Open access • Journal Article • DOI:10.1080/01690965.2013.801502

Anaphoric biases of null and overt subjects in Italian and Spanish: a cross-linguistic comparison — Source link

Francesca Filiaci, Antonella Sorace, Manuel Carreiras

Institutions: University of Edinburgh, University of the Basque Country

Published on: 26 Jun 2014 - Language, cognition and neuroscience (The University of Edinburgh)

Topics: Null (mathematics)

Anaphoric biases of null and overt subjects in Italian and Spanish: a cross-linguistic comparison

Running head: ANAPHORIC SUBJECTS IN ITALIAN & SPANISH

Anaphoric biases of Null and Overt Subjects in Italian and Spanish: a cross-linguistic comparison

Francesca Filiaci, Antonella Sorace

The University of Edinburgh, UK

Manuel Carreiras

Basque Center on Cognition, Brain and Language, Donostia, Spain

IKERBASQUE, Basque foundation for Science, Bilbao, Spain

University of the Basque Country (UPV/EHU), Bilbao, Spain

Corresponding author: Francesca Filiaci

Linguistics and English Language

The University of Edinburgh

Dugald Stewart Building

3 Charles St.

Edinburgh EH8 9AD, UK

e-mail: francesc@ling.ed.ac.uk

Abstract

The present study explores the cross-linguistic differences between Spanish and Italian in the anaphoric interpretation of null subjects and overt pronominal subjects. The availability of null subjects in a language is determined by the parametric settings of its syntax, but their felicitous use as an alternative to overt pronouns depends on contextual conditions affecting how different expressions retrieve their antecedents in the discourse. According to Accessibility Theory (Ariel, 1990, 1991), at least some of these principles must have universal validity (see also Carminati, 2002, for a discussion); however, up to now no experimental research has been carried out with the aim of comparing directly the interpretation of anaphoric dependencies in two typologically similar null subject languages. In this paper, we report the results of two pairs of self-paced reading experiments carried out in Spanish and in Italian. The results show a similar pattern for the resolution of null subjects, as predicted by Accessibility Theory, whereas the resolution of overt pronouns seems to diverge. This suggests that subtle differences restricted to the scope of the overt pronoun yield systematic variation between the two languages.

Anaphoric biases of Null and Overt Subjects in Italian and Spanish: a cross-linguistic comparison

The resolution of anaphoric dependencies has been the focus of investigation within several areas of linguistics and psycholinguistics, where research has shown that the felicitous use and interpretation of anaphoric expressions depends on the interplay of a number of syntactic, semantic, lexical and information structural factors (see for example Kaiser & Trueswell, 2008; Wilson, 2009). In the last decade anaphora resolution has received particular attention within the area of developmental linguistics, as an increasing body of evidence is showing that 'interface' phenomena, which require the integration of different sources of information, can be particularly demanding for both simultaneous and consecutive bilingual speakers (Hulk & Müller, 2000; Sorace, Serratrice, Filiaci, & Baldo, 2009; Sorace & Filiaci, 2006; Tsimpli, Sorace, Heycock, & Filiaci, 2004). However, in order to make generalizations about the cross-linguistic instability of such phenomena, it is necessary to ascertain whether they are governed by the same principles across languages. Recent psycholinguistic studies have started to investigate the question of which principles governing anaphora resolution may be considered universal and which may be more language-specific, or at least related to major typological differences between languages. Kwon and Sturt (2012), for example, compared findings on the use of morpho-syntactic cues for pronoun reference assignment in English with data on the resolution null subjects in Korean. The authors found that, while with English pronouns the parser tries to create a cataphoric dependency with suitable intra-sentential antecedents even if the

pronoun has already been assigned to a discourse topic (Liversedge and Van Gompel, ms), Korean null pronouns assigned to a discourse topic are not re-analysed even if the parser encounters a potential intra-sentential cataphoric antecedent. The authors suggest that this difference may be related to a typological difference between English and Korean, namely to the fact that English (like Italian and Spanish) is a 'syntactic language', in which morpho-syntactic cues play an important role in parsing dependencies compared to discourse cues, whereas Korean (like Chinese and Japanese) is a 'pragmatic language', lacking overt morphology and relying primarily on discourse-pragmatic cues for the interpretation of anaphoric expressions (see Huang, 1984; Huang 1994, 2000; Givón, 1979). Of course, as Huang (1994) points out, it is not the case that 'pragmatic' languages ignore syntax or 'syntactic' languages do not make use of discourse information, but it is more a question of which type information takes priority. If this explanation is on the right track, Kwon and Sturt predict, in languages like Italian and Spanish, which have null pronouns, but are typologically similar to English in their use of overt morphology, pronouns, both null and overt, should behave like in English. The authors also suggest that the difference they found may be due to the fact that overt and empty categories are processed in intrinsically different ways, in which case we would expect similarities across languages in the sensitivity of null and overt categories to different contextual cues.Italian and Spanish are two closely typologically related languages that have long been assumed to instantiate the same settings of the Null Subject Parameter (Jaeggli & Safir, 1989; Huang, 2000), and tend to be treated as having not only the same

inventory of pronominal forms, but also similar conditions governing their interpretation (e.g. Sorace et al., 2009; Sorace & Serratrice, 2009, but see also Alexiadou & Anagnostopoulou, 1998, for possible parametric differences). It is the validity of this assumption that is addressed in this paper through an experimental comparison of pronoun interpretation in the two languages. More precisely, we will test the cross-linguistic validity of the Position of Antecedent strategy, a parsing strategy proposed by Carminati (2002) for the processing of intra-sentential null and overt subject pronouns in Italian. In the next section we will describe the Position of Antecedent strategy and explain how this strategy is motivated by general pragmatic constraints. We will then turn our attention to the relevant characteristics of Italian and Spanish to determine if it is legitimate to compare directly the interpretation of null and overt subjects in the two languages, highlighting similarities and possible differences. In the following section we will review available data on the distribution of subject pronouns in Spanish and, based on this information, we will outline our predictions and present the results of two self-paced reading experiments.

The Position of Antecedent Strategy

Carminati (2002, p.33) proposed the following processing strategy for intra–sentential null and overt subject pronouns in Italian, that she names the Position of Antecedent Strategy: "The null pronoun prefers an antecedent which is in the SpecIP position […], while the overt pronoun prefers an antecedent which is not in the SpecIP position."

SpecIP (that is Spec(ifier) of the Inflectional Phrase) is defined according to standard

generative linguistic theory as the structural position occupied by the preverbal subject of a tensed verb, which is higher in the clause configuration than the position occupied by a direct or indirect object.

This processing strategy states that, in Italian, the null pronoun prefers an antecedent in the preverbal subject position, while the overt pronoun prefers an antecedent in a lower structural position. The idea is that the null subject prefers to co-refer with the most prominent discourse antecedent available while the overt pronoun prefers to skip it, and that the prominence of a potential antecedent is determined by its syntactic position. The validity of the Position of Antecedent Strategy was tested using a series of self-paced reading and sentence completion tasks. Sentence (1) provides an example of the construction tested by Carminati:

(1) a. Quando Maria$_i$ è andata a trovare Vanessa$_j$ in ospedale, lei$_{i(/j)}$ le ha portato un mazzo di fiori.

b. Quando Maria$_i$ è andata a trovare Vanessa$_j$ in ospedale, Ø$_{i(/j)}$ le ha portato un mazzo di fiori.

'When Maria went to visit Vanessa at the hospital, she brought her a bunch of flowers.'

c. Quando Maria$_i$ è andata a trovare Vanessa$_j$ in ospedale, lei$_{(i/)j}$ era già fuori pericolo.

d. Quando Maria$_i$ è andata a trovare Vanessa$_j$ in ospedale, Ø$_{(i/)j}$ era già fuori pericolo.

'When Maria went to visit Vanessa at the hospital, she was already out of

danger.'

As predicted by the Position of Antecedent strategy, longer reading times were found for the second clause of sentences such as (1a), where the overt pronoun is semantically forced to select a subject antecedent, and (1d), where the null subject is forced to co-refer with a non-prominent object antecedent. Carminati's data also suggest that overriding the overt pronoun bias appears to be less costly and more dependent on contextual factors than overriding the null subject bias.

The nature of the Position of Antecedent Strategy

As for the nature of the Position of Antecedent Strategy, Carminati argues that the *biases* it encodes are not grammatical in nature, since if they were their violation would yield outright ungrammaticality rather than infelicity. She suggests instead that they are motivated by universal pragmatic principles, as those outlined by Accessibility Theory (Ariel, 1990, 1991), which in the spirit of Relevance Theory (Sperber & Wilson, 1986, 1995) is based on the idea that the use of a referring expression depends on a balance between its cost and function. Following Almor (1999, 2000), such cost and discourse function of anaphors can be defined in terms of psychological processes. The function of anaphoric expressions, Almor argues, is to act as cues helping comprehenders to identify antecedents that can be more or less activated in their memory: semantically richer anaphors provide better memory cues, but activate more semantic information in the verbal working memory, and thus require more memory resources to be processed. Therefore, the use of a semantically rich anaphor is functionally justified only when the speaker believes that the mental

representation of its antecedent is not readily available (or accessible) to the addressee. As for the accessibility of a given antecedent in the memory, according to Almor, any factor, linguistic or non-linguistic, that draws the attention of the comprehender on a particular entity, increasing its *prominence* in a given context, should increase the activation of its mental representation.

As for the linguistic determinants of prominence, some theories have focused on semantic factors (e.g. Caramazza et al. 1977, Stevenson et al. 1994, 2000, Kehler et al. 2008), others on structural factors (e.g. Gernsbacher 1989, Smyth 1994, Crawlay et al. 1990, Grosz et al. 1983, 1995). Arnold (1998) suggests that several structural and semantic factors contribute jointly to determine the relative prominence of discourse entities, finally, Kaiser & Trueswell's (2008) *Form Specific Multiple Constraints Approach* suggests that, within a language, different but informationally equivalent anaphoric expressions may be sensitive to a different extent to several factors affecting prominence, and therefore reject the idea of a unified, monotonic salience hierarchy. More precisely, they found that while the Finnish pronoun *hän* seems to be biased towards referring to the antecedent that is most prominent in terms of syntactic position (the preceding subject) and is relatively insensitive to word order, the demonstrative *tämä* seems to prefer postverbal referents and this preference is modulated by the syntactic role of the postverbal constituent (it is greater for objects than for subjects). Carminati's Position of Antecedent strategy puts the emphasis on structural determinants of prominence. She argues in fact that Spec IP is associated with higher prominence than lower structural positions.

Carminati therefore assumes that the PAS is not itself a core-grammatical rule, but a processing preference motivated by general cognitive mechanisms, predicting felicity of an expression in context rather than its grammaticality, and that it operates on the basis of prominence relations encoded in the discourse through syntax. This point will be touched again in the last paragraph of the next section.

The linguistic determinants of prominence: are Italian and Spanish comparable?

Apart from Carminati's, other accounts associating the prominence of arguments to their structural position have suggested that prominence is associated to first mentioned participants (Gernsbacher 1989), syntactic subjects (Grosz et al. 1995, Arnold 1998), *focus* (left dislocation in a cleft construction; Arnold (1998)), and topichood (Gundel, Hedberg, & Zacharski, 1993). Gundel and colleagues also argue that the topic-comment structure is at least partly encoded into the syntactic structure of the utterance. This seems to suggest that there may be a relationship between the syntactic notion of subject, the structural and linear position of a constituent, and topichood as an information primitive.

Using data from Spanish and English, Casielles-Suárez (2001, 2004), suggests that utterances can fall into two types of information-structural articulations: the topic-comment articulation and the focus-background articulation. The topic-comment articulation is pragmatically, phonologically and syntactically unmarked, in the sense that it can be used at the beginning of a discourse or out of the blue and it does not

have a marked intonation pattern. The topic is defined as 'the point of departure of the sentence as a message'. In the following examples, 'Mark' is the topic and the rest of the sentence is the comment (the English and Spanish examples are from Casielles-Suárez 2001, the Italian example is our translation).

(2) Mark / took the children to the movies.

(3) Mark / llevó los niños al cine.

(4) Mark / ha portato i bambini al cinema.

In the focus-background articulation, the focus is the informative part of the sentence, it is intonationally marked and separated from the rest of the sentence containing the background, the information that is known to the hearer or recoverable from the context. This articulation is pragmatically marked, that is, it is felicitous only when everything but the focussed information is already known or recoverable from the context (English and Spanish examples are from Casielles-Suárez 2001, the capitalised part of the sentence indicates the focussed constituent).

(5) a. MARK bought a book for the children.

b. Mark BOUGHT a book for the children.

c. Mark bought A BOOK for the children.

(6) a. Un libro a los niños se lo compró MARK.

b. Mark un libro a los niños (sí) se lo COMPRÓ.

c. A los niños Mark les compró un LIBRO.

(7) a. Un libro ai bambini gliel'ha comprato MARCO.

b. Marco un libro ai bambini gliel'ha COMPRATO.

c. Ai bambini Marco ha comprato un LIBRO.

While English marks the focal part of the sentence prosodically, in Spanish and Italian the construction is marked syntactically by dislocating the non-focussed constituents, a type of movement known as clitic-left dislocation (Cinque, 1990).

Thus, Casielles-Suárez identifies two different 'topic-like' information primitives, sentence topics and backgrounds, with different discourse and syntactic characteristics. Backgrounds are necessarily discourse-old and unaccented; they are base generated adjuncts, not restricted to a unique element in the clause. Multiple backgrounds can take any relative order, they are not linked to a specific linear position (i.e. can be dislocated to the left or to the right) and are dislocated to escape the scope of focus (see Casielles-Suárez 2004 and references therein). In Spanish bare nominals are grammatical in this position. These properties are illustrated by the sentences below, where the background is in bold (from Casielles-Suárez 2001 and 2004):

(8) a. **Yo a él libros** no le dejaría.

b. **Libros yo a él** no le dejaría.

c. **A él yo libros** no le dejaría.

(9) a. **Los tomates**, Raquel los odia.

b. Raquel los odia, **los tomates**.

(10) **Libros** (los) hay en la biblioteca.

Sentence topics can be discourse old or discourse new; they are not necessarily unaccented (i.e. can bear contrastive stress); they are linked to a specific structural

position, a unique specifier position where they move via DP movement. This means that they are restricted to one element in the sentence and that bare nominals are excluded from this position (example (12) is from Casielles-Suárez, 2001):

(11) (Qué pasó con Juan y Maria?) Juan ha llegado, Maria se ha quedado en su cuarto.

(12) a. Los niños jugaban en el parque.

b. *Niños jugaban en el parque.

Adopting an analysis along the lines of Rizzi (1997), Casielles-Suárez argues that Spanish subjects optionally move to the preverbal Specifier position to receive a topic interpretation, that is to satisfy the Topic Criterion, or to check a topic feature, in terms of Chomsky (1993, 1995) (see Lozano & Mendikoetxea 2010 and references therein for theoretical accounts of subject movement in *pro* drop languages).

According to this analysis, preverbal subjects in Spanish are syntactically and informationally distinct from left (and right) dislocated constituents, contra Alexiadou & Anagnostopoulou (1998) or Ordóñez and Treviño (1999) who cannot account for the asymmetries in the distribution of preverbal bare nominals. Preverbal subjects are also distinct from arguments occupying lower structural positions (see Casielles-Suárez (2004) for further examples and a discussion of the informational status of the postverbal subjects of unaccusative verbs).

The examples below show that an analysis along these lines can be extended to Italian dislocated constituents and preverbal subjects (examples (13) (14) and (15) are our translation of sentences (8), (9) and (12)).

(13) a. **Io a lui libri** non gliene lascerei.

b. **Libri io a lui** non gliene lascerei.

c. **A lui io libri** non gliene lascerei.

(14) a. **I pomodori** Rachele li odia.

b. Rachele li odia, **i pomodori**.

(15) a. I bambini giocavano nel parco.

b. *Bambini giocavano nel parco.

To summarise, both Italian and Spanish seem to use syntax in a comparable way to encode the information structure of an utterance. Information structural roles impact prominence relations in that they direct the comprehender's attention by providing instructions about the relation between the information contained in the utterance and the object of thought (Vallduví 1990, Lambrecht 1994).

Here it should be pointed out that our assumptions depart slightly from Carminati's, in that we do not think that prominence depends on the structural and grammatical relations between constituents. We think that syntax impacts prominence because some languages, among them Italian and Spanish, use syntax to encode information structure at the level of the utterance (other languages may use different means, for example prosody or morphology), and the information structural status of a constituent determines whether or not such constituent is within the focus of attention of speaker and hearer (see Vallduví 1990).

Overt pronouns in Italian and Spanish: are they equivalent?

Within the framework of Accessibility Theory, anaphoric expressions are accessibility

markers, ranked along a continuous scale, the 'Accessibility Markers Scale', where their position is determined by three universal coding principles: *informativity* (amount of information they encode), *rigidity* (how uniquely they identify an entity), and *attenuation* (their phonological size).

Expressions that are highly informative, rigid and phonologically more conspicuous, are used to retrieve non accessible antecedents, whereas informationally poor, ambiguous and phonologically attenuated expressions are used to retrieve highly accessible antecedents.

Ariel argues that the ranking of referring expressions along the scale is universal, so if pronouns are less informative and rigid than NPs, there should be no language where NPs are used to refer to highly accessible antecedents while pronouns are used for inaccessible ones. However, there are at least two dimensions along which we may expect to find cross-linguistic variation. Firstly, the inventory of expressions can vary from language to language, for example some languages will have the null pronoun in their inventory, others will not. Secondly, the relative distance between expressions along the scale may vary from language to language, depending on the relative weight given to the three coding principles (see Ariel, 2006). This means that even if two expressions in two different languages seem equivalent in form, they could still refer to different levels of accessibility in each language, so for example overt unstressed pronouns can refer to topic antecedents in English, but they normally refer to non-topics in Japanese, a language that includes the null pronoun in its inventory.

A possible difference between the Italian and Spanish Accessibility Markers

scales is that Italian has two series of third person overt pronouns, the *lui*/*lei* series, analysed by Carminati, and the *egli*/*ella* series.[1] Based on a detailed analysis of the Italian pronominal system Cardinaletti and Starke (1999; Cardinaletti 1997) suggest that *lui* and *lei* have the morphological, prosodic, syntactic and referential characteristics of strong pronouns , whereas *egli*/*ella* belong to a separate class of deficient pronouns, called 'weak', sharing some of the characteristics of strong pronouns, for example they can bear semantic and contrastive stress, as well as some of the characteristics of clitics (the most deficient class of pronouns), in that they can only be interpreted through co-reference with a prominent discourse antecedent. Overall the *egli/ella* class is used less frequently in informal registers than the *lui//lei* class. This may be due in part to the different distributional properties of these pronouns, which, in addition to having different interpretational characteristics, are only grammatical in a restricted subset of contexts compared to their strong counterparts. This is illustrated by the following examples from Cardinaletti (1997):

(16) a. Lui /Egli/Gianni ha aderito.

b. Lui/*Egli/Gianni e suo fratello hanno aderito.

c. Ha aderito lui/*egli/Gianni.

d. Lui/*Egli/Gianni, Maria non l'ha appoggiata.

[1] A third series of third person subject pronouns in Italian is the *esso/essa* series. These pronouns are also very infrequent in oral and written Italian. According to the classification proposed by Cardinaletti and Starke (1999) they also belong to the class of weak pronouns, and descriptive grammars (Serianni, 1991) say that tend to be used to refer to animals and things rather than people. Like Carminati, we will not focus on these pronouns in this study.

e. Chi è venuto? Lui/*Egli/Gianni.

Example (16a) shows that both strong and weak pronouns as well as full NPs can appear in the preverbal position, but the weak pronoun cannot be coordinated (16b), it cannot appear in postverbal position (16c), in left dislocated position (16d), and it cannot be used in isolation, as an answer to a narrow focus question (16e). For a more detailed and principled explanation of these properties see Cardinaletti and Starke (1999).

Notice also that it is not uncommon for languages to have more than one series of personal pronouns with different strengths (see Cardinaletti and Starke, 1999 for Indo-European languages, Bresnan and Mchombo, 1987 on Chichewa). Like Carminati, here we will focus on the more frequently used class of strong pronouns *lui* and *lei.*

To sum up, there are differences between the Spanish and Italian pronominal system and these differences may have an impact on the properties of overt pronouns in each language, in terms of how they are used and interpreted in context, or more precisely in terms of their sensitivity to the accessibility of their antecedents. To our knowledge there has been no attempt to compare directly across these two languages the interpretation of formally equivalent anaphoric expressions.

Data on the distribution of Spanish pronouns

The empirical evidence for these distinctions in the scope of pronominal forms in Spanish has been ambiguous so far. Variationist studies investigating the distribution

of null and overt subjects in several varieties of Spanish, both in monolingual and in bilingual populations (usually Spanish-English) (Enriquez 1984, Hochberg 1986, Cameron 1992, Morales 1997, Silva-Corvalán 1994, Flores-Ferrán 2004, Montrul 2004, Otheguy et al. 2007) have consistently reported a correlation between the occurrence of pronominal subjects and a switch in subject reference from the previous tensed verb, as in (17), whereas null subjects tend to be used when reference to the same subject is maintained across clauses, as in example (18) (from Flores-Ferrán 2004, p. 119):

(17) **Yo** quiero que **tú** sepas que **nosotros** te íbamos a botar como bolsa.

'I want you to know that we were going to throw you out like a bag'

(18) Y de regreso ∅ me acordé que ∅ tenía un montón de correspondencia en casa de mi amigo José de los bancos, y eso porque ∅ tuve que [...].

'and upon returning, [I] remembered that [I] had a bunch of mail in the house of a friend, José of the banks, and that was because [I] had to [...].'

While this data seem to indicate at first that also Spanish obeys the biases encoded by the Position of Antecedent strategy, it should be noted that such pattern of distribution could also be found in a language where only one bias is at work, either the null subject or the overt pronoun bias. So this type of data alone cannot confirm directly that both null and overt pronouns in Spanish are constrained by the same processing biases as in Italian.

Moreover, while developmental studies on Italian have consistently shown that in

situations of language contact, speakers tend to weaken, or acquire latest, contextual restrictions on the use of overt pronouns (Tsimpli et al. 2004, Sorace & Filiaci 2006, Sorace et al. 2009) the evidence has not been very conclusive for studies on Spanish (Silva-Corvalán 1994, Flores-Ferrán 2004, but see also Montrul 2004 or Otheguy et al. 2007). One possible explanation for this asymmetry is that the Spanish overt pronoun may not obey the same contextual restrictions as the Italian one, and if the Spanish overt subject is not constrained by an antecedent preference in a given context, then such preference could not be lost as a result of language contact.

An experimental study by Alonso-Ovalle, Fernández-Solera, Frazier and Clifton (2002) suggests that Spanish speakers do not show any strong preference for one interpretation over the other when asked to interpret an overt anaphoric subject pronoun which could ambiguously refer to the previous subject or object, performing roughly at chance level. This again would suggest that Spanish overt pronouns do not obey the same pragmatic restrictions as Italian subject pronouns.

A recent study by Gelormini-Lezama & Almor (2011) looked at the processing of repeated names, overt pronouns and null pronouns in Buenos Aires Spanish. The authors claim that the distinctive features of this dialect are mainly lexical and phonological, nevertheless corpus data (Morales 1997) show for example that the proportion of pronominal preverbal subject realisation is different across Madrid Spanish (28% realisation for *yo* and *tú*, 8% for *él/la*, *ellos*/*as*, 22% for arbitrary pronouns (*uno*, *tú*)), San Juan Spanish (54%, 35% and 69% respectively) and Buenos Aires Spanish (30%, 18% and 48% respectively). This may suggest the presence of

different constraints, so it cannot be taken for granted that the results for the Buenos Aires variety can be generalised to other varieties. Another difference with the present study is that Gelormini-Lezama and Almor look at extra-sentential anaphora contexts. The first experiment in the study shows that, with a subject antecedent, sentences with overt pronouns and repeated names are read significantly more slowly than null subject sentences; when the antecedent is an object, null subject sentences are read slowest. One possible problem with this analysis is that sentences with a null anaphor were systematically shorter than sentences with an overt pronoun, which were systematically shorter than sentences with proper names. The data also suggest, but this is not discussed by the authors, that in the context analysed the null subject bias might be stronger than the overt subject bias, since the difference in reading times seems to be larger in this condition.

Finally, looking at language acquisition research, Bini (1990) found that native speakers of Spanish tended to overproduce overt subject pronouns in second language Italian. Similarly, and contrary to expectations, Sorace et al. (2009) found that bilingual children exposed to Spanish and Italian from birth accepted the use of overt pronouns in Italian to refer to a prominent antecedent significantly more often than their monolingual peers, and similarly to English-Italian bilinguals. On the other hand both bilingual groups performed like their monolingual peers in the null subject condition. Sorace and colleagues suggested that these results rule out the possibility that cross-linguistic influence may be the reason of the non-target performance, and that this was due instead to the fact that multilingual speakers need to use overt

pronouns as a default, when they are not able to coordinate different sources of information to chose the anaphoric expression that is most felicitous in the given discourse context (Sorace et al, 2009, Sorace, 2011). An alternative explanation could be that Spanish pronouns (similarly to the English ones) are not constrained by the same contextual bias that constrains their Italian counterparts (i.e. the syntactic position of the antecedent), and that cross-linguistic influence is indeed the reason for the different performance.To summarise, the data available for Spanish suggest that the resolution of subject pronouns in Spanish may obey biases similar to those encoded by the Position of Antecedent strategy, but it is not clear whether both biases (the null subject bias and the overt subject bias) apply and whether they and whether their strength is comparable to Italian.

Aim and predictions for the present study

To recap, we want to test the assumption that the contextual restrictions on the interpretation of null and overt subjects in Italian and Spanish are equivalent, and more precisely the validity in Spanish of the Position of Antecedent strategy, claiming that, for intra-sentential anaphors in Italian, the prominence of an antecedent depends on its structural position, so that reduced anaphors prefer an antecedent in the structurally prominent IP specifier. On the other hand, we acknowledge that prominence is likely to be determined simultaneously by multiple contextual factors and that within a language and across languages different expressions may be sensitive to a different

extent to several determinants of prominence.

We argue that preverbal subjects in Italian and Spanish occupy an equivalent position and that in both languages the particular prominence associated to this position depends on the information status of *topic* associated with it.

We also noticed there are differences between the Spanish and Italian pronominal system and point out that this might result in systematic cross-linguistic differences in the function of formally similar expressions across the two languages.

Finally we reviewed studies showing the presence of a division of labour between null and overt subjects in Spanish. On the other hand, evidence from second language acquisition, asymmetries found comparing studies on language contact (Spanish and English vs. Italian and English), and similarities between overt subject interpretation in bilingual learners of Spanish /Italian and English/Italian suggest that there may be differences across the two languages in the conditions determining the use of overt subjects (or at least the relative weight of these conditions).

We also need to point out here that varieties of Spanish spoken across the world may have different restrictions on *pro*-drop. The data of Morales (1997) cited above suggest that Buenos Aires Spanish may differ from both Caribbean (see Cameron (1992) on Puerto Rican), and Iberian varieties in terms of subject drop. Enríquez (1984), reports, in her corpus study of Madrid Spanish, a rate of pronoun expression around 20%, which is comparable to the rate of Italian (see Lorusso et al., 2005). Since our aim is to compare two languages assumed to be equivalent, we will focus on a comparison between Italian and Iberian Spanish, which we would expect to behave

most similarly. Obviously, we do not expect our findings to be generalisable to other varieties of Spanish, but more research is needed to establish the presence and extent of the differences across varieties.

In the next section we present two self-paced reading experiments in which the same set of materials was adapted and translated into the two languages. The first experiment used the same materials and methodology as Carminati (2002). The four experimental conditions were those illustrated in example (1) above: a subordinate clause introduced two antecedents, one in the subject and one in the object position, and was followed by a main clause containing a temporarily ambiguous null or overt anaphoric subject, disambiguated by the plausibility of the sentence towards one of the two antecedents. In these experiments the materials were presented clause by clause and the reading times for the clause containing the anaphoric subject were measured and analysed. Based on the studies on bilingual acquisition and on language contact cited above and on accessibility theoretic assumptions on the Accessibility Markers scale, we would expect to find no cross-linguistic differences in the conditions affecting the interpretation of null subjects, that is a similar processing penalty when null subjects are forced to retrieve a (syntactically) non-prominent antecedent. If in Spanish the overt pronouns are equivalent to Italian strong pronouns and constrained by the same contextual bias (a sensitivity to the syntactically encoded prominence of the antecedent and a preference for a less prominence), we should find a penalty in both languages when pronouns are forced to pick a prominent subject antecedent.

Experiment 2 presented the same four experimental conditions (null or overt anaphoric subjects with subject or object antecedents), but the materials were adapted to be presented phrase-by-phrase, to provide a more sensitive measure of any processing penalties and an insight on the time-course of the anaphora resolution. The sentences were disambiguated at the verb phrase following the anaphoric subject, which was followed by another phrase and by a final wrap up phrase. The results of both experiments confirm the presence of comparable biases in the resolution of null subjects, but highlight the presence of cross-linguistic differences in the anaphoric preferences of overt subjects.

Experiment 1

The aim of Experiments 1 and 2 was to replicate Carminati's (2002) results for Italian and compare them with equivalent data for Spanish. The expectation, based on Accessibility theoretical assumptions, is that we should not find cross-linguistic differences in the resolution of null subjects, but that differences may arise with respect to the overt subjects. Therefore, we expect to find in both languages a significant penalty when a null subject is forced to refer to an antecedent that is not the most prominent in the previous discourse (i.e. one that does not occupy the syntactic preverbal subject position). As for the overt pronoun, in Italian it should incur a processing penalty when it is forced to co-refer with the most prominent antecedent, in Spanish, it should incur a comparable penalty if it occupies an equivalent position along the Accessibility Markers scale, however there are at least some indications that it may actually occupy a position that is relatively closer to the

null subject in which case we expect to find no effect due to the antecedent position in this language, as well as a significant effect for the language variable.

Method

Participants. Two groups of participants, 32 monolingual speakers of Spanish and 32 monolingual speakers of Italian, were recruited among Spanish and Italian adults taking English summer courses in Edinburgh and Erasmus students at Edinburgh University. Participants had been living in Edinburgh (or in another English speaking country) only for a brief period at the time of the experiment (for the Italian group the mean number of months spent abroad was 2.4, SD = 4.3; for the Spanish group the mean was 3, SD = 4.9), the likelihood of attrition with English was therefore minimal as shown by previous studies on first language attrition at the level of the syntax-pragmatics interface (Tsimpli et al 2004). Spanish speakers were asked about their origin and only speakers from Spain were included in the study to control for dialectal variation.

Stimuli. The materials for this study were taken from Experiment 1 of Carminati (2002); they were adapted and translated into Spanish so that two equivalent sets of 16 items were created, one in each language. The filler sentences (n = 86) tested other aspects of anaphora resolution (manipulating the number of possible antecedents, the position of the subject antecedent, the type of disambiguation see Carminati, 2002, for the details). Like the experimental items, both in this experiment and in the next, the filler items were the same sentences translated across the two languages. The experimental sentences consisted of a subordinate clause followed by a main clause

containing an anaphoric subject. This anaphoric subject could be either null or overt (Overt Pronoun Anaphora condition and Null Subject Anaphora condition). The antecedent of the anaphoric subject was temporarily ambiguous and could correspond to either the subject or the object of the subordinate clause (Subject and Object Antecedent conditions). The temporarily ambiguous antecedent was disambiguated semantically by the plausibility of the sentence. Below is an example of experimental sentence in Italian and in Spanish, in the four experimental conditions originated crossing the Anaphora and Antecedent variables, Language was introduced as a between subjects variable:

(19) a. Dopo che Giovanni$_i$ ha criticato Bruno$_j$ così ingiustamente, / lui$_i$ si è scusato ripetutamente.

a'. Después de que Bernardo$_i$ criticó a Carlos$_j$ tan injustamente, / él$_i$ le pidió disculpas.

b. Dopo che Giovanni$_i$ ha criticato Bruno$_j$ così ingiustamente, / ∅$_i$ si è scusato ripetutamente.

b'. Después de que Bernardo$_i$ criticó a Carlos$_j$ tan injustamente, / ∅$_i$ le pidió disculpas.

'*After that John$_i$ has criticised Bruno$_j$ so unjustly, he$_i$ apologized repeatedly*.'

c. Dopo che Giovanni$_i$ ha criticato Bruno$_j$ così ingiustamente, / lui$_j$ si è sentito offeso.

c'. Después de que Bernardo$_i$ criticó a Carlos$_j$ tan injustamente, / él$_j$ se

sintió muy ofendido.

d. Dopo che Giovanni$_i$ ha criticato Bruno$_j$ così ingiustamente, / ∅$_j$ si è sentito offeso.

d'. Después de que Bernardo$_i$ criticó a Carlos$_j$ tan injustamente, / ∅$_j$ se sintió muy ofendido.

'*After that John$_i$ has criticised Bruno$_j$ so unjustly, he$_j$ felt offended.*'

Four experimental lists were created in each language. Half of the experimental items and fillers were followed by a comprehension question, asking to identify the antecedent of the anaphoric subject, to encourage participants to resolve the anaphors.

Procedure. The experiment consisted in a clause-by-clause self-paced reading task. Each clause appeared at the press of the space bar, in a moving window, on the screen of a 13" MacBook, using Psyscope X (Cohen, MacWhinney, Flatt, & Provost, 1993). When the sentence was followed by a comprehension question, the two possible answers appeared, together with the question, at the bottom of the screen, one on the left-hand side and one on the right-hand side. The participant was instructed to choose the correct answer by pressing the 'F' key on the keyboard (to chose the left-hand side answer) or the 'J' key (to chose the right-hand side). Each answer (subject or object antecedent) appeared half of the times on the left-hand side of the screen and half of the times on the right-hand side.

All responses were collected through the computer keyboard. The instructions were presented in written form, at the beginning of the experiment, in the native language of

the participant. The items were randomised at every run.

Results

In order to take into account the systematic differences in the length of the stimuli between the two languages, the RTs were adjusted for the number of characters, by computing for each participant the correlation between the RTs and the segment lengths and then calculating the residuals. The residuals of the RTs for the second clause, containing the anaphoric subject, were submitted to a 2 x 2 x 2 ANOVA, with Anaphora and Antecedent as within subject variables and Language as a between subjects variable. The RTs for the comprehension questions and the error rates were also analysed.

(figure 1 about here)

Main Clause RTs. The mean reading times and 95% confidence intervals for the main clause are shown in Figure 1(a) and (b). As we can see from the bar charts, the pattern of the RTs appears to be the same in both languages in the null subject condition, with a slightly larger effect in Italian. In contrast, the difference between the two languages becomes more apparent in the overt pronoun condition.

The ANOVA showed a significant main effect for Anaphora ($F_1(1, 62) = 6.23$; $p < .05$; $F_2(1, 30) = 8.72$; $p < .01$) with null subject sentences read overall faster than overt pronoun sentences (-60 msec. vs. 77 msec.); no main effect for Antecedent and no main effect for Language; a significant two-way interaction between Anaphora and Antecedent ($F_1(1, 62) = 37.81$; $p < .000$; $F_2(1, 30) = 42.85$; $p < .000$); no significant two-way interactions with the Language variable; and a significant three-way

interaction between Language, Anaphora and Antecedent ($F_1(1, 62) = 6.64$; $p < .05$; $F_2(1, 30) = 7.77$; $p = .01$).

In order to understand the interaction, we analysed the two Anaphora conditions (Null and Overt) separately. In the null subject condition there was a significant effect for Antecedent ($F_1(1, 62) = 21.91$; $p < .000$; $F_2(1, 30) = 21.47$; $p < .000$) but no effect for Language and no significant interaction between Language and Antecedent. This result confirms our prediction that in both languages RTs would be faster for null subject anaphors retrieving a subject antecedent (-276 msec. vs. 156 msec.).

In the overt pronoun Anaphora condition, there is a significant effect for Antecedent ($F_1(1, 62) = 16.12$; $p < .000$; $F_2(1, 30) = 20.21$; $p < .000$), with overall faster reading times for object antecedents (-92 msec. vs. 246 msec.), together with an interaction between Language and Antecedent that is marginally significant in the analysis by subjects and fully significant by items ($F_1(1, 62) = 3.76$; $p < .06$; $F_2(1, 30) = 4.83$; $p < .05$). No main effect for Language was found.

This interaction suggests that the Antecedent preferences with overt pronoun sentences vary depending on the language. More precisely, overt pronoun sentences referring to an object antecedent are read significantly faster than those referring to a subject antecedent only in Italian ($F_1(1, 31) = 34.82$; $p < .000$; $F_2(1, 15) = 18.77$; $p < .000$), whereas this effect is not significant in Spanish ($F_1(1, 31) = 1.45$; $p = .238$; $F_2(1, 15) = 3.27$; $p = .091$). If we compare across languages, we find that the overt pronoun sentences with an object antecedent are processed faster in Italian than in Spanish (-187 msec. vs. 3 msec.) although this effect is significant only in the analysis

by subjects (F_1 (1, 62) = 4.42; $p < .05$; F_2(1, 30) = 1.53; $p = .226$). There is no significant difference between the two languages when the overt pronoun retrieves a subject antecedent.

Finally, if we consider the sentences with a subject Antecedent, they are read significantly faster when they contain a null subject (-277 msec. vs. 246 msec.; F_1 (1, 62) = 31.78; $p < .000$; F_2(1, 30) = 4.37; $p < .05$), there is no main effect for Language and an interaction between Language and Anaphora is only significant in the analysis by items (F_1 (1, 62) = 2.36; $p = .13$; F_2(1, 30) = 56.92; $p < .000$). Similarly, sentences with object as an Antecedent yield a significant main effect for Anaphora, suggesting that they are read significantly faster when they contain an overt pronoun (-92 msec. vs. 157 msec.; F_1 (1, 62) = 11.68; $p < .001$; F_2 (1, 30) = 10.07; $p < .005$), but they also yield a fully significant interaction between Anaphora and Language (F_1 (1, 62) = 6.18; $p < .05$; F_2(1, 30) = 5.40; $p < .05$). This interaction shows that in the null subject anaphora condition there is no main effect for Language, that is both Spanish and Italian speakers encounter the same processing penalty when an object antecedent is retrieved by a null anaphor, but when the object antecedent is retrieved by an overt pronoun, as we have seen in the previous paragraph, Italian participants read the sentences significantly faster than Spanish participants, although this effect was only significant in the analysis by subjects.

(table 1 about here)

(figure 2 about here)

Comprehension questions. Figures 2(a) and 2(b) show the mean reaction times and

95% confidence intervals for the comprehension questions in the null subject and overt pronoun condition respectively. Table 1 shows the percentages of wrong answers.

Overall these data confirm the pattern found in the main clauses reading times. In both languages there were faster RTs and lower error rates for null anaphors retrieving an antecedent in the subject position. In contrast, there were opposite trends across the two languages in the overt anaphora condition, with larger effects in Italian than in Spanish.

An ANOVA performed on the RTs revealed a main effect for Language ($F_1(1, 62) = 6.08$; $p < .05$; $F_2(1,14) = 9.74$; $p < .01$), with overall shorter RTs for Spanish than for Italian; a significant interaction Anaphora by Antecedent ($F_1(1, 62) = 10.01$; $p < .005$; $F_2(1, 14) = 17.06$; $p < .001$) and a significant three-way interaction Anaphora by Antecedent by Language ($F_1(1, 62) = 8.01$; $p < .01$; $F_2(1, 14) = 12.86$; $p = .005$).

In the null subject anaphora condition the analyses reveal a significant effect for Language ($F_1(1, 62) = 4.8$; $p < .05$; $F_2(1, 14) = 6.75$; $p < .05$), with questions following null subject sentences answered significantly faster in Spanish than in Italian. There is also a significant effect for Antecedent ($F_1(1, 62) = 12.05$; $p < .000$; $F_2(1, 14) = 13.91$; $p < .005$), with questions following null subject clauses with a subject antecedent answered significantly faster, both in Italian and in Spanish. In the overt pronoun condition the analyses revealed no main effect for Language but a significant interaction Antecedent by Language ($F_1(1, 62) = 4.92$; $p < .05$; $F_2(1, 14) = 7.22$; $p < .05$). In Italian, questions following an overt pronoun sentence are answered

faster if the antecedent is in the object position, although this effect is fully significant only in the analysis by items ($F_1(1, 31) = 3.41$; $p < .08$; $F_2(1, 7) = 5.69$; $p < .05$). In Spanish no significant difference between the two Antecedent conditions was found ($F_1(1, 31) = 1.57$; $p = .22$; $F_2(1, 7) = 1.6$; $p = .25$).

As for the error rates, the data was analysed using logistic regression, as is suitable for the analysis of categorical data (Baayen 2008). An answer was coded as correct when the participant indicated as the referent of the anaphora the antecedent that was more plausible with the semantics of the sentence. Table 1 shows the percentages of errors to the comprehension questions. Once again, the trend seems to be the same across languages in the null subject condition, whereas with overt pronouns the error rate varies depending on the language. The predictors in the logit model were: Anaphora, Antecedent and Language; for each significant effect we report the coefficient β, its level of significance and the odds ratio (e^{β}) between the pair of levels the effect refers to. Overall the model yielded a significant main effect for Antecedent, with significantly more correct answers to questions following a sentence with a subject antecedent ($\beta = 1.99$; $p = .002$, $e^{\beta} = 7.35$); a significant main effect for Anaphora, with significantly more correct answers in the overt pronoun condition ($\beta = 1.08$; $p = .028$, $e^{\beta} = 2.94$); a significant interaction Anaphora by Antecedent ($\beta = -3.07$; $p < .000$, $e^{\beta} = .046$); and a significant three–way interaction between Antecedent, Anaphora and Language ($\beta = 3.03$; $p = .003$, $e^{\beta} = 20.8$). The likelihood ratio test for the model indicates that overall the model is explanatory. ($\chi^2(7) = 30.29$; $p < .000$), on the other hand the residual deviance is larger than expected ($\chi^2(248) = 342.31$; $p < .000$)

indicating a lack of goodness of fit.

Two other models were fitted to part of the data, the null subject and the overt pronoun data. In the null subject condition the chances of a correct answer increase significantly in both languages when the antecedent is a subject ($\beta = 1.99$; $p = .002$, $e^{\beta} = 7.35$). In the overt pronoun condition, there are slightly less chances to get a correct answer when the antecedent is a subject ($\beta = -1.08$; $p = .028$, $e^{\beta} = 0.339$); there are also slightly less chances to get a correct answer in Spanish than in Italian ($\beta = -1.08$; $p = .028$, $e^{\beta} = 0.339$); and the interaction between Antecedent and Language is significant ($\beta = 2.01$; $p = .003$, $e^{\beta} = 7.45$) showing that in Italian there are significantly less chances to get a correct answer after a subject antecedent ($\beta = -1.08$; $p = .028$, $e^{\beta} = 0.34$), whereas the opposite is true ($\beta = 0.93$; $p = .049$, $e^{\beta} = 2.53$) for Spanish.

Discussion

The results of the first pair of experiments provide some evidence that overt pronouns in Italian and Spanish do not respond to the same processing biases in intra-sentential anaphora contexts. While in Italian overt pronouns are resolved significantly faster when they refer to a syntactically non-prominent antecedent, in Spanish, there were no differences between overt pronouns referring to preverbal subjects and structurally lower object antecedents. However, for null subjects, there is no indication of cross-linguistic differences..

The RTs for the comprehension questions and the accuracy of the answers confirm the reading times data. In the null subject condition, questions are answered faster and

more accurately when the anaphora retrieves a prominent (subject) antecedent. In the overt pronoun condition, Spanish speakers answer more accurately when the pronoun retrieves a prominent antecedent, whereas in Italian answers are faster and more accurate when the overt pronoun refers to a non-prominent antecedent.

Nonetheless, the relevant interaction between the Language and Antecedent variables supporting the hypothesis of a cross-linguistic difference between the processing of Italian and Spanish overt pronouns, only yielded marginally significant results by subjects. We therefore wondered if the use of a more stringent methodology and more accurate measurements would yield fully significant results.

Furthermore, if there is a cross-linguistic difference in the interpretation of overt subjects in Italian and Spanish, we may want to try and identify its source. If equivalent anaphors are sensitive to different determinants of prominence (i.e. syntactic position, linear position, or discourse status of the antecedent), it is possible that these sources of information are evaluated at different stages by the parser. While there is evidence that processes leading to pronoun identification can be initiated early by the parser, as soon as the relevant information is encountered (Cacciari, Carreiras, & Cionini Barbolini, 1997; Arnold, Eisenband, Brown-Schmidt, & Trueswell, 2000), there is also evidence that at least some processing can be delayed until the end of the sentence (Greene, McKoon, & Ratcliff, 1992; McDonald & MacWhinney, 1995).

A second experiment was carried out to address these issues. Its aim was twofold: to see if the crucial interaction between Language and Antecedent with overt pronouns could reach full significance using a more stringent methodology and a finer grained

technique, and to get some information about the time course of anaphora resolution. In the second experiment the items were presented phrase by phrase, the plausibility information was provided early in the sentence, at the verb phrase, which was followed by another phrase and by a final sentence wrap up phrase. . A second methodological change was that a button box was used to collect the responses, allowing for more accurate timing. Finally, the type of fillers items used was changed to sentences with a different structure from the experimental sentences to avoid a possible habituation on the part of the participants.

Experiment 2

Experiment 2 used the same experimental design and technique as Experiment 1. The only difference is that the materials were designed to be presented phrase by phrase.

Method

Participants. Two groups of participants took part in this experiment: 32 adult monolingual speakers of Spanish, and 32 adult monolingual speakers of Italian. They were recruited among undergraduate and postgraduate students at the University of La Laguna (Spain) and at the University of Padua (Italy) respectively.

Stimuli. The design was the same as for Experiment 1; two variables were manipulated within subjects: the type of anaphora (null subject or overt pronoun) and its antecedent (the subject vs. the object of the previous sentence). The experimental sentences consisted of a subordinate clause, introducing the two antecedents, followed

by a main clause, containing the null or overt anaphoric subject. The antecedent was temporarily ambiguous between the subject and the object of the previous clause and was disambiguated semantically at the verb phrase. The same items (n = 48) were translated into Italian and Spanish to obtain two equivalent sets of materials. Below is an example of an item in the four experimental conditions:

(20) a. Quando Carlo$_i$ ha chiesto aiuto a Diego$_j$ per preparare l'esame, lui$_i$ lo ha superato con voti eccellenti.

a'. Cuando Carlos$_i$ pidió ayuda a Diego$_j$ para preparar el examen, él$_i$ aprobó con notas excelentes.

b. Quando Carlo$_i$ ha chiesto aiuto a Diego$_j$ per preparare l'esame, ∅$_i$ lo ha superato con voti eccellenti.

b'. Cuando Carlos$_i$ pidió ayuda a Diego$_j$ para preparar el examen, ∅$_i$ aprobó con notas excelentes.

'*When Carlo asked help to Diego to prepare the exam, he passed it with excellent marks.'*

c. Quando Carlo$_i$ ha prestato aiuto a Diego$_j$ per preparare l'esame, lui$_i$ lo ha superato con voti eccellenti.

c'. Cuando Carlos$_i$ ayudó a Diego$_j$ a preparar el examen, él$_i$ aprobó con notas excelentes.

d. Quando Carlo$_i$ ha prestato aiuto a Diego$_j$ per preparare l'esame, ∅$_i$ lo ha superato con voti eccellenti.

d'. Cuando Carlos$_i$ ayudó a Diego$_j$ a preparar el examen, ∅$_i$ aprobó con

notas excelentes.

'*When Carlo gave help to Diego to prepare the exam, he passed it with excellent marks.'*

The plausibility of the sentences was checked by native speakers in each language and across languages with the help of a Master student in translation native speaker of Spanish and specialised in Italian.

Four experimental lists were created for each language and 40 filler sentences were included in each list. To make sure that the participants engaged in the resolution of the anaphora, half of the experimental items and half of the fillers were followed by a comprehension question, asking to identify the antecedent of the anaphoric subject. The filler sentences were the same across the to languages (translated from Spanish into Italian) and tested ambiguities in relative clause attachment. We decided to use different filler items in this experiment so that the participants would engage with a different type of ambiguity and not focus on the same type of structure throughout the experiment.

Procedure. The experiment consisted in a phrase-by-phrase self-paced reading task. The sentences appeared at the press of a button of a USB button box, phrase by phrase, in a moving window, on the screen of a 13" MacBook; the experiment was run using Psyscope X software (Cohen et al., 1993).

When a sentence was followed by a comprehension question, two possible answers appeared, at the press of a button, at the bottom of the screen, one to the left and one

to the right. The participant was instructed to choose the correct one by pressing a button on the right side or on the left side of the button box. Each answer (subject or object antecedent) appeared half of the times on the left-hand side of the screen and half of the times on the right-hand side.

All responses were collected through the USB button box, this was changed from the previous experiment in order to get more accurate timings. The instructions were presented in written form, at the beginning of the experiment, in the native language of the participant. The items were randomised at every run.

Results

The regions, analysed, shown in the example below, were: the verb phrase of the main clause (i.e. the one following the anaphoric subject), which provided the semantic information to disambiguate the anaphora; the region following the verb phrase (VP+1); and the final wrap up region.

(21) Cuando /Antonio /pidió ayuda /a Diego /para preparar /el examen,/

(él) // aprobó / con notas / excelentes.

Anaphor Verb Phrase VP+1 WRAP UP

(figure 3 about here)

The data were submitted to a 2 x 2 x 2 ANOVA with Antecedent (Subject vs. Object) and Anaphora (null subject vs. overt pronoun) as within subject factors and Language (Spanish vs. Italian) as between subjects factor. The reaction times to the comprehension questions and the accuracy rate of the answers were also analysed.

Verb phrase. Figures 3(a) and 3(b) show the mean reading times and 95% confidence

intervals for the verb phrase region in the null subject and overt pronoun condition respectively.

The verb phrase region shows an overall significant effect for Anaphora ($F_1(1, 62) = 44.24$; $p < .000$; $F_2(1, 94) = 63.76$; $p < .000$) showing that this region is read significantly faster when it is preceded by an overt pronoun (51 msec. vs. -51 msec.). A main effect for Antecedent is marginally significant by subjects ($F_1(1, 62) = 2.99$; $p < .09$) and fully significant by items ($F_2(1, 94) = 8.41$; $p < .005$), showing that verb phrases tend to be read faster when the subject reference is maintained across clauses (-17 msec. vs. 17 msec.). No effects were found involving the Language variable and there were no interactions.

VP+1. Figures 4(a) and 4(b) show the mean reading times and 95% confidence intervals for the region following the verb phrase in the null subject and overt pronoun condition respectively.

(figure 4 about here)

The analysis of this region revealed a fully significant main effect for Anaphora ($F_1(1, 62) = 7.34$; $p < .05$; $F_2(1, 94) = 9.55$; $p < .005$), with faster RTs for sentences beginning with an overt pronoun (-19 msec. vs. 19 msec.) like in the verb phrase region; a fully significant effect for Antecedent ($F_1(1, 62) = 19.85$; $p < .000$; $F_2(1, 94) = 5.34$; $p < .05$), with overall faster RTs in sentences maintaining the subject reference across clauses (-19 msec. vs. 19 msec.); a fully significant interaction Anaphora by Antecedent ($F_1(1, 62) = 19.85$; $p < .000$; $F_2(1, 94) = 21.10$; $p < .000$); and a marginally significant three-way interaction Antecedent by Anaphora by Language

($F_1(1, 62) = 3.5$; $p < .07$; $F_2(1, 94) = 3.76$; $p < .06$). No main effect for Language was found.

In order to investigate the interaction, we analysed separately the null subject and the overt pronoun Anaphora conditions. In the null subject condition the ANOVA revealed a highly significant main effect for Antecedent ($F_1(1, 62) = 16.42$; $p < .000$; $F_2(1, 94) = 16.99$; $p < .000$) with faster reading times when the antecedent is the previous subject (72 msec. vs. -34 msec) and no main effects or interactions with the Language variable. In the overt subject condition a main effect for Antecedent is only significant by subjects ($F_1(1, 62) = 5.38$; $p < .05$; $F_2(1, 94) = 2.72$; $p = .10$), suggesting that in this condition there is a tendency for faster reading times with object antecedents (-34 msec. vs. -4 msec.). Like in the null subject condition, there are no main effects or interactions with the Language variable. The same interaction shows that when subject reference is maintained across clauses, there is a marginally significant effect of Anaphora ($F_1(1, 62) = 3.69$; $p < .06$; $F_2(1, 94) = 3.34$; $p < .08$), with marginally faster reading times in both languages for sentences with a null subject (-34 msec. vs. -4 msec). When the antecedent is in the object position, the reading times are significantly faster in the overt pronoun condition ($F_1(1, 62) = 17.76$; $p < .000$; $F_2(1, 94) = 22.97$; $p < .000$; -22 msec. vs. 43 msec.).

(figure 5 about here)

Wrap up. Figures 5(a) and 5(b) show the mean reading times and 95% confidence intervals for the wrap up region.

The ANOVA for this region shows a main effect for Anaphora that is marginally significant by subjects and fully significant by items ($F_1(1, 62) = 3.36$; $p < .08$; $F_2(1, 94) = 5.93$; $p < .05$), with overall faster reading times for the over pronoun condition than the null subject condition (-49 msec. vs. 51 msec.); a highly significant main effect for Antecedent ($F_1(1, 62) = 22.62$; $p < .000$; $F_2(1, 94) = 27.48$; $p < .000$), with faster reading times for sentences with a subject antecedent (-124 msec. vs. 126 msec.); a significant interaction Anaphora by Antecedent ($F_1(1, 62) = 18.89$; $p < .000$; $F_2(1, 94) = 50.95$; $p < .000$); and a significant three-way interaction between Anaphora, Antecedent and Language ($F_1(1, 62) = 5.41$; $p < .05$; $F_2(1, 94) = 14.78$; $p < .001$). In the null subject Anaphora condition, we find a highly significant effect of Antecedent ($F_1(1, 62) = 31.46$; $p < .000$; $F_2(1, 94) = 66.80$; $p < .000$), indicating that sentences containing null subjects are wrapped up significantly faster when the anaphora retrieves a prominent (subject) antecedent (-99 msec. vs. 151 msec.); and an interaction Antecedent by Language only significant in the analysis by items ($F_1(1, 62) = 1.9$; $p < .20$; $F_2(1, 94) = 4.04$; $p < .05$).

In contrast, overt pronoun anaphors showed no main effects, but a fully significant interaction Antecedent by Language ($F_1(1, 62) = 4.95$; $p < .05$; $F_2(1, 94) = 6.46$; $p < .05$). This interaction indicates that, with overt pronouns, the processing penalties encountered at the wrap up region vary between Italian and Spanish depending on the antecedent. More precisely, in Italian the analysis reveals no significant difference between subject or object antecedent condition, whereas in Spanish, there is a significant effect for Antecedent ($F_1(1, 31) = 6.85$; $p < .05$; $F_2(1, 47) = 4.82$; $p < .05$),

with significantly faster RTs for sentences with a subject antecedent.

If we compare across languages, when an overt pronoun retrieves an object antecedent the wrap up region is read significantly faster in Italian than in Spanish ($F_1(1, 62) = 5.35$; $p < .05$; $F_2(1, 94) = 4.12$; $p < .05$), whereas no significant difference between the two languages was found for an overt pronoun retrieving a subject antecedent.

Comprehension questions. Half of the experimental items were followed by comprehension questions. The charts in Figure 6(a) and 6(b) show the mean reaction times with the relative 95% confidence intervals for the answers and Table 2 the percentages of wrong answers.

(table 2 about here)

(figure 6 about here)

The RTs for the comprehension questions confirm the patterns found in the wrap up region. An ANOVA showed a main effect for Antecedent ($F_1(1, 62) = 14.72$; $p < .000$; $F_2(1, 46) = 12.87$; $p < .000$), showing that questions were answered overall significantly faster when they followed a sentence with a subject antecedent (1777 msec. vs. 2065 msec.). It also revealed a significant interaction Anaphora by Antecedent ($F_1(1, 62) = 12.68$; $p < .000$; $F_2(1, 46) = 14.1$; $p < .000$), and a significant interaction Anaphora by Language ($F_1(1, 62) = 8.22$; $p < .01$; $F_2(1, 46) = 4.94$; $p < .05$).

The interaction Anaphora by Antecedent indicates that questions following a null subject anaphora are answered significantly faster if the antecedent is the previous subject ($F_1(1, 62) = 19.35$; $p < .000$; $F_2(1, 46) = 26.87$; $p < .000$; 1675 msec. vs. 2252

msec.), similarly to what was found at sentence wrap up. In contrast, with overt pronoun sentences, no main effect for Antecedent was found and an interaction Antecedent by Language is only significant in the analysis by subjects ($F_1(1, 62) = 4.11$; $p < .05$; $F_2(1, 46) = 2.74$; $p < .20$). If we look at the other side of the interaction, when the antecedent is a subject, questions are answered significantly faster if the anaphora is null ($F_1(1, 62) = 5.02$; $p < .05$; $F_2(1, 46) = 4.63$; $p < .05$; 1675 msec. vs. 1879 msec.). When the antecedent is an object, questions are answered overall faster if the anaphora is an overt pronoun ($F_1(1, 62) = 11.83$; $p < .001$; $F_2(1, 46) = 9.82$; $p < .01$; 1878 msec. vs. 2251 msec.), with a significant interaction Anaphora by Language ($F_1(1, 62) = 7.86$; $p < .01$; $F_2(1, 46) = 6.20$; $p < .05$).

The Anaphora by Language interaction shows that in Italian questions that follow a sentence with an object antecedent are answered significantly faster if the sentence contains an overt pronoun ($F_1(1, 31) = 12.88$; $p < .005$; $F_2(1, 23) = 11.73$; $p < .005$), whereas such effect does not obtain in Spanish. When a question follows a sentence with a subject antecedent, no effect for Anaphora was found in Italian, whereas in Spanish a main effect for anaphora is significant in the analysis by subjects and marginally significant by items ($F_1(1, 31) = 4.62$; $p < .05$; $F_2(1, 23) = 3.01$; $p < .10$) showing significantly faster reaction times if the question follows a sentence with a null subject.

Table 2 shows the error percentages. They increase in both languages when a null subject retrieves an object antecedent, whereas the pattern differs across languages

in the overt pronoun condition.

The data were analysed using logistic regression, as is suitable for categorical data. An answer was coded as correct when the antecedent that was more plausible with the semantics of the sentence was indicated as the referent of the anaphora. The predictors were: Anaphora, Antecedent and Language. The model shows that overall there are more correct answers when the antecedent is a subject ($\beta = 1.34$; $p < .000$, $e^{\beta} = 3.83$); there are also significantly more correct answers in the overt pronoun condition than in the null subject condition ($\beta = 0.82$; $p < .000$, $e^{\beta} = 2.28$). Furthermore, the results show a significant interaction Antecedent by Anaphora ($\beta = -1.76$; $p < .000$, $e^{\beta} = 0.17$) and a significant three-way interaction Antecedent by Anaphora by Language. The overall model can be considered explanatory, as it is revealed by the likelihood ratio test ($\chi^2(7) = 60.59$; $p < .000$); on the other hand the model residual deviance is larger than expected, ($\chi^2(248) = 464.21$; $p < .000$), which suggests that the model does not fit the data very well. If we fit two separate models to the two Anaphora conditions, we see that after null subject sentences there are significantly more correct answers when the antecedent is a subject ($\beta = 1.34$; $p < .000$, $e^{\beta} = 3.83$); following overt pronoun sentences there are slightly less correct answers in Spanish than in Italian ($\beta = -0.47$; $p = .047$, $e^{\beta} = 0.62$) and a significant interaction Antecedent by Language ($\beta = 1.026$; $p = .002$, $e^{\beta} = 2.79$). This interaction shows that in Italian there are slightly less correct answers when a question follows an overt pronoun sentence referring to a subject antecedent, and this difference is marginally significant ($\beta = -0.42$; $p < .08$, $e^{\beta} = .65$), whereas in Spanish there are

significantly more correct answers after an overt pronoun referring to a subject antecedent ($\beta = 0.60$; $p < .05$, $e^{\beta} = 1.83$).

Discussion

Experiment 2 confirms the pattern of processing biases found in Experiment 1. More precisely, that in both languages the processing of null subjects incurs similar processing penalties when the pronouns are forced to co-refer with a syntactically non-prominent antecedent. In the overt pronoun condition, the interaction between Language and Antecedent was now fully significant at the wrap up region, confirming that, in Spanish, overt pronouns that are associated to a syntactically prominent antecedent do not incur a processing penalty comparable to Italian, and when they are associated to a shift in reference from the previous preverbal subject, they are processed with more ease in Italian than in Spanish. This pattern of cross-linguistic differences, arising only late in the sentence, is confirmed by the answers to the comprehension questions, with Spanish participants giving significantly more accurate answers when the antecedent of an overt pronoun is a subject, and Italian participants answering significantly faster and marginally more accurately when the antecedent is an object. This suggests that such processing penalties may be related to a late discourse integration stage and possibly to strategic processing. Some effects related to the resolution of the subject anaphors, though, are detected earlier in the sentence. At the verb phrase, there is a main effect of Anaphora, with a processing advantage for verb phrases preceded by an overt pronoun. This advantage may be due to the fact that the presence of a nominative pronoun makes the appearance of a verb in the next

display highly predictable.

A main effect for Antecedent is marginally significant at the verb phrase and reaches full significance in the following region and at the final phrase, indicating a processing advantage for sentences that maintain subject reference across clauses, a result that is in keeping with the idea that the default expectation on the part of hearers/readers is for a topic to be maintained across sentences (see Gundel et al. 1993). The reason why this effect becomes significant later than the Anaphora effect may be that, in order for the antecedent to be identified, the semantic information contained into the verb phrase needs to be processed and checked against the comprehender's knowledge of the world, before it can be used to identify the most plausible antecedent and this process necessarily takes some time.

General Discussion

Overall the results from both experiments provide support for the hypothesis that overt pronouns in intra-sentential anaphora contexts obey different processing constraints in Italian and Spanish. In these contexts, null subjects are resolved more easily in both languages when the anaphor refers to a subject antecedent. In contrast, while Italian overt pronouns are associated with a change in subject reference, Spanish overt pronouns do not seem to be associated with such constraint.

Remember that we are assuming that the prominence of an antecedent is determined by its informational status, and that in Spanish and Italian this information is syntactically encoded. We also assume, following Carminati (2002) that the syntactic structure of the subordinate clause, containing the antecedent, is still available in the

memory when the anaphor is encountered in the main clause.

Therefore our data suggest that Spanish and Italian overt pronouns are not sensitive to the same extent to syntactically encoded determinants of prominence and in particular that Spanish overt pronouns are relatively insensitive to syntactic prominence compared to Italian pronouns but also to null subjects in both languages. This opens up the question of which determinants of prominence may Spanish overt subjects be sensitive to. Preliminary results from Filiaci (2011) and Chamorro (2012) suggest that it may be the case that Spanish pronouns (but also Italian pronouns to a certain extent) are in fact more sensitive to the linear distance of the antecedent rather than its syntactic position, as observed in the context of extra-sentential anaphora. The assumption is that, in these contexts, the clause in which the antecedents appear can be fully parsed before the clause containing the anaphors is encountered and that the memory for the exact syntactic structure of this clause decays abruptly at sentence boundary (Garnham et al., 1998). The data shows that, when the exact syntactic information about the position of the antecedents is no longer available, a penalty arises when the Spanish overt pronoun is forced to refer to the first mentioned antecedent (the preverbal subject), compared to when it refers to the last mentioned antecedent (the postverbal object). This possible difference between intra and extra-sentential anaphora may also explain the difference between our results and those of Gelormini-Lezama and Almor (2011), although, as we pointed out in the introduction, it may not be appropriate to compare processing biases of subject pronouns across varieties of Spanish when there are indications that such varieties have different rates

of overt subject expression, which may indicate the presence of different contextual restrictions on the use of overt pronouns.

In relation to the question of the cross-linguistic validity of the processing strategies involved in pronoun resolution, our results suggest that even equivalent anaphoric expressions in closely typologically related languages might be sensitive to different processing cues. In the introduction we mentioned a study by Kwon and Sturt (2012), who found a cross-linguistic difference between the processing of null anaphors in Korean and the functionally equivalent English overt pronouns, in that the former appear to rely on discourse information cues to a larger extent than the latter. The authors attempt to relate this cross-linguistic difference to a major typological difference between Korean and English, that is the absence of morphology in the former language, which, they argue, may lead to a stronger reliance on discourse cues. According to the authors, if this suggestion is correct, in languages like Italian and Spanish, which are typologically related and similar to English, we should expect to find both null and overt anaphors to be sensitive to a similar extent to the same type of contextual cues and should behave similarly to English anaphors in the sense of giving priority to syntax over discourse. The authors also acknowledge that another possibility that remains open is that null and overt anaphors may be processed in intrinsically different ways.

We think that our data does not support the latter hypothesis, since at least in Italian, both null and overt subjects show to be sensitive to a similar extent to syntactic determinants of prominence. On the other hand, our results suggest that formally

similar expressions may not give the same weight to the same contextual cues, here more precisely to syntactic cues, even across typologically related languages. In this case, though, it does not seem possible to relate the cross-linguistic difference found to any obvious typological difference.

We have also mentioned in the introduction that with the *Form Specific Multiple Constraint Approach*, Kaiser and Truswell (2008) suggested that informationally equivalent anaphoric expressions within the same language may be sensitive to different factors determining the prominence of a potential antecedent, for example syntactic position or informational status as opposed to linear order. This idea leads Kaiser and Truswell to reject a monotonic prominence hierarchy, as the one suggested by Accessibility Theory. Following this idea, we could then allow that formally similar and informationally equivalent anaphoric expressions, both across languages and within the same language, may be sensitive to different linguistic determinants of prominence, or at least assign a different weight to the same type of cues, without this difference being associated to typological characteristics of the language.

Finally, in the context of the differences between Spanish and Italian pronominal system outlined in the introduction, it should be noted that the behaviour of Italian *lui* and *lei* is typical of strong pronouns (Cardinaletti and Starke, 1999; Bresnan and Mchombo, 1987), while the behaviour of Spanish overt pronouns is cross-linguistically associated with weak pronouns. In other words, Spanish and Italian subject pronouns, in spite of being formally similar, may actually be two

different types of lexical items. This difference between the two languages may provide new information to be taken into account in the interpretation of data on cross-linguistic influence in situations of language contact, in the sense that the absence of certain contextual restrictions on the use of overt pronouns in Spanish may explain the lack of conclusive evidence for first language attrition as a result of contact with English (Silva-Corvalán, 1994; Montrul, 2004; Flores-Ferrán, 2004): if the use of overt pronouns is not restricted by a bias based on syntactic prominence, then contact with English cannot result in a loss of pragmatic restrictions (see Sorace & Filiaci 2006). Similarly our findings suggest that in the acquisition data of Bini (1999) and in the bilingual data of Sorace et al. (2009) the over-extension of the scope of Italian overt pronouns on the part of Spanish speakers may not just be the result of the use of a default processing strategy but also an outcome of cross-linguistic influence.

In conclusion, this study has shown that despite the close typological and morpho–syntactic similarities between Italian and Spanish, anaphoric expressions that look superficially equivalent in the two languages have different anaphoric preferences. Whether this cross–linguistic difference could be related to other morpho-syntactic features of the two languages, such as the inherent ambiguity of the verbal morphology (that is, the number of entries in the paradigm that are overtly and uniquely marked for person features), or whether they extend to other anaphoric expressions or other null-subject languages are questions open for future investigation.

References

Alexiadou, A., & Anagnostopoulou, E. (1998). Parametrizing Agr: Word order, V-movement and EPP-checking. *Natural Language and Linguistic Theory*, *16*, 491-539.

Almor, A. (1999). Noun-Phrase Anaphora and Focus: The Informational Load Hypothesis. *Psychological Review*, *106*(4), 748-765.

Almor, A. (2000). Constraints and Mechanisms in Theories of Anaphor Processing. In M. Crocker, M. Pickering, & C. Clifton (Eds.), *Architectures and Mechanisms for Language Processing.* England: Cambridge University Press.

Alonso-Ovalle, L., Fernández-Solera, S., Frazier, L., & Clifton, C. (2002). Null vs. Overt Pronouns and the Topic-Focus Articulation in Spanish. *Rivista di Linguistica*, *14*(2).

Ariel, M. (1990). *Accessing Noun Phrase Antecedents.* London: Routlege.

Ariel, M. (1991). The function of accessibility in a theory of grammar. *Journal of Pragmatics*, *16*, 443-463.

Ariel, M. (2006). Accessibility Theory. In K. Brown (Ed.), *Encyclopaedia of language and linguistics* (second ed., pp. 15-18). Oxford: Elsevier.

Arnold, J. E. (1998). *Reference form and discourse patterns.* PhD Stanford University.

Arnold, J. E., Eisenband, J. G., Brown-Schmidt, S., & Trueswell, J. C. (2000). The rapid use of gender information: evidence of the time course of pronoun resolution from eyetracking. *Cognition*, *76*, B13-B26.

Baayen, H. (2008). *Analyzing Linguistic Data. A Practical Introduction to Statistics Using R.* Cambridge University Press.

Bini, M. (1993). La adquisición del italiano: más allá de las propiedades sintácticas del parámetro pro-drop. In J. M. Liceras (Ed.), *La lingüística y el análisis de los sistemas no nativos* (pp. 126-139). Ottawa: Dovehouse.

Bresnan, J., & Mchombo, S. A. (1987). Topic, pronoun, and agreement in Chichewa. *Language*, *63*(4), 741-782.

Cacciari, C., Carreiras, M., & Cionini Barbolini, C. (1997). When Words Have Two

Genders: Anaphor Resolution for Italian Functionally Ambiguous Words. *Journal of Memory and Language*, *37*(4), 517 - 532. doi:DOI: 10.1006/jmla.1997.2528

Cameron, R. (1992). *Pronominal and null subject variation in Spanish: Constraints, dialects, and functional compensation.* PhD University of Pennsylvania.

Caramazza, A., Grober, E., Garvey, C., & Yates, J. (1977). Comprehension of Anaphoric Pronouns. *Journal of Verbal Learning and Verbal Behavior*, *16*(5), 601-609.

Cardinaletti, A. (1997). Subjects and clause structure. In L. Haegeman (Ed.), *The New comparative Syntax* (pp. 33-63). New York: Longman.

Cardinaletti, A., & Starke, M. (1999). The typology of structural deficiency: A case study of the three classes of pronouns. In H. van Riemsdijk (Ed.), *Clitics in the languages of Europe* (pp. 145-233). The Hague: Mouton de Gruyter.

Carminati, M. N. (2002). *The Processing of Italian Subject Pronouns.* Doctor of Philosophy University of Massachusetts Amherst.

Casielles-Suárez, E. (2001). The syntax and semantics of preverbal topical phrases in Spanish. In J. Gutiérrez-Rexach & L. Silva-Villar (Eds.), *Current issues in Spanish syntax and semantics* (pp. 65-82). Berlin/New York: Mouton De Gruyter.

Casielles-Suárez, E. (2004). *The Syntax-Information Structure Interface: Evidence from Spanish and English.* New York/London: Routledge.

Chamorro, G. (2012). The effect of L1 exposure on Spanish attrition: an eye-tracking study. *Poster presented at the AMLaP 2012 Conference*.

Chomsky, N. (1995). *The Minimalist Program.* Cambridge, MA: MIT Press.

Cinque, G. (1990). *Types of A'-Dependencies.* Cambridge: Massachusetts: MIT Press.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). Psyscope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers*, *25*(2), 257-271.

Crawlay, R. A., Stevenson, R. J., & Kleinman, D. (1990). The use of heuristic

strategies in the interpretation of pronouns. *Journal of Psycholinguistic Research*, *19*, 245-264.

Enríquez, E.-V. (1984). *El pronombre personal sujeto en la lengua española hablada en Madrid* (4). Madrid: Consejo Superior de Investigaciones Científicas, Instituto Miguel de Cervantes.

Filiaci, F. (2011). *Anaphoric Preferences of Null and Overt Subjects in Italian and Spanish: a Cross-linguistic Comparison.* The University of Edinburgh.

Flores-Ferrán, N. (2004). Spanish subject personal pronoun use in New York City Puerto Ricans: Can we rest the case of English contact? *Language Variation and Change*, *16*, 49-73.

Garnham, A., Oakhill, J. V., & Cain, K. (1998). Selective retention of information about the superficial form of text: Ellipses with antecedents in main and subordinate clauses. *The Quarterly Journal of Experimental Psychology*, *51 A*, 19-39.

Gelormini-Lezama, C., & Almor, A. (2011). Repeated names, overt pronouns, and null pronouns in Spanish. *Language and Cognitive Processes*, *26*(3), 437-454.

Gernsbacher, M. A. (1989). Mechanisms that improve referential access. *Cognition*, *32*(2), 99-156.

Givón, T. (1979). From discourse to syntax: grammar as a processing strategy. In T. Givón (Ed.), *Syntax and Semantics: vol.12: Discourse and syntax* (pp. 81-112). New York: Academic Press.

Greene, S. B., McKoon, G., & Ratcliff, R. (1992). Pronoun resolution and discourse models. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *18*(2), 266-283. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=1532820

Grosz, B. J., Joshi, A. K., & Weinstein, S. (1983). Providing a unified account of definite noun phrases in discourse. *Proceedings of the 21st Annual meeting of the association of computational linguistics,* 44-50.

Grosz, B. J., Joshi, A. K., & Weinstein, S. (1995). Centering: A Framework for Modelling the Local Coherence of Discourse. *Computational Linguistics*, *21*(2), 203-225.

Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive Status and the Form of Referring Expressions in Discourse. *Language*, *69*(2), 274–307. Retrieved from http://www.jstor.org/stable/416535

Huang, Y. (2000). *Anaphora: a cross-linguistic study.* Oxford: Oxford University Press.

Huang, C.-T. J. (1984). On the distribution and reference of empty pronouns. *Linguistic Inquiry*, *15*(4), 531-574.

Huang, Y. (1994). *The syntax and pragmatics of anaphora.* Cambridge: CUP.

Huang, Y. (2000). *Anaphora: a cross-linguistic study.* Oxford: Oxford University Press.Hulk, A., & Müller, N. (2000). Bilingual first language acquisition at the interface between syntax and pragmatics. *Bilingualism: Language and Cognition*, *3*(3), 227-244.

Jaeger, T. F. (2008). Categorical Data Analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models. *Journal of Memory and Language*, *59*(4), 434-446. doi:10.1016/j.jml.2007.11.007

Jaeggli, O., & Safir, K. J. (Eds.). (1989). *The Null Subject Parameter.* Dordrecht: Kluwer Academic Publishers.

Kaiser, E., & Trueswell, J. (2008). Interpreting pronouns and demonstratives in Finnish: Evidence for a form-specific approach to reference resolution. *Language and Cognitive Processes*, *23*(5), 709-748. doi:10.1080/01690960701771220

Kehler, A., Kertz, L., Rohde, H., & Elman, J. L. (2008). Coherence and Coreference Revisited. *Journal of Semantics*, *25*, 1-44.

Kwon, N., & Sturt, P. (2012). Null pronominal (*pro*) resolution in Korean, a discourse-oriented language. *Language and Cognitive Processes*, *27*.

Lambrecht, K. (1994). *Information structure and sentence form* (71). Cambridge, UK:

Cambridge University Press.

Liversedge , S. P., & Van Gompel, R. P. G., ms. Resolving anaphoric and cataphoric pronouns.

Lorusso, P., Caprin, C., & Guasti, M. T. (2005). Overt Subject Distribution in Early Italian Children. In Brugos, A., Clark-Cotton, M. R., Ha, Seugwan (eds), *BUCLD 29 Online Proceedings Supplement*.

Lozano, C., & Mendikoetxea, A. (2010). Interface conditions on postverbal subjects: A corpus study of L2 English. *Bilingualism: Language and Cognition*, *13*(4), 475-497.

McDonald, J. L., & MacWhinney, B. (1995). The time course of anaphor resolution: effects of implicit verb causality and gender. *Journal of Memory and Language*, *34*(4), 543-566.

Montrul, S. (2004). Subject and object expression in Spanish heritage speakers: A case of morphosyntactic convergence. *Bilingualism: Language and Cognition*, *7*(2), 125-142.

Morales, A. (1997). La hipótesis funcional y la aparición de sujeto no nominal: el español de Puerto Rico. *Hispania*, *80*(1), 153-165. Retrieved from

Ordóñez, F., & Treviño, E. (1999). Left dislocated subjects and the pro-drop parameter: A case study of Spanish. *Lingua*, *107*, 38-68.

Otheguy, R., Zentella, A. C., & Livert, D. (2007). Language and dialect contact in Spanish in New York: toward the formation of a speech community. *Language*, *83*(4), 770-802.

Rizzi, L. (1997). The fine structure of the left periphery. In L. Haegeman (Ed.), *Elements of Grammar.* Dordrecht: Kluwer Academic Publishers.

Serianni, L., & Castelvecchi, A. (1991). *Grammatica Italiana, Italiano comune e lingua letteraria.* Torino: UTET.

Silva-Corvalán, C. (1994). *Spanish in Los Angeles.* New York: Oxford University Press.

Smyth, R. (1994). Grammatical determinants of ambiguous pronoun resolution.

Journal of Psycholinguistic Research, *23*(3), 197-229.

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

Sorace, A., & Filiaci, F. (2006). Anaphora resolution in near-native speakers of Italian. *Second Language Research*, *22*(3), 339-368.

Sorace, A., & Serratrice, L. (2009). Internal and external interfaces in bilingual language development: Beyond structural overlap. *International Journal of Bilingualism*, *13*(2), 195-210.

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*(3), 460-477.

Sperber, D., & Wilson, D. (1995). *Relevance: communication and cognition.* Oxford, UK: Blackwell.

Stevenson, R., Knott, A., Oberlander, J., & McDonald, S. (2000). Interpreting pronouns and connectives: Interaction among focusing, thematic roles and coherence relations. *Language and Cognitive Processes*, *15*(3), 225-262.

Stevenson, R. J., Crawlay, R. A., & Kleinman, D. (1994). Thematic roles, focus and the representation of events. *Language and Cognitive Processes*, *9*, 519-548.

Tsimpli, I. M., Sorace, A., Heycock, C., & Filiaci, F. (2004). First language attrition and syntactic subject: The role of discourse pragmatics in the acquisition of subjects. *Applied Psycholinguistics*, *8*(3), 257-277.

Vallduví, E. (1990). *The informational component.* PhD University of Pennsylvania.

Wilson, F. (2009). *Processing at the syntax-discourse interface in second language acquisition.* PhD University of Edinburgh.

Author Note

This work is part of the doctoral dissertation of the first author, funded by the AHRC grant 06/123038 and by a Small Project Grant of the Development and Alumni fund of the University of Edinburgh. Manuel Carreiras is supported by grants PSI2009-08889/PSIC and CONSOLIDER-INGENIO2010 CSD2008-00048 from the Spanish Government.

Table 1 Percentages of Wrong Answers to the Comprehension Questions (Experiment 1).

		Wrong Answers (%)	
		Italian	Spanish
NULL ANAPHORA	Object Antecedent	27	33
	Subject Antecedent	5	16
OVERT ANAPHORA	Object Antecedent	11	27
	Subject Antecedent	27	13

Table 2 Percentages of Wrong Answers to the Comprehension Questions (Experiment 2).

		Wrong Answers (%)	
		Italian	Spanish
NULL ANAPHORA	Object Antecedent	37	40
	Subject Antecedent	13	20
OVERT ANAPHORA	Object Antecedent	21	30
	Subject Antecedent	29	19

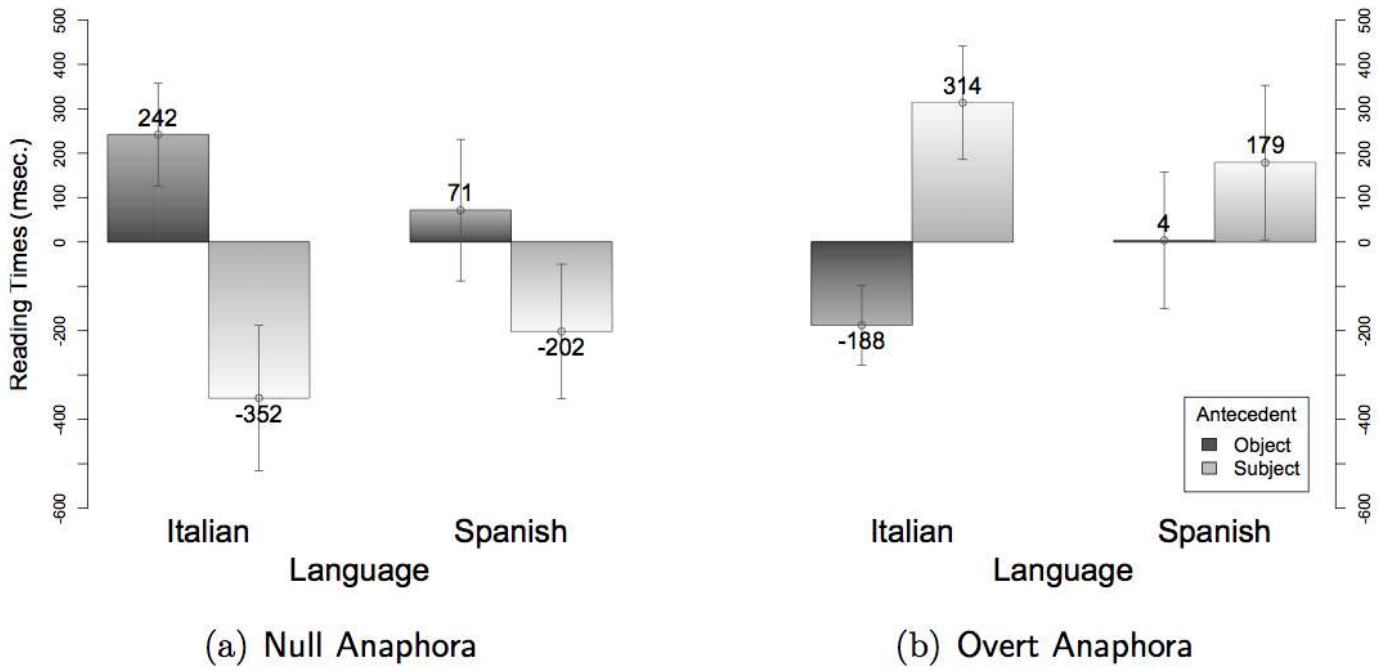


Figure 1 Experiment 1: reading times for the second clause.

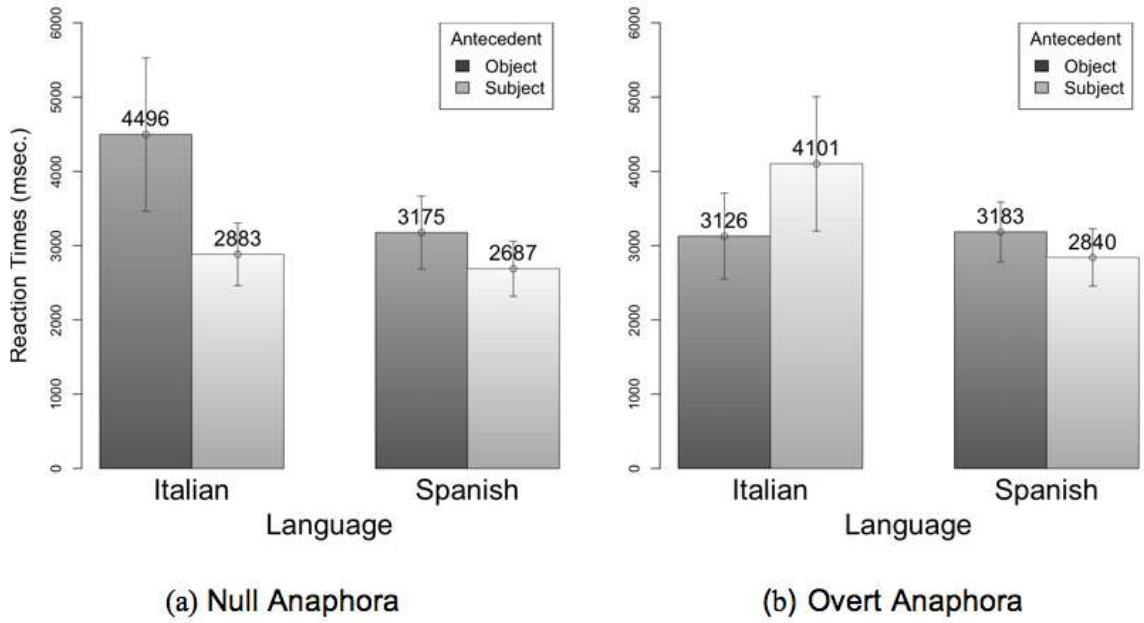


Figure 2 Experiment 1: RTs for the comprehension questions.

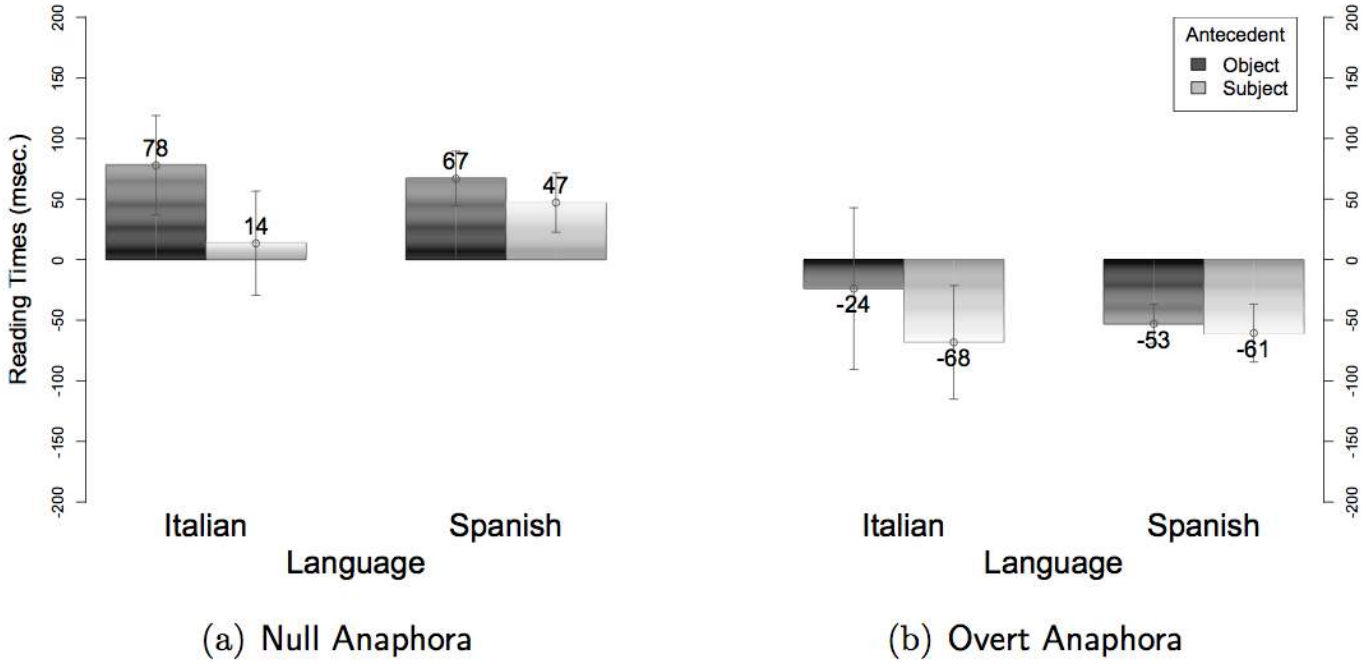


Figure 3 Experiment 2: reading times for the VP region.

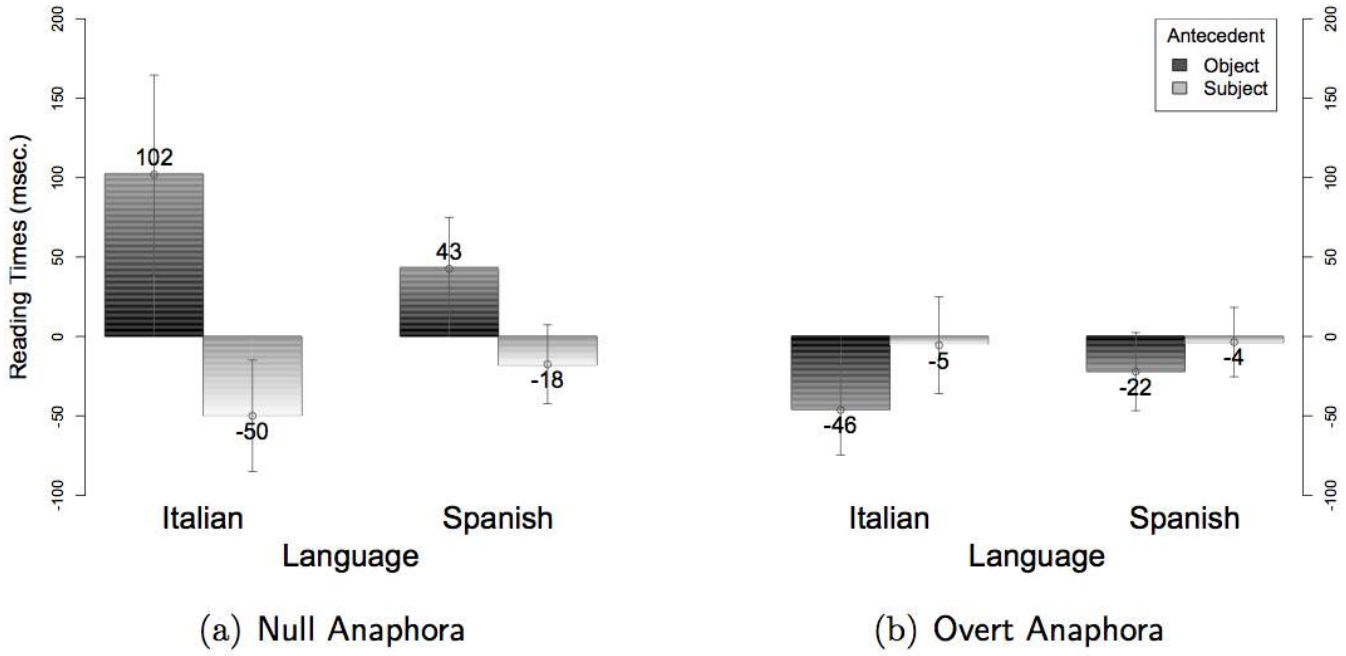
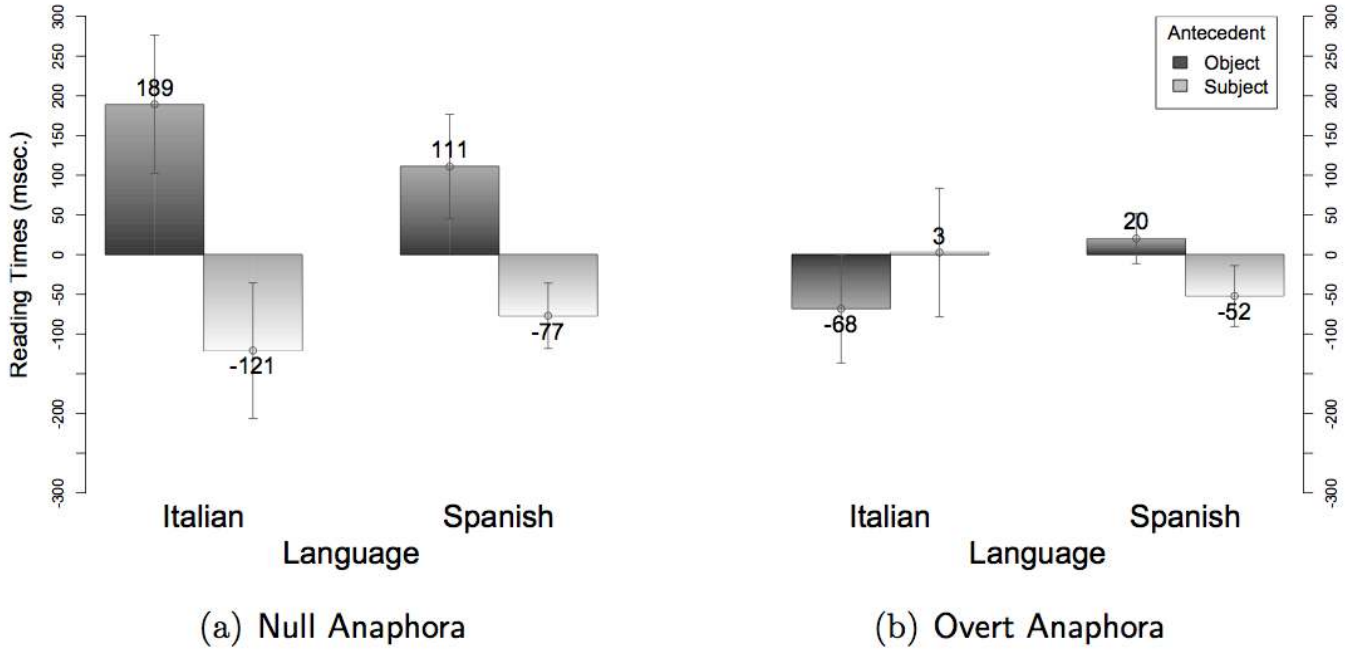


Figure 4 Experiment2: reading times for the region following the VP.



(a) Null Anaphora

(b) Overt Anaphora

Figure 5 Experiment 2: reading times for the final wrap up region.

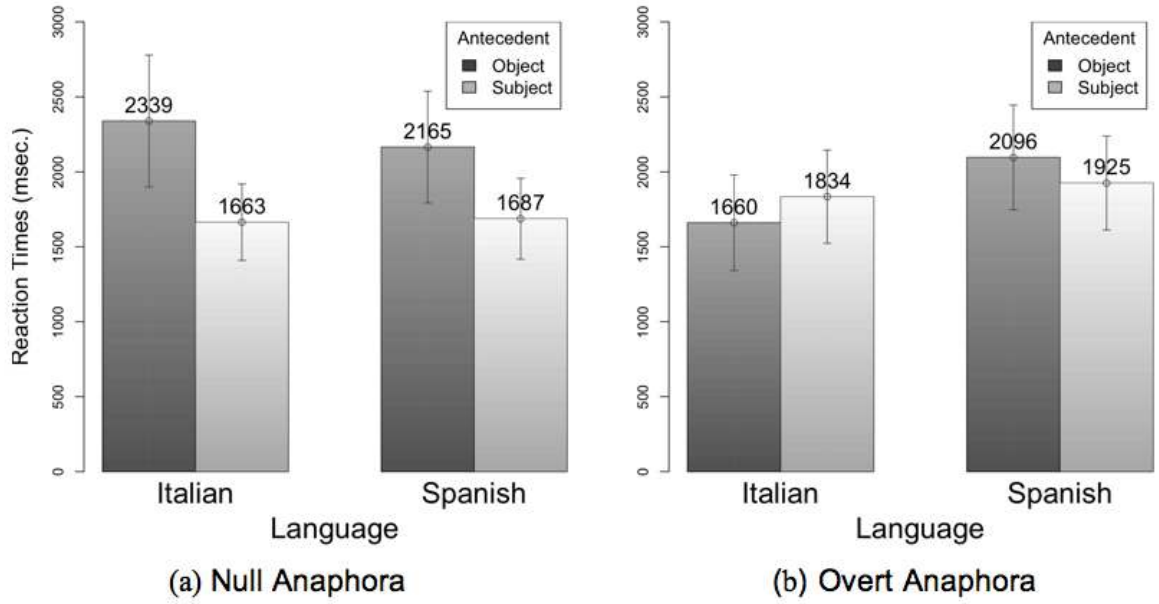


Figure 6 Experiment 2: RTs for the comprehension questions