AUTHOR QUERY – S0272263119000111

QA The distinction between surnames can be ambiguous, therefore to ensure accurate tagging for indexing purposes online (eg for PubMed entries), please check that the highlighted surnames have been correctly identified, that all names are in the correct order and spelt correctly.

AU1 Please provide the publisher location for reference 'Wilson, 2009'.

AU2 should this be possessive?

AU3 Please note that references 'Baumann et al., 2010; Sturt, 2003' are not listed in the reference list. Please add it to the list or delete the citation.

AU4 is there publication information for this?

AU5 Please provide the volume number and page range for references 'Chamorro, 2018; Fernandes et al., 2018'.

AU6 I there an URL for this?

AU7 city and publisher?

AU8 capitalize "group" in figure 6

Studies in Second Language Acquisition **0** (2019), 1–28
doi:10.1017/S0272263119000111

Research Article

ANAPHORA RESOLUTION IN L2 ENGLISH

AN ANALYSIS OF DISCOURSE COMPLEXITY AND CROSS-LINGUISTIC INTERFERENCE

Carla Contemori*

University of Texas, El Paso

Ohood Asiri

University of Texas, El Paso

Elva Deida Perea Irigoyen

Centro de Lengua, Universidad Autónoma de Ciudad Juarez

QA

Abstract

We test the interpretation of pronominal forms in L2 speakers of English whose L1 is Spanish. Previous research on learners of nonnull subject languages has shown conflicting results. The aim of the present study is to reconcile previous evidence and shed light on the factors that determine learners' difficulty to interpret pronominal forms in the L2. In six comprehension experiments, we found that intermediate L2 speakers did not show increased difficulty compared to native speakers in integrating multiple sources of information (syntactic, discourse, pragmatic) to resolve ambiguous pronouns in intrasentential anaphora and cataphora conditions. However, we also found that when two referents with equal prominence are introduced using a conjoined noun phrase in the preceding context, the learner's performance is significantly different than the performance of the native speakers, both in intrasentential and intersentential anaphora. We suggest that L2 speakers may encounter difficulties evaluating the salience of the antecedents during pronoun resolution.

INTRODUCTION

It has been shown that learners of a null subject language whose first language (L1) is a nonnull subject language show some optionality in the interpretation of null and overt

*Correspondence concerning this article should be addressed to Carla Contemori, Department of Languages and Linguistics, University of Texas at El Paso, 500 W. University Ave, El Paso, Texas 79968. E-mail: carla.contemori@gmail.com

subjects in the second language (L2) (e.g., Belletti, Bennati, & Sorace, 2007; Keating, VanPatten, & Jegerski 2011; Montrul & Rodríguez Louro, 2006; Rothman, 2008; Sorace & Filiaci, 2006; Sorace, Serratrice, Filiaci, & Baldo, 2009). To account for these results, the Interface Hypothesis (IH; e.g., Sorace, 2011) proposed that interface structures between syntax and pragmatics (as in the case of anaphoric expressions) require an increasing use of cognitive resources and are therefore less likely to be successfully acquired by bilinguals in comparison to structures without this interface. So far, research on learners of nonnull subject languages has shown conflicting results (Contemori & Dussias, 2015; Cunnings, Fotiadou, & Tsimpli, 2017; Roberts, Gullberg, & Indefrey 2008; Schimke & Colonna, 2016). For example, studies on English have demonstrated no differences between learners of English and native speakers (Contemori & Dussias, 2015; Cunnings et al., 2017). Other studies have found reliance on L1 strategies for the interpretation of pronominal forms in the L2 (Roberts et al., 2008) or more reliance on discourse-level cues in L2 speakers than in native speakers (Schimke & Colonna, 2016). Here, we test the interpretation of anaphoric and cataphoric pronouns in learners of English (nonnull subject language) whose L1 is Mexican Spanish (null subject language). The aim is to reconcile previous evidence and shed light on how the structure of the discourse impacts the learners' interpretation of subject pronouns. In the following section, we first give an introduction about anaphora and cataphora in English and Spanish, the languages tested here. Second, we present previous findings on L2 comprehension of anaphoric expressions. Then, we describe the aims of the study and summarize the experiments.

REFERRING EXPRESSIONS IN ENGLISH AND SPANISH

The felicitous interpretation of anaphoric expressions is based on the interaction of a number of factors that determine their accessibility, including saliency, semantic, syntactic, and lexical and discourse structural information (e.g., Arnold, 2010 for a review). For instance, anaphora resolution is known to be influenced by the saliency of the antecedents presented in the previous discourse, for example, pronouns in nonnull subject languages typically refer to entities that are highly prominent in the discourse (e.g., Arnold, 2010). Among the key factors identified to describe the notion of salience, the grammatical role of the antecedent is known to play a role, with a referent in subject position being more likely to be interpreted as the antecedent for a pronoun in subject position than referents that are in nonsubject positions (e.g., Ariel, 1990). For example, in English, in a sentence such as (1), native speakers are more likely to interpret the pronoun *she* as referring to *Yolanda*, rather than *Josefina.*

(1) When Yolanda met Josefina, she was in high school.

According to this bias (also called *first mention*),[1] English speakers are expected to refer the pronoun *she* to the previous subject (*Yolanda*), which is the more accessible referent in the previous sentence given its syntactic role (e.g., Arnold, 2010).

While English speakers have the possibility to choose a referring expression between either an explicit referential form (e.g., *Yolanda*, *Josefina*) or a reduced lexical form (e.g., *she*), languages may differ in the inventory and function of referring expressions

available. For instance, in null subject languages like Italian, Spanish, and Greek, null and explicit pronouns can be used. The Position of Antecedent Strategy (PAS) first proposed by Carminati (2002) claimed that while a null pronoun typically refers to an antecedent that is in the highest syntactic position (i.e., the subject), explicit pronouns tend to refer to an antecedent that is in a syntactic position lower than the subject position (i.e., the object). The PAS, which was first proposed to explain anaphora resolution preferences in Italian, has been extended to other null subject languages like Spanish (e.g., Sorace et al., 2009). Under this approach, for the sentence in (2), Spanish native speakers would likely interpret *Yolanda* as the subject of the verb *estaba*. Conversely, native speakers would interpret an explicit pronoun as referring to *Josefina*, the object of the first clause, as illustrated in (2).

(2) Quando Yolanda conoció a Josefina, *pro*/ella estaba en la preparatoria.

When Yolanda met Josefina, pro/she was in high school.

Recent research comparing anaphora resolution in Italian and Spanish has found differences among the two null subject languages (Filiaci, Sorace, & Carreiras 2014). Filiaci et al. (2014) demonstrated that Italian and Peninsular Spanish do not differ on the interpretation of null subjects, that is both languages prefer a subject interpretation as predicted by the PAS (e.g., *pro* = Yolanda in (2)). However, with respect to the interpretation of explicit pronouns, while Italian native speakers prefer an object interpretation (e.g., *ella* = Josefina in (2)), Spanish native speakers accept both a subject and an object interpretation.

So far we have presented cases of intrasentential anaphora in which an overt/null pronoun follows the potential antecedents (1)–(2). We now present the case of intrasentential cataphora, in which an overt/null pronoun precedes the potential antecedents, as illustrated in the English (3) and Spanish (4) examples.

(3) While she was in high school, Yolanda met Josefina.
(4) Mientras que *pro*/ella estaba en la preparatoria, Yolanda conoció a Josefina.

While pro/she was in high school, Yolanda met Josefina.

For Spanish, no previous study has looked at native speakers' interpretations of cataphoric pronouns. Previous studies on Italian have demonstrated a subject preference for null cataphoric pronouns and an extralinguistic referent preference for overt cataphoric pronouns (e.g., Serratrice, 2007). The extralinguistic referent preference consisted of interpreting the overt pronoun as an external referent (i.e., "someone else") that is not mentioned in the discourse.

In English, the preferred interpretation for the pronoun *she* in (5) is the subject of the following sentence (Van Gompel & Liversedge, 2003). To our knowledge, the interpretation of cataphora has never been investigated in learners of nonnull subject languages. As demonstrated by Van Gompel and Liversedge, English native speakers show an overwhelming preference to assign the cataphoric pronoun to the following subject rather than to the object, in comparison to the higher variability observed in the interpretation of anaphoric pronouns. Here we investigate the interpretation of intrasentential cataphora to verify if L2 speakers have acquired the constraints on the

interpretation of this type of pronoun, despite its infrequent occurrence in the input (Trnavac & Taboada, 2016).

To summarize, in the present study we examine pronoun resolution in English and Mexican Spanish (intrasentential anaphora and cataphora; intersentential anaphora). The aim is to understand if L2 learners of English (L1 Mexican Spanish) encounter difficulties in the comprehension of referring expressions in the L2 or if they can acquire the interpretation biases of English successfully. Additionally, we manipulate the structure of the discourse in which pronouns are presented to assess how the context influences L2 leaners' pronoun interpretation. For each discourse context manipulation, we first measure the preferences of Mexican Spanish speakers and we compare them to the performance observed in English in the L2 and native speakers.

ANAPHORA RESOLUTION IN L2 LEARNERS

Studies on anaphora resolution have mainly focused on learners of null subject languages. The findings show that while L2 speakers can master the interpretation of null pronouns in a nativelike way, they may interpret overt pronouns to index topic continuity more often than native speakers (e.g., Belletti et al., 2007; Keating et al., 2011; Montrul & Rodríguez Louro, 2006; Rothman, 2008; Sorace & Filiaci, 2006). To account for these results, the IH (e.g., Sorace, 2011) proposed that interface structures between syntax and pragmatics (as in the case of anaphoric expressions) require an increased use of cognitive resources and are therefore less likely to be successfully acquired by L2 speakers in comparison to structures without this interface.

Studies on anaphora resolution and choice of referring expressions in learners of nonnull subject languages are much more limited and have shown mixed evidence (Contemori & Dussias, 2015, 2016; Cunnings et al., 2017; Ellert 2013; Roberts et al., 2008; Schimke & Colonna, 2016; Wilson, 2009).

The studies by Contemori and Dussias (2015), Cunnings et al. (2017), and Roberts et al. (2008) are particularly relevant for the present research. Cunnings et al. (2017) investigated the interpretation and processing of explicit pronouns in English native speakers and L2 speakers of English whose L1 is Greek, a null subject language. Cunnings et al.'s results indicate that L1 Greek learners of English comprehend pronouns in the L2 in the same way as English native speakers, demonstrating a preference for interpreting an ambiguous pronoun as referring to the sentence subject, according to the first-mention bias. Similar results as in Cunnings et al. were found by Contemori and Dussias (2015) using the eye-tracking methodology and investigating similar sentence structures. In Contemori and Dussias (2015), the population tested was highly proficient in the L2 (English) and the first language was Spanish.

In Roberts et al. (2008), L2 Dutch learners whose L1 is a null subject language (Turkish) and learners whose L1 is a nonnull subject language (German) participated in a reading eye-tracking study and an offline sentence comprehension task. Roberts et al. found a difference between the two L2 groups in the offline interpretation of Dutch subject pronouns, showing that L1 Turkish–L2 Dutch participants interpreted pronouns as signaling a topic shift more often than native speakers of Dutch. According to Roberts et al., this result is a consequence of cross-linguistic interference that can be attributed to the difficulty with coordinating information from different sources (e.g., pragmatic,

syntactic). In the online measures, both L2 groups incurred in a processing cost when reading sentences in which they had to integrate discourse pragmatic contextual information when interpreting a potentially ambiguous pronoun (e.g., *Hans and Peter are in the office. While Peter is working, he is eating a sandwich.*). This effect again suggests that integrating information from multiple sources may be more difficult for L2 learners than for native speakers.

Notice that previous studies have used a variety of experimental materials. As previously mentioned, the accessibility of an antecedent is influenced by a range of factors at the discourse level (e.g., saliency, semantic, syntactic, and lexical and discourse structural information), and the presence of multiple antecedents in a context can have an impact on the amount of attention that an individual can dedicate to each antecedent within a discourse representation (e.g., Arnold & Griffin, 2007). Additionally, similar degrees of accessibility in potential antecedents can increase the complexity in the pronoun resolution process, reducing an individual's ability to calculate discourse status (e.g., Vogels, Krahmer, & Maes, 2014), which may affect L2 learners to a higher extent than native speakers (Cunnings, 2017). As recently proposed by Cunnings (2017), L2 learners may be less successful at retrieving information from memory during sentence processing. As a result, the differences between L1 and L2 anaphora resolution may stem from an increased susceptibility to retrieval interference. Increasing complexity in the discourse may affect L2 processing during anaphora resolution, leading to nonnative patterns of interpretation. Thus, the fact that previous studies tested the comprehension of pronouns in discourse contexts with varying levels of complexity may be one of the reasons why mixed results were found (e.g., Cunnings et al., 2017; Roberts et al., 2008).

AIMS AND RESEARCH QUESTIONS

Previous studies on learners of nonnull subject languages have shown mixed evidence, and it is not yet clear (a) if L2 speakers can reach nativelike competence on anaphora resolution and (b) how discourse complexity may have impacted previous results. The aim of the present study is to address these open questions. L2 speakers of English (L1 Spanish) are tested here on anaphora resolution and compared to a group of native speakers of English. In three experiments, the structure of the discourse is manipulated to uncover the factors underlying the difficulties experienced by the learners. In each study, a comparison with the pronoun resolution preferences in Mexican Spanish (L1) sheds light on the potential role of cross-linguistic interference on the interpretation of anaphoric dependencies in L2 English.

In Experiment 1, we test the interpretation of intrasentential anaphora and cataphora with a group of native speakers of Mexican Spanish. In Experiment 2, the interpretation of anaphora and cataphora is tested in English L2 speakers and native speakers of English. In Experiment 3, we examine the interpretation of intrasentential anaphora further with native speakers of Mexican Spanish, by increasing the complexity of the discourse. In Experiment 4, we translate Experiment 3's materials in English and compare the performance of L2 and native speakers of English. In Experiment 5, we investigate the interpretation of intersentential anaphora in a group of native speakers of Mexican Spanish. In Experiment 6, we test the interpretation of the same contexts as in Experiment 5 with L2 and native speakers of English.

Based on previous studies on nonnull subject languages, we expect that L2 speakers may show a similar pattern of interpretation as L1 speakers in simple contexts that require the use of the first-mention bias, as in Contemori and Dussias (2015) and Cunnings et al. (2017) (Experiment 2; e.g., *Yolanda met Josefina while she was in high school*). However, L2 speakers may experience difficulties when there is an increase in discourse complexity (e.g., if the referents presented in the preceding context have similar prominence, as in Roberts et al., 2008). In this case, we may expect differences between the performance of L2 and native speakers in conditions in which the structure of the discourse is manipulated (Experiments 4 and 6), as predicted by the IH (Sorace, 2011).

EXPERIMENT 1

In Experiment 1, we test the interpretation of intrasentential anaphora and cataphora in a group of native speakers of Mexican Spanish. The aim is to understand the interpretation biases for explicit pronouns in the L1 of the learners. In Experiment 2, we translate the experimental materials into English and test the interpretation of anaphora and cataphora in L2 English speakers whose L1 is Spanish.

PARTICIPANTS

Twenty-four native speakers of Mexican Spanish were recruited at a large Mexican university (14 females; mean age = 28; SD = 9) in a city on the border with the United States. Participants had some exposure to English in school, but none of them reported being fluent in English or any other language. They were all born and living in Mexico at the time of testing.

MATERIALS

The data were collected using a sentence comprehension task. The experimental sentences consisted of a main clause and a subordinate clause, as shown in (5) and (6). The two referents in the main clause were always proper nouns and had the same gender. The subordinate clause included an ambiguous pronoun that could either refer to the subject (*Yolanda*) or object (*Josefina*) noun phrase (NP).

Twelve experimental sentences were constructed that included a pronoun. In the anaphoric condition, the main clause was presented first, and the pronoun appeared in the second clause, as shown in (5). Twelve experimental sentences were constructed that included a cataphoric pronoun. In the cataphoric condition, the subordinate clause appeared first, followed by the main clause, as shown in (6). Only conditions with explicit pronouns were used in the experimental sentences because we were interested in understanding potential effects of cross-linguistic interference between Spanish and English. Therefore, null pronouns were not included in the experimental design.

(5) Yolanda conoció a Josefina mientras que ella estaba en la preparatoria.

Yolanda met Josefina while she was in high school.

(6) Mientras que ella estaba en la preparatoria, Yolanda conoció a Josefina.

While she was in high school, Yolanda met Josefina.

Each sentence in the experiment was followed by a multiple-choice comprehension question, as illustrated in (7).

(7) ¿Quién estaba en la preparatoria?
Who was in high school?
(a) Yolanda
(b) Josefina
(c) Otra persona (someone else)

The three choices were presented with the question, one corresponding to the NP1 (Yolanda), one corresponding to the NP2 (Josefina), and one corresponding to an external referent (someone else). As mentioned in the introduction, previous studies looking at the interpretation of cataphora in Italian (e.g., Serratrice, 2007), a null subject language, have found that the explicit cataphoric pronoun is often interpreted as an external referent. Thus, we included an external referent choice in the Spanish task. Additionally, the external referent choice was presented in the English task to maintain similarity in the three choices across the two tasks. The position of the three referents in the multiple-choice question was counterbalanced across the experiment.

The variable manipulated in the experiment is Condition, with two levels (anaphora vs. cataphora).

Two counterbalanced lists containing 12 experimental sentences (each list containing six items per condition) were created. Additionally, 24 filler sentences were included in each list that had similar syntactic structure and length as the experimental sentences. Each list was presented in a pseudorandomized order, and the order was inverted to create an additional two lists.

PROCEDURE

Spanish monolingual participants were tested in a computer lab at their university. Participants were asked to read the sentences and answer the questions by choosing one of three provided options. The task was designed as a Qualtrics survey.

RESULTS

Figure 1 illustrates the proportion of NP1 interpretations given by native speakers of Mexican Spanish in the anaphora and cataphora conditions. Error bars represent 95% Confidence Intervals in all the figures. Table 1 illustrates the amount of NP1, NP2, and External referent interpretations.

Mixed-effects logistic regression (Jaeger, 2008) was used in all the analyses in this article. In each analysis, the dependent variable per each subject and item is coded as 1 or 0 and analyzed using glmer (lme4 library; Bates & Sarkar, 2007). Factors with two levels were centered and factors with three levels were two-contrasts coded. For each model, the interactions between the fixed-effect factors and maximal random effects were

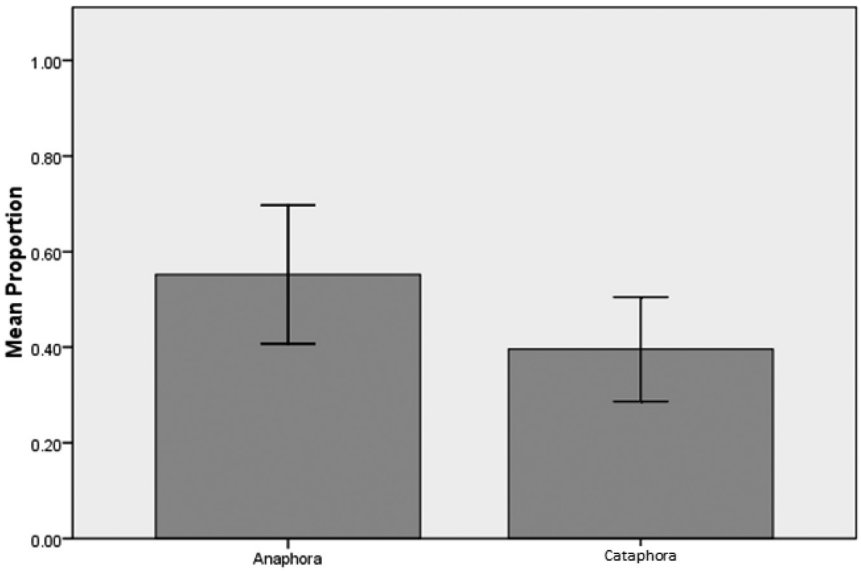

FIGURE 1. (color online) Experiment 1: Proportion of NP1 chosen by native speakers of Mexican Spanish in the anaphora and cataphora conditions.

included (e.g., Barr, Levy, Scheepers, & Tily, 2013). If convergence was not reached, the random effects were simplified, by first removing the random correlation parameters. If convergence was not reached after refitting the model, random effect parameters that accounted for the least amount of variance were removed until convergence was reached. For each statistical analysis, we report in the table the estimate (s), standard error, *t* value, and *p* value for all fixed effects and the random-effects structure.

In Experiment 1, Condition was included in the model as a fixed effect. The dependent variable is the proportion of NP1 interpretations. The model included by subject and item random intercepts and by subject and item random slopes for the effect of Condition.

In the analysis, no main effect of Condition was found (ß = –0.8, SE = 0.57, t = –1.426, p = .7; Intercept: ß = –0.1, SE = 0.37, t = –0.35, p = .7), indicating that participants gave a comparable amount of NP1 responses for the anaphora and cataphora condition. To test if the Spanish speakers' choice of NP1 is above chance, we conducted a one-sample t-test comparing the native speaker's mean on the two conditions against a chance level accuracy (50%). The analysis revealed that native speakers' NP1 choices were not different than chance in the anaphora (t(23) = .718; p = .48) and cataphora condition (t(23) = –1.908; p = .07).

TABLE 1. Experiment 1: Distribution of responses and SDs in parentheses

	NP1	NP2	External referent
Anaphora	0.56(0.29)	0.38(0.27)	0.06(0.25)
Cataphora	0.39(0.21)	0.28(0.27)	0.33(0.27)

INTERIM DISCUSSION

The results of Experiment 1 show a preference for NP1 interpretations in the anaphora and cataphora condition in native speakers of Mexican Spanish. Additionally, the speakers' preferences in both conditions are not different than chance. The results presented in Table 1 indicate that speakers chose the NP2 in the anaphora condition when a NP1 was not selected. For the cataphora condition, the results are more evenly distributed between NP1, NP2, and External referent. The findings seem in line with the study by Filiaci et al. (2014) that investigated similar anaphoric contexts in Peninsular Spanish as the ones tested here. Filiaci et al. demonstrated that Spanish speakers' preferences are equally divided between the subject and the object antecedent when interpreting an explicit pronoun.

For the cataphora condition, to our knowledge no previous study has looked at Spanish native speakers' interpretations. Our results show that in this condition responses are equally distributed among NP1, NP2, and External referent interpretations.

Notice that most of the research on Peninsular and Mexican Spanish has shown that native speakers have a preference for interpreting an explicit pronoun as referring to a nontopic antecedent (e.g., Chamorro, 2018; Keating, Jegerski, & VanPatten, 2016; Shin & Cairns, 2009). However, these studies have tested different anaphoric contexts and or have used different methodologies, for which native speakers may have different interpretation biases (e.g., a main clause followed by a subordinate clause was tested in Chamorro 2018; sentences introduced by different connectors were tested by Keating et al., 2016; and grammaticality judgments were used in Shin & Cairns, 2009). Furthermore, differently from previous studies, our participants were recruited in a large city on the U.S.-Mexican border. We cannot exclude that the variety of Mexican Spanish spoken in the region may present some differences in comparison to the population tested in previous research, due to contact with English (but see corpus research on varieties of Spanish spoken in the United States; e.g., Los Angeles: Silva-Corvalán, 1994).

Another factor that may have influenced our results is that in our experiment the majority of sentences contained an explicit pronoun. Because our focus was the interpretation of explicit pronouns in native speakers, we did not counterbalance contexts containing explicit and null pronouns, as in previous research (e.g., Chamorro, 2018; Keating et al., 2016). Recent work on European and Brazilian Portuguese has demonstrated that the frequency of overt/null pronouns in the input (in a situation in which overt/null pronouns can co-refer either with a subject or an object referent) can change participants' interpretation preferences (Fernandes, Luegib, Correa Soares, de la Fuente, & Hemforth, 2018). In particular, Fernandes et al. found that participants were more likely to interpret an overt pronoun as referring to an object when there were fewer overt pronouns in the input. We tested the possibility that native speakers of Mexican Spanish may have adapted to the high probability of overt pronouns occurring in the experiment, and they may have changed their preference for interpreting the overt pronoun as referring to a nontopic antecedent. To do so, we added the Order of Presentation of the items as a continuous variable in the analysis. If an effect of adaptation emerged, we would expect a decrease in NP2 choices throughout the course of the experiment as participants encounter more instances of explicit pronouns. We analyzed the proportion of interpretations Order of Presentation as fixed effect. The dependent variable is the AU2

number of NP1 responses chosen by the participants. The model included by-subject and by-item random intercepts and by-subject and by-item random slopes for the effect of Order of Presentation.

No main effect emerged for Order of Presentation (ß = –0.01, SE = 0.01, t = –1.353, p = .1; Intercept: ß = –0.7, SE = 0.54, t = –1.396, p = .1). Therefore, we hypothesize that exclusion of null subjects from the experimental design should not influence participants' interpretations (see also Experiment 3 for additional results and discussion). We also considered a model that included the factor Condition (anaphora vs. cataphora), the factor Order of Presentation, and the interaction between the two. However, the model fit did not improve significantly, indicating that Order of Presentation of the items did not affect the two pronoun conditions differently.

EXPERIMENT 2

PARTICIPANTS

Twenty-four native English speakers (7 females; mean age = 31; SD = 2.5) and 31 intermediate proficient learners of English (L1 Mexican Spanish) (12 females; mean age = 20; SD = 3.3) participated in the study.

Native English speakers were recruited using the Amazon Mechanical Turk platform. Only participants who indicated no prior experience with a language other than English were included in the native speakers' group.

The L2 participants were recruited in intermediate-level English as a Second Language (ESOL) classes at a large American university. They were born in Mexico and were exposed to English at different times during childhood, with some participants having early exposure to English in school. L2 participants completed a subsection of the English Language test (MELICET), containing 50 multiple-choice questions testing grammar and reading comprehension. The results of the MELICET confirm that L2 participants have intermediate proficiency in English (mean = 29.1 over 50; SD = 8.2). Table 2 shows the language background information for L2 learners collected with a Language History Questionnaire (LHQ; Marian, Blumenfeld, & Kaushanskaya, 2007).

TABLE 2. Experiment 2: Participant information: Mean (SD)

	Spanish-L1	English-L2
Age of exposure (years)	0.84(0.8)	7.47(4.9)
Became fluent speaking (age in years)	3.06(2.8)	11.81(6.0)
Length of residence in a country where the language is spoken (years)	16(6.3)	4.72(5.9)
Speaking proficiency (1–10)	8.96(2.9)	7.30(2.6)
Listening proficiency (1–10)	9.03(2.9)	7.80(3.1)
Reading proficiency (1–10)	9.06(2.9)	7.77(3)
Average daily exposure %	62.94(15)	40.31(16)
Language dominance (number of participants)	31/31	0/31

MATERIALS

The English sentences used in this experiment were translations of the Spanish sentences used in Experiment 1.

PROCEDURE

Native English participants were tested remotely. They were asked to read sentences and answer questions about the sentences by choosing one of three provided options. The task was designed as a Qualtrics survey.

L2 participants were tested in the ESOL classroom. They completed a pen-and-pencil version of the sentence comprehension task. Additionally, they completed a copy of the LHQ and the MELICET.

RESULTS

Figure 2 illustrates the proportion of NP1 interpretations given by the native speakers and the L2 participants. Table 3 illustrates the proportion of NP1, NP2, and External referent interpretations.

In the analysis, Group and Condition were included as fixed effects, and the number of NP1 interpretations was the dependent variable (Table 4). In the analysis, a main effect of Condition was found, indicating that participants gave more NP1 responses in the

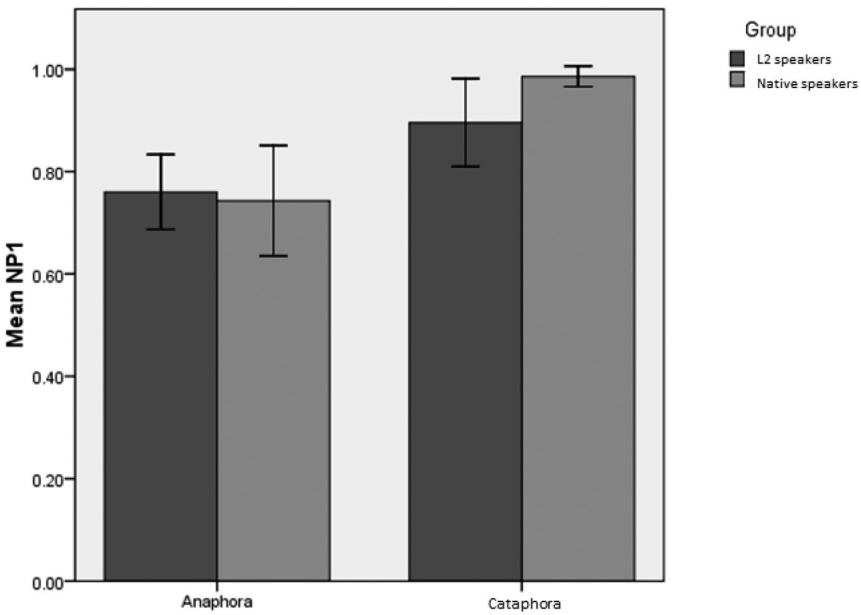

FIGURE 2. (color online) Experiment 2: Proportion of NP1 responses in the anaphora and cataphora condition for the English native speaker and L2 group.

TABLE 3. Experiment 2: Distribution of responses and standard deviations in parentheses

	NP1		NP2		External referent	
	L2	Native	L2	Native	L2	Native
Anaphora	0.76(0.42)	0.74(0.43)	0.24(0.43)	0.26(0.43)	0(0)	0(0)
Cataphora	0.89(0.30)	0.98(0.11)	0.04(0.17)	0.02(0.11)	0.07(0.26)	0(0)

cataphora condition (mean = 0.93) compared to the anaphora condition (mean = 0.75). No other main effect or interaction emerged from the analysis.

INTERIM DISCUSSION

The results indicate that L2 speakers of English interpret pronouns in a nativelike manner and have successfully acquired the first-mention bias. Native and L2 English speakers interpreted ambiguous pronouns as referring to the sentence subject and current discourse topic in the case of anaphora. Notice that when participants did not select a NP1 interpretation, both groups interpreted the pronoun as referring to the NP2. For the cataphoric pronouns, native speakers of English exhibited an overwhelming preference to interpret the pronoun as referring to the first NP (i.e., the subject of the main clause), confirming results from previous studies on monolingual speakers (e.g., Van Gompel & Liversedge, 2003). Despite its infrequency (Trnavac & Taboada, 2016), L2 speakers demonstrated that they have acquired the constraints of cataphora interpretation in the L2.

The results confirm the findings from Cunnings et al. (2017) and Contemori and Dussias (2015), suggesting similar preferences for L1 and L2 speakers in pronoun resolution in English. As the explicit pronoun is the default referring expression signaling topic continuity in English, our results support findings from previous research on null subject languages, in which the default form (null subject) is typically interpreted like in native speakers (e.g., Sorace, 2011).

In Experiments 3 and 4, we examine the interpretation of pronouns in different discourse contexts, to test the hypothesis that discourse complexity may have an impact on L2's interpretation biases.

TABLE 4. Experiment 2: Full model statistics

	Estimate	Std. Error	t-value	P-value
(Intercept)	0.8	0.02	34.715	0.0001
Group	0.03	0.04	0.781	0.4
Condition	0.1	0.05	3.449	0.001
Group*Condition	0.1	0.08	1.246	0.2

Note: The model included by-subject and by-item random intercepts, together with by-subject random slopes for the effects of Condition and by-item random slopes for the effects of Condition and Group and their interaction.

EXPERIMENT 3

In Experiments 3 and 4, stimuli similar to Roberts et al. (2008) are used to observe if any difference in anaphora interpretation emerges when two entities are introduced in the preceding context that are equally salient. Additionally, the saliency of the two referents is manipulated to further investigate what type of discourse structure impacts the interpretation of ambiguous pronouns in English and Spanish. In Experiment 3, we test the interpretation biases of native speakers of Mexican Spanish. In Experiment 4, we test the same L2 English speakers as in Experiment 2 and we compare them to a group of native English speakers.

PARTICIPANTS

Twenty-two native speakers of Mexican Spanish were recruited at a large Mexican university (14 females; mean age = 28; SD = 11) in a city on the border with the United States. Participants had some exposure to English in school, but none of them reported being fluent in English or any other language. They were all born and living in Mexico at the time of testing.

METHOD

Twenty-four sets of experimental sentences were constructed in three conditions. In the first condition (Equal prominence), prominence was operationalized by making two referents (Carlos, Martín) equally salient in the preceding context, as shown in (8). We refer to the local antecedent of the pronoun as the NP1 (Carlos) and to the second referent mentioned in the first sentence as the NP2 (Martín). The Equal Prominence condition has a similar sentence structure as the experimental sentences in Roberts et al. (2008).

In the second condition (NP2 Repetition), the second referent (Martín) was made more prominent, by using a repeated NP2 reference in the discourse, as shown in (9).

In the third condition (NP2 pronoun), the second referent (Martín) is not only repeated but also referred to by using a pronoun, as illustrated in (10), therefore making the NP2 more salient in comparison to the other two conditions.

(8) Equal prominence condition: Carlos y Martín están en la oficina. Mientras que Carlos está trabajando, él se está comiendo un lonche.

Carlos and Martín are at the office. While Carlos is working, he is eating lunch.

(9) NP2 Repetition condition: Carlos y Martín están en la oficina. Martín es uno de los mejores empleados de la compañía. Mientras que Carlos está trabajando, él se está comiendo un lonche.

Carlos and Martín are at the office. Martín is one of the best employees in the company. While Carlos is working, he is eating lunch.

(10) NP2 pronoun condition: Carlos y Martín están en la oficina. Martín es uno de los mejores empleados de la compañía, él ganó el reconocimiento del mejor empleado del mes. Mientras que Carlos está trabajando, él se está comiendo un lonche.

Carlos and Martín are at the office. Martín is one of the best employees in the company, he won the best employee of the month award. While Carlos is working, he is eating lunch.

Each experimental sentence was followed by a comprehension question testing the interpretation of the ambiguous explicit pronoun, as illustrated in (11).

(11) ¿Quién se está comiendo un lonche?
Who is eating lunch?
(a) Carlos
(b) Martín
(c) Otra persona (someone else)

Three choices were given for each question, one corresponding to the NP1 interpretation of the pronoun (*él* = Carlos), one corresponding to a NP2 interpretation (*él* = Martín), and one corresponding to an External referent interpretation (*él* = someone else). The position of the three referents in the multiple-choice question was counterbalanced across the experiment.

The variable manipulated in the experiment is Condition (Equal prominence vs. NP2 repetition vs. NP2 pronoun) as within-subjects factor.

Three counterbalanced lists containing 24 experimental sentences (each list containing eight items per condition) were created. Additionally, 40 filler sentences were included in each list that had similar syntactic structure and length as the experimental sentences. Each list was presented in a pseudorandomized order, and the order was inverted to create an additional three lists.

PROCEDURE

Same as in Experiment 1.

RESULTS

Figure 3 illustrates the proportion of NP1 interpretations chosen by native speakers of Mexican Spanish in the three conditions. Table 5 illustrates the proportion of NP1, NP2, and External referent interpretations.

In the statistical analysis, Condition was included as a fixed effect in the model, and number of NP1 interpretations was used as dependent variable. The model included by-subject and by-item random intercepts and by-subject and by-item random slopes for the effect of Condition.

No main effect of Condition emerged from the analysis (Condition 1: ß = 0.06, SE = 0.22, t = 0.280, p = .7; Condition 2: ß = 0.1, SE = 0.28, t = –0.557, p = .5; Intercept: ß = –0.2, SE = 0.3, t = –0.749, p = .4). To test if the Spanish speakers' choice of NP1 was above chance, we conducted a one-sample t-test comparing the native speaker's mean on the three conditions against a chance level accuracy (50%). The analysis revealed that native speakers' NP1 choices were not different than chance in the three conditions (Equal Prominence: t(21) = .234; p = .81; NP2 Repetition condition: t(21) = .924; p = .36; NP2 Pronoun condition: t(21) = .892; p = .38).

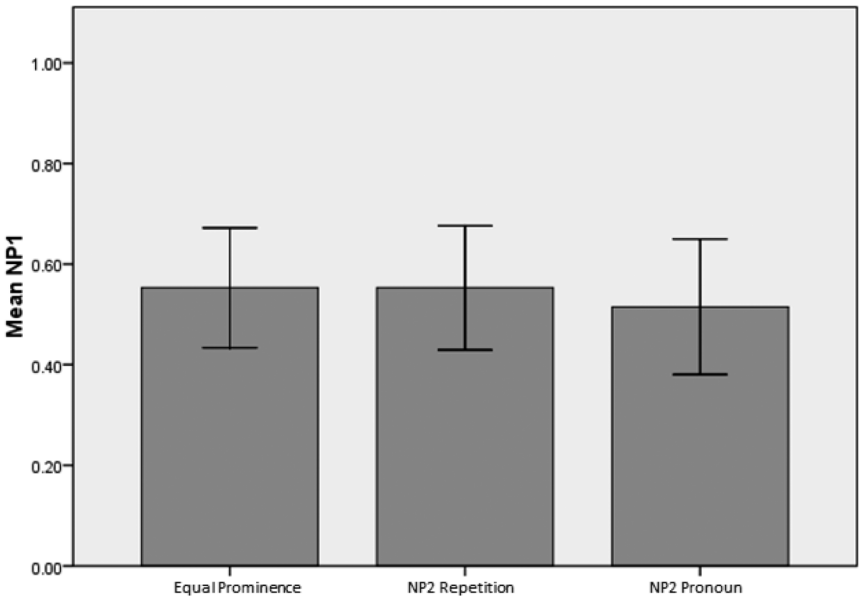

FIGURE 3. (color online) Experiment 3: Proportion of NP1 chosen by native speakers of Mexican Spanish in the three conditions.

INTERIM DISCUSSION

The results of Experiment 3 show that native speakers' NP1 preference did not differ across the three conditions, and the choice of a NP1 response was at chance. We found a considerable amount of nonlocal antecedent interpretations, showing that in Mexican Spanish there is no clear preference for interpreting an explicit pronoun as referring to a local antecedent and subject of the preceding subordinate clause. This result is in line with Filiaci et al. (2014), but not with other studies on Mexican Spanish, showing that native speakers have a preference for interpreting an explicit pronoun as referring to a nontopic antecedent (e.g., Keating et al., 2016; Shin & Cairns, 2009). As in Experiment 1, we hypothesize that, in comparison to previous research, the structure of the sentences and the ambiguity of the pronoun may have had an impact on the native speakers' interpretations. Additionally, as mentioned in Experiment 1, we cannot exclude that the variety of Mexican Spanish spoken in the border region may present some differences

TABLE 5. Experiment 3: Distribution of responses and standard deviations in parentheses

	NP1	NP2	External referent
Equal prominence condition	0.55(0.49)	0.40(0.49)	0.05(0.20)
NP2 Repetition condition	0.55(0.49)	0.39(0.49)	0.05 (0.22)
NP2 pronoun condition	0.52(0.5)	0.45(0.49)	0.03(0.17)

with respect to pronoun use and interpretation in comparison to the population tested in previous research, due to contact with English.

As in Experiment 1, we tested again the possibility that native speakers of Mexican Spanish may have adapted to the high probability of overt pronouns occurring in the experiment. Thus, we added Order of Presentation of the items as a continuous variable in the analysis. The dependent variable is the number of NP1 responses chosen by the participants. The model included by-subject and by-item random intercepts and by-subject and by-item random slopes for the effect of Order of Presentation.

If an effect of adaptation emerges, we may expect a decrease in NP2 choices throughout the course of the experiment as participants encounter more instances of explicit pronouns. The analysis did not show a main effect of Order of Presentation, excluding the possibility of adaptation (ß = 0.1, SE = 0.07, t = 1.421, p = .1; Intercept: ß = –0.1, SE = 0.28, t = –0.476, p = .6). We also analyzed a model that included the factors Condition and Order of Presentation and their interaction. However, the model fit did not improve significantly, indicating that Order of Presentation of the items did not affect the two pronoun conditions differently.

To conclude, native speakers of Mexican Spanish did not demonstrate a clear preference for interpreting explicit pronouns as referring to a topic antecedent in the contexts tested in Experiment 3. Participants also showed a high percentage of NP2 responses throughout the experiment. We will discuss this result further in Experiment 4, as it may have had an influence on L2 participants' interpretations in English.

EXPERIMENT 4

PARTICIPANTS

A new group of 30 native speakers of English was recruited through Amazon Mechanical Turk (mean age = 35.4; SD = 7).

Twenty-eight intermediate proficiency learners of English (L1 Spanish) (mean age = 20; SD = 3.3) participated. The L2 participants were a subgroup of the participants tested in Experiment 2 (11 females). L2 participants were tested at least two months apart.

MATERIALS

Same as in Experiment 3. Materials were translated from Spanish into English.

PROCEDURE

Same as in Experiment 2.

RESULTS

Figure 4 illustrates the proportion of NP1 interpretations chosen by the native speakers and the L2 participants in the three conditions. In Table 6, we show how the responses of the two groups are distributed across the three conditions. When L2 speakers do not

choose the NP1, they prefer the nonlocal antecedent interpretation rather than the External referent interpretation.

In the statistical analysis, Condition and Group were used as fixed effects, and the NP1 interpretations were included as dependent variable (Table 7). In the analysis, a main effect of Group was found, indicating a significantly higher amount of NP1 responses in native speakers in all three conditions compared to L2 speakers.

INTERIM DISCUSSION

In the present experiment, pronoun interpretation was tested when the discourse context and the salience of the two referents are manipulated. In one condition, topichood was operationalized in the preceding context by making the two referents (*Eric* and *John*) in the discourse equally prominent (Equal Prominence condition). In the second condition (NP2 Repetition Condition), the second referent (*John*) was made increasingly prominent with repeated reference in the preceding discourse. In a third condition (NP2 Pronoun Condition), the second referent is referred to by using a pronoun, hence making it more salient in comparison to the other two conditions. A group of native speakers and a subgroup of the same ESOL students who participated in Experiment 2 participated in the sentence comprehension task. While native speakers overwhelmingly chose NP1 as a reference for the pronoun in all three conditions, L2 learners' preference for the NP1 interpretation was less robust.

The results of the current study are in line with Roberts et al. (2008). Roberts et al. used a reading task during eye-tracking to examine local versus disjointed interpretation of

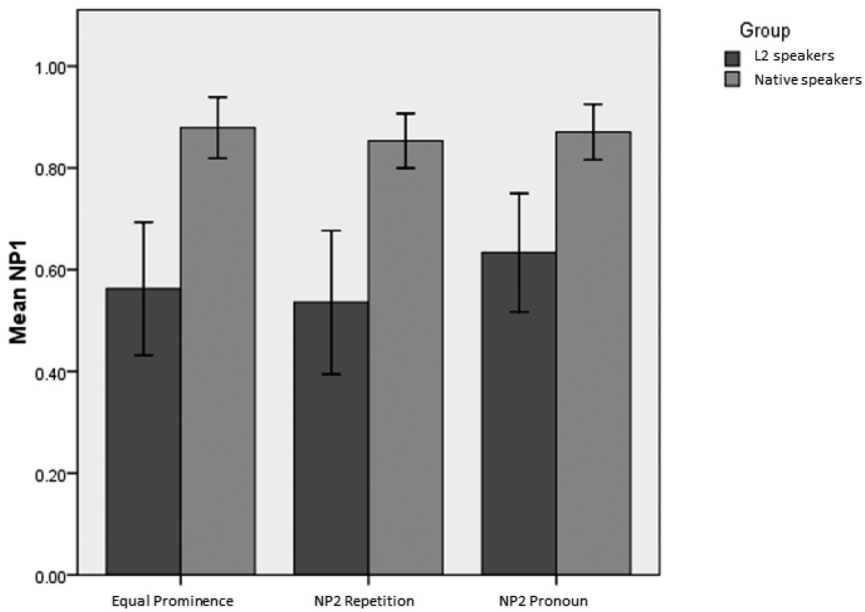

FIGURE 4. (color online) Experiment 4: Proportion of NP1 responses in the three conditions for the native speaker and L2 group.

TABLE 6. Experiment 4: Distribution of responses and SDs in parentheses

	NP1		NP2		External referent	
	L2	Native	L2	Native	L2	Native
Equal prominence condition	0.59(0.49)	0.87(0.32)	0.33(0.49)	0.05(0.20)	0.08(0.27)	0.07(0.26)
NP2 Repetition condition	0.53(0.50)	0.85(0.35)	0.38(0.48)	0.07(0.25)	0.09(0.28)	0.08(0.26)
NP2 Pronoun condition	0.59(0.49)	0.87(0.33)	0.30(0.45)	0.06(0.22)	0.11(0.31)	0.07(0.26)

Dutch subject pronouns. In Roberts et al., the condition in which L2 speakers had more difficulties contained an ambiguous pronoun and two entities in the preceding discourse (e.g., *Peter and Hans are in the office. While Peter is working, he is eating a sandwich*.). While native speakers of Dutch resolved the pronoun locally (i.e., he = Peter), native speakers of Turkish whose L2 is Dutch chose the NP2 interpretation (i.e., he = Hans) significantly more often. The authors interpreted this result as a problem with the integration of multiple types of information (syntactic, discourse, pragmatic) for L2 learners, resulting in interference from the L1. Interestingly, our results show that the presence of an equally prominent referent in the preceding context created a less robust NP1 interpretation in the L2 speakers in comparison to the anaphoric contexts tested in Experiment 2. However, the increased prominence of the NP2 in the NP2 Repetition Condition and in the NP2 Pronoun Condition did not differentially affect the L2 speakers' performance, as no difference was found between these two conditions and the Equal Prominence Condition (i.e., no Condition*Group interaction emerged). This result suggests that L2 speakers' NP1 preference decreased in the present experiment as a result of the presence of an equally salient second referent in the preceding context. As indicated by the results on Mexican Spanish in Experiment 3, Spanish native speakers had no clear preference for the interpretation of explicit pronouns in similar contexts, choosing the NP2 to a high extent (0.41 on average across the three conditions). We can speculate that the high percentage of NP2 choices observed in the L2 speakers in the present experiment is suggestive of cross-linguistic interference. As the complexity of the discourse increased in the three conditions tested here, compared to the anaphora and

TABLE 7. Experiment 4: Full model statistics

	Estimate	Std. Error	t-value	P-value
(Intercept)	1.5	0.53	2.920	0.003
Group	1.2	0.14	8.536	0.0001
Condition 1	–0.2	0.17	–1.487	0.1
Condition 2	0.1	0.18	0.647	0.5
Group*Condition 1	–0.03	0.17	–0.204	0.8
Group*Condition 2	-0.08	0.18	–0.485	0.6

Note: The model included by-subject and by-item random intercepts, together with by-subject random slopes for the effects of Condition and by-item random slopes for the effects of Condition and Group and their interaction.

cataphora conditions tested in Experiment 2, participants interpreted the pronoun significantly less often as referring to the local antecedent, and adopted instead a nonlocal antecedent interpretation. In Experiments 5 and 6, we examined again the possibility that L2 differ from L1 speakers on the interpretation of English pronouns when the complexity of the discourse increases. More specifically, we tested contexts of intersentential anaphora in which the referents introduced in the discourse have similar salience.

EXPERIMENT 5

Previous studies have pointed out that pronoun resolution preferences can be affected by the presence of the pronoun and its antecedent in the same sentence (intrasentential pronoun resolution) or in different sentences (intersentential pronoun resolution) (e.g., Baumann, Konieczny, & Hemforth, 2014; Carminati, 2002; Miltsakaki, 2002). In particular, previous studies testing intersentential pronoun resolution demonstrated a preference for interpreting an ambiguous pronoun as referring to the subject antecedent in German, French, English, and Portuguese (e.g., Baumann et al., 2014). In Experiment 5, we aimed at testing the interpretation of explicit pronouns in Mexican Spanish in contexts of intersentential anaphora to observe if a subject antecedent preference is found in this language. Additionally, we manipulated the context to understand how the salience of the referents introduced in the discourse impacts the interpretation of ambiguous pronouns. In Experiment 6, we used a similar test as in Experiment 5 to investigate the interpretation preferences of L2 speakers in English.

PARTICIPANTS

Twenty-six native speakers of Mexican Spanish were recruited at a large Mexican university (16 females; mean age = 28; SD = 8) in a city on the border with the United States. Participants had some exposure to English in school, but none of them reported being fluent in English or any other language. They were all born and living in Mexico at the time of testing.

MATERIALS

A sentence comprehension task was designed in which participants had to choose the referent of an ambiguous pronoun. The stimuli contained: (a) a sentence in which two referents were introduced; (b) a sentence in which either one of the two referents or both referents were present; and (c) a sentence containing a potentially ambiguous pronoun. The structure of the second sentence was manipulated across conditions in three different ways to resemble the complexity of the discourse created in Experiment 3 for intrasentential anaphora. The labeling of the three conditions is based on a description of the second sentence presented in the stimuli.

In the NP1 Only condition, the first referent introduced in the previous sentence (NP1; i.e., Rafael) is the only referent mentioned in the second sentence. In this case, the NP1 is the topic, and the more likely referent for the ambiguous pronoun "he," as illustrated in (12).

In the NP1 and NP2 condition, the second sentence contains one of the two referents presented in subject position (NP1; i.e., Rafael) and the other referent in object position (NP2; i.e., Pablo), as shown in (13). In this condition, the likely referent for the ambiguous pronoun is the NP1, as it is more salient in the second sentence than the NP2.

In the Subordinate NP2–Main NP1 condition, the second sentence contains a subordinate clause in which the NP2 (Pablo) is the subject, followed by a main clause in which the NP1 (Rafael) is the subject, as illustrated in (14). In this condition, the NP1 is again the most salient referent for the ambiguous pronoun as it is the subject of the main clause and the last-mentioned referent.

(12) NP1 Only: Rafael y Pablo fueron admitidos al hospital. Rafael estuvo en un accidente de automóvil. Después de unos pocos días, él fue dado de alta.

Rafael and Pablo were admitted at the hospital. Rafael was in a car wreck. After a few days, he was discharged.

(13) NP1 and NP2: Rafael y Pablo fueron admitidos al hospital. Rafael estuvo en un accidente de automóvil con Pablo. Después de unos pocos días, él fue dado de alta.

Rafael and Pablo were admitted at the hospital. Rafael was in a car wreck with Pablo. After a few days, he was discharged.

(14) Subordinate NP2- Main NP1: Rafael y Pablo fueron admitidos al hospital. Mientras Pablo se rompió la pierna, Rafael tenía una lesión del brazo. Después de unos pocos días, él fue dado de alta.

Rafael and Pablo were admitted at the hospital. While Pablo broke his leg, Rafael had an injury to his arm. After a few days, he was discharged.

Each experimental sentence was followed by a comprehension question testing the interpretation of the ambiguous pronoun, as illustrated in (15).

(15) ¿Quién fue dado de alta después de unos pocos días?
Who was discharged after a few days?
(a) Rafael
(b) Pablo
(c) Otra persona (someone else)

One of the three choices corresponded to the NP1 interpretation (Rafael), one corresponded to the NP2 interpretation (Pablo), and one corresponded to an External referent interpretation (someone else). The position of the three referents in the multiple-choice question was counterbalanced across the experiment.

Three counterbalanced lists containing 24 experimental sentences (each list containing eight items per condition) were created. Forty filler sentences had similar syntactic structure and length as the experimental sentences. Each list was presented in a pseudorandomized order, and the order was inverted to create an equal number of additional lists. The variable manipulated is Condition (three levels) as within-subjects factor.

PROCEDURE

Same as in Experiment 1.

RESULTS

Figure 5 illustrates the proportion of NP1 interpretations chosen by the native speakers of Mexican Spanish in the three conditions. Table 8 illustrates the proportion of NP1, NP2, and External referent choices produced by the native speakers of Mexican Spanish.

In the statistical analysis, we analyzed the proportion of NP1 choices produced by the native speakers of Mexican Spanish, including Condition as a fixed effect. The model included by-subject and by-item random intercepts and by-subject and by-item random slopes for the effect of Condition. The analysis did not reveal a main effect of Condition, indicating that participants' NP1 choices did not differ significantly (Condition 1: ß = –0.02, SE = 0.10, t = –0.231, p = .8; Condition 2: ß = –0.003, SE = 0.10, t = –0.033, p = .9; Intercept: ß = –0.08, SE = 0.07, t = –1.222, p = .2).

As in Experiments 1 and 3, we used Order of Presentations of the trials as a continuous variable in the model to evaluate whether participants may have adapted their responses during the task, due to the absence of null subjects in the experimental design. In the model, we included Order of Presentation of the trials as fixed factor. The model included by-subject and by-item random intercepts and by-subject and by-item random slopes for the effect of Order of Presentation.

A main effect of Order of Presentation emerged from the analysis (ß = –0.1, SE = 0.03, t = –2.701, p = .006; Intercept: ß = 6.0, SE = 1.63, t = 3.703, p = .0002). To

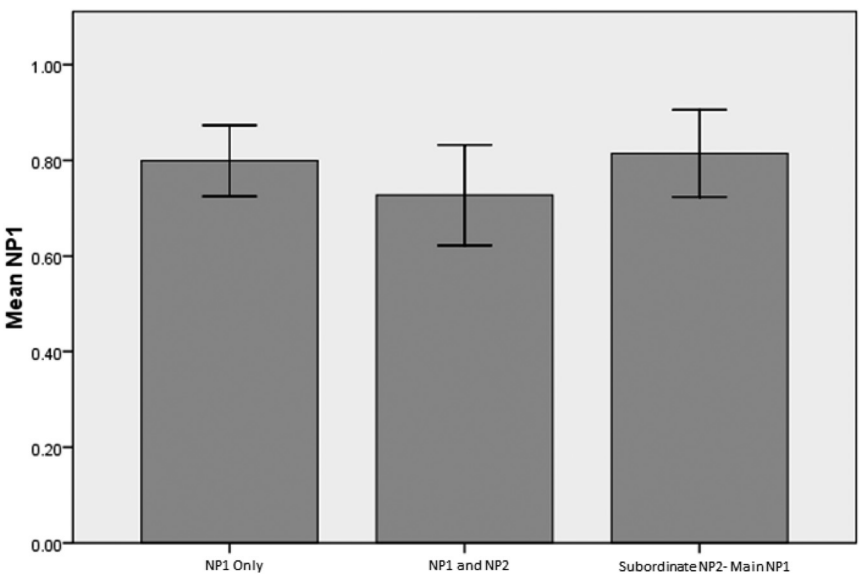

FIGURE 5. (color online) Experiment 5: Proportion of NP1 chosen by native speakers of Mexican Spanish in the three conditions.

TABLE 8. Experiment 5: Distribution of responses and SDs in parentheses

	NP1	NP2	External referent
NP1 only	0.8(0.40)	0.17(0.36)	0.03(0.19)
NP1 and NP2	0.73(0.44)	0.22(0.40)	0.05(0.23)
Subordinate NP2–Main NP1	0.81(0.38)	0.12(0.30)	0.07(0.26)

follow up on the main effect, we did a median split of the Order of Presentation of the items (experimental sentences presented in the first half vs. second half of the experiment). The median split showed that the amount of NP1 choices decreased from the first (mean = 0.90) to the second half of the experiment (mean = 0.66), while the amount of NP2 choices increased from 0.07 to 0.22. This result is not in line with the hypothesis that participants may have adapted to the increasing presence of explicit pronouns in the experiment as proposed by Fernandes et al. (2018). According to the results presented in Fernandes et al., we should have found fewer (not more) NP2 choices throughout the course of the experiment as participants encounter more instances of explicit pronouns.

INTERIM DISCUSSION

The results of Experiment 5 indicated no significant difference between the NP1 choices selected in the three conditions. As it emerged from the data in Table 8, we found a preference for interpreting an explicit pronoun as the subject/local antecedent introduced in the previous sentence. This result is in line with previous studies on intersentential anaphora showing a subject preference in the interpretation of ambiguous pronouns in German, French, English, and Portuguese (e.g., Baumann et al., 2014).

Creating similar contexts as in Experiment 3 was not possible in Experiment 5, due to the difference in sentence structure between intra- and intersentential anaphora. However, the discourse structure designed in the present experiment resembled as much as possible the complexity of the intrasentential anaphora tested in Experiment 3. Thus, the results demonstrate that the variety of Mexican Spanish tested here displayed a higher subject/local antecedent preference in intersentential anaphora compared to intrasentential anaphora with similar complexity.

EXPERIMENT 6

PARTICIPANTS

Twenty-eight native English monolingual speakers (15 females; mean age = 38; SD = 6) and 37 intermediate proficient learners of English (L1 Spanish) (17 females; mean age = 20; SD = 4) participated in the study. Native English speakers were recruited using the Amazon Mechanical Turk platform. Only participants who indicated no prior experience with a language other than English were included in the group.

The L2 participants were recruited in intermediate level ESOL classes at a large American university. L2 participants were born in Mexico. Table 9 shows the language background information for L2 learners collected with an LHQ (Marian et al., 2007).

To test proficiency in English, L2 participants completed a subsection of the MELICET. The results of the MELICET confirmed that L2 participants had intermediate proficiency in English (mean = 21/50; SD = 6.1). A t-test was used to compare the mean scores on the MELICET for participants in Experiment 2 and participants recruited in Experiment 6. L2 participants recruited in the present experiment had significantly lower scores on the MELICET compared to participants in Experiment 2 ($t(67) = 4.523$; $p < .0001$). Hence, even though the L2 learners in the present experiment were recruited from the same ESOL classes, we can assume that they had lower proficiency in English than participants tested in Experiment 2.

MATERIALS

Same as in Experiment 5. Materials were translated into English from Spanish.

PROCEDURE

Same as in Experiment 2.

RESULTS

Figure 6 shows the proportion of NP1 choices for the two groups in the three conditions and Table 10 shows the distribution of responses for the two groups. Both groups preferred a NP2 interpretation (rather than an external referent) when a NP1 was not selected.

In the model, we analyzed the number of NP1 interpretations, including Condition (three levels) and Group (two levels) as fixed effects (Table 11).

The model shows a main effect of Condition and a main effect of Group. The main effect of Group indicates that L2 speakers chose the NP1 less often than native English speakers across the three conditions. The main effect of Condition indicates that

TABLE 9. Experiment 6: Participant information

	Spanish L1	English L2
Age of exposure (years)	0.8(0.7)	8.5(5.6)
Became fluent speaking (age in years)	6.2(5.8)	11.6(6)
Length of residence in a country where the language is spoken (years)	3.7(5.6)	18.3(3.3)
Speaking proficiency (1–10)	9.01(2.6)	7.1(3.6)
Listening proficiency (1–10)	9.05(2.8)	7(4.1)
Reading proficiency (1–10)	9.07(2.8)	7(4)
Average daily exposure %	79(24.4)	50.4(23.7)
Language dominance (number of participants)	37/37	0/37

participants produced more NP1 interpretations in the NP1 Only condition compared to the NP1 and NP2 (ß = –1.2, SE = 0.28, t = –4.37, $p < .0001$) and the Subordinate NP2–Main NP1 condition (ß = –1.9, SE = 0.83, t = –2.306, $p < .02$). No significant difference was found between the NP1 and NP2 and the Subordinate NP2–Main NP1 condition (ß = –0.25, SE = 0.37, t = –0.687, $p < . 0.4$).

INTERIM DISCUSSION

The results of Experiment 6 showed that L2 leaners did not choose the NP1/local antecedent as often as native speakers. However, notice that the L2 speakers in the present experiment were less proficient than the L2 participants tested in previous experiments. Thus, the lower level of proficiency of L2 participants in Experiment 6 may have contributed to the observed difference between L1 and L2 speakers.

In comparison to Experiment 4, we observed an increase in NP1/local antecedent preferences in the L2 speakers. As shown in Experiment 5, Spanish displays a subject/local antecedent preference, which may have contributed to increase the subject/local antecedent preference observed in the L2 group. Although Mexican Spanish and English adopt similar interpretation in the intersentential anaphora tested in Experiment 5 and 6, the NP1/local antecedent interpretation seems stronger in English than in Mexican Spanish. This difference between the two languages may explain why L2 leaners still experienced some difficulty in the interpretation of pronouns in the intersentential anaphora context. We discuss potential cross-linguistic effects further in the discussion

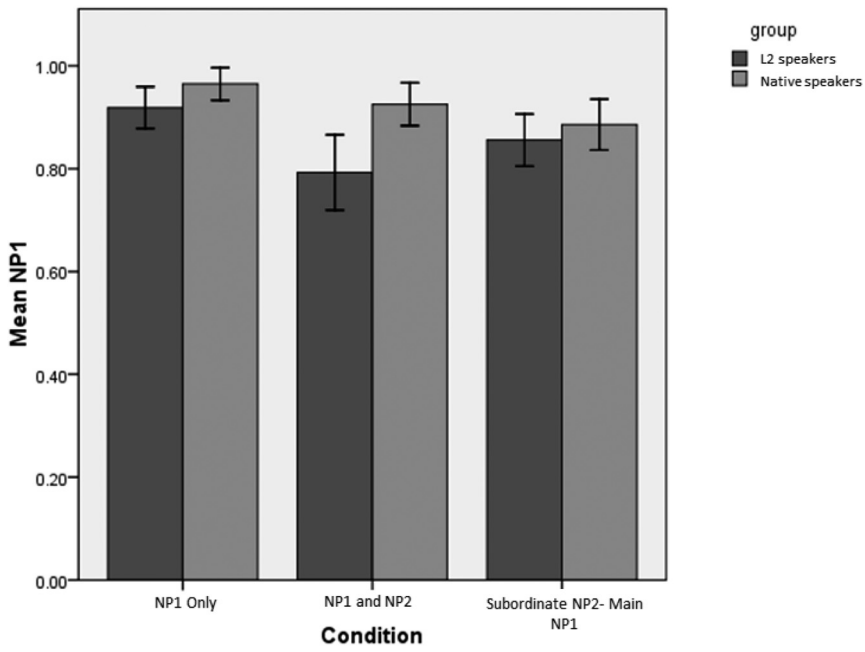

FIGURE 6. (color online) Experiment 6: Proportion of NP1 choices in L2 and native speakers. AU8

TABLE 10. Experiment 6: Distribution of responses and standard deviations in parentheses

	NP1		NP2		External referent	
	L2	Native	L2	Native	L2	Native
NP1 Only condition	0.91(0.27)	0.96(0.18)	0.08(0.25)	0.03(0.13)	0.01(0.06)	0.01(0.08)
NP1 and NP2 condition	0.79(0.40)	0.92(0.26)	0.19(0.38)	0.08(0.26)	0.02(0.14)	0(0)
Subordinate NP2-Main NP1 condition	0.85(0.35)	0.88(0.31)	0.14(0.33)	0.11(0.30)	0.01(0.13)	0.008(0.09)

section. We speculate that the underlying factor determining the difficulty may be the complexity of the discourse structure, in which two referents with equal salience are initially introduced.

DISCUSSION

In three experiments we measured the preferences for interpreting explicit pronouns in Mexican Spanish in the following contexts: (a) intrasentential anaphora and cataphora (Experiment 1); (b) intrasentential anaphora in which the salience of the antecedents was manipulated (Experiment 3); and (c) intersentential anaphora in which the salience of the antecedents was manipulated (Experiment 5). The results of the three experiments revealed that the variety of Mexican Spanish tested here does not follow the predictions of the PAS, as found by Filiaci et al. (2014). In Experiments 1 and 3, we did not find an object/nonlocal antecedent preference for the interpretation of explicit pronouns, although participants did choose the object/nonlocal antecedent as a possible referent to a high extent. In Experiment 5, we found an overwhelming subject/local antecedent preference for intersentential anaphora, as it has been reported for other languages (e.g., Baumann et al., 2010). The results on Mexican Spanish were particularly revealing for AU3
the interpretation of the L2 results, as explained in the following text.

For L2 speakers, we demonstrated that participants at the intermediate proficiency level can achieve nativelike preferences on the use of the first-mention bias, replicating

TABLE 11. Experiment 6: Full model statistics

	Estimate	Std. Error	t value	P value
(Intercept)	2.8	0.27	10.333	0.0001
Condition 1	0.7	0.17	4.246	0.0001
Condition 2	–0.3	0.14	–2.495	0.01
Group	–0.4	0.14	–3.114	0.001
Group*Condition 1	0.001	0.17	0.006	0.9
Group*Condition 2	–0.2	0.14	–1.694	0.09

Note: The model included by-subject random intercepts and by-subject random slopes.

previous research conducted with different L2 populations and using a different methodology (Contemori & Dussias, 2015; Cunnings et al., 2017). This result demonstrates that L2 speakers, like native speakers, can interpret the default referential form signaling topic continuity (i.e., the explicit pronoun; Sorace, 2011). However, L2 speakers' interpretations are comparable to those of native speakers only when the context is relatively simple and one of the antecedents is highly salient. As it clearly emerged in Experiment 4 and 6, L2s' pronoun interpretation is susceptible to the complexity of the discourse. L2 speakers encountered difficulties when the discourse structure was manipulated and two referents with similar salience were presented in the preceding context, as in Roberts et al. (2008). L2 participants were significantly different than native speakers in choosing a subject/local antecedent to interpret both intrasentential (Experiment 4) and intersentential anaphora (Experiment 6). The results of Experiment 4 and 6 are suggestive of cross-linguistic interference. However, our study did not test a group of participants in which the L1 and L2 have a similar set of referring expressions and interpretation biases. Therefore, we cannot make strong conclusions about the interference of the L1. Further research is needed to provide a more detailed description of anaphora resolution across L2 groups with different language pairs. Additionally, in the present research we tested intermediate-proficiency learners of English. We cannot exclude that L2 speakers may achieve nativelike performance on pronoun resolution even in the more complex contexts when they develop higher proficiency in the L2. Future research should investigate how proficiency modulates the effects observed in the present research.

The results of our experiments do not contradict the assumptions of the IH (e.g., Sorace, 2011). According to the IH, interface structures between syntax and pragmatics (as in the case of anaphoric expressions) require an increase use of cognitive resources and are therefore less likely to be successfully acquired by L2 speakers in comparison to structures without this interface. However, in its current formulation the IH does not account for how the complexity of the context may affect the leaners interpretation of subject pronouns.

Cunnings (2017) proposed that the differences observed between L1 and L2 anaphora resolution may result from an increased susceptibility to retrieval interference in L2 processing. Upon reaching the pronoun during online anaphora resolution, a memory-retrieval operation triggers the retrieval of the antecedent. However, competition between the accessible antecedent and another discourse-salient referent that matches the gender of the pronoun can occur (e.g., Roberts et al., 2008; Sturt, 2003). In this case, a processing cost and nonnative interpretation of the pronoun may emerge in L2 learners, as observed in Roberts et al. (2008). In this article we only analyzed offline pronoun interpretation and not online processing. However, under Cunnings's approach we can explain our results as a consequence of retrieval interference. L2 learners may have experienced higher competition between the two salient antecedents that were introduced in the context through coordination. As a result, L2 speakers may have failed to evaluate the saliency of the referents, choosing the nonlocal antecedent more often than native speakers.

To conclude, we propose that our results may best be explained within Cunnings's (2017) L2 memory interference model. Our findings indicate that anaphora resolution may present a challenge for L2 speakers of English under specific circumstances (i.e.,

when the complexity of the discourse increases). This finding is important if compared to the large body of research looking at anaphora resolution in L2 learners of null subject languages, as it reveals the importance of discourse complexity in L2 pronoun resolution.

CONCLUSION

The L2 speakers in the present study performed like native speakers of English in integrating multiple sources of information (syntactic, discourse, pragmatic) to resolve ambiguous pronouns in the intrasentential anaphora and cataphora conditions, demonstrating successful use of the first-mention bias. With a set of six studies, we reconciled findings from previous research, showing that L2 speakers performed differently than native speakers in conditions in which two referents that are equally salient are presented in the previous discourse using a conjoined NP. Using Cunnings's L2 memory interference model, we proposed that L2 speakers may encounter difficulties evaluating the salience of the antecedents in contexts with increased discourse complexity.

NOTES

[1]Although the bias may be determined by various aspects of the saliency of the preceding referent, we adopt here the term *first-mention bias* consistently with existing psycholinguistic literature that looked at the bias in English monolingual speakers (e.g., Arnold, 2010).

REFERENCES

Ariel, M. (1990). *Accessing noun phrase antecedents.* New York, NY: Routledge.

Arnold, J. (2010). How speakers refer: The role of accessibility. *Language and Linguistics Compass*, *4*, 187–203.

Arnold, J., & Griffin, Z. M. (2007). The effect of additional characters on choice of referring expression: Everyone counts. *Journal of Memory and Language*, *56*, 521–536.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278.

Bates, D. M., & Sarkar, D. (2007). *lme4: Linear mixed-effects models using S4 classes*. R package version 0.99875-6. AU4

Baumann, P., Konieczny, L., & Hemforth, B. (2014). Conversational implicatures in anaphora resolution: Alternative constructions and referring expressions. In B. Hemforth, B. Mertins, & C. Fabricius-Hansen (Eds.), *Psycholinguistic approaches to meaning and understanding across languages* (pp. 197–212). Heidelberg, Germany: Springer.

Belletti, A., Bennati, E., & Sorace, A. (2007). Theoretical and developmental issues in the syntax of subjects: Evidence from near-native Italian. *Natural Language and Linguistic Theory*, *25*, 657–689.

Carminati, M. N. (2002). *The processing of Italian subject pronouns* (Unpublished doctoral dissertation). Amherst: University of Massachusetts.

Chamorro, G. (2018). Offline interpretation of subject pronouns by native speakers of Spanish. *Glossa: A Journal of General Linguistics*. doi.org/10.5334/gjgl.256. AU5

Contemori, C., & Dussias, G. (2015). Pronouns in L2 discourse: Evidence from eye-tracking and production. BUCLD 39 Online Proceedings Supplement. AU6

Contemori, C., & Dussias, G. (2016). Referential choice in a second language: Evidence for a listener-oriented approach. *Language, Cognition and Neuroscience*, *31*, 1257–1272.

Cunnings, I. (2017) Parsing and working memory in bilingual sentence processing. *Bilingualism: Language and Cognition*, *20*, 712–721.

Cunnings, I., Fotiadou, G., & Tsimpli, I. (2017). Anaphora resolution and reanalysis during L2 sentence processing: Evidence from the visual world paradigm. *Studies in Second Language Acquisition*, *39*, 621–652.

Ellert, M. (2013). Resolving ambiguous pronouns in a second language: A visual-world eye-tracking study with Dutch learners of German. *International Review of Applied Linguistics*, *51*, 171–197.

Fernandes, E., Luegi, B., Soares, C., de la Fuente, I., & Hemforth, B. (2018). Adaptation in pronoun resolution: Evidence from Brazilian and European Portuguese. *Journal of Experimental Psychology: Learning Memory and Cognition*. doi: 10.1037/xlm0000569. AU7

Filiaci, F., Sorace, A., & Carreiras, M. (2014). Anaphoric biases of null and overt subjects in Italian and Spanish: A cross-linguistic comparison. *Language, Cognition and Neuroscience*, *29*, 825–843.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*, 434–446.

Keating, G. D., Jegerski, J., & VanPatten, B. (2016). Online processing of subject pronouns in monolingual and heritage bilingual speakers of Mexican Spanish. *Bilingualism: Language and Cognition*, *19*, 36–49.

Keating, G., VanPatten, B., & Jegerski, J. (2011). Who was walking on the beach: Anaphora resolution in Spanish heritage speakers and adult second language learners. *Studies in Second Language Acquisition*, *33*, 193–221.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q). *Journal of Speech, Language, and Hearing Research*, *50*, 940–967.

Miltsakaki, E. (2002). Towards an aposynthesis of topic continuity and intrasentential anaphora. *Computational Linguistics*, *28*, 319–355.

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). Amsterdam, The Netherlands: Benjamins.

Roberts, L., Gullberg, M., & Indefrey, P. (2008). Online pronoun resolution in L2 discourse: L1 influence and general learner effects. *Studies in Second Language Acquisition*, *30*, 333–357.

Rothman, J. (2008). How pragmatically odd! Interface delays and pronominal subject distribution in L2 Spanish. *Studies in Hispanic and Lusophone Linguistics*, *1*, 317–339.

Schimke, S., & Colonna, S. (2016). Native and non-native interpretation of pronominal forms: Evidence from French and Turkish. *Studies in Second Language Acquisition*, *38*, 131–162.

Serratrice, L. (2007). Cross-linguistic influence in the interpretation of anaphoric and cataphoric pronouns in English–Italian bilingual children. *Bilingualism: Language and Cognition*, *10*, 225–238.

Shin, N. L., & Cairns, H. S. (2009). Subject pronouns in child Spanish and continuity of reference. In J. Collentine, M. García, B. Lafford, & F. M. Marín (Eds.), *Selected proceedings of the 11th Hispanic Linguistics Symposium* (pp. 155–164). Somerville, MA: Cascadilla Proceedings Project.

Silva-Corvalán, C. (1994). *Language contact and change: Spanish in Los Angeles.* Oxford. UK: Oxford University Press.

Sorace, A. (2011). Pinning down the concept of "interface" in bilingualism. *Linguistic Approaches to Bilingualism*, *1*, 1–33.

Sorace, A., & Filiaci, F. (2006). Anaphora resolution in near-native speakers of Italian. *Second Language Research*, *22*, 339–368.

Sorace, A., Serratrice, L., Filiaci, F., & Baldo, M. (2009). Discourse conditions on subject pronoun realization: Testing the linguistic intuitions of older bilingual children. *Lingua*, *119*, 460–477.

Trnavac, R., & Taboada, M. (2016) Cataphora, backgrounding and accessibility in discourse. *Journal of Pragmatics*, *93*, 68–84.

Van Gompel, R. P. G., & Liversedge, S. P. (2003). The influence of morphological information on cataphoric pronoun assignment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 128–139.

Vogels, J., Krahmer, E., & Maes, A. (2014). How cognitive load influences speakers' choice of referring expressions. *Cognitive Science*, *39*, 1396–1418.

Wilson, F. (2009). *Processing at the Syntax-Discourse Interface in Second Language Acquisition*(Unpublished doctoral dissertation). University of Edinburgh. AU1