

[Draft version]

Vázquez Rozas, Victoria & Marcos García Salido. 2012. [A discourse-based analysis of object clitic doubling in Spanish](#). In Davidse, Kristin, Tine Breban, Lieselotte Brems y Tanja Mortelmans (eds.). *Grammaticalization and Language Change: New reflections*, 271–297. Amsterdam: John Benjamins.

A discourse-based analysis of object clitic doubling in Spanish

Victoria Vázquez Rozas & Marcos García Salido

Universidade de Santiago de Compostela

Abstract

The paper provides a functional, usage-based analysis of the genesis and diffusion of the object clitic doubling in Spanish, which is seen as an agreement phenomenon, and hence as a working example of grammaticalization. It has been claimed that doubling arises from the so-called topic-shift construction (Givón 1976), but historical and contemporary data support an alternative proposal that pays more attention to frequency effects and takes into account the often neglected high proportion of clitic-only objects in discourse. Our analysis suggests a strong correlation between being an object encoded frequently by means of a clitic/affix and developing agreement, which points to the accessibility of the discourse referents as the main factor in determining the grammatical form of direct and indirect objects in Spanish.

1. Introduction¹

The examples below show various possible encodings of direct (1a-c) and indirect (2a-c) objects in Spanish:²

¹ We are indebted to the editors of the volume and to two anonymous referees for useful comments on a previous version of this paper. This research has been funded by the Spanish Ministry of Science and FEDER funds (grants HUM2006-05776 and FFI2010-17417) and The Spanish Ministry of Education (AP 2006-02002).

² The particular constructions represented by these examples are not evenly spread in current Spanish, i.e. the pattern of (1c) is used mostly in the Río de la Plata variety and the

[Draft version]

Vázquez Rozas, Victoria & Marcos García Salido. 2012. [A discourse-based analysis of object clitic doubling in Spanish](#). In Davidse, Kristin, Tine Breban, Lieselotte Brems y Tanja Mortelmans (eds.). *Grammaticalization and Language Change: New reflections*, 271–297. Amsterdam: John Benjamins.

(1a) [...]y de pronto abrió los ojos [...])
And of sudden open-PST.3SG the eyes
'And suddenly he opened his eyes' (SONRISA: 91, 31)³

(1b) [...]y poco después lo-s-invite-ó
And few after 3M.ACC-PL-invite-PST.3SG
al Park a almorzar
to.the Park to lunch
'And shortly after he invited them to the Park for lunch'
(HISTORIAS: 14, 30)

(1c) Me-fui al baño
REFL.1SG-go[PST.1SG] to.the bathroom
para acompañar-la a la Dilia
for accompany-INF.3F.ACC.SG to the Dilia

'I went to the bathroom to accompany Dilia' (GLENDA: 147, 8)

(2a) El banco [...] liquida-ba los correspondientes intereses
The bank pay.off-PST[3SG] the respective interests
a los clientes
to the clients

'The bank paid off the clients' interest' (2VOZ: 28, 1, 2, 2)

(2b) le-gusta mucho la medicina
3SG.DAT-like[PRES.3SG] much the medicine
'He likes medicine very much' (SEVILLA: 82, 15)

indirect object encoding in (2a) is nowadays more common in written texts than in spoken discourse.

³ Complete references for the cited texts are provided in appendix 2.

[Draft version]

Vázquez Rozas, Victoria & Marcos García Salido. 2012. [A discourse-based analysis of object clitic doubling in Spanish](#). In Davidse, Kristin, Tine Breban, Lieselotte Brems y Tanja Mortelmans (eds.). *Grammaticalization and Language Change: New reflections*, 271–297. Amsterdam: John Benjamins.

pronominal element to be reanalyzed as an agreement affix, as will be shown in section 3. However, diachronic data suggest a path of emergence of the object agreement that does not support the TS hypothesis (see section 4), hence the necessity of examining a comprehensive corpus of Spanish texts, which we will do in section 5. The consideration of extensive discourse data of the relevant constructions requires an alternative analysis. While previous studies (e.g. Poston 1953; Barrenechea & Orechia 1970; Silva-Corvalán 1984; Rini 1991; Becerra Bascuñana 2006) have focused on lexical and strong pronominal objects, accompanied or not by co-indexed clitics, as in (1a) vs. (1c) and (2a) vs. (2c)–, there are some works that take a more inclusive view on the data, such as Weissenrieder (1995) and Dufter & Stark (2008). A short survey of these proposals will be given in section 5.1. Next, section 5.2. will provide detailed information on the sample we have used in our study. In section 5.3., besides full-fledged constituents, our analysis will pay attention to those very frequent cases in which objects are represented only by a clitic, as in (1b) and (2b), and accordingly we will put forward a new hypothesis for the origin and spread of object agreement in Spanish (Section 6). In the final section we will summarize the advantages of this proposal, which relates the discourse-pragmatic function of the constructions to phonological, morphological and syntactic properties of the linguistic units involved in this process of grammaticalization leading to object agreement.

2. The status of so-called clitics in Spanish

Before addressing the question of the origin of object agreement, it is useful to justify the assigning of affix status to those elements in bold type in (1b) and (2b) from contemporary Spanish, since these are generally considered to be clitics in the literature. The morphological status of such forms nowadays is not the same as in initial stages of the language. In this respect, we agree with Enrique-Arias (2003), who maintains that, in contemporary Spanish,

[Draft version]

Vázquez Rozas, Victoria & Marcos García Salido. 2012. [A discourse-based analysis of object clitic doubling in Spanish](#). In Davidse, Kristin, Tine Breban, Lieselotte Brems y Tanja Mortelmans (eds.). *Grammaticalization and Language Change: New reflections*, 271–297. Amsterdam: John Benjamins.

they are verbal inflections. Taking into account a number of previous works (Greenberg 1954; Carstairs-McCarthy 1981; Anderson 1982; Zwicky & Pullum 1983; Bybee 1985; Hopper & Traugott 1993; Bybee, Perkins & Pagliuca 1994), Enrique-Arias suggested a set of properties that define the prototype of inflectional expressions. The so-called ‘clitics’ of Spanish comply with almost every feature proposed:

a) Inflectional expressions appear obligatorily in their relevant structural position. With pronominal and preverbal definite objects, object affixes (‘clitics’) are obligatory in contemporary Spanish.

(3a) *lo-vieron* *a él*
3SG.ACC-see-PAST3PL to he
‘they saw him’

(3b) **vieron* *a él*⁵
see-PAST3PL to him
‘they saw him’

The fact that obligatoriness does not spread to every context cannot be taken as evidence against the existence of object agreement, since, as Enrique-Arias notices, constraints in agreement like the ones displayed by in Spanish are attested in other languages (cf. also Siewierska 2004: Ch. 4).

b) Inflectional expressions have grammatical meanings (gender, case) and do not fill any NP slot (they are not in complementary distribution with co-referential NPs). This is the case with the so-called object clitics of Spanish: they have the meanings of person, gender, number and case (though not all

⁵ The examples are taken from Enrique-Arias (2003).

[Draft version]

Vázquez Rozas, Victoria & Marcos García Salido. 2012. [A discourse-based analysis of object clitic doubling in Spanish](#). In Davidse, Kristin, Tine Breban, Lieselotte Brems y Tanja Mortelmans (eds.). *Grammaticalization and Language Change: New reflections*, 271–297. Amsterdam: John Benjamins.

the members of the paradigm exhibit these meanings),⁶ and they are not in complementary distribution with object NPs.

c) Spanish object affixes appear fused with the verb, and, following the criteria that define the inflectional expression prototype as proposed by Enrique-Arias, inflections are bound. As unstressed object pronouns became verbal affixes it is no longer possible to insert other elements in between them and the verb. Rivero (1986: 775) and Rini (1992: 133ff) observe that “interpolation” was common in Old Spanish, as it is shown in (4), where the indefinite stressed pronoun *algo* ‘something’ is placed between the clitic *te* ‘you’ and the verb *tomar* ‘take’. Such a distribution has disappeared in Present-day Spanish:

- (4) *por alguna parte dél [tu reino] te=entra-rá*
for some part of.it 2SG.DAT=enter-FUT[3SG]
alguno para te=algo toma-r
someone to 2SG.DAT=something take-INF
‘in some part of it [your kingdom] someone will enter to take
something from you’ (*Amadís*)

d) Spanish object affixes only attach to the verb, so they exhibit strong selectional restrictions for their hosts.

⁶ According to Enrique-Arias (2003: 73) the loss of some or all the grammatical meanings encoded by clitics reflects “a process towards a higher degree of semantic generalization of the grammatical marker”. This phenomenon can be observed in an example like (i), taken from Enrique Arias (*ibid.*): the dative clitic *le*, that does not make gender distinctions, in some instances also loses the number distinctions (*le* vs. *les*). Moreover, and again according to Enrique Arias, the distinction of case is also lost, since the verb *temer* takes an accusative object:

- (i) *Juan le-teme a las brujas*
Juan 3SG-M-ACC-fears to the witches
‘Juan is scared of witches’

[Draft version]

Vázquez Rozas, Victoria & Marcos García Salido. 2012. [A discourse-based analysis of object clitic doubling in Spanish](#). In Davidse, Kristin, Tine Breban, Lieselotte Brems y Tanja Mortelmans (eds.). *Grammaticalization and Language Change: New reflections*, 271–297. Amsterdam: John Benjamins.

e) While clitics tend to be affected by discourse features, the expression of inflectional affixes is influenced by features of the host word. In contemporary Spanish, the distribution of object affixes is determined by the form of the verbal stem they are attached to: with imperatives and non-finite forms, they are in final position, with finite forms in initial position. This was not the case in Old Spanish: object clitics were subject to a number of restrictions depending on discourse or syntactic features (e.g. they never appeared at the beginning of the sentence; the presence of an indefinite pronoun in the clause conditioned their position, etc.).

d) Inflectional affixes exhibit morphological idiosyncrasies. This criterion proposed by Carstairs-McCarthy (1981) and Zwicky & Pullum (1983) is questioned in Enrique-Arias (2003), as it is not a necessary condition to establish the inflectional status for agreement markers. Inflectional affixes are less likely to develop morphological idiosyncrasies in their combination with the verbal stem than other inflections (tense or mood), which tend to undergo a higher degree of fusion with the verb. However, Enrique-Arias identifies some phonic readjustments that take place in combinations of verb plus object affix:

(5) *va-mos* + *nos* > *vámonos*
go[PRES]-1PL + 1PL.REFL > ‘let’s go’

(6) *venid* + *os* > *veníos*
come[IMP]2PL + 2PL.REFL > ‘come’

In (5) and (6) the endings *-s* and *-d* disappear when personal reflexive affixes are added to the verb.

Another phenomenon mentioned by Enrique-Arias as an argument supporting the analysis of these elements as clitics is attested in non-

[Draft version]

Vázquez Rozas, Victoria & Marcos García Salido. 2012. [A discourse-based analysis of object clitic doubling in Spanish](#). In Davidse, Kristin, Tine Breban, Lieselotte Brems y Tanja Mortelmans (eds.). *Grammaticalization and Language Change: New reflections*, 271–297. Amsterdam: John Benjamins.

standard dialects: the verbal person affix that marks subject agreement –or encodes the referent of the subject– occasionally appears separated from the verb by what is traditionally considered a clitic (It has to be noted that in these dialects, there is no distinction between second and third person plural).

- (7) *sient-e-se-n*
sit-IMP-REFL-2PL
'sit down'

This further supports the contention that, in Spanish, the elements considered clitics should be analyzed as affixes. This is moreover in keeping with a criterion put forward by Zwicky & Pullum (1983: 504) to distinguish both categories according to which “[c]litics can attach to material already containing clitics, but affixes cannot”.

3. TS hypothesis

The most widely accepted explanation for the origin of agreement from pronouns is that proposed by Givón (1976), according to which the TS construction is the context that allows the pronominal element to be reanalyzed as an agreement affix. Givón maintains that agreement has its origin in the TS construction, a device for re-introducing into the discourse an entity that has not been referred to in the immediately preceding context, as illustrated in the following example:

- (8) Context: Once there was a *wizard*. He was very wise, rich, and was married to a beautiful witch. They had two sons. The first was tall and brooding, he spent his days in the forest hunting snails, and his mother was afraid of him. The second was short and vivacious, a bit crazy but always game.

[Draft version]

Vázquez Rozas, Victoria & Marcos García Salido. 2012. [A discourse-based analysis of object clitic doubling in Spanish](#). In Davidse, Kristin, Tine Breban, Lieselotte Brems y Tanja Mortelmans (eds.). *Grammaticalization and Language Change: New reflections*, 271–297. Amsterdam: John Benjamins.

AP [anaphoric pronominalization]: ? *He* lived in Africa.

TS [topic-shift]: *Now the wizard*, he lived in Africa (Givón 1976: 153)

In such a context, if the speaker goes on to mention *the wizard* again, given that several referring expressions that might serve as antecedent to an anaphoric element have occurred in the interim, anaphoric pronominalization becomes too weak a resource, and the speaker must resort to a stronger referring mechanism, i.e. the TS construction. In this construction, the topic is coded by a NP which is detached from the core clause by a pause, and, within the clause, the topic is referred to again by an anaphoric pronoun.

However, the TS construction in fact seems to be too strong as a referring expression in contexts such as that illustrated by (9):

(9) “Context: Once there was a wizard.

AP: He lived in Africa.

TS: ?*Now the wizard*, he lived in Africa” (Givón 1976: 153).

In (9), the anaphoric pronoun *He* is sufficient for the act of reference to be successful. Yet, according to Givón (1976: 153-156), under certain conditions—he mentions the “heavy communicative stress” of Pidgins and Creoles—the TS pattern may be over-used, appearing in contexts such as (9), where, in principle, it would not be likely to appear. If this is the case, the TS construction may be interpreted as unmarked by speakers: the topic-referring NP is no longer seen as an extra-clausal constituent and is reanalysed as an argument of the predicate, so the erstwhile anaphoric pronoun is reinterpreted as an agreement marker.

Following Givón (1976), the TS theory has been used to account for the origin and extent of object agreement in several languages. Some well-

[Draft version]

Vázquez Rozas, Victoria & Marcos García Salido. 2012. [A discourse-based analysis of object clitic doubling in Spanish](#). In Davidse, Kristin, Tine Breban, Lieselotte Brems y Tanja Mortelmans (eds.). *Grammaticalization and Language Change: New reflections*, 271–297. Amsterdam: John Benjamins.

known works are those of Duranti & Ochs (1979) on spoken Italian, Lambrecht (1981) on non-standard French, Fassi Fehri (1987) on Arabian, and Bresnan & Mchombo (1987) on Chichewa.

As for Spanish, Silva-Corvalán (1984) presents a study based on historical and contemporary data that she claims “support Givón’s (1976) proposal that *grammatical agreement arises from topic-verb agreement*” (Silva-Corvalán 1984: 556; emphasis original). The author relates the use and diffusion of object agreement (or clitic doubling) to the “degree of topicality of the object noun phrase” (Silva-Corvalán 1984: 555), conceiving of topicality as a gradual and compositional property made up of features such as animacy and definiteness, and also of formal features such as preverbal order, which Silva-Corvalán identifies as “topic position” (Silva-Corvalán 1984: 555).

One problematic point in Silva-Corvalán’s account, and in the literature dealing with object topicalization in Spanish, is that two partially different structures are lumped together as ‘topicalized objects’, ‘objects in topic position’, etc. The former is made up by a detached noun phrase, without prepositions marking its functional or semantic relation with the verb, and a clause; the latter is a clause with a fronted object marked as such and non-detached. The former resembles the TS construction as described by Givón the most. In practice, however, preverbal, definite, and human objects, such as *A María* in (10), are equated with detached referential expressions such as *Los mancebos de la çibdat* in (11), which are seldom found in the texts, as noted by Riiho (1988), in spite of their differences.⁷

⁷ “El objeto dativo repetido lleva *normalmente* la preposición *a*, como marca de su función. *En algunos casos*, sin embargo, aparece al principio de la oración, como una especie de tópico general, que se comenta luego, y sin presentar ninguna señal externa de su papel.” (Riiho 1988:39; emphasis added).

‘the repeated dative object usually takes the preposition *a* as an index of its function. In certain cases, however, it appears at the beginning of the sentence, like a sort of general topic, which is then talked about and does not show any external marking of its role’.

[Draft version]

Vázquez Rozas, Victoria & Marcos García Salido. 2012. [A discourse-based analysis of object clitic doubling in Spanish](#). In Davidse, Kristin, Tine Breban, Lieselotte Brems y Tanja Mortelmans (eds.). *Grammaticalization and Language Change: New reflections*, 271–297. Amsterdam: John Benjamins.

- (10) *A María le-dio un libro*
To María 3SG.DAT-give[PST.3SG] a book
'To Mary gave (her) a book' (Silva-Corvalán 1984: 555)

- (11) *Los mancebos de la çibdat, tanto*
The boys from the city so.much
les=plaze de la beldat
3PL.DAT=please[PRES.3SG] of the beauty
'As for the boys of the city, they are very pleased with beauty'
(Egipciana 172; cited in Riiho 1988: 40)

The principal problem of Silva-Corvalán's study in connection with the TS hypothesis is that it does not deal with the TS construction as a motivation for the origin of object agreement, but with formal manifestations of topicality and their relation with object agreement. Thus, though her data seem to confirm that preverbal position, animacy, and definiteness of the objects are relevant factors in accounting for the extent of the object agreement, her analysis does not offer new insights into the relation of object agreement and the construction discussed.

4. Shortcomings of TS hypothesis

In spite of some positive evidence, accounting for the origin of object agreement in Spanish by means of the TS hypothesis poses certain problems and leaves several questions unsolved.

Some studies suggest that there is a relation between constructions that seem to fit Givón's description of the TS construction and the presence of an anaphoric element retrieving the reference of a topicalized NP:

[Draft version]

Vázquez Rozas, Victoria & Marcos García Salido. 2012. [A discourse-based analysis of object clitic doubling in Spanish](#). In Davidse, Kristin, Tine Breban, Lieselotte Brems y Tanja Mortelmans (eds.). *Grammaticalization and Language Change: New reflections*, 271–297. Amsterdam: John Benjamins.

There is [...] some evidence which would indicate that, when a strongly stressed object or indirect object precedes the verb, it is followed by a pause, after which a new breath-group follows an unstressed personal pronoun then becomes necessary in order to link the isolated object with the main clause (Keniston 1937: 83-84)

It appears, however, that in Old Spanish the unstressed pronoun alluded to here by Keniston was in fact used to refer anaphorically or cataphorically to any constituent detached from the clause, regardless of the topicality of that constituent. That is why Riiho (1988) finds that the most frequent context for the doubling of postverbal direct objects in his corpus is one in which a subordinate clause is detached from the main clause:

El caso típico de la duplicación de un objeto directo pospuesto al verbo lo constituye el uso del pronombre neutro para anticipar cualquier clase de frase subordinada, después de un adverbio, (p. ej. *bien, ya mucho, assi*), una locución adverbial fija (p. ej. *por tu boca, con uuestro conseio, en buen ora*), un complemento circunstancial o una conjunción. (Riiho 1988: 35)

[The typical case of doubling of a postverbal object concerns the use of the neuter pronoun to anticipate any kind of subordinated clause after an adverb (e. g. *bien, ya mucho, assi*), a fixed adverbial phrase (e. g. *por tu boca, con uuestro conseio, en buen ora*), an adjunct or a conjunction]

- (12) Ca todos lo=saben que
For all.M.PL 3SG.M.ACC=know[PRES]3PL that

[Draft version]

Vázquez Rozas, Victoria & Marcos García Salido. 2012. [A discourse-based analysis of object clitic doubling in Spanish](#). In Davidse, Kristin, Tine Breban, Lieselotte Brems y Tanja Mortelmans (eds.). *Grammaticalization and Language Change: New reflections*, 271–297. Amsterdam: John Benjamins.

percentages of more than a 90% in some genres irrespective of their distribution, while, among Direct Objects (henceforth DO), it takes place almost only with preverbal objects (with the exception of Argentinian Spanish).⁸ In turn, with stressed personal pronouns functioning as objects, verbal agreement is obligatory.

5. A new approach to discourse data

5.1. A note on previous literature

Most of the previous research on object agreement in Spanish has focused on full-fledged objects (with a comparison to without clitic agreement) and has disregarded the occurrences of object clitics on their own in the analyses. As far as we know, in the body of literature on clitic doubling only two pieces of research have based their analyses on the various types of object expressions found in the selected discourse samples, i.e., non-clitic (NP and stressed pronoun), non-clitic+clitic (NP/stressed pronoun + clitic), and clitic in isolation. These are Weissenrieder (1995) and Dufter & Stark (2008). Although the aim of these papers, the selection of data and the conclusions drawn are different from each other, and are also different from ours, Weissenrieder's and Dufter & Stark's proposals are worth a short reference/review here, for the sake of comparison with the data and the analysis that will put forward in the next sections.

First of all, it should be pointed out that both articles focus on very restricted sets of constructions. Weissenrieder limits her research to 3rd person indirect objects, whereas Dufter & Stark analyze just the expression of “the indirect objects of a small set of highly frequent verbs” (Dufter & Stark 2008: 114), which are *dar* ‘give’, *decir* ‘say’, *parecer* ‘seem’, *gustar*,

⁸ Argentinian Spanish has developed the agreement with animate (and even inanimate) and definite postposed DO, and indeed with specific and given inanimate DO. For a recent functional analysis based on Buenos Aires' conversational data, see Belloro (2007).

[Draft version]

Vázquez Rozas, Victoria & Marcos García Salido. 2012. [A discourse-based analysis of object clitic doubling in Spanish](#). In Davidse, Kristin, Tine Breban, Lieselotte Brems y Tanja Mortelmans (eds.). *Grammaticalization and Language Change: New reflections*, 271–297. Amsterdam: John Benjamins.

encantar ‘please’. Dufter & Stark’s paper has a contrastive and diachronic bias, so they also provide data from Italian and they take a closer look at the historical evolution of “preverbal pronominal indirect object doubling” in Spanish (Dufter & Stark 2008: 120). As for our proposal, every occurrence of both direct and indirect objects in the sample has been accounted for (see section 5.2. below).

The size and characteristics of the samples used in the two papers is also fairly reduced when compared to ours, since Weissenrieder relies on a sole novel from the Argentinian writer Manuel Puig, and Dufter & Stark base their analysis on a sample of contemporary spoken Spanish of approximately 300,000 words, and also on some very selective searches in the *Corpus del español* of Mark Davies.

Both papers give us new insights into the phenomenon of object agreement, but also both leave some problems unsolved. Weissenrieder summarizes the main lines of research taken in the past, points out the shortcomings of analyses limited to the sentence level, and argues for adopting a discourse perspective. Her proposal relates verbal agreement to the topical character of the discourse entities encoded as subject, indirect object and direct object, though only empirical data on the indirect object are given in the article. The main drawback of this view lies in the elusiveness of the notion of ‘topicality’, which Weissenrieder relates to such heterogeneous properties as the NP characteristics (animacy and definiteness), the syntactic function (S>IO>DO) and also the “form” of the constituent (“noun phrase” vs. “pronominal phrase” vs. “agreement particle”).

As for Dufter & Stark’s (2008) paper, it provides us with some interesting suggestions based on corpus data, for instance, the existence of a correlation between clitic doubling and lexically-driven thematic roles (e.g., experiencer of *parecer*, *gustar*, *encantar*). Nevertheless, their analysis presents a serious shortcoming, since it is based on an peculiar interpretation

[Draft version]

Vázquez Rozas, Victoria & Marcos García Salido. 2012. [A discourse-based analysis of object clitic doubling in Spanish](#). In Davidse, Kristin, Tine Breban, Lieselotte Brems y Tanja Mortelmans (eds.). *Grammaticalization and Language Change: New reflections*, 271–297. Amsterdam: John Benjamins.

of what “clitic doubling” is. In the view of Dufter & Stark, in constructions like *a mí non me pesa* ‘It does not grieve me’ (c. 1140, *Cantar de Mio Cid*; quoted in Dufter & Stark 2008: 117), the element considered as “doubled” is the clitic *me*. As a consequence of this analysis, they calculate “the percentage of clitic-doubled *a mí* (or *a mi*) among all occurrences of *me*” (Dufter & Stark 2008: 121-122; see also Table 6 p. 121). And as the proportion of this doubling does not grow over time, they conclude that there are no arguments to see the effects of a grammaticalization process in this construction. Nevertheless, the usual view of the clitic doubling as grammaticalization is just the opposite, as this view is based on the gradual increasing over time of agreeing clitics when a stressed nominal or pronominal object is present in the clause. In other words, whereas the usual path of grammaticalization of the object agreement is

(13) *A mí plaze* > *A mí me=plaze*
To me please[PRES.3SG] > To me 1SG.DAT=please[PRES.3SG]
‘It pleases to me’ ‘It pleases to me’

(14) *Die-ro-n el libro a María*
Give-PST-3PL the book to María
‘they gave the book to Mary’
> *Le-die-ro-n el libro a María*
> 3SG-DAT-give-PST-3PL the book to María
‘they gave the book to Mary’;

in Dufter & Stark's analysis, however, the grammaticalization is supposed to follow a very different course:

(15) *Me=plaze* > *A mí me=plaze*
1SG.DAT=please[PRES.3SG] > To me 1SG.DAT=please[PRES.3SG]

[Draft version]

Vázquez Rozas, Victoria & Marcos García Salido. 2012. [A discourse-based analysis of object clitic doubling in Spanish](#). In Davidse, Kristin, Tine Breban, Lieselotte Brems y Tanja Mortelmans (eds.). *Grammaticalization and Language Change: New reflections*, 271–297. Amsterdam: John Benjamins.

‘It pleases to me’

‘It pleases to me’

Adopting such a view on the phenomenon of clitic doubling, quantitative historical data cannot give any evidence of change along centuries.⁹ Dufter & Stark (2008: 121) do not notice that the “unexpected” results arise from a unsound approach to the data.¹⁰ Starting from erroneous premises, they finally come to an inadequate conclusion:

Therefore, it does not appear justified to consider spoken varieties of Spanish or Italian to be further advanced towards the grammaticalization of object clitics into agreement markers. With respect to linguistic factors influencing the likelihood of double indirect pronouns, the evolution of Spanish does not provide evidence for any change at all. In particular, we failed to find a marked increase in clitic-doubling for the most frequent type of pronominal indirect objects (Dufter & Stark 2008: 127)

5.2. *The data*¹¹

The sample we use here contains roughly 1,936,000 words. It comprises texts ranging from the 12th to the 17th century. It is at the beginning of 17th century that the grammaticalization process by means of which pronominal clitics became affixes is more or less consolidated (cf. Rini 1991: 282; Enrique-Arias 2003). In addition, we have added data extracted from

⁹ It would be as expecting a grow of the use of subject (stressed) pronouns *yo* ‘I’ and *tú* ‘you’ as a proof of the grammatical character of the subject agreement, when it goes without discussion that subject agreement is fully grammaticalized in Spanish no matter a subject pronoun appears or not in particular clauses.

¹⁰ “Given the claim that indirect object clitic-doubling has been undergoing a process of grammaticalization in Modern Spanish, one might expect the proportion of clitic-doubled *a mí* to grow over time. However, the data from the *Corpus del español* show precisely the opposite development” (Dufter and Stark 2008: 121)

¹¹ A complete relation of the texts for 13th-17th centuries is provided in appendix 1.

[Draft version]

Vázquez Rozas, Victoria & Marcos García Salido. 2012. [A discourse-based analysis of object clitic doubling in Spanish](#). In Davidse, Kristin, Tine Breban, Lieselotte Brems y Tanja Mortelmans (eds.). *Grammaticalization and Language Change: New reflections*, 271–297. Amsterdam: John Benjamins.

corpora of Present-day Spanish in order to get a more complete perspective of this process.¹²

The non-contemporary texts are all prose with one exception (*Cantar de mio Cid*) and the majority of them are narratives. There are two instances of doctrinal prose (*Corbacho* and *Las Moradas*), two are dialogues, after the Latin *comoedia humanistica* (*La Celestina* and *La Dorotea*), and there is an instance of legal texts (*Andalusian documents*). The contemporary ones include the texts of the ARTHUS corpus (made up of newspapers, literary prose and transcriptions of oral productions) and three spontaneous conversations from Briz and Val.Es.Co (2002).

We have analysed direct and indirect objects and divided them into three different categories:

- i. We use the label *clitic* for unstressed personal forms like *me, te, se, lo, le*, etc. For the sake of economy and ease of comparison, in table 1 below, we apply this term to those forms even in periods when they should be considered verbal affixes.
- ii. Under the term *non-clitic* we group all those objects encoded by forms other than those described in i. and without doubling or verbal agreement. This group encompasses a wide range of units: relative pronouns, stressed pronouns, bare nouns, noun phrases, complement clauses and relative clauses.¹³
- iii. The third category comprises those objects encoded by a clitic/affix plus a non-clitic unit.

¹² Although in the case pronominal objects the conditions of agreement are practically the same in the present and in the 17th century (agreement is obligatory here), in other contexts there have been changes: for instance, the agreement of indirect objects has undergone an increase, so much so that in contemporary Spanish conversation indirect objects without verbal agreement are virtually non-existent.

¹³ The label ‘non-clitic’ is not entirely accurate, since some unstressed units like the relative pronoun *que* are covered by it.

[Draft version]

Vázquez Rozas, Victoria & Marcos García Salido. 2012. [A discourse-based analysis of object clitic doubling in Spanish](#). In Davidse, Kristin, Tine Breban, Lieselotte Brems y Tanja Mortelmans (eds.). *Grammaticalization and Language Change: New reflections*, 271–297. Amsterdam: John Benjamins.

For centuries 13th and 14th we had access to a database made previous to this study by researchers of the Spanish Department of the University of Santiago de Compostela. This part of the sample comprises ca. 345,000 tokens. From the 15th century onwards, we have collected data from excerpts of three different texts per century, each excerpt made up of ca. 10,000 tokens (i.e. 30,000 tokens per century).

The Present-day data include 1,470,046 words in total, of which 21,041 tokens correspond to conversation and 1,449,005 tokens are from ARTHUS, a heterogeneous corpus made up of texts published between 1980 and 1991. Syntactic and semantic features for each clause in ARTHUS (158,954 in total) are recorded in a complex database (BDS/ADESSE).¹⁴

Table 1 summarizes the analysis of the whole sample, providing an overview of the encoding of the objects across centuries

		IO			DO		
		Clitic	Non-clitic	Clitic+non-clitic	Clitic	Non-clitic	Clitic+non-clitic
13th cent.		1044 (72.1%)	336 (23.2%)	67(4.7%)	1427(19%)	5915 (78.3%)	207 (2.7%)
14th cent.		1738 (75.9%)	519 (22.7%)	33 (1.4%)	970 (13.5%)	6145 (85.6%)	66 (0.9%)
15th cent.		302 (64.5%)	154 (32.9%)	12 (2.6%)	424 (20.6%)	1602 (78%)	28 (1.4%)
16th cent.		579 (82%)	112 (15.9%)	15 (2.1%)	628 (24.7%)	1870 (73.6%)	43 (1.7%)
17th cent.		466 (75.4%)	106 (15.2%)	46 (7.4%)	563 (23.8%)	1771 (74.9%)	31 (1.3%)
20th cent.	BDS	9727 (78.6%)	1085 (8.8%)	1554 (12.6%)	13333 (21%)	49131 (77.4%)	1005 (1.6%)
	Conv.	361 (88.5%)	2 (0.5%)	45 (11%)	302 (27.6%)	740 (67.5%)	54 (4.9%)

Table 1: Object encoding across centuries

¹⁴ For further information, the reader is referred to <http://www.bds.usc.es/> and <http://adesse.uvigo.es/>.

[Draft version]

Vázquez Rozas, Victoria & Marcos García Salido. 2012. [A discourse-based analysis of object clitic doubling in Spanish](#). In Davidse, Kristin, Tine Breban, Lieselotte Brems y Tanja Mortelmans (eds.). *Grammaticalization and Language Change: New reflections*, 271–297. Amsterdam: John Benjamins.

These results are quite consistent across centuries and genres: indirect objects overwhelmingly tend to be encoded by clitics/affixes, while direct objects are their mirror image in this respect. They are mostly encoded by means of a non-clitic unit. Finally, the doubled/agreeing objects, which include all the exemplars of TS construction, but not only, are consistently a minority of cases.

5.3. *The analysis*

Some earlier studies (Poston 1953, Fish 1968, Barrenechea & Orecchia 1970, Silva-Corvalán 1984 or García-Miguel & Vázquez Rozas 1994, among others) have supported their findings with discourse data. Yet the way in which data in such studies have been used is, to our mind, a skewed one: only the contrast between non-doubled and doubled NP objects is addressed and a considerable set of objects is neglected, that is, those encoded by means of a clitic or object affix alone. In contrast, if we consider the whole range of possible encodings, this overview can shed considerably more light on the issue in question and can also serve to avoid the shortcomings of the TS hypothesis.

Figures 1 and 2 present the data displayed in table 1 regarding the encoding of direct and indirect objects in Spanish:

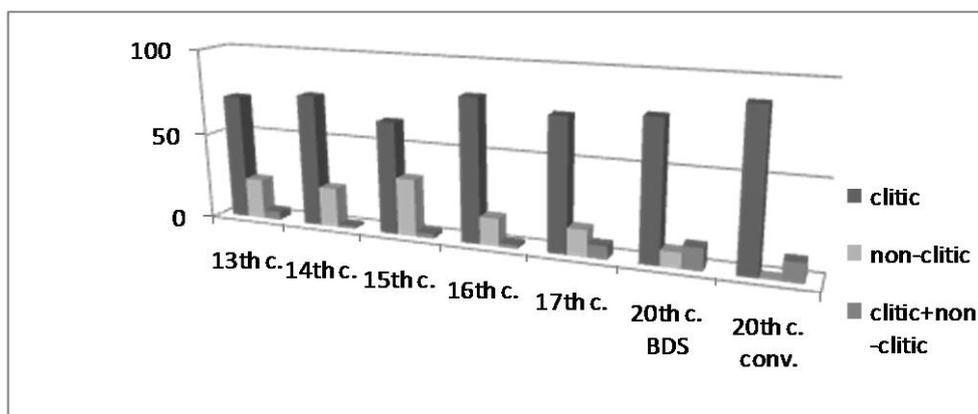


Figure 1: Coding devices for IO across time

[Draft version]

Vázquez Rozas, Victoria & Marcos García Salido. 2012. [A discourse-based analysis of object clitic doubling in Spanish](#). In Davidse, Kristin, Tine Breban, Lieselotte Brems y Tanja Mortelmans (eds.). *Grammaticalization and Language Change: New reflections*, 271–297. Amsterdam: John Benjamins.

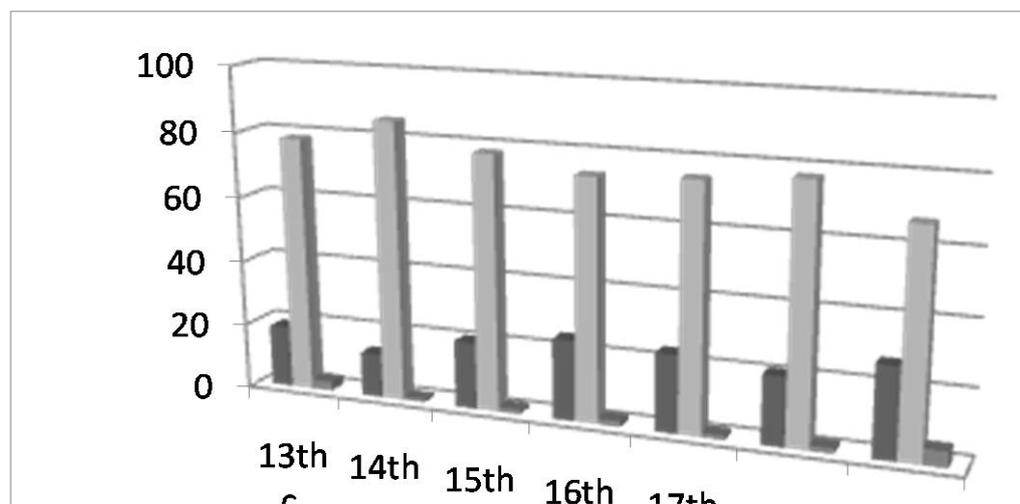


Figure 2: Coding devices for DO across time

From the first texts of the sample, it can be seen that objects are encoded according to patterns that are recurrent over time and across the different genres represented in our sample.

In the first place, and as already pointed out with respect to the TS hypothesis, doubled objects are a marginal object encoding, although in contemporary Spanish a considerable increase of agreeing IO's is attested. Secondly, there is a clear split between IO and DO what concerns preferred categories: IO's exhibit a tendency to be encoded by clitics alone, whereas DO's tend to be encoded by more complex constituents. In addition, the data suggest a strong correlation between an object which is encoded frequently by means of a clitic/affix and developing agreement. Thus, in Present-day Spanish, the agreement of non-clitic IO's oscillates between 59% of the cases in a multi-genre corpus and 95,7% in conversation, while DO's agreement is quite restricted in most of its dialects.

What is the meaning of the different linguistic encodings an object can receive and why do these encodings seem to follow a pattern that can be verified across different periods and discourse genres?

With regard to the first question, it is generally agreed upon in the literature that a relation exists between the accessibility that a referent has

[Draft version]

Vázquez Rozas, Victoria & Marcos García Salido. 2012. [A discourse-based analysis of object clitic doubling in Spanish](#). In Davidse, Kristin, Tine Breban, Lieselotte Brems y Tanja Mortelmans (eds.). *Grammaticalization and Language Change: New reflections*, 271–297. Amsterdam: John Benjamins.

for speakers and the linguistic encoding that this referent requires (cf. Givón 1983, 1992; Ariel 1990): broadly, the more accessible a referent is, the less semantic and phonic content is required for its linguistic expression. This relation has been empirically tested in the cited works.

In the light of this, we can draw the conclusion that objects encoded by means of a clitic are highly accessible to the hearer, since a clitic has a minimum semantic content (person and, not always, gender and number features) and quite a weak phonic structure (clitics are unstressed monosyllabic units). The following examples illustrate this point.

- (16) [...] pero muchas veces acontece que los que tenían méritamente granjeada y alcanzada gran fama por sus escritos, en dándolos a la estampa la perdieron del todo [...] (*Quijote*, 2nd part, ch. 3)
‘but many times it happens that those who have conquered and merited fame by means of their writings, they lost it completely by having them printed’
- (17) pero la Santa Escritura, que no puede faltar un átomo en la verdad, nos muestra que los hubo [gigantes], contándonos la historia de aquel filisteazo de Golías (*Quijote*, 2nd part, ch. 1)
‘but the Holy Scriptures, which cannot lie, show us that there were some [giants], telling us the story of Goliath, that big philistine’

In (16), the object affix *los* ‘them’ refers to a highly accessible entity, since it has been mentioned in the previous clause (the NP *sus escritos* ‘their writings’). By contrast, the NP *la estampa* ‘the printing’ encodes a referent that has not been mentioned before, hence its more complex encoding. In (17), we again have a noun phrase encoding a newly introduced referent (*la historia de aquel filisteazo de Golías* ‘the story of Goliath, that big philistine’), while the first person affix (*-nos* ‘us’) is used

[Draft version]

Vázquez Rozas, Victoria & Marcos García Salido. 2012. [A discourse-based analysis of object clitic doubling in Spanish](#). In Davidse, Kristin, Tine Breban, Lieselotte Brems y Tanja Mortelmans (eds.). *Grammaticalization and Language Change: New reflections*, 271–297. Amsterdam: John Benjamins.

to encode the speech act participants (SAP; here writer and reader). In this case, it is the fact of being prominent elements of the communicative situation that licenses a high accessibility encoding. The examples differ, however, in the syntactic function each form is associated with: in (16) the clitic is the DO and the NP the IO, whereas example (17) illustrates the opposite situation.

Following this argument, and with the data presented above in mind, we can conclude that, in Spanish, IO's in general are more accessible than DO's, since the latter are encoded mostly by units which are semantically and phonically more elaborated than clitics —so that (16) above exemplifies an exceptional context, while (17) is the default case. This leads us to our second question, the existence of regular encoding patterns for different syntactic functions.

Du Bois (2003a, b) has put forward a hypothesis according to which O arguments are preferred loci for introducing referents which constitute new information or, in the terms we have been used up until now, whose accessibility is low, while A arguments exhibit more restrictions in this regard. The letters A and O refer to agent-like and patient-like arguments, respectively, which, in nominative languages, usually correspond to the subject and object of a transitive clause.¹⁵ This preference for the distribution of referents according to their accessibility has received the name of Preferred Argument Structure (PAS). There is an amount of work that has proved the cross-linguistic validity of the predictions made by Du Bois (see Du Bois, Kumpf and Ashby 2003 and references therein), that he himself extends to ditransitive clauses (cf. Du Bois 2003b: 42 and further).

In the case of Spanish a similar patterning of informational distribution can be proposed for IO's and DO's, the latter being preferred loci for low-accessibility referents.

¹⁵ There is a third category, S arguments, which stands for the arguments of intransitive predicates.

[Draft version]

Vázquez Rozas, Victoria & Marcos García Salido. 2012. [A discourse-based analysis of object clitic doubling in Spanish](#). In Davidse, Kristin, Tine Breban, Lieselotte Brems y Tanja Mortelmans (eds.). *Grammaticalization and Language Change: New reflections*, 271–297. Amsterdam: John Benjamins.

Examples (16) and (17) above illustrate two different sources of high accessibility: recency of mention and salience in the communicative situation. In the case of IO, these two contextual properties can be related to a pragmatico-semantic feature recurrently associated with this semantic function: its high animacy. Like A arguments, IO's overwhelmingly tend to be highly animated and human, as the data in table 2, extracted from BDS show:

	Animate		Inanimate	
Object affix	9503	81,21%	224	27,76%
Object affix + NP	1404	12,15%	150	18,59%
NP	652	5,64%	433	53,65
Total	11559		807	

Table 2: Frequencies and percentages of animate vs. inanimate IO's (data BDS)

DO objects, on the other side, typically show lower animacy (table 3):

	Animate		Inanimate	
Object affix	8027	65,87%	5306	10,35%
Object affix + NP	429	3,52%	576	1,12%
NP	3731	30,61%	45400	88,53%
Total	12187		51282	

Table 3: Frequencies and percentages of animate vs. inanimate DO's (data BDS)

The ultimate reason behind the link between high animacy/humanness and IO's can be found in the semantic roles encoded by this syntactic function (i.e. experiencer, recipient). DO's, on the other hand,

[Draft version]

Vázquez Rozas, Victoria & Marcos García Salido. 2012. [A discourse-based analysis of object clitic doubling in Spanish](#). In Davidse, Kristin, Tine Breban, Lieselotte Brems y Tanja Mortelmans (eds.). *Grammaticalization and Language Change: New reflections*, 271–297. Amsterdam: John Benjamins.

semantically tend to be themes or patients, which favour an interpretation in which they are viewed as entities with a low degree of animacy. Since it is generally acknowledged that human discourse deals mostly with human referents (cf. Hawkinson & Hyman 1974; Givón 1983; Dahl and Fraurud 1996; Dahl 1997; 2000; 2008; Thompson & Hopper 2001: 53), i.e. its topics are frequently human referents, it is plausible that in Spanish, the syntactic function IO encodes highly continuous or prominent topics, by virtue of its association with human and highly animated referents.¹⁶

Thus, a parallel can be drawn between the semantic and pragmatic properties related to each syntactic function and the differences in agreement advancement: figure 3 illustrates the advancement of agreement with IO across centuries, whereas figure 4 illustrates how agreement with DO is similarly restricted in Old Spanish and in 20th century.

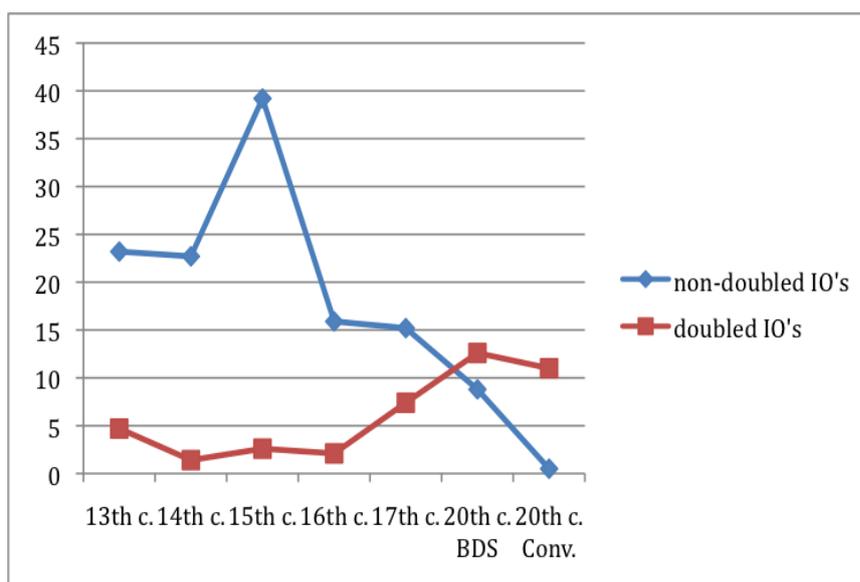


Figure 3: Advancement of IO agreement across centuries

¹⁶ Likewise, the tables above suggest a strong association between minimal linguistic encodings (affixes) and a high degree of animacy.

[Draft version]

Vázquez Rozas, Victoria & Marcos García Salido. 2012. [A discourse-based analysis of object clitic doubling in Spanish](#). In Davidse, Kristin, Tine Breban, Lieselotte Brems y Tanja Mortelmans (eds.). *Grammaticalization and Language Change: New reflections*, 271–297. Amsterdam: John Benjamins.

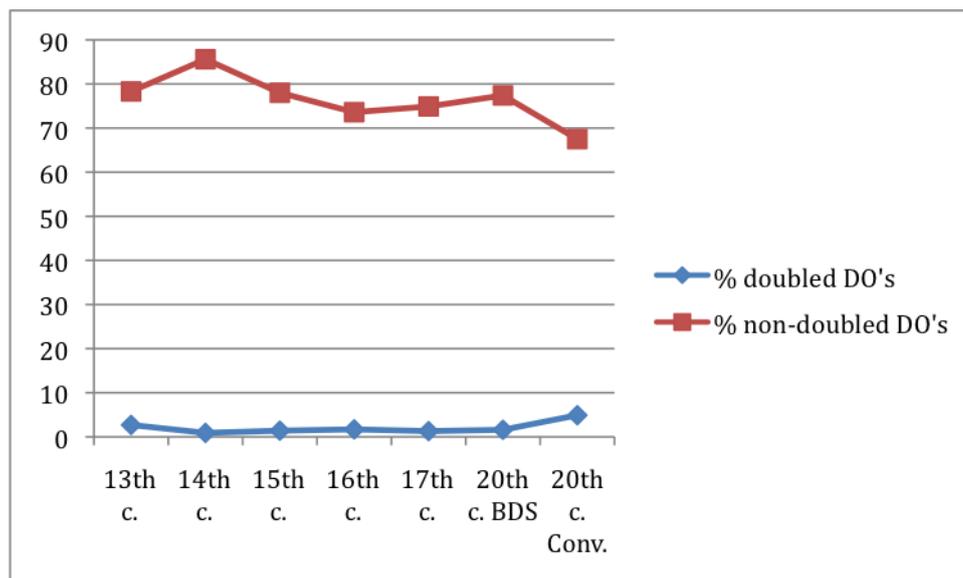


Figure 4: Frequency of doubled and non-doubled DO's across centuries

So far, we have dealt with the differences between IO's and DO's in terms of their respective agreement patterns are concerned. However, stressed personal pronouns exhibit peculiar patterns irrespective of being IO's or DO's. When an independent personal pronoun occurs as an object (*a mí, a ti, a ella*, and so on) it always agrees with the verb. This state of affairs has remained the same since around the end of 16th century. It is plausible to assume that when speakers make use of person markers (clitics/affixes or independent/stressed personal pronouns), the intended referent possesses a high degree of accessibility. Most of the times they would use simply an unstressed form (an affix in Present-day Spanish; a clitic before the grammaticalization of these units into affixes was completed). Only under certain circumstances (e. g., in order to convey a contrastive meaning), would they employ independent/stressed personal pronouns.¹⁷ But these circumstances arise sporadically, hence the data in Table 4, showing that

¹⁷ That is the pragmatic function that Rini (1991) attributes to the use of a stressed personal pronoun.

[Draft version]

Vázquez Rozas, Victoria & Marcos García Salido. 2012. [A discourse-based analysis of object clitic doubling in Spanish](#). In Davidse, Kristin, Tine Breban, Lieselotte Brems y Tanja Mortelmans (eds.). *Grammaticalization and Language Change: New reflections*, 271–297. Amsterdam: John Benjamins.

SAP's are encoded by an affix at a proportion of approximately nine to one:¹⁸

	Affix	Affix + pers. pronoun
DO	6749 (91.7%)	559 (8.3%)
IO	3505 (95%)	173 (5%)

Table 4: Frequencies and percentages of syntactic encoding of 1st and 2nd pers. SG. objects (data from BDS)

The situation depicted in this table is easily projectable on preceding stages of Spanish and very similar to that of indirect objects: highly accessible referents are preferably encoded by attenuated forms.

The frequent clitic encoding seems to have been crucial for the development of agreement both in IO's and in independent/stressed pronominal objects. As for the latter, from the first Spanish texts, stressed personal pronouns doubled by clitics are attested. Melis & Flores (2005), studying the diachrony of agreement in IO's, point out that, until the 17th century, clitic doubling seems to be favoured by objects occurring in preverbal position. From then on agreement is associated with the object being a personal pronoun, irrespective of its distribution.

6. Discussion

The data we have discussed so far show that the types of objects that, because of their high accessibility, exhibit a tendency (diachronically and synchronically) to be encoded by a clitic or an affix, are those that have developed agreement.

¹⁸ Speech Act Participants, that is, speakers and hearers, which are almost only encoding by means of 1st and 2nd person forms (clitics, affixes or pronouns).

[Draft version]

Vázquez Rozas, Victoria & Marcos García Salido. 2012. [A discourse-based analysis of object clitic doubling in Spanish](#). In Davidse, Kristin, Tine Breban, Lieselotte Brems y Tanja Mortelmans (eds.). *Grammaticalization and Language Change: New reflections*, 271–297. Amsterdam: John Benjamins.

The different object agreement types in contemporary Spanish can be regarded as a case of fixation of highly frequent patterns. Clitics were the preferred form to encode objects in contexts where their referents were highly accessible. In such contexts, they have undergone a morphologization process through which they have become verbal affixes. This evolution was favoured because of the scant phonic substance of clitics (related also to the high accessibility of their referents) and the fact that they must attach to other words. That they have become fixed as verbal morphemes suggests that the verb+clitic combination was more frequent than other possible ones. In fact, other combinations attested in earlier stages of Spanish have disappeared.

On the other hand, at a certain point of the language's history the complementary distribution of clitics and full-fledged objects was lost. But this process is relatively independent of the morphologization of clitics into affixes; while the co-occurrence of clitics and full object phrases is attested in the earliest documents of Spanish, clitics do not exhibit the features that classify them as verbal affixes at least until the 17th century (cf. Rini 1990).

The non-complementary distribution between stressed pronouns and clitics before the latter had reached morpheme status could lead again to the vexed question of whether the doubled object construction was a redundant one before the 17th century. Probably there was no such redundancy, since clitics and stressed pronouns fulfilled different informative functions. The following contemporary examples may perhaps serve to illustrate this point.

(18) [...] digo *no no/yo/lo-pag-o yo que / la dueña soy yo* [...]

I 3M.SG.ACC-pay[PRES]-1SG I

‘I say: “no, no. I pay it myself because I am the owner”’

(VALESCO: S.65.A.1; 493)

[Draft version]

Vázquez Rozas, Victoria & Marcos García Salido. 2012. [A discourse-based analysis of object clitic doubling in Spanish](#). In Davidse, Kristin, Tine Breban, Lieselotte Brems y Tanja Mortelmans (eds.). *Grammaticalization and Language Change: New reflections*, 271–297. Amsterdam: John Benjamins.

syntactic slot, even before clitics had fulfilled their morphologization path. From the initial stages of Spanish, stressed pronouns and some clitics were already split into two different series –*me* vs. *a mí*, *te* vs. *a ti*, etc.–, so that the phonological differences between the two did not consist simply of a stress contrast. Moreover, clitics and other object constituents did not have the same syntactic freedom (e.g. clitics could not constitute utterances on their own, they could not be modified, they could not be coordinated, etc.), and the factors that affected the distribution of clitics and other objects were different (the distribution of clitics depended on features such as the polarity of the clause, the presence of indefinites and so on). Hence, a paradigmatic relation between clitics and full-fledged objects was not felt, which facilitate their co-occurrence in the same clause. Both of these facts, the morphologization of clitics and the loss of complementary distribution with syntactic objects, have contributed to a reanalysis of the doubled object construction as an agreement relation.

The hypothesis presented here avoids the problem of the lack of a documented over-use of the TS construction, and also solves a question that would otherwise remain unanswered, viz. that of the unequal spread of agreement across different types of objects. This is the result of the automatization of different encoding patterns related to different properties of object referents.

The fact that an over-use of the TS construction cannot be attested is hardly surprising: the scant use of TS constructions is a consequence of its discourse function. TS is defined by Givón as a conversational device to re-introduce an inactive topic. TS is hence related to a kind of discontinuity in discourse, but discourse topics, usually human, remain continuous over long stretches and are referred to by grammatical markers such as agreement and clitics, associated with their high accessibility. The TS hypothesis, though, is based on a syntactic pattern hardly found in texts.

[Draft version]

Vázquez Rozas, Victoria & Marcos García Salido. 2012. [A discourse-based analysis of object clitic doubling in Spanish](#). In Davidse, Kristin, Tine Breban, Lieselotte Brems y Tanja Mortelmans (eds.). *Grammaticalization and Language Change: New reflections*, 271–297. Amsterdam: John Benjamins.

Furthermore, the evidence in favour of the relation between topicality and object agreement can also serve as support for the hypothesis suggested here. High topicality has much to do with high accessibility. It has been noted that discourse is mostly about human referents: these are more continuous, hence more topical, and, at the same time, more easily accessible to speakers.

7. Conclusion

The object agreement patterns of contemporary Spanish can be viewed as the result of a morphologization process due to the recurrence of certain constructions associated with certain referential properties of their arguments. Clitics have been reanalyzed as verbal affixes as a result of their scant phonic substance and their frequent attachment to verbs with highly accessible objects. At the same time, and to a certain extent independently of the fixation of clitics into affixes, the complementary distribution between affixes/clitics and phrasal or clausal objects disappeared.

This view has some advantages over the most widely accepted hypothesis explaining the emergence of agreement, i.e. the TS hypothesis. Whereas the TS hypothesis may provide a good account for the occurrence of clitics in clauses with topicalized arguments, since it postulates an over-use in order to reanalyze clitics as agreement affixes, it fails to explain the morphologization of clitics in Spanish, because in Spanish such an over-use has never been attested. According to the alternative hypothesis proposed here, the frequent recurrence of the complex clitic+verb is enough for the reanalysis to take place. This is supported by the fact that those objects which, due to their semantic and informational properties, tend to be encoded overwhelmingly by clitics, are those that tend to develop agreement. Furthermore, our hypothesis explains why, in Spanish, object agreement is restricted to certain types of objects.

[Draft version]

Vázquez Rozas, Victoria & Marcos García Salido. 2012. [A discourse-based analysis of object clitic doubling in Spanish](#). In Davidse, Kristin, Tine Breban, Lieselotte Brems y Tanja Mortelmans (eds.). *Grammaticalization and Language Change: New reflections*, 271–297. Amsterdam: John Benjamins.

References

- Academia, Real ____ Española. 1973. *Esbozo de una nueva gramática de la lengua española*. Madrid: Espasa-Calpe.
- Anderson, Stephen. 1982. Where's Morphology? *Linguistic Inquiry* 13: 571-612.
- Ariel, Mira. 1990. *Accessing noun-phrase antecedents*. London: Routledge.
- Barrenechea, Ana María & Teresa Orecchia. 1970. La duplicación de objetos directos e indirectos en el español hablado en Buenos Aires. *Romance Philology* 24(1): 58-83
- Becerra Bascuñana, Silvia. 2006. *Estudio diacrónico y sincrónico del objeto indirecto en español peninsular y de América*. Copenhagen: Museum Tusulanoum Press.
- Belloro, Valeria A. 2007. Spanish Clitic Doubling: a Study of the Syntax-Pragmatics Interface. PhD dissertation, State University of New York at Buffalo.
- Bossong, Georg. 1980. Aktanten Funktionen in romanischen Verbalsystem. *Zeitschrift für romanische Philologie* 96: 1-22.
- Bresnan, Joan & Sam Mchombo. 1987. Topic, pronoun, and agreement in Chichewa. *Language* 63: 741-782.
- Bybee, Joan. 1985. *Morphology*. Amsterdam: John Benjamins.
- Bybee, Joan, Revere D. Perkins & William Pagliuca. 1994. *The evolution of grammar*. Chicago: Chicago University Press.
- Carstairs-McCarthy, Andrew. 1981. *Notes on affixes, clitics and paradigms*. Bloomington: Indiana University Linguistics Club.
- Company, Concepción. 2006. El objeto indirecto. In *Sintaxis histórica de la lengua española. Primera parte: La frase verbal*, Concepción Company (ed.), 1479-1574. México: UNAM-FCE.
- Dahl, Östen. 1997. Egocentricity in Discourse and Syntax.
<

[Draft version]

Vázquez Rozas, Victoria & Marcos García Salido. 2012. [A discourse-based analysis of object clitic doubling in Spanish](#). In Davidse, Kristin, Tine Breban, Lieselotte Brems y Tanja Mortelmans (eds.). *Grammaticalization and Language Change: New reflections*, 271–297. Amsterdam: John Benjamins.

texto.net/marges/marges/Documents%20Site%206/doc0008_dahl_o/e
goc.pdf>

Dahl, Östen. 2000. Egophoricity in discourse and syntax, *Functions of Language* 7: 33-77.

Dahl, Östen. 2008. Animacy and egophoricity: Grammar, ontology and phylogeny. *Lingua* 118(2): 141-150.

Dahl, Östen & Kari Fraurud. 1996. Animacy in grammar and discourse. In *Reference and referent accessibility*, Torsten Fretheim & Jeannette K. Gundel (eds), 47-64. Amsterdam: John Benjamins.

Dufter, Andreas & Elisabeth Stark. 2008. Double indirect object marking in Spanish and Italian. In *Theoretical and Empirical Issues in Grammaticalization*, Elena Seoane & María José López-Couso (eds), 111-129. Amsterdam: John Benjamins.

Du Bois, John W. 2003a. Discourse and Grammar. In *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure vol. 2*, Michael Tomasello (ed.), 47-87. Mahwah (New Jersey): Erlbaum.

Du Bois, John W. 2003b. Argument structure. In *Preferred Argument Structure. Grammar as architecture for function*, John W. Du Bois, Lorraine E. Kumpf & William J. Ashby (eds), 11-60. Amsterdam: John Benjamins.

Duranti, Alessandro & Elinor Ochs. 1979. Left-Dislocation in Italian Conversation. In *Syntax and Semantics. Discourse and Syntax, vol. 12*, Talmy Givón (ed.), 377-416. New York: Academic Press.

Enrique-Arias, Andrés. 2003. From clitics to inflections: diachronic and typological evidence for affixal object agreement marking in Spanish. In *Forum de Morphologie (3e. rencontres)*, Bernard Fradin (ed.), 67-75. Lille: Université de Lille.

[Draft version]

Vázquez Rozas, Victoria & Marcos García Salido. 2012. [A discourse-based analysis of object clitic doubling in Spanish](#). In Davidse, Kristin, Tine Breban, Lieselotte Brems y Tanja Mortelmans (eds.). *Grammaticalization and Language Change: New reflections*, 271–297. Amsterdam: John Benjamins.

- Fassi-Fehri, Abdelkader. 1987. Agreement in Arabic, binding and coherence. In *Agreement in natural language*, Michael Barlow & Charles A. Ferguson (eds), 107-158. Stanford: CSLI.
- Fernández Soriano, Olga. 1999. El pronombre personal. Formas y distribuciones. In *Gramática descriptiva de la lengua española* vol.1, Ignacio Bosque & Violeta Demonte (eds), 1208-1272. Madrid: Espasa-Calpe.
- Fish, Gordon T. 1968. The indirect object and the redundant construction. *Hispania* 51: 862-866.
- Franco, Jon. 2000. Agreement as a Continuum: The Case of Spanish Pronominal Clitics. In *Clitic Phenomena in European Languages*, Frits Beukema & Den Dikken (eds), 147-190. John Benjamins: Amsterdam.
- García-Miguel, José M^a. 1991. La duplicación de complemento directo e indirecto como concordancia. *Verba* 18: 375-410.
- García-Miguel, José M^a and Victoria Vázquez Rozas. 1994. Lingüística de corpus y Lingüística descriptiva: el caso de la duplicación de objetos. *Boletín de la Sociedad Española para el Procesamiento del Lenguaje Natural* (SEPLN) 14: 47-62.
<<http://www.sepln.org/revistaSEPLN/revista/14/>>
- Givón, Talmy. 1976. Topic, pronoun and grammatical agreement. In *Subject and Topic*, Charles Li (ed.), 151-188. New York: Academic Press, 151-188.
- Givón, Talmy. 1983. Topic continuity in discourse: an introduction. In *Topic Continuity in Discourse. A Quantitative Cross-language Study*, Talmy. Givón (ed.), 5-41. Amsterdam: John Benjamins.
- Greenberg, Joseph H. 1954. A quantitative approach to the morphological typology of language. In *Method and perspective in anthropology. Papers in honor of Wilson D. Dallis*, R. F. Spencer (ed.), 192-220. Minneapolis: The University of Minnesota Press.

[Draft version]

Vázquez Rozas, Victoria & Marcos García Salido. 2012. [A discourse-based analysis of object clitic doubling in Spanish](#). In Davidse, Kristin, Tine Breban, Lieselotte Brems y Tanja Mortelmans (eds.). *Grammaticalization and Language Change: New reflections*, 271–297. Amsterdam: John Benjamins.

Hawkinson, Annie K. & Larry M. Hyman. 1974. Hierarchies of natural topic in Shona. *Studies in African Linguistics* 5: 147-170.

Heger, Klaus. 1966. La conjugaison objective en français et en espagnol. *Langages* 3: 19-39.

Hopper, Paul J. & Elizabeth C. Traugott. 1993. *Grammaticalization*. Cambridge: Cambridge University Press.

Keniston, Hayward. 1937. *The Syntax of Castilian Prose. The Sixteenth Century*, Chicago: University of Chicago Press.

Lambrecht, Knud. 1981. *Topic, anti-topic and verb agreement in Non-standard French*. Amsterdam: John Benjamins.

Lenz, Rodolfo. 1920. *La oración y sus partes*. 3rd. Ed. Madrid: Centro de Estudios Históricos.

Llorente, Antonio & José Mondéjar. 1974. La conjugación objetiva en español. *RSEL* 4(1): 1-60.

Melis, Chantal & Marcela Flores. 2005. Acercamiento diacrónico a la duplicación del objeto indirecto. In *Homenaje a Rubén Bonifaz Nuño. 30 años del Instituto de Investigaciones Filológicas*, Alejandra Viguera Ávila (ed.), 481-503. México: UNAM.

Poston, Lawrence. 1953. The redundant object pronoun in contemporary Spanish. *Hispania* 36: 263-273.

Riiho, Timo. 1988. *La redundancia pronominal en el iberorromance medieval*. Tübingen: Max Niemeyer.

Rini, Joel. 1990. Dating the grammaticalization of the Spanish clitic pronoun. *Zeitschrift für romanische Philologie* 106: 354-370.

Rini, Joel. 1991. The Redundant Indirect Object Constructions in Spanish: A New Perspective. *Romance Philology* 45(2): 269-286.

Rini, Joel. 1992. *Motives for Linguistic Change in the Formation of the Spanish Object Pronouns*. Newark, Delaware: Juan de la Cuesta.

Rivero, María Luisa. 1986. Parameters in the Typology of Clitics in Romance and Old Spanish. *Language* 62: 774-807.

[Draft version]

Vázquez Rozas, Victoria & Marcos García Salido. 2012. [A discourse-based analysis of object clitic doubling in Spanish](#). In Davidse, Kristin, Tine Breban, Lieselotte Brems y Tanja Mortelmans (eds.). *Grammaticalization and Language Change: New reflections*, 271–297. Amsterdam: John Benjamins.

Rothe, Wolfgang. 1966. Romanische Objektkonjugation. *Romanische Forschungen* 78: 530-547.

Siewierska, Anna. 2004. *Person*. Cambridge: Cambridge University Press

Silva-Corvalán, Carmen. 1984. Semantic and pragmatic factors in syntactic change. In *Historical Syntax*, Jacek Fisiak (ed.), 555-573. Berlin: Mouton.

Tesnière, Lucien. 1959. *Eléments de syntaxe structural*. Paris: Klincksieck.

Thompson, Sandra A. & Paul J. Hopper. 2001. Transitivity, clause structure, and argument structure: Evidence from conversation. In *Frequency and the emergence of linguistic structure*, Joan L. Bybee & Paul J. Hopper (eds), 27-60. Amsterdam: John Benjamins.

Weissenrieder, Maureen. 1995. Indirect Object Doubling: Saying Things Twice in Spanish. *Hispania* 78: 169-177.

Zwicky, Arnold M. & Geoffrey L. Pullum. 1983. Cliticization vs. Inflection: English *n't*. *Language* 59(3): 502-13.

Corpora

Briz, Antonio and Val.Es.Co Group. 2002. *Corpus de conversaciones coloquiales*. Madrid: Arco/Libros.

Appendix 1: Composition of 13th-17th c. corpus

13th century

Cid: *Cantar de mio Cid*. Alberto Montaner (ed.). Barcelona: Crítica, 2000.

PCG: Alfonso X: *Primera Crónica General*. Ramón Menéndez Pidal (ed.). Madrid: Gredos, 1979³.

14th century

CL: Don Juan Manuel: *Conde Lucanor*. José Manuel Blecua (ed.). Madrid: Gredos, 1983.

[Draft version]

Vázquez Rozas, Victoria & Marcos García Salido. 2012. [A discourse-based analysis of object clitic doubling in Spanish](#). In Davidse, Kristin, Tine Breban, Lieselotte Brems y Tanja Mortelmans (eds.). *Grammaticalization and Language Change: New reflections*, 271–297. Amsterdam: John Benjamins.

15th century

DA: *Electronic texts and concordances of Andalusian documents (1324-1500) selected from the collection of The Hispanic Society of America*. Cynthia Kauffeld (ed.). Madison: HSMS, 1999.

Corbacho: Alfonso Martínez de Toledo: *Arcipreste de Talavera o Corbacho*. Digital edition based on Cristóbal Pérez Pastor's Edition, *Arcipreste de Talavera, Corbacho, o Reprobación del amor mundano*, Madrid, Sociedad de Bibliófilos Españoles, 1901.

Alicante: Biblioteca virtual Miguel de Cervantes, 2004.

Celestina: *Early "Celestina" electronic texts and concordances*. Ivy A. Corfis and John O'Neill (eds). Madison: HSMS, 1997.

16th century

Amadís: *The Electronic text and concordances of the "Amadís de Gaula", 1539 Seville Edition*. Steven R. Fondow (ed.). Madison: HSMS, 1999.

Lazarillo: Anonymous: *Vida de Lazarillo de Tormes y sus fortunas y adversidades*. Digital edition based on those of Burgos, Juan de Junta, 1554; Alcalá de Henares, Salzedo, 1554; Antwerp Martín Nucio, 1554 and Medina del Campo, Mateo & Francisco del Canto, 1554. Alicante: Biblioteca Virtual Miguel de Cervantes.

Moradas: Teresa de Jesús: *Las Moradas*. Alicante: Biblioteca Virtual Miguel de Cervantes, 2000.

17th century

Quijote: Miguel de Cervantes: *Segunda parte del ingenioso hidalgo don Quijote de la Mancha*. Francisco Rico (ed.). Instituto Cervantes, 1998.

Dorotea: Lope de Vega: *La Dorotea*. Edwin S. Morby (ed.). Madrid: Castalia, 1968.

[Draft version]

Vázquez Rozas, Victoria & Marcos García Salido. 2012. [A discourse-based analysis of object clitic doubling in Spanish](#). In Davidse, Kristin, Tine Breban, Lieselotte Brems y Tanja Mortelmans (eds.). *Grammaticalization and Language Change: New reflections*, 271–297. Amsterdam: John Benjamins.

NE: Alonso de Castillo Solórzano: *Teresa de Manzanares o La niña de los embustes*. Madrid: Viuda de Rico, 1906.

Appendix 2: Complete references of the sources of contemporary examples

From the *Archivo de Textos Hispánicos de la Universidad de Santiago de Compostela (Arthus)*:

2VOZ: *La Voz de Galicia* (22/11/1991)

GLENDIA: Cortázar, Julio. 1981. *Queremos tanto a Glenda*. Alfaguara: Madrid (4th ed.).

HISTORIAS: Bioy Casares, Adolfo. 1986. *Historias desafortunadas*. Alianza: Madrid.

MADRID: Manuel Esgueva and Mariola Cantarero (eds.). 1981. *El habla de la ciudad de Madrid. Materiales para su estudio*. Madrid: CSIC.

SEVILLA: Pineda, M. A. de (ed.). 1983. *Sociolingüística andaluza 2. Materiales de encuestas para el estudio del habla urbana culta de Sevilla*. Sevilla: Universidad de Sevilla.

SONRISA: José Luis Sampedro. 1985. *La sonrisa etrusca*. Madrid: Alfaguara.

Conversational discourse:

VALESCO: Briz, Antonio and Val.Es.Co Group. 2002. *Corpus de conversaciones coloquiales*. Madrid: Arco/Libros.