 2: Do MA-Spanish early sequential bilinguals display discourse constraints in pronoun choice like those observed in the Spanish control group? Alternatively, do they show instances of ambiguous or redundant use of pronouns (felicitous condition of usage)? In addition, can the findings in bilingual production allow the current linguistic theory and description of pronoun discourse constraints to be refined?

Hypothesis 2: As their proficiency level is very high and both languages share discourse features in pronoun production, bilinguals will mirror natives in relation to the discourse constraints that govern pronoun distribution. If, on the other hand, handling two languages has consequences for bilingual production (and processing) as some approaches propose (i. e. the IH), pronoun usage patterns will show optionality and will result in the overextension of OPs (i.e. the learner-default form) for more complex pragmatic contexts. Moreover, it could be that strategies and preferences attested in bilingual production can be extended to native production, but to a lesser extent (less used or, even, hidden), given that managing reference tracking through pronouns entails complexity even for natives. This would highlight how the study of bilingualism can contribute to formal and descriptive linguistics.

3 Method

3.1 Participants

Thirty-four MA-Spanish sequential bilinguals and thirty L1 Spanish control speakers participated in the study. MA-Spanish bilinguals are teenagers recruited from a public secondary school in Barcelona. They were born in Spain or in Morocco and moved to Spain at a very young age. Their family language is MA, and Spanish and Catalan are the environmental languages. They all recognize Spanish as the most used language in social and school contexts (see Table 1). According to Montrul (2008), they are all early sequential bilinguals as they were intensively and regularly exposed to Spanish from the age of three (and, in all cases before six).

Table 1: Sociolinguistic features of MA-Spanish bilingual speakers.

	Mean	Range	SD
Age	13.94	12–17	1.44
Arrival Age in Spain	1.34	0–5.5	2.02
Self-reported MA level (out of 100)	88.23	50–100	14.08
Self-reported Spanish level (out of 100)	91.91	75–100	11.87
Average daily use of Spanish in family context (out of 100)	40.44	25–50	12.33
Average daily use of Spanish in school context (out of 100)	92.64	75–100	11.56
Average daily use of Spanish in social context (out of 100)	77.94	50–100	17.14

Participants were asked to produce two texts in oral and written modalities. A judge with specialized expertise in the assessment of Spanish as a second language evaluated the two texts according to the Common European Framework of Reference for Languages (CEFR). Only MA-Spanish bilingual speakers with a C1–C2 Spanish level (advanced user) were included in the study.

L1 Spanish control speakers were university students from the same sociolinguistic context (mean age = 20.68; range 19–25; SD = 3.28) and all were born and raised in Catalonia. They were exposed to Spanish before 3 and self-evaluated their Spanish oral and written skills as excellent. Both MA-Spanish bilinguals and L1 Spanish controls were exposed to the same variety of Spanish. Spanish and Catalan are two languages with similar discourse constraints to determine subject pronoun choice (Bel et al. 2010; Mayol and Clark 2010) so that the fact that both groups also speak Catalan does not seem to interfere in this specific domain.

3.2 Production task

We analyze written (semi) spontaneous narrative production. According to Lozano (2009), both written and oral spontaneous production reflects native and non-native linguistic intuitions regarding subject pronoun choice and expression. Likewise, previous research studying the co-reference patterns of REs in oral and written discourses show parallel distribution of NPs and OPs in both text modalities (Bel et al. 2010; García-Alcaraz and Bel 2011). Our corpus is a sample of a wider corpus that includes oral and written narrative production texts from MA-Spanish and Amazigh-Spanish bilinguals with different proficiency levels of Spanish (Bel 2013).

Following the methodology and procedure of Berman's (2008) international project on cross-linguistic comparison in literacy development, participants were shown a three-minute silent film twice. The film, used as a video

prompt to elicit production, depicted different scenes of interpersonal conflicts at school. They were then asked to write down a similar story that had happened to a friend. By focusing on discourse non-present referents and not on personal stories is in line with the desire to elicit a large number of third person subject pronouns.

3.3 Data treatment: transcription and codification

The texts elicited were transcribed in the CHILDES system (MacWhinney 2000). The original spelling was preserved as well as all the lexical, morphological and spelling errors and were marked according to CHAT conventions (see Bel 2013). Once transcribed, the texts were divided into clauses. According to Berman and Slobin (1994), each clause involves a unified predicate that expresses a unique situation (activity, event or state) and, therefore, it necessarily includes a subject (null or overt).

All third-person (animate) REs were isolated and codified with the following exceptions: REs containing universal or existential quantifiers, REs with generic interpretation, relative clauses where the relative pronoun functioned as a subject and emphatic and contrastive OPs. In all these cases NPs and OPs do not adjust to the recovery conditions analyzed in this study. For each analyzable subject, a dependent line was created and specific codes were introduced (see Table 2). Since only forward anaphora was addressed –the occurrence of backward anaphora was almost non-existent–, the searching for an antecedent started in the preceding clause, and then it continued up through the different clauses until it was found. To control for inter-rater reliability, two independent transcribers coded 20% of the narratives, corresponding to six full transcriptions (three from each group). For the control group, Cohen's kappa was $\kappa=0.876$ ($p<0.001$) for discourse function and $\kappa=0.788$ ($p<0.001$) for access to antecedent. Discordant rates were discussed to get consensus.

Following Zubizarreta's (1999) proposal, we understand topic as the theme -i.e. what the sentence is about and appears in sentence initial position. There is a high degree of overlapping between a subject and a topic, as shown in (1). However, looking carefully at 'real' discourse may provide insights into the intricate interaction between these two roles and help us realize that they are not always interchangeable. For example, in example (2), the topic, the dislocated object ('*Carmen*'), introduced by the typical Spanish accusative marker '*a*', and the subject ('*Manuel*') are dissociated. In this case

Table 2: Coding criteria.

Morphosyntactic form of the RE	
DP	Determiner Phrase
OP	Overt Subject Pronoun
NP	Null Subject Pronoun
Discourse function of the RE	
–TS	Topic continuity (or topic maintenance)
+TS	Topic shift (or topic change)
Sentence relationship between the clauses that contain the pronoun and the antecedent	
TRA	Intrasentential
TER	Intersentential
Type of clause where the anaphoric pronoun was found	
M	Main clause
S	Subordinate clause
C	Coordinate clause
Clause order (applies only to intrasentential contexts)	
MS	Main-Subordinate order
SM	Subordinate-Main order
Access to the antecedent	
CA	Clear Antecedent
AA	Ambiguous Antecedent

the *pro* refers to an object antecedent (Carmen) and its discourse function is [–TS], since *a Carmen* is the topic of the previous clause.

- (1) Raquel_i nunca corregía a Julián_j porque *pro*_i era muy vergonzosa.
‘Raquel never corrected Julian because *pro* was very shy’
- (2) A Carmen_i, Manuel_j siempre la consolaba cuando *pro*_i estaba triste.
‘Carmen, Manuel always consoled her when *pro* was sad’

Examples of topic continuity and change as well as felicitous and infelicitous (ambiguous or redundant) pragmatic contexts from our corpus are explained in detail in Section 5 through examples (3)–(5).

3.4 Corpus description

The corpus consists of 64 written narrative texts (30 from Spanish speakers and 34 from bilinguals) whose general features are summarized in Table 3.

Table 3: General features of narratives texts.

	Spanish L1			MA-Spanish bilinguals		
	Mean	Range	SD	Mean	Range	SD
Number of <i>tokens</i> ^a per text	258.50	91–419	77.92	114.73	56–201	38.01
Number of <i>types</i> ^b per text	151.50	61–232	40.37	67.09	34–110	19.35
Number of clauses	41.33	17–79	13.96	21.73	11–39	7.46
Type/Token Ratio	0.59	0.47–0.70	0.05	0.58	0.44–0.71	0.07

^aTotal number of words.

^bUnrepeated number of words.

L1 speakers produced longer texts than bilinguals did as shown by two independent t-test comparing number of tokens ($t(62) = 9.553$; $p = 0.001$) and number of clauses ($t(62) = 7.118$; $p = 0.001$). Text richness, as measured by number of types, yielded a significant difference ($t(62) = 10.867$; $p = 0.001$); however, Type/Token Ratio, which is also a measurement of lexical richness, did not reveal differences between groups ($t(62) = 0.108$; $p = 0.914$).

4 Results

4.1 Production of referring expressions

The total number of third-person REs in subject position were 876, a substantial number showing that narratives and the specific instruction given promoted a large amount of target REs. Comparing groups, L1 speakers are more productive than bilinguals (they produced 518 and 358 REs, respectively; Table 4). This finding correlates with differences in text length, as measured by the average number of

Table 4: Frequencies (and percentages) of third person subject REs.

	Determiner Phrase	Null pronoun	Overt pronoun	Total subject productions
L1 Spanish	194 (37.45 %)	302 (58.30 %)	22 (4.25 %)	518 (100 %)
Bilingual Spanish	124 (34.64 %)	208 (58.10 %)	26 (7.26 %)	358 (100 %)

clauses by text (L1 speakers = 41.33 vs. bilinguals = 21.73; Table 3). On the other hand, the average realization and the overall distribution of the three REs considered at this level of the analysis are similar for both groups as confirmed by a two-way independent Chi-square analysis ($\chi^2 = 3.977$, $df = 2$, $p = 0.137$).

Data from Table 4 also shows that NPs are the most represented RE and that OPs are uncommon, resulting in a very limited number of occurrences. These findings fit with descriptive studies (Luján 1985; Fernández-Soriano 1999) that have identified NPs as the non-marked pronominal option (the linguistic-default option), while OPs are considered marked and the learner-default option (Tsimpli 2011) as explained earlier. In between both pronouns, we find DPs. A Chi-square analysis comparing NPs and overt REs (OPs and DPs) revealed no differences between groups ($\chi^2 = 0.003$, $df = 1$, $p = 0.953$). The rationale behind the grouping of these two overt REs was to check IH's predictions, according to which bilingual speakers rely on the overuse of explicit REs to override processing load when they face the task of selecting the appropriate anaphor; however, this pattern is not evident in our corpus at this level of analysis.

4.2 Discourse trends of null and overt pronouns in L1 and bilingual Spanish

In this section, we ask whether NPs are specialized to express continuity, understood as [-TS], and OPs to express change of reference, i.e. [+TS], across clauses. Due to the low number of OPs and, hence, the impossibility to perform parametric tests, the statistics for each pronoun are displayed separately.

4.2.1 Null pronouns

Tables 5 and 6 include the results for each discourse function, [\pm Topic-Shift], separately for intrasentential and intersentential contexts. As can be seen, NPs in L1 and bilingual Spanish entail [-TS] (always higher than 64%) regardless of the dependence relationship between the clauses containing the anaphor and its antecedent. Surprisingly, both groups also use NPs to convey topic reintroduction, i.e. [+TS], at around 30–35% of the cases in both inter- and intrasentential contexts; such discourse function is unexpected for these pronouns, as many authors have claimed, and appears to entail an infelicitous, or ambiguous, usage. We will discuss this matter later in further detail.

To confirm the earlier observations, the values for each discourse function in both sentence relationships were submitted to a repeated-measures ANOVA with

Table 5: Frequencies (and percentages) of NPs' discourse functions at intrasentential level by group.

Group	Intrasentential contexts	
	-Topic Shift	+Topic Shift
L1 Spanish	49 (69.01 %)	22 (30.99 %)
Bilingual Spanish	37 (69.81 %)	16 (30.19 %)

Table 6: Frequencies (and percentages) of NPs' discourse functions at intersentential level by group.

Group	Intersentential contexts	
	-Topic Shift	+Topic Shift
L1 Spanish	149 (64.50 %)	82 (35.50 %)
Bilingual Spanish	101 (65.16 %)	54 (34.84 %)

Discourse function ([-TS] vs. [+TS]) and Sentence relationship (intrasentential vs. intersentential) as within-subject factors and Group (L1 Spanish vs. bilingual Spanish) as between-subject factor. All parametric inferential statistics were performed based on arcsine transformed mean proportions. Results revealed a main effect of Discourse function ($F(1, 49) = 9.565$; $p = 0.003$) and Sentence relationship ($F(1, 49) = 26.558$; $p = 0.001$), meaning that NPs are specialized for [-TS] in both groups and that the type of sentence dependency plays a role in how NPs are internally distributed but does not determine a different discourse specialization.

Furthermore, as stated above, a percentage of NPs higher than expected were used for [+TS] (around 30–35%), which might lead to ambiguity. To go deeper into this issue, a more detailed analysis of NPs within [+TS] contexts was developed. In Tables 7 and 8, we distinguish the number of NPs used for [+TS] that

Table 7: Frequencies (and percentage) of NPs for [+TS] depending on the access to the antecedent at intrasentential level by group.

Group	Intrasentential contexts	
	Clear antecedent	Ambiguous antecedent
L1 Spanish	22 (100 %)	0 (0 %)
Bilingual Spanish	15 (93.75 %)	1 (6.25 %)

Table 8: Frequencies (and percentage) of NPs for [+TS] depending on the access to the antecedent at intersentential level by group.

Group	Intersentential contexts	
	Clear antecedent	Ambiguous antecedent
L1 Spanish	82 (100 %)	0 (0 %)
Bilingual Spanish	45 (83.33 %)	9 (16.67 %)

have a clear antecedent or an ambiguous antecedent for sentence relationship and group. In the former case, the existence of only one antecedent –or several antecedents, of which only one matches the phi-features of the NP at issue–in the preceding discourse leaves no room for any ambiguity as the referential content of the NP can be unequivocally retrieved by linking it to that antecedent (see [3] in Section 5 for an example). In the latter case, the presence of several referents in the current discourse that match phi-features of the NP being solved may lead to ambiguity if, additionally, the interlocutor (i.e. the reader) cannot rely on semantic or pragmatic clues that allow clear selection and identification of the appropriate antecedent (see [4] in Section 5 for an example of this scenario). From the data on Tables 7 and 8 we see, in fact, that the total number of true ambiguous cases is rather low and only affects the bilingual group.

Data from Tables 7 and 8 shows that L1 Spanish speakers resort to NPs for [+TS], regardless of the sentence relation, only when the reader can recover the antecedent unambiguously since no ambiguous reference is attested. However, NPs for [+TS] are not always used by bilinguals to recover an unequivocal antecedent. A repeated-measures ANOVA with two within-subject factors (Type of antecedent [clear vs. ambiguous] and Sentence relationship [intrasentential vs. intersentential]) and Group as between-subject factor confirmed the main effect of Type of antecedent ($F(1, 26) = 374.319$; $p = 0.001$) and the interaction Type of antecedent \times Group ($F(1, 26) = 4.518$; $p = 0.043$).

To summarize, L1 speakers employ NPs for [+TS] only when the antecedent is easily accessible and clearly retrievable; in contrast, bilingual speakers appear not to have a complete command of the whole set of specific discourse conditions that allow NPs in topic-change contexts. Rather bilingual speakers seem to display an overproduction of NPs for [+TS] with non-accessible antecedents, which results in referential ambiguity that could lead to misunderstandings. This issue will be exemplified and discussed further in Section 5 with example 5.

At the intrasentential level, main-subordinate and subordinate-main clause orders were analyzed to ascertain whether variable order plays a role in the discourse function that NPs convey. Data in Tables 9 and 10 show that, in both

Table 9: Frequencies (and percentages) of NPs and discourse function at intrasentential level by group (main-subordinate).

Group	Main-subordinate clause order	
	-Topic Shift	+Topic Shift
L1 Spanish	33 (62.26 %)	20 (37.74 %)
Bilingual Spanish	30 (66.67 %)	15 (33.33 %)

Table 10: Frequencies (and percentages) of NPs and discourse function at intrasentential level by group (subordinate-main).

Group	Subordinate-main clause order	
	-Topic Shift	+Topic Shift
L1 Spanish	16 (88.89 %)	2 (11.11 %)
Bilingual Spanish	7 (87.5 %)	1 (12.5 %)

clause orders, NPs are mainly used to maintain a topic. However, while in subordinate-main clause order there is a sharp trend favoring [-TS] function, the results are more flexible in main-subordinate order since NPs (more than one-third) are even used for topic shifting.

A repeated-measures ANOVA performed with Discourse function ([-TS vs. +TS]) and Clausal (main-subordinate vs. subordinate-main) order as within subject factors and Group as a between-subject factor yielded only a main effect of Discourse function ($F(1, 9) = 16.645$; $p = 0.003$). These findings bring to light that NPs, in both L1 and bilingual Spanish, are mainly used for [-TS] at the intrasentential level despite the relative order of clauses.

4.2.2 Overt pronouns

Due to the low number of occurrences of OPs, data was analyzed through the non-parametric Fisher exact test and only regarding sentence linking. According to data in Tables 11 and 12, a division of work among OPs emerges as they are used for [-TS] at intrasentential level and for [+TS] at intersentential level both in L1 and bilingual Spanish.

The number of OPs produced by both groups for [+TS] and [-TS] at both levels were submitted to inferential statistics that did not reveal significant differences between L1 and bilingual speakers neither at intrasentential

Table 11: Frequencies (and percentages) of OPs and discourse function at intrasentential level by group.

Group	Intrasentential contexts	
	–Topic Shift	+Topic Shift
L1 Spanish	3 (100%)	0 (0%)
Bilingual Spanish	6 (85.72%)	1 (14.28%)

Table 12: Frequencies (and percentages) of OPs and discourse function at intersentential level by group.

Group	Intersentential contexts	
	–Topic Shift	+Topic Shift
L1 Spanish	4 (21.05%)	15 (78.95%)
Bilingual Spanish	5 (26.32%)	14 (73.68%)

($p = 0.999$) nor intersentential ($p = 0.723$) level. Therefore, in the subsequent analysis both samples were grouped together. The frequency of OPs for [+TS] and [–TS] in both sentence relationships were compared and a significant difference ($p = 0.001$) was obtained, meaning that the type of sentence dependency that mediates between the pronoun and the antecedent influences OPs' discourse specialization. Briefly, OPs are mainly used for [+TS] at intersentential level and for [–TS] at intrasentential level. Although data from OPs is very limited and does not allow us to draw strong conclusions, it does show some discourse tendencies.

5 Discussion

RQ 1: Do NPs and OPs have complementary discourse functions in Spanish narrative discourse as has been shown in experimental and theoretical approaches? Do discourse configuration features such as sentence linking (intersentential vs. intrasentential) and clause order (main-subordinate vs. subordinate-main within intrasentential contexts) play a role?

Inferential statistics showed that NPs are mainly used to maintain a topic in L1 Spanish, regardless of sentence dependency and clause order, and that OPs display divergent discourse functions in intersentential [+TS] and

in intrasentential [-TS] contexts. Our data also reveals that NPs are used for [+TS] at around 31–35 % (intra- and intersentential, respectively), a percentage higher than expected according to previous studies (Tsimplici et al. 2004; Sorace 2011). A closer inspection to NPs conveying [+TS] in L1 Spanish reveals that this usage is not illicit since it does not generate referential ambiguity when the antecedents are accessible and can be easily retrieved. As Ariel (1990) predicts, the more salient a referent is the more likely speakers tend to use minimally phonetic forms, which explains the “overuse” of NPs detected. These findings are consistent with those in Licerias et al. (2010), as they argue that NPs can reintroduce a topic in specific discourse conditions that rule out ambiguity (see our discussion in Section 4.2.1). Bel and Albert (2016) also find non-illicit NPs for the same discourse function in spontaneous production among Spanish speakers of different ages and it is a recurrent pattern reported in production that should not be disregarded in further study of anaphora resolution. In a similar vein, Lozano (2016: 255) notices that “natives sporadically ‘violate’ pragmatic rules by being redundant and ambiguous”, mainly when they can rely on alternative mechanisms, such as contextual information, to facilitate anaphor interpretation. Not all authors agree on the different roles that NPs can convey; for example, Montrul and Rodríguez-Louro (2006) consider all NPs for [+TS] anomalous, but according to our data, this proposal is quite a simplification. Example 3 provides a discourse excerpt with an unambiguous NP expressing topic shift given that antecedent identification can rely on morphological cues.

- (3) Example of non-ambiguous NP for [+TS]. L1 female Spanish speaker (21 years old)
- (1) *STU: [...] El día anterior había llovido.
‘The day before was raining’
 - (2) *STU: el suelo estaba resbaladizo.
‘the floor was slippery’
 - (3) *STU: y alguna que otra vez pro_i se había resbalado.
‘and occasionally *pro* slipped’
 - (4) *STU: mojada, pero satisfecha, pro_i subió al coche.
‘wet, but satisfied, *pro* got into the car’
 - (5) *STU: y pro_i apresuró a mis padres_j.
‘and *pro* hurried my parents’
 - (6) *STU: para que pro_i la_i llevaran a la ciudad.
‘in order *pro* drove her into the city’

In this passage, a female protagonist of the narrated story is reintroduced in clause 3 by means of a NP (*pro*) expressing the current sentence topic, different from the previous one. Given that in the immediate previous sentences no other characters are presented, the protagonist is still active in the writer and, allegedly, the reader's mind. Hence, the NP reference is clear and easily retrievable albeit it conveys a topic shift. In the following clauses 4 and 5, the topic is maintained through an NP, by realizing the [-TS] feature typical of this kind of pronoun. Interestingly, in subordinate clause 6 another NP embodies a new change in sentence topic but, again, it is not ambiguous because its phi-features (evidenced from verb plural marking) unequivocally identify a referring expression ('*my parents*') in the preceding clause as its antecedent. As the antecedent occurs in a non-topic position and the NP in a topic position, then the NP realizes a topic shift.

- (4) Example of non-ambiguous NP for [+TS]. L1 female Spanish speaker (20 years old)
- (1) *STU: [...] Una vez este chico_i se metió con mi hermano_i.
'Once this guy teased my brother'
- (2) *STU: y mi hermano_j lo_j estampó contra la pared.
'and my brother crashed him into the wall'
- (3) *STU: y pro_j le_j rompió la muñeca.
'and *pro* broke his wrist'
- (4) *STU: bueno, pro_k no sé.
'well, *pro* don't know'
- (5) *STU: si pro_j se_j la rompió.
'if *pro* broke it'
- (6) *STU: pero al día siguiente pro_i la llevaba enyesada [...].
'but the following day *pro* had the wrist in plaster'

Example (4) provides an instance of a non-ambiguous NPs used for [+TS] in clause 5 since semantic clues play an essential role in reference retrieving. Clauses 1 and 2 alternate between two different topics, '*the guy*' and '*the brother*' respectively. In clause 3 a NP, *pro*, maintains the preceding topic; clause 4 contains a NP whose referential content is unambiguously retrieved by means of morphological cues (the verb ending, '*sé*', signals first person), as in example (3). Then, in clause 5, another NP introduces an unexpected change in topic sentence but its referential content is not compromised as it is easily interpreted as co-referent with only one of the competing antecedents given the predicate

semantic information: ‘*the brother*’, who broke the wrist. Again, in clause 6, a NP represents a new shift in topic sentence: in this case, the predicate semantics visibly signals ‘*the guy*’, who is the only individual with a broken wrist, as co-referent with the NP.

In short, despite the main pragmatic function of NPs to maintain the topic, they can also be used to felicitously change the topic under specific discourse conditions that do not disturb antecedent retrieving over shifts of topic across clauses. Pragmatic plausibility may modulate the interpretative preferences of NPs -usually used to convey topic continuity- in that they can also be engaged in topic shift contexts. To summarize, the conditions that ‘cancel’ the pragmatic [-TS] feature of NPs, thus favoring an unexpected twist in topic tracking, are mainly two: i) the existence of intervening competing antecedents of different (gender or number, or even person) phi-features to those of the anaphor, allowing linkage to the appropriate existing antecedent matching NP phi-features, and ii) the semantic plausibility provided by the predicate that comes with the anaphor being solved by the inferences from contextual information that better fit with a reasonable reference resolution. By this means, ambiguity is prevented as shown by the results in Tables 7 and 8.

Lozano (2016) provides results along the same lines for NPs but he goes one step further in claiming that NPs are the preferred RE for topic change, so our findings are not so surprising. He eventually observes that speakers are sporadically redundant and use OPs as a disambiguating device when two different gender antecedents are at play. Again, such a strategy does not occur in our corpus. It is worth mentioning that speakers can resort to a full RE, a DP, to bypass referential uncertainty. This broader tracking reference device goes beyond our aims.

Because the number of occurrences of OPs is very limited, we need to be cautious when drawing conclusions. In brief, OPs are typically used for [+TS], as observed in previous corpus studies (Bentivoglio 1987; Lozano 2009). This applies mainly to the intersentential level as the number of instances in intrasentential contexts is negligible. Given this picture, redundancy linked to OPs is absent from our corpus.

In its original form, Carminati’s (2002) PAH was proposed as a processing strategy, at a structural level in intrasentential contexts, to explain the division of labor between NPs, used for [-TS], and OPs, used for [+TS]. Although our results do not exactly replicate these findings in the case of OPs, concurrent patterns of usage are attested across contexts in the case of NPs, as they appear to be very steady and constant in conveying topic continuity. This could mean that the principles that regulate NPs’ distribution and co-reference preferences are not constrained by structural requirements but rather by

pragmatic-contextual properties and processing issues. In our corpus study, similar findings have been found in both sentence relations (intrasentential and intrasentential) as well as in both clause orders (main-subordinate-subordinate-main) suggesting that the PAH goes beyond the structural level and that it may be possible to extrapolate to a larger discourse level. It is revealing that evidence coming from careful and reliable experimentation, using a continuation story task, has shown that pronoun choice “is influenced by the likelihood that the referent is the current topic” (Rohde and Kehler 2014). In sum, we draw on the relevant theoretical assumptions of intersentential and null and overt pronouns’ constraints but it is evident that our findings based on empirical data from a language corpus go one step further, uncovering some regularities across types of sentence linking that might benefit the understanding of Spanish pronouns and by extension have implications for linguistic theory.

RQ 2: Do MA-Spanish early sequential bilinguals display discourse constraints in pronoun choice like those observed in the Spanish control group? Alternatively, do they show instances of ambiguous or redundant use of pronouns (felicitous condition of usage)? In addition, can the findings in bilingual production allow the current linguistic theory and description of pronoun discourse constraints to be refined?

Our results indicate that Spanish native speakers and MA-Spanish early sequential bilinguals use NPs and OPs in a similar way. Thus, overall, these findings illustrate that bilingual speakers are sensitive to the discourse constraints that regulate the distribution of NPs and OPs and mirror native preferences. Both groups use NPs for [-TS], regardless of type of sentence relationship and clause order, and OPs for [+TS], especially in intersentential contexts, where more instances were attested in our corpus. Additionally, but slightly different from natives, bilingual speakers use NPs in an unexpected flexible way that gives rise to their emergence in topic changing contexts, both at intrasentential and intersentential levels. These results contrast with Blackwell and Quesada’s (2012: 162) findings, who found evidence that native Spanish speakers choose more minimal forms than L2 speakers, who gradually replace elaborate forms with pronouns and, then, OPs with NPs. A possible explanation for this divergence could be language proficiency, since they studied beginner, intermediate and advanced speakers of Spanish, whereas we included near-native speakers of Spanish (born and/or raised in Spain), who seems to perform more native-like. However, additional analyses have shown that our bilinguals have not gained full mastery in this area yet: bilinguals produced more ambiguous NPs than Spanish natives did

(see the excerpt in [5]). This lack of control¹ of subtle discourse-pragmatic features corroborates previous findings on bilinguals (García-Alcaraz and Bel 2011; Bel et al. 2016; for the current language pair, or Kras 2008 for Croatian-Italian bilinguals, crucially all null subject languages).

- (5) Example of ambiguous NP for [+TS]. MA-Spanish bilingual female (15 year old)
- (1) *STU: una vez pro_i me enfade [: enfadé] con una amiga mia_i [: mía].
‘Once *pro* got upset with a friend of mine’
- (2) *STU: porque pro_i le_k iba [: iba] diciendo a mi mejor amiga_k.
‘because *pro* was telling to my best friend’
- (3) *STU: que yo_i era una falsa.
‘that I was insincere’
- (4) *STU: y pro_{2j/2k} me insultaba.
‘and *pro* insulted me’
- (5) *STU: y yo_i no me di cuenta.
‘and I didn’t notice it’

In this excerpt, three REs are involved: the narrator (1st person *pro* in clause 1), ‘a friend of mine’ and ‘my best friend’. In clause 2, the NP changes the topic unambiguously and in clause 3 the OP introduces a new topic shift. In clause 4, the NP in subject position is predicted to be indicating topic continuity (interpretation provided by the subindex _i) but this reading contradicts the inference encoded in the event described by the predicate information and, then, it is ruled out. Alternatively, it is interpreted as instantiating a topic change and this seems to be the case; however, there are two potential antecedents (‘a friend of mine’ and ‘my best friend’) but clues are not clear enough to guide disambiguation towards one of the two competing elements and co-reference cannot be solved; the subsequent discourse does not help to disambiguate either.

As discussed earlier, the constraints that allow NPs for [+TS] are very restrictive since they must prevent full recovery of referential content; it is worth remembering that morphological, semantic and contextual cues can act as disambiguating cues, as Pešková (2014) highlights.

¹ As proposed by an anonymous reviewer, the hypothesis that Spanish speakers use subject pronouns in accord with both syntactic and pragmatic constraints and that bilingual speakers do not, could be influenced, somehow, by their age and writing skills: university students (Spanish speakers) vs. high school students (bilingual speakers).

Considering that bilingual speakers produce more illicit NPs as [+TS] than native speakers, our results would seem to support instability in pronominal choice, a syntax-pragmatic phenomenon. Nevertheless, and curiously enough, the locus of indeterminacy, or residual optionality, in bilingual discourse is not directly associated with OPs, as predicted by the IH, but with NPs. Note that the IH, as it is formulated to date, could not explain our findings. Consequently, we would like to contribute to linguistic theory proposing that the phenomenon of indeterminacy in pronominal anaphora choice –and resolution– do not depend on the type of pronoun, as defended up to now, but rather on the [+TS] discourse feature. In our proposal, the feature [+TS], in comparison with [–TS], would be marked and, therefore, could represent a challenge for acquisition. Hendriks (2014) argue that changing the topic is pragmatically more challenging in cognitive terms since this activity seems to consume more cognitive resources than maintaining the topic. So, in parallel with NPs as the linguistic-default form, and OPs as the learner-default option (Tsimplici 2011), the [–TS] discourse feature would represent the discourse-default choice.

Focusing on OPs, the analyses performed showed that they are mainly used for [+TS] in bilingual and L1 discourse at the intersentential level, the most common contexts in our corpus (the number of occurrences in intrasentential contexts is insignificant). Thus, the overproduction of OPs for [–TS] entailing redundancy, as the IH and previous studies predict (Lozano 2009; García-Alcaraz and Bel 2011), is not present in our bilingual written corpus, and then the alleged learner-default option proposed to supersede processing pressure in production processes in L2 learners (Tsimplici 2011) does not appear to be operative here. Lozano's (2016) Pragmatic Principles Violation Hypothesis is not confirmed either, as our bilingual corpus includes more ambiguous NPs than redundant OPs. In contrast, our findings are in line with those in Montrul and Rodríguez-Louro (2006), Domínguez (2013) and Bel et al. (2016), which show that bilingual and L2 speakers can achieve a native-like management of OPs' discourse constraints. Of course, these observations must be taken cautiously as they are based on a small, and maybe heterogeneous, sample. This is of particular concern when dealing with infrequent language facts in corpus studies, as is the case of OPs; as noticed by some variationist studies, the proportion of third subject pronoun expression in Peninsular Spanish is not higher than 10–11% (Silva-Corvalán 2014). Nonetheless, despite the low number of OPs, tendencies can be perceived.

Finally, since one of the goals of this article is to connect linguistic theory and empirical data from bilingual speakers, we would like to point out that, although the discourse properties that regulate the distribution of pronouns in Spanish have been identified in the linguistic literature, some specific

constraints have still not been evaluated. The most important and novel finding in relation to the organization of pronouns in bilingual discourse production in Spanish is that the switch-reference function is not only restricted to OPs -or other full REs- but that NPs can also convey reference change; as pointed out in a previous paragraph this result is not in accordance with theoretical assumptions. This finding, together with our proposal that [-TS] is the unmarked option of the [\pm TS] feature both in native and bilingual discourse, opens new avenues of inquiry for language description and illustrates how empirical data from bilinguals, as well as natives, can further linguistic theory.

6 Conclusions

Using a corpus of written narrative texts, we have found an unnoticed flexible usage of NPs overall both in L1 and bilingual Spanish linked to [-TS], which contradicts previous findings and led us to speculate that this function is the least costly and unmarked option in discourse. We argue that anaphora resolution in MA-Spanish early sequential bilinguals, a language combination unexplored so far, is mostly native-like, except for the control of ambiguity. We also propose that investigating corpus data of language production and developing a fine-grained codification system might shed light on how NPs and OPs' discourse constraints work in real and extensive texts and might help to discover uses that might not have been noticed otherwise. In other words, the study of Spanish subject pronouns is an extremely fruitful topic that clearly demonstrates how important it is to build bridges between linguistic theory and research on (bilingual) language acquisition. In the current study, only written data was analyzed. Hence, future research should complement it with the analysis of spoken corpus.

Acknowledgments: We sincerely thank the editor of this special volume as well as the three anonymous reviewers of this paper for their valuable comments, which helped to improve the final version of this manuscript. Special thanks are due to our participants, since without them this work would not have been possible.

Funding: This research was funded by the Spanish Ministry of Education through the FPU program (Estela García-Alcaraz) and two research projects VARIAD FFI2012-35058 and CONTACT FFI2016-75082-P (Principal Investigator: Aurora Bel).

References

- Alonso-Ovalle, Luis, Susana Fernández-Solera, Lyn Frazier & Charles Clifton. 2002. Null vs. overt pronouns and the topic-focus articulation in Spanish. *Rivista di Linguistica* 14. 151–169.
- Ariel, Mira. 1990. *Accessing noun phrase antecedents*. London: Routledge.
- Bel, Aurora. 2013. *BCN-L2 Spanish corpus (Moroccan Arabic/Berber-Spanish)*. Talkbank database. <http://talkbank.org/access/SLABank/Spanish/BCN.html>.
- Bel, Aurora & Milagros Albert. 2016. The development of referential choice in Spanish narratives among school-age children and adolescents. In Joan Perera, Melina Aparici, Elisa Rosado & Naymé Salas (eds.), *Written and spoken language development across the lifespan*, 251–269. Frankfurt: Springer.
- Bel, Aurora & Estela García-Alcaraz. 2015. Subjects in the L2 Spanish of Moroccan Arabic speakers: Evidence from bilingual and second language learners. In Tiffany Judy & Silvia Perpiñán (eds.), *The acquisition of Spanish as a second language: Data from understudied languages pairings*, 201–232. Amsterdam: John Benjamins.
- Bel, Aurora, Estela García-Alcaraz & Elisa Rosado. 2016. Reference comprehension and production in bilingual Spanish: The view from null subject languages. In Anahí Alba de la Fuente, Elena Valenzuela & Cristina Martínez-Sanz (eds.), *Language acquisition beyond parameters. Studies in honour of Juana M. Liceras*, 37–70. Amsterdam: John Benjamins.
- Bel, Aurora, Joan Perera & Naymé Salas. 2010. Anaphoric devices in written and spoken narrative discourse: Data from Catalan. *Written Language and Literacy* 13(2). 236–259.
- Bentivoglio, Paola. 1987. *Los sujetos pronominales de primera persona en el habla de Caracas*. Caracas: Central University of Venezuela.
- Benveniste, Émile. 1966. *Problèmes de linguistique générale*. Paris: Gallimard.
- Berman, Ruth A. 2008. The psycholinguistics of developing text construction. *Journal of Child Language* 35. 735–771.
- Berman, Ruth A. & Dan I. Slobin. 1994. *Relating events in narrative: A cross-linguistic developmental study*. Hillsdale: Lawrence Erlbaum.
- Blackwell, Sarah E. & Margaret Lubbers Quesada. 2012. Third-person subjects in native speakers' and L2 learners' narratives: Testing (and Revising) the givenness hierarchy for Spanish. In Kimberly Geeslin & Manuel Díaz-Campos (eds.), *Selected proceedings of the 14th Hispanic Linguistics Symposium*, 142–164. Somerville, MA: Cascadilla Proceedings Project.
- Carminati, Maria Nella. 2002. *The processing of Italian subject pronouns*. Amherst, MA: University Massachusetts Amherst dissertation.
- Carvalho, Ana Maria, Rafael Orozco & Naomi Lapidus Shin. 2015. Introduction. In Ana Maria Carvalho, Rafael Orozco & Naomi Lapidus Shin (eds.), *Subject pronoun expression in Spanish: A cross-dialectal perspective*, xiii–xxvi. Washington, DC: Georgetown University Press.
- Domínguez, Laura. 2013. *Understanding interfaces: Second language acquisition and native language attrition of Spanish subject realization and word order variation*. Amsterdam: John Benjamins.
- Fernández Soriano, Olga. 1999. El pronombre personal. Formas y distribuciones. Pronombres átonos y tónicos. In Ignacio Bosque & Violeta Demonte (dir.), *Gramática Descriptiva de la Lengua Española*, 1275–1315. Madrid: Espasa-Calpe.
- Filiaci, Francesca. 2011. *Anaphoric preferences of null and overt subjects in Italian and Spanish: A cross-linguistic comparison*. Edinburgh: University of Edinburg dissertation.

- García-Alcaraz, Estela & Aurora Bel. 2011. Selección y distribución de los pronombres en el español L2 de los hablantes de árabe. *Revista de Lingüística y Lenguas Aplicadas* 6. 165–179.
- Hendriks, Petra. 2014. *Asymmetries between Language Production and Comprehension*. Dordrecht: Springer.
- Kras, Tihana. 2008. Anaphora resolution in near-native Italian grammars: Evidence from native speakers of Croatian. In Leah Roberts, Florence Myles & Annabelle David (eds.), *EUROSLA Yearbook 8*, 107–134. Amsterdam: John Benjamins.
- Lardiere, Donna. 2011. Who is the interface hypothesis about?. *Linguistic Approaches to Bilingualism* 1. 48–53.
- Liceras, Juana M. 1989. On some properties of the pro-drop parameter: Looking for missing subjects in non-native Spanish. In Susan M. Gass & Jacquelyn Schachter (eds.), *Linguistic perspective on second language acquisition*, 109–133. Cambridge: Cambridge University Press.
- Liceras, Juana M., Anahí Alba de la Fuente & Cristina Martínez Sanz. 2010. The distribution of null subjects in non-native grammars: Syntactic Markedness and Interface Vulnerability. In Michael Iverson, Ivan Ivanov, Tiffany Judy, Jason Rothman, Roumyana Slabakova & Marta Trizna (eds.), *Proceedings of the mind/context divide workshop*, 84–95. Somerville, MA: Cascadilla Proceedings Project.
- Lozano, Cristóbal. 2009. Selective deficits at the syntax-discourse interface: Evidence from the CEDEL2 corpus. In Neal Snape, Yan-kit Ingrid Leung & Michael Sharwood-Smith (eds.), *Representational deficits in SLA*, 127–166. Amsterdam: John Benjamins.
- Lozano, Cristóbal. 2016. Pragmatic principles in anaphora resolution at the syntax-discourse interface: Advanced English learners of Spanish in the CEDEL2 corpus. In Margarita Alonso Ramos (ed.), *Spanish learner corpus research: State of the art and perspectives*, 236–265. Amsterdam: John Benjamins.
- Luján, Marta. 1985. Binding properties of overt pronouns in null pronominal languages. *Chicago Linguistic Society* 21. 424–438.
- MacWhinney, Brian. 2000. *The childe project: Tools for analyzing talk*. Mahwah, NJ: Lawrence Erlbaum.
- Margaza, Panagiota & Aurora Bel. 2006. Null subjects at the syntax-pragmatics interface: Evidence from Spanish interlanguage of Greek speakers. In Mary Grantham O'Brien, Christine Shea & John Archibald (eds.), *Proceedings of the 8th generative approaches to second language acquisition conference*, 88–97. Somerville, MA: Cascadilla Press.
- Mayol, Laia & Robin Clark. 2010. Overt pronouns in Catalan: Games of partial information and the use of linguistic resources. *Journal of Pragmatics* 42. 781–799.
- Montrul, Silvina. 2008. *Incomplete acquisition in Bilingualism. Re-examining the age factor*. Amsterdam: John Benjamins.
- Montrul, Silvina & Celeste Rodríguez Louro. 2006. Beyond the syntax of the null subject parameter: A look at the discourse-pragmatic distribution of null and overt subjects by L2 learners of Spanish. In Vincent Torrens & Linda Escobar (eds.), *The acquisition of syntax in romance languages*, 401–418. Amsterdam: John Benjamins.
- Paradis, Johanne & Samuel Navarro. 2003. Subject realization and crosslinguistic interference in the bilingual acquisition of Spanish and English: What is the role of input?. *Journal of Child Language* 30. 1–23.
- Pérez-Leroux, Ana T. & William R. Glass. 1999. Null anaphora in Spanish second language acquisition: Probabilistic versus generative approaches. *Second Language Research* 15. 220–249.

- Pešková, Andrea. 2014. Information structure and the use of pronominal subjects in Spanish. In Dina El Zarka & Steffen Heidinger (eds.), *Methodological issues in the study of information structure*, 43–67. Graz: Grazer Linguistische Studien.
- Prada Perez, Ana. 2009. *Subject expression in Minorcan Spanish: Consequences of contact with Catalan*. Pennsylvania, PA: Pennsylvania State University dissertation.
- Quesada, Margaret L. 2015. *The L2 acquisition of Spanish subjects*. Berlin: De Gruyter.
- Rohde, Hanna & Andrew Kehler. 2014. Grammatical and information-structural influences on pronoun production. *Language, Cognition, and Neuroscience* 29(8). 912–927.
- Serratrice, Ludovica, Antonella Sorace & Sandra Paoli. 2004. Subjects and objects in Italian-English bilingual and monolingual acquisition. *Bilingualism: Language and Cognition* 7. 183–206.
- Silva-Corvalán, Carmen. 2014. *Bilingual language acquisition: Spanish and English in the first six years*. New York: Cambridge University Press.
- Sorace, Antonella. 2011. Pinning down the concept of “interface” in bilingualism. *Linguistic Approaches to Bilingualism* 1. 1–33.
- Sorace, Antonella, Ludovica Serratrice, Francesca Filiaci & Michela Baldo. 2009. Discourse conditions of subject pronoun realization: Testing the linguistic intuitions of older bilingual children. *Lingua* 119. 460–477.
- Tsimpli, Ianthi. 2011. External interfaces and the notion of “default”. *Linguistic Approaches to Bilingualism* 1. 101–103.
- Tsimpli, Ianthi, Antonella Sorace, Caroline Heycock & Francesca Filiaci. 2004. First language attrition and syntactic subjects: A study of Greek and Italian near-native speakers of English. *International Journal of Bilingualism* 8. 257–277.
- Zubizarreta, María Luisa. 1999. Las funciones informativas: Tema y foco. In Ignacio Bosque & Violeta Demonte (dir.), *Gramática Descriptiva de la Lengua Española*, 4215–4245. Madrid: Espasa-Calpe.

Copyright of Applied Linguistics Review is the property of De Gruyter and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.