Encoding Information Structure via Object Agreement in Spanish Interactions

VALERIA A. BELLORO
Columbia University

0. Introduction
The goal of this paper is to examine the information structure of traditionally named “direct object doubling” (DO-doubling) constructions in Spanish, based on interactional data from the Buenos Aires dialect. Even though Buenos Aires Spanish (often also referred to as Porteño, Río de la Plata, and ‘River Plate’ Spanish) is generally considered as the most permissive, these structures have also been attested in the dialects of Corrientes, Argentina (Colantoni 2002), Santiago, Chile (Silva-Covalán 1981), Lima, Perú (Sánchez 2006), Quito, Ecuador (Suñer 1989), and some areas of México (Parodi and Santa Ana, 2002; Alarcón and Orozco 2004) and Spain (Suñer 1989, Franco 2000).

In so-called DO-doubling, an accusative pronominal clitic and a coreferential lexical phrase in canonical posverbal position co-occur inside the clause. Following the proposals put forth by Suñer (1988) and Franco (1993, 2000), I consider the pronominal clitic in “doubling” constructions to be functioning as an object agreement marker. The sentence in (1) serves as an illustration:

(1) El año pasado yo la invitó a Mabel.
‘Last year I invited Mabel.’ (hc:xiv)

1 For ease of presentation, I will often use here the traditional label and refer to the cases where the clitic and the lexical phrase co-occur as “doubling” structures, and to the lexical phrase triggering object agreement as the “doubled” phrase. The use of these labels does not entail any particular commitment with respect to the syntactic role of these forms. See Belloro (2007) for an account of the grammatical aspects of these constructions within the framework of Role and Reference Grammar (Van Valin 2005).

2 The relevant coreferential structure is marked in boldface. The letters ‘hc’ followed by a roman numeral mean that the example is taken from the corpus El Habla Culta de la Ciudad de Buenos Aires (Barrenechea 1987). The number represents the particular interaction from which the example is taken.

391
Valeria A. Belloro

The occurrence of an accusative clitic in connection with a lexical phrase in DO function not only is optional, but also relatively constrained. Within formal approaches it has been claimed, for instance, that the clitic is allowed in the structure only if the coreferential phrase is marked by the pseudo-preposition a (Jaeggli 1981, 1986), or if the referent it denotes is specific (Suñer 1988, 1999), or presupposed (Franco 2000). Corpora-based studies, on the other hand, have arrived at seemingly contradictory conclusions, correlating the occurrence of accusative doubling with reference to topical (Silva-Corvalán 1981) and new (Colantoni 2002) discourse participants.

With respect to Buenos Aires Spanish, even conceived of as the paradigmatic doubling variety, there is a lack of corpus-based analyses which could confirm or belie the different hypotheses proposed so far. One of the central aims of this research is to address this gap, and to propose an alternative analysis of the interpretive import of these constructions based on naturally occurring data.3

With these goals in mind, the organization of this paper is as follows: In section 2, I review the main claims that have been made with respect to the semantics/pragmatics of DO-doubling constructions. Each proposal is evaluated with respect to a corpus of oral interactions among native speakers of Buenos Aires Spanish (Barrenechea 1987). This corpus consists of 33 samples, amounting to about 24 hours of recording (250,000 words). Within it, all sentences containing a post-verbal nominal direct object co-occurring with a coreferential accusative clitic were identified. Although the number of relevant instances is relatively small (119 tokens), this analysis constitutes, to my knowledge, the first attempt to evaluate the existing hypotheses and quantify their relative accuracy based on naturally occurring examples from a homogeneous corpus. The data show that none of the existing proposals can provide a comprehensive account of the context in which doubling constructions emerge. However, in section 3 I suggest that they are in fact consistent with an explanation of accusative agreement as marking the cognitive accessibility of the referents involved, which is proposed as a more accurate alternative for understanding not only under which conditions these structures are allowed, but also the discourse-pragmatic meaning they convey. Further, it is suggested that this approach makes it possible to provide a principled way of distinguishing between the information-structure of “doubling” constructions and its closely related “allosentences” (Lambrecht 1994), namely those structures where only the pronominal clitic or the lexical phrase occur. A brief summary of the analysis and its general conclusion are presented in section 4.

3 The analysis I develop here follows a proposal originally presented in Belloro (2004) and elaborated in Belloro (2007). See also Estigarribia (2006) for a similar approach.
Encoding Information Structure via Object Agreement

1. Direct Object Doubling in Buenos Aires Spanish

1.1. A-marking and Animacy

In one of the seminal studies of clitic doubling in Buenos Aires Spanish, Jaeggli (1981, 1986) proposed that accusative doubling is only possible if the lexical phrase is introduced by the preposition-like particle a, as in (1) above (cf. also Lyons 1999, Belletti 2005, Anagnostopoulou 2006). The a-marking of direct objects is generally triggered by the animacy of the referent, although it may occur with inanimates if they are specific and topical (Torrego 1999, Leonetti 2003). The kind of features associated with a-marked direct objects is relevant here inasmuch as it coincides with the ones posited to correlate with “doubled” ones; namely animacy and topicality (cf. Silva-Corvalán 1981, Suñer 1988). Under this perspective, it is claimed that “accusative clitics in Spanish are in strict complementary distribution with direct objects which are not preceded by the marker a” (1986:19). This prediction posits the ungrammaticality of sentences such as those in (2) (from Jaeggli 1986:19):

(2) a. La compré (*la casa).
   ‘I bought it (*the house).’

   b. Lo vendí (*el libro).
   ‘I sold it (*the book).’

The analysis of the corpus suggests, on the contrary, that at least in the Buenos Aires dialect doubling is also possible with non a-marked DOs, as originally argued for by Suñer (1988). These cases account for 35% of the tokens in the corpus. Two examples are presented in (3):

(3) a. Ahora tiene que seguir usandoló el apellido.
   ‘Now she has to keep using the last name.’ (hc:xvi)

   b. …lo han aprendido el predicativo.
   ‘…they have learnt the predicative.’ (hc:xi)

In fact, the data show not only that non a-marked phrases can co-occur with the clitic, but also that the animacy of the referent is relatively irrelevant as well, since almost 40% of the “doubled” phrases (some of them marked by a) refer to inanimate entities.4

---

4 It is important to note that there is no evidence for claiming that these structures involve either a pause or an intonation break before the doubled phrase, and that the doubled phrase does not encode topical participants, facts that would support an analysis in terms of dislocation. Moreover, structures like (3) encode the same pragmatic status, with respect to the referent of the “doubled” phrase, as the a-marked counterparts, a fact that would remain mysterious if each were the instantiation of a different construction (cf. Belloro 2008).
1.2. Specificity and Presuppositionality

It has been suggested that DO-doubling may only occur if the target referent can be conceived of as specific. The hypothesis was originally presented in Suñer (1988) and subsequently adopted, in general terms, by Sportiche (1995), Bleam (1999) and Gutiérrez Rexach (2000), among others.

Suñer's proposal crucially assumes that specific indefinites allow doubling (4a), but nonspecific definite (4b) do not, thus stressing the relevance of specificity over definiteness, as illustrated in the following examples (from Suñer 1988:396):

(4) a. Diariamente, la escuchaban a una mujer que cantaba tangos.
   ‘Daily, they listened to a woman who sang[indicative] tangos.’

b. (*Lo) alabarán al niño que termine primero.
   ‘They will praise the boy who finishes first.’

The specificity approach is not problem-free. At least 10% of the instances of the Buenos Aires corpus are difficult to accommodate in terms of specificity, regardless of the particular interpretation this notion is given (cf. Farkas 1994). Some examples involving generic reference are presented below:

(5) a. …hay que ver las cosas para aprenderlas un poquito mejor ¿no?
   …one has to see things to understand them a little better, no?’ (hc:i)

b. Hasta en el exterior uno inmediatamente, casi a veces esté… viéndolo de lejos, lo ve al porteño.
   ‘Even abroad you immediately, almost sometimes eh… watching him from far away, you spot the porteño.’ (hc:i)

As an alternative to specificity, it was proposed that the relevant criterion for DO-doubling is the presuppositionality of the target referent (Franco 2000). The clearest formulation of the sense in which the notion of presupposition is used in this connection appears in Franco and Mejías-Bikandi (1999), based on examples such as the following:

(6) a. Lo he visto a un marinero.
   ‘I have seen one of the sailors.’

b. He visto a un marinero.
   ‘I have seen a sailor.’

Most of their data come from Basque Spanish in which, as in other “leísta” varieties, animate accusative clitics appear as le(s) instead of lo/a(s). In these examples the canonical accusative clitics are substituted for the leísta variants occurring in Basque Spanish. Nothing in Franco and Mejías-Bikandi’s analysis hinges on this distinction. The same hypothesis summarized here is defended in Franco (2000) with respect to what is there labeled “Southern Cone Spanish.”
In consistence with the glosses they provide, the authors argue that the noun *un marinero* ‘one sailor’ in (6a) is presuppositional. On the other hand, the object in (6b), may be interpreted either as presuppositional, or as existential non-presuppositional. That is, like (6a) the object in (6b) may refer to a previously introduced set of entities, but it could also be used to introduce a new entity in the discourse.

The presuppositionality of the target referent approach is not problem free either. The analysis of the Buenos Aires corpus shows that over 57% of the tokens in the corpus involve referents which do not have a previously introduced antecedent and have to be considered, in strict terms, “discourse-new” (Prince 1992).

1.3. **Topicality**

In relation to the presence of an antecedent for the “doubled” phrase, it was mentioned above that existing corpora-based studies arrive at contradictory conclusions. In a study of the Spanish of Santiago de Chile, Silva-Corvalán (1981) suggests that doubling is triggered by the high topicality of the certain referents in DO function (in particular, humans) and correlates with continuing topics. On the other hand, based on an analysis of the Spanish of Corrientes, Colantoni (2002) sees doubling as the result of the addition to the structure of a lexical phrase in those cases where exclusive pronominal reference would have been insufficient, as it is the case, for instance, with entities newly introduced in the discourse context.

With respect to the Buenos Aires data, each hypothesis makes it possible to account for the tokens with which the other fails, as doubling occurs in relationship with both “discourse-old” and “discourse-new” referents. The question is, then, whether there is a way of reinterpreting the trends found in Silva-Corvalán and Colantoni’s corpora so that they may be not only compatible among themselves but also with the data collected from the Buenos Aires dialect. I would like to propose that such as reinterpretation is indeed possible by acknowledging the relevance to these constructions of the speaker’s assumptions about the hearer’s knowledge and attentional state.

2. **Marking Referents’ Accessibility**

The Buenos Aires data show that the best grammatical correlate of doubling is in fact definiteness (95%), suggesting that these constructions target referents that the speaker assumes the hearer can identify (Lambrecht 1994, Lyons 1999). Since the identifiability of a referent is not exclusively dependent on its prior mention in the discourse context, from the perspective of the encoding of identifiability it is possible to explain why DO-doubling targets discourse-new participants almost as much as discourse-old ones. In effect, the referent of a definite NP may be considered identifiable by virtue of its anaphoric relation with a discourse-antecedent, but also due, for instance, to its associative connection with a different participant evoked in the exchange. In fact, inferential links of this type have been shown to pattern with anaphoric relations in a number of “marked” syntactic constructions.
Valeria A. Belloro

The occurrence of DO-doubling in connection with new participants is therefore not surprising, as long as the participants belong to the set reasonably assumed to be inferrable from some other mentioned entities in the given discourse context. This is exactly what the Buenos Aires data shows. Two typical examples are presented in (7):

(7) a. Y... y cuando se toma el taxi lo mira al taximetrista...
   ‘And when she takes the taxi she looks at the taxi driver...’ (hc:xxxii)

   b. …yo la invito a hacer este experimento: tomar todo lo que es sacando
   Piazzolla, tomar... tomar todo lo que es eh... yo no diría--- nueva ola, diría--
   -- vanguardia en tango;...en todos los movimientos musicales siempre nos
   interesa la vanguardia; es lo que va quedando, por supuesto. Tomémoslo a
   Troilo, tomémoslo a Berlingieri, tomémoslo a Salgán…
   ‘…I invite you to do this experiment: take everything except Piazzolla
   take all that is eh…I wouldn’t say--- new wave, I would say--- avant garde
   in tango; in all musical movements there is always an interest in the avant
   garde; is what remains, of course. Let’s take Troilo, let’s take Berlingeri,
   let’s take Salgán…’ (hc:ii)

In the example in (7a), the mention of the taxi determines the “inferrability” (Prince 1981) of the taxi driver, given our natural expectations about the described scene. In (7b) the participants targeted by the doubling construction are more culturally specific, they are salient members of the avant garde tango music, and they are all introduced in the same turn. Although the different tokens found in the corpus exhibit certain particularities with respect to the inferential processes involved in each case, there is also a common factor relating all the different instances where doubling targets new discourse participants: in all cases they are assumed to be relatively easy to conjure up in the mind of the addressee given the particular discourse-context in which they are mentioned. This suggests that the occurrence of DO-doubling is not dependent on the existence of a coreferential link between the doubled object and a discourse antecedent, but on the existence of some “cluster of interrelated expectations” or “schema” (Chafe 1987:29) which allows the speaker to treat the relevant referents as identifiable.

Let us consider now the case of doubling of discourse-old participants. Two typical examples are presented in (8):

(8) a. Mire el otro día sube al colectivo un un porteño, bien bien calibradamen-
   te porteño. Y el que manejava el colectivo--- era otro bien calibradamen-
   te porteño, éso que manejan de costado--- en ángulo de cuarenta y cinco
   grados con respecto al volante. Éste no sé si le pagó con cien pesos o
   con quinientos y el otro le dio un vuelto... esté... marcadamente en mo-
   nedas--- cualquier cantidad. Entonces éste cuando recibió ese impacto de
   todas las monedas que no se lo esperaba, quedó ahí con la mano todavía
   en forma de balanza como pesándola y mirando ese paquete brutal de
Encoding Information Structure via Object Agreement

monedas. Y lo miró fijo y provocativamente al colectivo, y todo lo que le dijo es: ‘Mucho, ¿no?’
‘Look, the other day it gets in the bus a porteño, a very typical porteño. And the guy who drove the bus---was another very typical porteño, those that drive on the side---in a forty five degree angle with respect to the wheel. And this guy I don’t know if he paid him with one hundred or five hundred pesos and the other guy gave him the change... eh... notoriously in coins---a huge amount of them. Then this guy, when he received the impact of all those coins, that he wasn’t expecting, stayed there with his hand still in the shape of a scale like weighing them and staring at this brutal pack of coins. And he looked at the bus driver straight and provocatively, and all he said was: ‘A lot, no?’ (hc:ii)

b. …tenemos un problema porque ese abrigo suyo vino una clienta y dijo que le quedaba muy bien y se lo quería para ella. Es una clienta hace mucho nuestra, así que se lo vamos a tener que dar." "Ah, no", dice. "Si quiere le hacemos otro, y después se lo mandamos a Bahía Blanca.” Y Betty le dijo: "No–dice- yo lo elegí primero. Si ustedes no me dan ese abrigo, yo no compró nada…
‘…we have a problem because, that coat of yours, a client came and said that it fit her very nicely and that she wanted it for herself. She has been our costumer for a long time, so we’ll have to give it to her’. ‘Ah, no’, she says. ‘If you want we’ll make you another one, and then we’ll send it to Bahía Blanca’ And Betty said: ‘No–she says- I chose it first. If you don’t give me that coat I won’t buy anything…’ (hc:xxvii)

In (8a) the speaker is recounting an exchange he witnessed between a bus driver and the person who was buying the ticket from him. The bus driver is introduced in the second clause as “the guy who drove the bus”, together with brief description of his physical appearance connoting his general character. Then the attention shifts to the traveler, the paper bill he used to pay for the ticket, and the coins he received in exchange. By the time the speaker wants to focus again on the bus driver, this participant is no longer the focus of attention. It is in this context that this referent appears encoded in a doubling construction.

Consider now the example in (8b). Here the speaker is telling something that happened to a friend of hers with a coat this person had bought. The coat, which the protagonist goes to pick up at the store, is introduced in the second clause as “that coat of yours”. Then, several other participants intervene: another client, the clerk, Bahía Blanca and, crucially, another coat. When the relevant coat, the one the protagonist wanted, has to be referred to again, a doubling construction is chosen. These sentences exemplify a general trend in the data. If the doubled DO has a discourse antecedent, then between this antecedent and the doubled phrase there has been an attentional shift, so that by the time the doubling construction is used the referent it denotes is no longer the center of attention. It can be assumed to be, however, peripherally active in the mind of the addressee, inasmuch as even
if its level of activation decayed as the discourse proceeded, it has not rendered it fully inactive. Crucially, the same “peripheral” activation status with respect to the hearer’s attention can also be assumed in relation with those cases involving discourse-new participants. In this later case it is not that their activation decayed from a previous active state, but on the contrary that they have acquired their relative accessibility by virtue of their association with an active conceptual “schema” (Chafe 1987, 1994).

Following Chafe, we can distinguish three levels of activation relevant for grammar: a conceptual representation of a referent may be “active” if it is in the interlocutor’s focus of consciousness; “accessible”, if it is only peripherally active; or “inactive” if it is in the interlocutors’ long-term memory, and neither focally nor peripherally active. This typology makes it possible to capture the contexts where both discourse-old and discourse-new DO-doubling constructions occur, since in both cases these constructions not only target referents which are identifiable, but further those which are assumed to be neither “active” nor “inactive” but specifically “accessible” in the mind of the addressee.

The hypothesis that doubling constructions are sensitive to the hearer’s knowledge and attentional state is in accordance with the structural properties of the lexical phrases found to co-occur with the clitic. If doubling were sensitive only to the speaker’s knowledge (as some interpretations of the specificity hypothesis would predict) we should find at least some tokens where descriptively rich, formally complex lexical phrases were used to convey the speaker’s relative familiarity with a referent presumably unknown to the addressee. In other words, where the referent could be assumed to be specific but not identifiable. The example presented in (4a), and repeated below for convenience, is a good candidate for occurring in this kind of context:

(9) Diariamente, la escuchaban una mujer que cantaba tangos.
‘Daily, they listened to a woman who sang tangos.’

Notably, however, the doubled phrases in the Buenos Aires corpus are consistently simpler than the one in (9); over 70% consist of just a proper noun or a definite determiner and a common noun (Belloro 2007:112). Given the correlation between form of encoding and identifiability (Ariel 1990, 2001; Gundel et al. 1993) this pattern, even if not presenting negative evidence against specificity, does suggest that at least in the Buenos Aires dialect doubling is not used in connection with referent which is identifiable for the speaker (i.e. specific) but not for the addressee.

Besides shifting the stress from the speaker to the addressee, another difference between identifiability and specificity or presuppositionality is that whereas the latter are typically understood as binary notions, identifiability is inherently gradual, since the identification of a referent may require more or less processing effort, depending on its relative activation in the mind of the interlocutors. What this perspective affords us, therefore, is the possibility of placing doubling con-
Encoding Information Structure via Object Agreement

Instructions alongside its closest grammatical alternatives, uncovering the particular pragmatic function that binds them; namely, the encoding of the pragmatic properties of the target referents. From this perspective clitic doubling can be interpreted as the formal correlate of an intermediate level of referent accessibility, along a continuum that has clitics and lexical phrases at either end. The proposed relationship between formal encoding and cognitive states for DO arguments can be illustrated as in (11):

(11) Correlation between Formal encoding and cognitive states:

<table>
<thead>
<tr>
<th>Form:</th>
<th>Clitic-only</th>
<th>Clitic-doubling</th>
<th>NP-only</th>
</tr>
</thead>
<tbody>
<tr>
<td>Example:</td>
<td>Lo mira</td>
<td>Lo mira al taximetrista</td>
<td>Mira al taximetrista</td>
</tr>
<tr>
<td>Activation:</td>
<td>Active</td>
<td>Accessible</td>
<td>Inactive</td>
</tr>
</tbody>
</table>

3. Conclusions

This paper examined previous accounts of so-called DO-doubling constructions in Spanish, and in particular the claims that the doubled lexical phrase has to be a-marked (which indirectly predicts that doubling of inanimate referents should be virtually unattested), that it needs to be interpreted as specific, and that it has to retrieve an antecedent from the discourse context. These claims were evaluated with respect to a corpus of interactional speech from a paradigmatic accusative doubling variety (Buenos Aires Spanish). The analysis showed that even though current proposals tend to capture general correlations, none of their claims can completely account for the empirical data.

The instances of accusative doubling in the corpus fall within two general categories: those which refer to a discourse antecedent that is no longer the focus of attention in the interactional segment, and those which involve an entity typically associated with some other entity just mentioned or the general issue being discussed. Resorting to Chafe’s three-way distinction of cognitive accessibility levels, and his claim that accessible referents acquire this state via deactivation from an earlier active state or association with a semantic schema, I proposed that accusative doubling in Spanish is used to mark accessible referents and, in this sense, could be analyzed in light of its non-doubled alternatives (i.e. NP-only and clitic-only structures) as an extra point along the continuum for encoding the pragmatic properties of discourse-referents.

References


399
Valeria A. Belloro


Encoding Information Structure via Object Agreement


Valeria A. Belloro
Columbia University
Department of Spanish and Portuguese
612 W. 116th Street #309
New York, NY 10027

vb2237@columbia.edu