Language Variation and Change, **21** (2009), 381–412.
© Cambridge University Press, 2009 0954-3945/09 $16.00
doi:10.1017/S0954394509990111

Cross-dialectal variation in propositional anaphora: Null objects and propositional *lo* in Mexican and Peninsular Spanish

Asela Reig Alamillo
LIDILEM–Université Stendhal Grenoble III

ABSTRACT

Despite the interest in null direct objects in Spanish, the case of direct objects with propositional antecedents, which complement cognition and communication verbs, remains mostly uninvestigated. This article investigates, from a comparative variationist perspective, null direct objects with propositional antecedents, variably coded by the clitic *lo*, in Mexican and Peninsular Spanish. Variable rule analysis of six Spanish corpora reveals a big difference between the two dialects in the frequency of overt vs. null neuter pronoun yet shows that some of the linguistic constraints conditioning the variation are shared by both dialects (presence of a dative pronoun, type of antecedent, sentence type), suggesting that the null pronoun has the same grammatical role in both dialects. Some divergences in the conditioning of the null pronoun also emerge from the analysis and the sociodemographic information available suggests the existence of a change in progress in Mexican Spanish.

Spanish is traditionally described as a non-object-drop language, one which encodes anaphoric direct objects (DOs) by means of overt pronominal elements, such as the clitics *lo(s)/la(s)*, or the demonstrative *esto/e/a, eso/e/a*, and *aquel/lla/llo*. Null objects (a.k.a. "object drop" or "null direct object pronominalization") in Spanish have been traditionally perceived as restricted to nonreferential and noncountable referents, that is, mass nouns and bare plurals (Campos, 1986; Clements, 1994, 2006). Nevertheless, null objects in Spanish have lately awakened the interest of several scholars who have shown that in Spanish contact dialects null objects with definite and specific antecedents are commonly found (1) (Paraguayan [Choi, 1998, 2000; Palacios Alcaine, 1998], Quiteño [Suñer & Yépez, 1988; Yépez, 1986], and Basque Spanish [Eguía, 2002; Landa, 1995]); in the Spanish of River Plate, a noncontact dialect, null objects are also said to be possible if the referent is recoverable from the immediate context of utterance (2) (Masullo, 2003).

(1) *¿Sabes si Pedro trajo los libros?*
Sí, Ø trajo. (from Palacios Alcaine, 1998:432)

I am grateful to Scott Schwenter, Terrell Morgan, Rebeka Campos-Astorkiza, and three anonymous referees for their insightful comments on previous versions of this article. I alone am responsible for any remaining shortcomings.

https://doi.org/10.1017/S0954394509990111 Published online by Cambridge University Press

'Do you know if Pedro brought the books?'
'Yes, he brought Ø.'

(2) *¿Le Ø retiro, señor?* (from Masullo, 2003)
'Shall I take Ø for you?' (e.g., a tray)

The null pronoun is also found in monolingual varieties of Spanish when the antecedent of the DO pronoun is a proposition instead of a noun phrase (NP) and the hosting verb is a cognition or communication verb (Schwenter, 2006:27). This case of variation between the null pronoun and the clitic *lo* is illustrated in (3) and (4):

(3) *Bueno, mi trabajo consiste en arreglar za … en … reparación de zapatos; ya le Ø dije anteriormente*. (Mex Cult)
'Well, my job consists of repairing … repairing shoes; I told you Ø before.'

(4) *Yo creo que a las cuatro, yo es que a las cuatro me es imposible venir. Ya se lo dije a él*; (Mad Cult)
'I think that at four, for me, at four it is impossible to come. I already told him LO.'

The null pronoun in (3) and the clitic *lo* in (4), referring to "my job consists of repairing shoes" and "at four it is impossible for me to come," respectively, are two distinct concrete realizations of a linguistic variable, the anaphoric DO with a propositional antecedent. They are equivalent in referential meaning. The propositional *lo* in (3) and the null pronoun in (4) would not change the referential meaning of the sentences.

Although this phenomenon has been mostly ignored until now (with the exception of Reig Alamillo, 2007, 2008; Reig Alamillo & Schwenter, 2007), variation between the two pronominal forms seems to exist in every dialect of Spanish. The distribution and productivity of the null pronoun, however, are at first glance different among dialects, and the similarities and differences between dialects have never been analyzed in the literature.

In this article, I will compare the distribution of the propositional clitic *lo* and the null pronoun in two dialects of Spanish, Mexican and Peninsular Spanish, which show an important difference in the frequency of the two variants. The null object is much more frequent than the clitic *lo* in Mexican Spanish, whereas the presence of the clitic *lo* is clearly preferred in the Peninsular variety. This difference makes it interesting to compare the two dialects, because the factors that condition the distribution of the two forms remain, until now, unexplored. The purpose of this article is to identify the constraints on the variation between the DO pronoun *lo* and the null pronoun with propositional antecedents in Peninsular and Mexican Spanish using the variationist comparative method (Poplack & Tagliamonte, 2001).[1]

https://doi.org/10.1017/S0954394509990111 Published online by Cambridge University Press

THE VARIANTS: PROPOSITIONAL *LO* AND NULL PRONOUN

The null pronoun studied in this article has the following characteristics, which differentiate it from other phonetically null elements in Spanish. First, it is anaphoric, that is, it refers to an element that can be retrieved from the discourse context, and the entity referred to by means of the phonetically null pronoun—the proposition—can be the object of subsequent references, unlike contexts such as (5), where it is not clear whether the verb *saber* has a DO or what its content might be.

(5) *Yo tengo un cuñado también ingeniero y, sin embargo, pues es un hombre bastante preparado, con bastantes inquietudes, un hombre que opina, un hombre que sabe, pero, es un poco excepción a la regla.* (Mad Cult)
'I have a brother in law who is also an engineer and, nevertheless, he is a pretty well qualified man, with interests on the side, a man who has opinions, a man who knows, but he is the exception to the rule.'

Second, the content of the null pronoun is propositional and would be expressed, if it were explicitly stated, by means of a sentence as opposed to a noun phrase, and specifically by means of a finite clause and not an infinitival complement, unlike null complement anaphora (6) (Brucart, 1999; Depiante, 2001).

(6) *Luis fue al acto; María, en cambio, no pudo (ir).* (from Brucart, 1999:2838)
'Luis attended the event; Maria, on the other hand, couldn't (go).'

The third characteristic is that this null pronoun is the DO of a cognition or communication verb (and not a modal or aspectual verb, as in null complement anaphora), and the verb is explicit in the discourse (as opposed to verb phrase ellipsis, in (7)).

(7) *Luis gana mucho dinero y María también.* (from Brucart, 1999:2822)
'Luis makes a lot of money, and so does Maria.'

Finally, this null object could also be coded in Spanish with the clitic pronoun *lo* (which also differentiates it from null complement anaphora as studied in Brucart, 1999; Depiante, 2001).

The pronoun that is in variation with this null object will be called "propositional *lo*" to differentiate it from other so-called neuter *lo* forms that are not considered in this study. The notion of a neuter *lo* is commonly found in traditional Spanish grammars (Bello & Cuervo, 1945:119; Gili y Gaya, 1964:237) to refer to the *lo* occurring in a range of functions and syntactic positions—DO pronoun, pronoun in a cleft construction, preceding an adjective or adjective phrase and a predicate attribute—even though this terminology has not been unanimously accepted (Bosque & Moreno, 1990; Leonetti, 1999;

https://doi.org/10.1017/S0954394509990111 Published online by Cambridge University Press

Ojeda, 1984; Otheguy, 1978). In this article, I will avoid using "neuter *lo*," in agreement with the observation that such terminology is inappropriate in light of the fact that there is no morphological neuter gender in Spanish. I will instead use "propositional *lo*," which allows me to distinguish the DO pronoun *lo* studied here from the so-called neuter *lo* in different syntactic constructions.

PROPOSITIONAL NULL OBJECTS IN SPANISH AND OTHER LANGUAGES

Even though the phenomenon of null objects with propositional antecedents has been rarely studied in depth in the literature, it has nevertheless been attested in some descriptions of different Spanish varieties. Kany (1945:146) affirmed that *lo* with cognition and communication verbs is frequently omitted in American Spanish, almost always when there is an explicit indirect object pronoun, due to the tendency to avoid double clitics. Solé & Solé (1977:41) also pointed out that the DO clitic is often omitted in Spanish with verbs such as *decir* (to say), *preguntar* (to ask), and *pedir* (to ask for) if an indirect object pronoun is present, as in *Pregúntale* (Ask him/her) or *Le diré* (I will tell him/her). Null pronouns with propositional antecedents can be found in examples included in studies that observe the use of null objects referring to definite NPs, mainly as the result of language contact (the Spanish of Ecuador [Súñer & Yépez, 1988; Toscano, 1953; Yépez, 1986], Paraguayan Spanish [Palacios Alcaine, 2000] and Basque Spanish [Landa, 1995], but this subtype of objects is always included under the category of "inanimate" and no distinction is made between inanimate entities introduced into the discourse by NPs and those introduced by propositions.

The use of null pronouns to refer to propositions has not been studied in Spanish in spite of the fact that propositions seem to be more likely referred to with a null pronoun than first-order entities in other languages. Meyer-Lübke (1923) mentioned the use of the null pronoun that refers to an idea already introduced in the discourse in Italian, Romanian, Portuguese, Patois, and French, especially when there is a co-occurring dative pronoun, and stated that this null pronoun was not unusual in Latin (Meyer-Lübke, 1923:417). More recently, several studies have pointed out that the use of null objects, especially with propositional antecedents, is common in spoken French (8) (Fónagy, 1985; Lambrecht & Lemoine, 1996; Larjavaara, 2000; Noailly, 1997).

(8) [*La femme de la victime au commissaire qui l'interroge sur sa liaison avec le secrétaire de son mari*]: *Mon mari savait Ø* [= *que nous avons eu une liaison*]. (from Fónagy, 1985:6):
'[The victim's wife to the police captain who questions her about her affair with her husband's secretary]: My husband knew Ø [that we had an affair].'

Propositional null objects co-occur with dative pronouns in first and second person—and not only in third person, unlike null objects referring to concrete

https://doi.org/10.1017/S0954394509990111 Published online by Cambridge University Press

entities. Their distribution is, therefore, less restricted than null objects with first-order referents, and they cannot be explained as a case of haplology (Armary, 1997:379; Yaguello, 1998:270).

In Brazilian Portuguese (BP), the propositional objects were crucial in the development of the null object system that BP has today, because this was the context in which null objects first became more frequent. In the 18th century, 46.3% of DOs with sentential antecedents were realized as zeros, whereas only 7.5% of the DOs with a specific NP and 6.1% of the DOs with a nonspecific NP were coded as zeros. In the 19th century, DOs of propositional verbs such as *saber* or *dizer* were realized as zeros 87.1% of the time and only in 12.9% of the cases was the neuter clitic *o* used (Cyrino, 1997:250).

Iliescu (1988) noted in Romanian the potential lack of an overt anaphoric DO when the neuter pronoun would be expected to refer to an idea, a fact, or an event. According to her observations, the use of the null pronoun is more frequent with verbs that can be used transitively and as "absolute verbs" (i.e., *to know, to learn, to imagine, to understand*).

Outside of the Romance languages, Meyerhoff studied the occurrence of null objects in Bislama and found that only inanimate entities favor the use of a null object (whereas animate referents disfavor it) and, specifically, the referent type "event or proposition" has a very high favoring effect on the null object (.935 on the Goldvarb results) (Meyerhoff, 2002:333).[2] In English, Fillmore (1986) presented the null anaphor as being lexically determined. Some verbs, with some of their meanings, accept null objects whereas similar verbs or the same verb with different meanings require an overt object (I forgot that she had fixed it: I forgot vs. I forgot my keys: *I forgot), and Cornish (2006) pointed out that the verbs that allow a null pronoun in Fillmore's list frequently have a propositional complement.

(9) John saw the "No Entry" sign. But Bill didn't see it/*Ø
(10) A: You'll have to wait till Monday, sir. The Council offices are closed today.
B: I see Ø/*it

To summarize, the aforementioned references suggest that the difference between DOs referring to noun phrases and those referring to clausally introduced entities is a pertinent one in many languages, at least regarding DO realization as a null or overt pronoun. Nevertheless, the choice of anaphoric forms taking into consideration this difference in the referent has not, to the best of my knowledge, been thoroughly studied in any language so far.

METHODOLOGY

The linguistic material analyzed in this study consists of naturally occurring data extracted from six Spanish corpora: for Mexican Spanish, Habla Culta de México (Mex Cult, complete corpus, approximately 167,000 words), Habla

https://doi.org/10.1017/S0954394509990111 Published online by Cambridge University Press

Popular de México (Mex Pop, complete corpus, approximately 172,000 words) (Lope Blanch, 1971, 1976) and 120 interviews—chosen by a system of sampling based on quotas of sex, age, and level of education—from Habla de Monterrey (Rodríguez Alfano, 2004). For Peninsular Spanish, the data comes from the Habla Culta de Madrid (Mad Cult, complete corpus, over 140,000 words) (Esgueva & Cantarero, 1981), the "conversation" and "debate" sections (approximately 404,000 words) of the Corpus de Referencia de la Lengua Española Contemporánea: Corpus Oral Peninsular, género conversacional (COREC) (Marcos Marín, 1992), and the first three sections of Alicante-Corpus Oral del Español (ALCORE), Huertas-San Juan, Tómbola-San Agustín, and Pla-Garbinet, which add up to over 260,000 words (Azorín, 2002). All are corpora of non-task-oriented spoken language and are intended to be representative of natural speech.

The envelope of variation was determined inductively. I first searched for the overt clitic pronoun *lo* and then for the verbs that appeared with this clitic in the corpora, noting whether *lo* was present or not in all occurrences of the same verb with anaphoric reference to a previously mentioned proposition. This method ensures that each verb included in the analysis allows both variants as its DO and allows for a more replicable procedure. With this methodology, the analysis includes the DOs of a limited group of cognition and communication verbs: *decir* (to say), *entender, comprender* (to understand), *saber* (to know), *contar* and, in Mexico, *platicar* (to tell), *explicar* (to explain), *imaginar* (to imagine), and *preguntar* (to ask).

Contexts where variation is not possible (one of the variants is either necessarily present or absent), as well as other contexts where both anaphors seem to be acceptable but which are considered distinct from the context of variation defined here were excluded from the analysis. These include cataphoric uses of the pronouns, topicalization/left dislocation (*Eso me lo dijo mamá* [That my mom told me that] or *No, que eso no* digas [No, that don't tell]), impersonal and passive *se* sentences (*no se sabe* [It is not known]), duplicated DOs (*Lo vamos a preguntar, si se puede* [We are going to ask it, if it's possible ('it' coreferential with 'if it's possible')]), and null complement anaphora with the verb *saber* meaning 'be able to'. Excluded also were fixed or idiomatic expressions in which the object is necessarily present or absent. The decision of whether an expression should be excluded from the envelope of variation was based on whether that expression showed variation in the corpora.[3] For example, although it should not be impossible to find in language cases of *¡No me lo digas!* (Don't tell me LO), the fact that the construction *¡No me digas!* (Don't tell me) is perceived as an idiomatic expression, and, crucially, the fact that the expression with the overt pronoun *¡no me lo digas!* was never documented in any of the six corpora analyzed led me to exclude it as a site of potential variation. With this criterion, the following expressions were excluded: *¿comprende(s)?, ¿entiendes?* (do you understand?), *¿Cómo diré/ cómo diría?* (How should I say?), *no me diga(s)* (Don't tell me), *ya (te/le) digo* (I'm telling you), *por decirlo así* (to say it that way), *quién lo iba a decir* (who would have said), *¿sabe(s)?* (you know?),

https://doi.org/10.1017/S0954394509990111 Published online by Cambridge University Press

quién sabe (who knows), *digo* (I say) (discourse marker, Mexico), *no sé* (I don't know). It should be clarified that *no sé* was only excluded in those uses where there was no available identifiable referent of the object of *saber*, and these cases were generally uses of *no sé* that interrupt the discourse, frequently preceding a reformulation. The following example illustrates this use.

(11) *Luego, por otra parte, una cosa muy interesante para mí, desde mi punto de vista al menos, es el intercambio; porque ... eso es ... no sé, entonces sí que es vivir plenamente en un ambiente ... de otro país. Por tanto, la misma familia donde se esté pues ya ... hablando otro idioma. Y el ... mismo, no sé, el mismo instinto de ... de conservación pues hace ... hace intentar un poco más seriamente hablar el idioma.* (Mad Cult)
'Then, on the other hand, a very interesting thing for me, at least from my point of view, is the exchange; because … that's … I don't know, then it is really living in an environment … from another country. Therefore, the same family where one is speaks another language. And the same … I don't know, the same instinct of conservation makes one try more seriously to speak the language.'

The same decision was made regarding cases where the referent of the DO was not clearly identifiable in the discourse, that is, was not explicitly mentioned in the previous discourse and, therefore, required some inferential reasoning to be identified. By way of illustration, consider the following example.

(12) *Y al final me dicen: ... iba yo, claro, co ... ¿quien iba a ir?, iba yo por los resultados y me dicen: ≪ No, tienen que venir tus padres ≫. Dije: ≪ ¿Mis padres? Vamos, ¡si no lo saben!* (Mad Cult)
'And, in the end they say … I was going, of course, who else would go?, I was going to get the results and they say, "No, your parents have to come." I said 'My parents? Come on, they don't even know LO!'

Although in some cases it could be inferred from the discourse what the content of the DO is (for instance, in (12), based on the previous discourse, we can infer that what the parents did not know is that the girl went to have this test done), more than one option is often available. To avoid second-guessing and because several of the factors included in the quantitative analysis could not be tested if the antecedent is not clearly identifiable, such contexts were excluded as sites of potential variation. Following these exclusions, the number of cases in which the antecedent was not clearly identifiable was higher than the number of tokens that could be finally included in the analysis.

HYPOTHESES AND CODING SCHEME

A total of 1,324 tokens of the dependent variable were collected from the corpora—669 from Mexican Spanish and 655 from Peninsular Spanish. Each of these tokens was coded for the value of the linguistic variable (i.e., the presence or absence of the

https://doi.org/10.1017/S0954394509990111 Published online by Cambridge University Press

DO clitic pronoun *lo*) as well as for 11 linguistic and 2 (for Peninsular Spanish) or 3 (for Mexican Spanish) extralinguistic variables.

A strong hypothesis was drawn from previous literature on anaphora resolution—mainly anaphora to first-order entities (Ariel, 1994, 1996; Givón, 1983; Gundel, Hedberg, & Zacharski, 1993), but not exclusively (Borthen, Fretheim, & Gundel, 2003; Eckert & Strube, 2000; Gundel, Hegarty, & Borthen, 2003; Hegarty, 2003). These studies share the idea that shorter anaphoric expressions are used when the antecedent is highly accessible, and longer, more complex anaphoric expressions refer typically to less accessible antecedents. Based on this observation, the logical hypothesis to test is whether the two variants under study—the overt pronoun *lo* and the null object—are distributed in the data relative to the accessibility of the antecedent. To operationalize the notion of accessibility, three factor groups were used: referential distance (Givón, 1983); the number of times the proposition is referred to (Borthen et al., 2003; Gundel et al., 2003; Hegarty, 2003); and turn, that is, whether the last mention of the antecedent is in the same or in a different speech turn (Schegloff, 2007). Based on intuitions or observations expressed in previous studies dealing with null objects in different languages, I included the internal factors verb tense (Masullo, 2003), verb person (Noailly, 1997), polarity of the host sentence (Iliescu, 1988), sentence type (Noailly, 1997), and presence/absence of a dative pronoun (Kany, 1945; Lambrecht & Lemoine, 1996; Landa, 1995; Solé & Solé, 1977). Finally, the factor groups kind of antecedent, presence of the adverb *ya* (already) and presence of a manner adverbial were included based on my own observations of the data.

QUANTITATIVE RESULTS

Presented in Table 1 is the overall distribution of the variants in Mexican and Peninsular Spanish.

Anaphoric DOs referring to propositions are realized as the "canonical" clitic *lo* 70% of the time in Spain but only 17% of the time in Mexico, whereas the null variant is used 30% of the time in Spain and 83% of the time in Mexican Spanish. The results corroborate the initial observation that there is a highly significant difference between the dialects and, in fact, the rate of null DOs in Mexican Spanish is even higher than expected.

Only a small number of verbs admit the null pronoun in the corpora analyzed, and it seems to be the case that some cognition and communication verbs do not accept both variants (for example, *rogar* [to beg], *asumir* [to assume], *sugerir* [to suggest], or *decidir* [to decide]).[4] As Figure 1 shows, not all the verbs are equally represented in the data. Two verbs, *saber* and *decir*, are much more frequent than the remaining verbs and together account for 85% of the data in Spain and 86% in Mexico (Spain: *saber* 51%, *decir* 34%; Mexico: *saber* 45%, *decir* 41%), and the remaining 15% in Spain and 14% in Mexico is distributed over seven or eight verbs. However, in neither of the two dialects was a

https://doi.org/10.1017/S0954394509990111 Published online by Cambridge University Press

TABLE 1. *Overall frequency of* lo *and null DO*

	LO	Null DO
Spain (656)	70% (459)	30% (197)
Mexico (669)	17% (113)	83% (556)

Chi square = 380.359, $p < .0001$, $df = 1$.

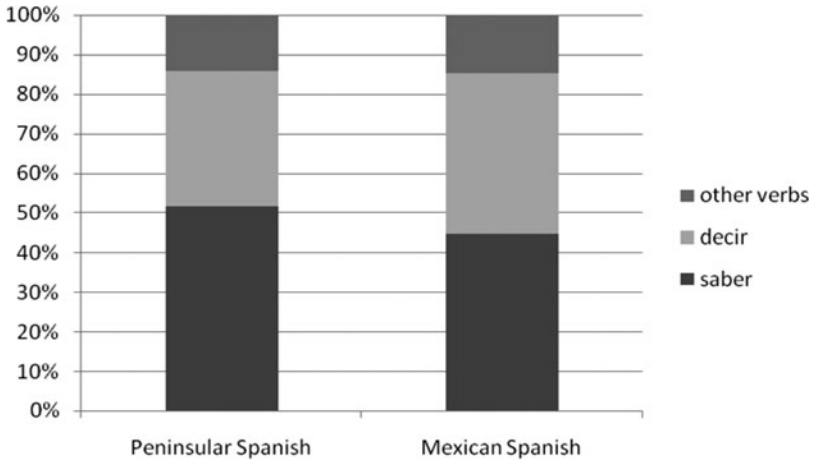


FIGURE 1. Distribution of verbs in Peninsular and Mexican data.

verb-type factor group distinguishing *saber*, *decir*, and other verbs selected as significant when included in the multivariate analyses.[5]

Tables 2, 3, and 4 show the results of two different multivariate analyses of the data from Mexico and Spain. In these tables, each factor group is presented in the leftmost column, and the individual factor values for each group are presented immediately below. The factor groups are presented in decreasing order of strength (as indicated by the range between the highest and lowest factor values for each group). Also included are the percentages of null DOs for each factor value, the total number of tokens per factor value, and the percentage of the data represented by each value. Statistically insignificant factor groups are listed at the bottom of the tables.[6] Although the internal and external factor groups selected as significant in Mexico are presented in different tables for the sake of simplicity, the analyses whose results are presented in Tables 2, 3, and 4, including both internal and external factor groups, in both dialects.[7]

The factor group selected with the greatest magnitude of effect in both dialects is dative pronoun, which included three factors: presence of a dative pronoun (dp), absence of a dative pronoun with a ditransitive verb (ditransitive w/o dp)— *decir* (to say), *contar*, *platicar* (to tell), *explicar* (to explain), *preguntar* (to ask)—and lack of a dative pronoun because of the monotransitivity of the verb

https://doi.org/10.1017/S0954394509990111 Published online by Cambridge University Press

TABLE 2. *Factors contributing to the choice of the null DO in Peninsular Spanish*

	Probability	Percent Ø	*n*	Percentage of data
Dative pronoun				
Monotransitive	.61	44%	368	56%
Dative pronoun (dp)	.54	19%	166	25%
Ditransitive w/o dp	.17	4%	122	19%
Range	44			
Type of antecedent				
Interrogative	.66	54%	220	33.5%
Declarative	.41	18%	436	66.5%
Range	25			
Polarity				
Negative	.64	50%	285	43%
Affirmative	.39	14%	371	57%
Range	25			

$N = 656$, input: 0.235 (30% null), log likelihood = –320.430, $p = 0.00$.
Factor groups not selected: referential distance, verb tense, verb person, sentence type, manner adverbial, age, and sex.

TABLE 3. *Internal factors contributing to the choice of the null DO in Mexican Spanish*

	Probability	Percent Ø	*n*	Percentage of data
Dative pronoun				
Monotransitive	.55	89%	338	51%
Dative pronoun (dp)	.54	84%	268	40%
Ditransitive w/o dp	.15	48%	63	9%
Range	40			
Manner adverbial				
Absent	.53	86%	623	93%
Present	.13	37%	46	7%
Range	40			
Type of antecedent				
Interrogative	.67	95%	368	55%
Declarative	.30	68%	301	45%
Range	37			
Sentence type				
Nondeclarative	.72	91%	81	12%
Declarative	.47	82%	588	88%
Range	25			
Polarity				
Negative	.59	93%	357	53%
Affirmative	.40	72%	312	47%
Range	19			

$N = 669$, input: 0.91 (83.1 % null), log Likelihood = -205.545, $p = .078$.
Factor groups not selected: referential distance, verb tense, verb person. Other factor groups included in analysis: speaker sex, age, education (see Table 4).

https://doi.org/10.1017/S0954394509990111 Published online by Cambridge University Press

TABLE 4. *Social factors contributing to the choice of the null DO in Mexican Spanish*

	Probability	Percent Ø	*n*	Percentage of data
Sex				
Women	.61	88%	391	60%
Men	.34	75%	265	40%
Range	27			
Age, years				
≤34	.62	87%	303	46%
≥55	.45	82%	123	19%
35–54	.37	78%	229	35%
Range	25			
Education				
Low	.58	87%	369	55%
High	.40	78%	299	45%
Range	18			

Other groups included in analysis: dative pronoun, manner adverbial, type of antecedent, sentence type, polarity, referential distance, verb tense, verb person (Table 3).

(monotransitive). Even though it is a common assumption that the overt clitic DO is absent exclusively or more frequently when it co-occurs with a dative pronoun (Grevisse, 1993; Kany, 1945; Lambrecht & Lemoine, 1996; Landa, 1995; Solé & Solé, 1977; Suñer & Yépez, 1988), marginal results show that null objects are above average in both dialects when the hosting verb is a monotransitive verb—in this study, *saber* (to know), *imaginar* (to imagine), *entender* and *comprender* (to understand) (89%, 301 of 338 in Mexico and 44%, 160 of 368 in Spain), and the multivariate analyses show that monotransitive verbs favor the null pronoun at least as much as the presence of a dative pronoun (.55 in Mexico and .61 in Spain). Example (13) illustrates a null object with a monotransitive verb.

(13) *Luego le decía a mi hermana: "Oyes, oyes, Agripina" dice- que ... este ... "¿Por dónde salió el sol, por dónde salió el sol?", dice. —"Pos quien sabe, mamá. Solamente Juana sabe Ø*". (Mex Pop)
'Then she would tell my sister: "listen, listen, Agripina," she said …, "where did the sun rise, where did the sun rise?," she said. —"Who knows, mom. Only Juana knows Ø."'

Taking into consideration only ditransitive verbs, the distribution of the variants would seem to agree with the observation that null objects are linked to the presence of a dative pronoun. In Spain, the null occurs 19% (32 of 166) with a co-occurring dative pronoun but only 4% (5 of 122) with ditransitive verbs that do not take a dative pronoun, and in Mexican Spanish, null objects are also more frequent when there is a dative pronoun (84%, 225 of 268) than when the ditransitive verb lacks one (48%, 30 of 63). The statistical analysis shows that a dative pronoun in the sentence only slightly favors the use of a null object (.54 in both

https://doi.org/10.1017/S0954394509990111 Published online by Cambridge University Press

dialects). Interestingly, when the propositional DO occurs with a ditransitive verb (such as *decir* or *preguntar*) that does *not* host a dative pronoun, the null object is highly disfavored in both corpora (.15 in Mexico and .17 in Spain). This is illustrated in (14).

(14) *le 'ije mira, para que se den cuenta que tú no trabajas and no hay necesidad que yo lo diga*. (Monterrey)
'I told him, look, for them to realize that you don't work it is not necessary that I say LO'.

The factor group manner adverbial was selected in the multivariate analysis of Mexico with a magnitude of effect similar to that of co-occurring dative pronoun and not selected as significant in the Spanish data, but the same direction of effect is found in both dialects (see Table 6). In Mexican Spanish, when manner adverbial is present in the sentence, only 37% (17 of 46) of the propositional DOs are coded as null pronouns. The multivariate analysis reveals that the presence of a manner adverbial strongly disfavors the null DO in Mexico (.13), but when no adverbial occurs, there is little effect on the variation (.53). This tendency to use an overt clitic when a manner adverbial expression co-occurs in the sentence, clearly found in the corpora even though the number of tokens is quite small, is illustrated in the following examples.

(15) *porque así le pu ... le puse: "Hermosa Luna." Porque hay muchas canciones relacionadas a la luna, ¿no? Tú lo sabes perfectamente* (Mex Pop)
'because I named it "Beautiful Moon." Because there are a lot of songs related to the moon, you know LO perfectly well'

(16) *La última vez que fui, llorando, llorando, me dijo que ya se habían ido las madres de ahí, de su colonia: llorando me lo dijo*, (Mex Cult)
'The last time I went, crying, crying she told me that the nuns had left the neighborhood, crying she told me LO'

In both (15) and (16), a manner adverbial (*perfectamente* and *llorando*) is present in the sentence hosting the propositional DO, and this context favors the use of the overt *lo*. Even though there are few tokens with a manner adverbial in the Mexican corpora (46), we find a wide variety of manner expressions[8] and no special collocations are identified in the corpora.

The factor group type of antecedent[9] was selected as significant in both dialects and the direction of effect within this factor group is also the same. In Mexican Spanish, null objects are used 95% (350 of 368) of the time when the antecedent is interrogative and 68% (206 of 301) with a declarative antecedent, and in Peninsular Spanish, we find 54% (118 of 220) of null objects with interrogative antecedents and 18% (79 of 436) with declarative antecedents. The Varbrul analysis reveals that the type of antecedent has a very similar effect in both dialects. With interrogative antecedents, the null

https://doi.org/10.1017/S0954394509990111 Published online by Cambridge University Press

object is favored in Mexican and Peninsular Spanish (.67 and .66, respectively) (17), and a declarative antecedent disfavors the null pronoun (.30 in Mexico and .41 in Spain) (18).

(17) A: *Yo es que, como no sabía dónde estaba ... tenía una idea, ¿no?, pero ...*
B: *Sí, pero no Ø sabías exactamente* (Mad Cult)
A: 'Since I didn't know where it was … I had an idea, but …'
B: 'yes, but you didn't know Ø exactly.'

(18) *Yo sabía que el cadáver aparecería. Yo lo sabía, porque …* (Corec)
'I knew that the body would turn up. I knew LO, because …'

The factor group polarity of the host sentence is selected as significant in both dialects with the same effect. Negative polarity of the host sentence favors null objects (.59 in Mexican Spanish [93%, 357] and .64 [50%, 285] in Peninsular Spanish), and affirmative polarity disfavors it (.40 [72%, 312] in Mexican and .39 [14%, 371] in Peninsular Spanish). However, the favoring effect of negation seems to be linked to a particular construction repeated in the corpora, namely *no sé* (I don't know). Table 5 shows that 66% of the null objects in Peninsular Spanish occur in the *no sé* construction, and only the remaining 34% ($n = 67$) are used in constructions other than *no sé*. In Mexican Spanish, the construction *no sé* accounts for 28% of the null objects in the data set, and 72% of the null objects are found in other constructions.

Furthermore, the 130 tokens of *no sé* in Peninsular Spanish account for most (90%) of the null objects in a negated hosting sentence (144 null objects) and, in fact, only 4 out of these 144 null objects occur in a negated sentence with a verb other than *saber* in Peninsular Spanish. In Mexican Spanish, the 153 tokens of *no sé* represent only 46% of the null objects in a negated hosting sentence (153 of 331). Finally, within this construction, the null object is almost categorical in Mexican Spanish (98% of *no sé* vs. 2% [3] of *no lo sé*), whereas in Peninsular Spanish we find 67% (130) of *no sé* and 33% (65) of *no lo sé*. To test whether the selection of some factor groups as statistically significant could be a consequence of a skewing effect in terms of this one construction, an analysis of the data excluding the *no (lo) sé* construction was conducted (see below).

Sentence type was selected as significant in the Mexican analysis and not selected in the analysis shown in Table 3 for Peninsular Spanish, but the direction of effect of the two factors is similar in both dialects, as shown in Table 6. The tokens were recoded as either declarative and nondeclarative sentences in order to capture more general tendencies in the data, and the marginal results show that when the host sentence is a declarative, 82% of the DOs are null in the Mexican corpora (482 of 588), and when the host sentence is a nondeclarative, the use of the null pronoun is higher, 91% (74 of 81). The multivariate analysis reveals that, in fact, declarative sentences slightly disfavor the use of the null pronoun (.47), but when the host sentence is a nondeclarative

https://doi.org/10.1017/S0954394509990111 Published online by Cambridge University Press

TABLE 5. *Null objects with* no sé *vs. other constructions*

	No sé	Other construction
Mexican Spanish (556)	28% (153)	72% (403)
Peninsular Spanish (197)	66% (130)	34% (67)

TABLE 6. *Internal factors contributing to the choice of the null DO in Mexican and Peninsular Spanish*

	Mexico	Spain
Dative pronoun		
Monotransitive	.55	.61
Dative pronoun (dp)	.54	.54
Ditransitive w/o dp	.15	.17
Range	40	44
Manner adverbial		
Absent	.53	[.51]
Present	.13	[.38]
Range	40	
Type of antecedent		
Interrogative	.67	.66
Declarative	.30	.41
Range	37	25
Sentence type		
Nondeclarative	.72	[.55]
Declarative	.47	[.49]
Range	25	
Polarity		
Negative	.59	.64
Affirmative	.40	.39
Range	19	25

Other groups included in analyses: referential distance, verb tense, verb person, speaker sex, age, education (Mexico).

sentence,[10] the null pronoun is clearly favored (.72) in Mexican Spanish. Example (19) illustrates this tendency.

(19) *No, porque andaban, supimos que andaban dando estos terrenos, vinimos y, ya aquí nos acomodamos y aquí nos quedamos*
¿Cómo Ø supo?, ¿o quién le Ø dijo?, ¿o cómo le Ø contaron? (Monterrey)
'No, because they were, we found out that they were giving these lands, we came, and here we made ourselves comfortable and here we stayed'
'How did you find out Ø? or who told Ø you? or how did they tell Ø you?'

https://doi.org/10.1017/S0954394509990111 Published online by Cambridge University Press

DIALECT COMPARISON AND THE EFFECT OF *'NO (LO) SÉ'*

In spite of the very different overall frequency of null objects in both dialects and the fact that more factor groups were selected as significant in the analysis of Mexican Spanish than in Peninsular Spanish, the results of the multivariate analyses show very interesting similarities between both dialects. These are summarized in Table 6.[11]

This table shows that the variation under study responds to similar patterns in both dialects. Three of the factor groups selected as significant coincide, which means that the factor groups that are playing a role in the distribution of the two variants in Peninsular Spanish also have an effect in the variation in Mexico. Moreover, the relative strength of these factor groups is also similar. The factor group dative pronoun is strongest (shows the largest range), sentence type follows, and polarity has the weakest effect (smallest range). Crucially, the constraint hierarchy within each of these factor groups is also parallel in both dialects, that is to say, the linguistic factors included in each of the groups have the same effects on the choice of the null object and these factors are ordered in the same way.

Regarding the vast difference in input (Spain: .235; Mexico: .91)—despite the similar grammatical patterning just described—it should be noted that, in comparison with Peninsular Spanish, Mexican Spanish has more tokens in contexts that favor the null pronoun (Tables 2 and 3). Ditransitive verbs with a dative pronoun (a context favoring the null object) account for 40% of the Mexican data vs. 25% of the Peninsular tokens, and ditransitive verbs without a dative pronoun (disfavoring null objects) are more frequent in the Peninsular corpora. This is not surprising given that the dative pronoun is used more in Mexico than in Spain, especially in the double dative construction (Company, 2006:542). Tables 2 and 3 also show that interrogative antecedents (a context favoring null objects) are more frequent in the Mexican than in the Peninsular data (55% and 33.5%, respectively). It is not entirely clear why the two data sets should have such different distributions, but a possible explanation could be the fact that one of the Peninsular corpora (COREC) did not consist of sociolinguistic interviews but of conversations, and fewer question-answer pairs were found in this corpus.

To test whether the results presented in Tables 2 and 3 are biased by the high occurrence of null objects in the construction *no sé* in either or both dialects, a multivariate analysis was conducted excluding all the tokens of *no sé* and *no lo sé* from the data. The overall distribution of the two variants in the new data set is shown in Table 7. Tables 8 and 9 show the results of the independent analyses of Mexican[12] and Peninsular Spanish[13] without the *no (lo) sé* tokens.

In these analyses, interestingly, the factor group polarity is no longer selected in either of the two dialects,[14] showing that its significance was epiphenomenal, due to the effect of the *no (lo) sé* tokens now excluded from the data.[15] Another important difference in this analysis is that the factor group type of antecedent is no longer selected in Peninsular Spanish, but it is selected in Mexican Spanish. This can be explained because, in the corpus, many of the interrogative

https://doi.org/10.1017/S0954394509990111 Published online by Cambridge University Press

TABLE 7. *Overall frequency of* lo *and null DO in data without* no (lo) sé

	Overt *lo*	Null DO
Spain (461)	85.5% (394)	14.5% (67)
Mexico (513)	21% (110)	79% (403)

TABLE 8. Factors contributing to the choice of null pronoun in the data without *no (lo) sé* in Peninsular Spanish

	Probability	Percent Ø	*n*	Percentage of data
Dative pronoun				
Dative pronoun (dp)	.61	19	166	36%
Monotransitive	.60	17	173	37.5%
Ditransitive w/o dp	.23	4	122	26.5%
Range	38			
Sentence type				
Nondeclarative	.70	13	403	87%
Declarative	.47	28	58	13%
Range	23			

$N = 461$, log likelihood = –178.343, input = 0.12 (14.5%), $p = .008$.
Factor groups not selected: type of antecedent and polarity.

antecedents are licensing the DO of a *no (lo) sé* construction, and when those tokens are excluded, the number of interrogative antecedents in the Peninsular data is reduced to 79. Although it was not selected as significant, the weights of the factors show the same tendency, or direction of effect, in this analysis of the Peninsular data: interrogative antecedent: .61; declarative antecedent: .48.

There are other changes in the multivariate analysis when the *no (lo) sé* tokens are excluded. The factor group sentence type was selected as significant in Peninsular Spanish when the *no (lo) sé* tokens were excluded from the data; as in Mexican Spanish, nondeclarative host sentences clearly favor the null pronoun, whereas declarative sentences slightly disfavor it. Finally, the factor group dative pronoun now shows the third greatest magnitude of effect and the constraint hierarchy within this factor group also changes. The presence of a dative pronoun favors the null object more than a monotransitive verb (which would have included *no (lo) sé* tokens) in both dialects.

INTERNAL CONDITIONING OF THE VARIATION

The results just presented offer interesting insights into the distribution of the DO null pronoun and the propositional clitic *lo* and the remarks found in previous literature on null objects.

https://doi.org/10.1017/S0954394509990111 Published online by Cambridge University Press

TABLE 9. Factors contributing to the choice of null pronoun in the data without *no (lo) sé* in Mexican Spanish

	Probability	Percent Ø	*n*	Percentage of data
Manner adverbial				
Absent	.54	83	471	92%
Present	.12	33	42	8%
Range	42			
Type of antecedent				
Interrogative	.72	93	222	43%
Declarative	.33	67	291	56%
Range	39			
Dative pronoun				
Dative pronoun (dp)	.57	84	268	52%
Monotransitive	.53	81	182	36%
Ditransitive w/o dp	.18	48	63	12%
Range	39			
Sentence type				
Nondeclarative	.72	91	81	16%
Declarative	.45	76	432	84%
Range	27			

$N = 513$, log likelihood = –190.867, input = 0.86 (79%), $p = 0.009$.
Factor groups not selected: referential distance, polarity, verb person.

The only internal factor repeatedly mentioned in previous studies on null objects in Spanish (and other languages) that would favor the use of a null object was the presence of a dative pronoun. In this study, the presence or absence of dative pronoun, now refined with the distinction between monotransitivity and ditransitivity of the verb, is revealed as an important factor contributing to the distribution of the two variants, but not in the way suggested in previous studies. The existence of a dative pronoun is not required for a null pronoun to occur, because monotransitive verbs (*entender, comprender* [to understand], *saber* [to know], *imaginar* [to imagine]) favor null objects in any case, and this is more clearly the case when all the tokens of *saber*, including *no sé*, are analyzed. Dative pronouns do favor null objects, but the most striking effect in all the analyses is that ditransitive verbs (*decir* [to say], *contar*, *platicar* [to tell], *explicar* [to explain], *preguntar* [to ask]) that do not occur with a dative pronoun strongly disfavor the null object. It looks like, regarding their phonetically overt arguments, the verbs with propositional objects observed here reduce their valency by one. When the verb is monotransitive, the only pronoun available to be zero is the DO; when the verb is ditransitive, only one of the two objects can be phonetically null. The DO is realized as zero as long as there is another object explicit in the sentence, that is, a dative pronoun. Yet, if there is not a second object in the sentence (the dative pronoun), the verb seems to require one explicit object, to maintain at least one of its two objects explicit in the sentence,

https://doi.org/10.1017/S0954394509990111 Published online by Cambridge University Press

reducing by one its expressed arguments, but not by two. The propositional DO is coded then as an overt pronoun and the null object is highly disfavored in this case.[16]

The results of the factor group manner adverbial are interesting inasmuch as they suggest the existence of a pragmatic/discursive constraint on the variation, not very well defined at this point but that will be worth exploring in future research. Recall that in Mexican Spanish, the presence of a manner adverbial clearly disfavored the null object and that the same tendency was found in Peninsular Spanish, although it was not statistically significant. Looking closer at these results, and even though there are only 46 tokens of manner adverbials in the Mexican data, it seems to be the case that the presence of a manner adverbial is really affecting pronouns with declarative antecedents. From these 46 tokens, 36 (78%) occur when the antecedent is a declarative sentence, and when the antecedent is a declarative and there is a manner adverbial, only 9 of 36 objects are nulls (25%) and the *lo* becomes much more frequent than average in Mexico (75% vs. the 17% of overt *lo* in the complete data set in Mexico). When the antecedent is an interrogative sentence, even if there is a manner adverbial, the null is still preferred in the data that we have (80%), but the number of object pronouns with interrogative antecedents co-occurring with a manner adverbial is too low ($n = 10$) to draw any conclusions. The disfavoring effect of the manner adverbial on null DOs seems to be indicating that a modifying element (here, the manner adverbial) needs to have some overt linguistic material to modify (here, the DO is part of the modified predicate).[17]

A last observation regarding the distribution of the manner adverbials in the data has to do with the presence or absence of the adverb *ya* (already), excluded from the analyses in Tables 2 and 3 because it was not orthogonal with the factor group modal adverbial. There were no sentences in which the propositional DO co-occurred with both a manner adverbial and with the adverb *ya*.[18]

A new analysis was conducted in which both manner adverb and adverb *ya* were combined. Like the factor group manner adverbial, this new factor group was not selected as significant in the Peninsular analysis but was selected in Mexican Spanish, with the results shown in Table 10.[19] The presence of *ya* clearly favors the null object, as exemplified in (20).

(20) *Pero yo no sabía dónde vivías. Ahora ya Ø sé.*
'But I didn't know where you lived. Now I know Ø'

Out of the 63 occurrences of *ya* in Mexican Spanish, 8 occur in the collocation *ya le digo* (I'm telling you), 11 in *ya me imagino* (I [can] imagine), and 6 in *ya sé* (I know). Although we have very few tokens of both manner adverbials ($N = 46$) and adverb *ya* ($N = 63$), the fact that both do not co-occur in the same sentence is of interest and supports the intuition that the adverb *ya* and the manner adverbials have an opposite role in the discourse. The adverb *ya* tends to co-occur with telic predicates and, therefore, to focus the verbal action (Torres Cacoullos &

https://doi.org/10.1017/S0954394509990111 Published online by Cambridge University Press

TABLE 10. Results of the factor group adverb in Mexican Spanish

	Probability	Percent Ø	*n*	Percentage of data
Adverb				
Ya	.76	90%	63	9%
No manner/ya adverbial	.50	86%	560	84%
Manner adverbial	.13	37%	46	7%
Range	63			

Log likelihood = –202.475, $p = .007$.
Also selected as significant: dative pronoun, type of antecedent, sentence type, polarity.

Schwenter, 2008), and seems to defocus the rest of the predicate, in this case the (propositional) object. As García & Portero (2002:12) pointed out, it is commonly the case that when the focus of the sentence is turned to the verbal process itself, the object is more likely to be omitted and, in the case of propositional objects in Spanish, the presence of the adverb *ya* has precisely the effect of favoring a null object. On the other hand, in sentences with a manner adverb, the (propositional) DO seems to be part of the focused element, because the whole predicate is being modified by the manner adverbial, and therefore the object tends to merit pronominal remention and the overt clitic *lo* is favored in the statistical analysis.[20]

The effect of the factor group kind of antecedent is also worth discussing. Because the distribution of null and overt pronouns referring to first-order entities in object and subject position in Spanish is clearly constrained by certain semantic characteristics of the referent, such as animacy, person, or definiteness, the factor group type of antecedent was included in the statistical analysis to test whether the semantics of the referent would also condition the use of the null pronoun and the propositional *lo*. The results in Tables 6 and 9 show that the distribution of the two variants is constrained by the characteristics of the proposition that is introduced into the discourse by a declarative vs. an interrogative sentence. A declarative sentence typically introduces in the discourse a complete, saturated proposition. An interrogative sentence, on the other hand, can be described as an incomplete proposition (Peterson, 1997:39). The denotation of a question has been described as an incomplete or open proposition or as a set of propositions, corresponding to the set of its possible (true) answers (Hamblin, 1973; Karttunen, 1976). The "completeness of the proposition," understood as the existence or availability in the discourse of all the elements needed to verify the truth conditions of the proposition, would be an important constraint on the distribution of the two anaphors. In both dialects, the null pronoun is favored when the antecedent is an interrogative sentence, that is, when the referent of the pronoun is an incomplete proposition[21] (21). On the other hand, declarative antecedents in both dialects disfavor the null pronoun (22), which can be explained by saying that when the referent of the DO pronoun is a complete, saturated sentence, the null is disfavored and *lo* is favored.[22]

https://doi.org/10.1017/S0954394509990111 Published online by Cambridge University Press

(21) *¿Cuál es el [nombre] del calendario? ¿No me Ø dices*? (Mex Pop)
'What is the (saint's) name from the calendar? You don't tell me Ø?'

(22) A: *Es que yo quiero que venga un maestro*
B: *¿Y hasta ahorita me lo dices*? (Mex Cult)
A: 'I want a teacher to come'
B: 'And you tell me LO now?'

Acceptability judgments suggest that when the antecedent is an incomplete proposition, the null pronoun is accepted by Peninsular speakers and preferred by Mexican speakers, whereas the overt clitic *lo* in Mexico is less accepted than in contexts in which the anaphoric DO refers to a complete proposition (Reig Alamillo, 2008). In sentences that would be excluded from the statistical analysis because the anaphoric DO has two potential referents in the discourse, one incomplete and one complete proposition, the overt propositional *lo* is preferred to refer to the complete proposition and the null object is preferably interpreted as the incomplete proposition:

(23) *Nos preguntaron qué casa preferíamos y yo no Ø sabía /y yo no lo sabía*
'They asked us which house we preferred, and I didn't know Ø / I didn't know LO.'

In (23), the null refers to the interrogative antecedent, *yo no sabía qué casa preferíamos* (I didn't know which house we preferred), whereas the preferred interpretation of *lo* is *yo no sabía que nos lo habían preguntado* (I didn't know they had asked us that), referring to a complete proposition introduced in the previous discourse, for the Mexican and Peninsular speakers surveyed.

Regarding lexical effects on the variation, it was previously mentioned that the factor group verb type was not selected as significant when included in the analysis, indicating that it is not the lexical type that is affecting the variation. The question that arises is whether any particular configurations of person/tense or specific collocations are affecting the variation. Interestingly, in the data analyzed such particular configurations tend to occur with or without *lo* in an almost categorical manner. Recall that there were several fixed constructions that were excluded from the data because the presence or absence of *lo* was categorical and they could be considered discourse markers based on other criteria. In the remaining data, some candidates for collocations—not as clearly classified as "fixed expressions" or discourse markers due to a greater variability in their form and lack of a clear pragmatic function—can be noticed: in Spain, 23 *ya lo sé* (I know LO) and only 1 *ya sé* (I know) are used, whereas in Mexico, we find 6 *ya sé* and no tokens of *ya lo sé*; in Spain, *lo* is always present in *no lo entiendo* (I don't understand LO) (8) and *me lo imagino* (I can imagine LO) (4), and in Mexico we don't find the overt pronoun in *me lo imagino*, but we find 13 tokens of *(ya) me imagino*, and, more importantly, 59 tokens of *no saber decir* (cannot tell).[23] Nevertheless, the analysis shows that, even though the null object is more frequent than average in the *no sé* construction, the variation between the null

https://doi.org/10.1017/S0954394509990111 Published online by Cambridge University Press

object and the propositional *lo* is not restricted to fixed expressions or collocations, but grammatically active in both dialects (cf. Torres Cacoullos & Schwenter, 2008; Torres Cacoullos & Walker, 2009).

PROPOSITIONAL ANAPHORA AND ACCESSIBILITY

One of the questions posed in this study is whether the accessibility of the antecedent would condition the variation. The theories dealing with anaphoric distribution share the idea that anaphors with more phonetic, morphological, and semantic content are typically used to refer to less accessible entities, and vice versa (Ariel, 1988, 1990; Fox, 1987; Givón, 1983; Gundel et al., 1993), but the results of the quantitative analysis show that the variation of the null and the propositional *lo* is not conditioned by the factor groups included to operationalize the idea of accessibility.

One of these factor groups is referential distance, proposed by Givón (1983) to measure topic continuity. According to Givón, the more continuous a topic is, the more accessible the antecedent and, therefore, the less complex the anaphoric expression. This measure was calculated counting the distance from the target anaphoric expression (*lo* or zero) to the most recent prior mention of the same referent, counting the number of clauses and (arbitrarily) establishing an upper limit of 10 clauses, a distance quite larger than what is expected to be found in propositional anaphora according to previous literature (Dahl & Hellman, 1995; Schiffman, 1985).

Figure 2 shows the percentages of null pronouns according to the number of clauses since the last mention of the proposition in the discourse (1, 2, 3, and 4 or more sentences, along the *x* axis) in Mexican and Peninsular Spanish. As Figure 2 shows, the overall rate of null pronouns is very different but the gradation pattern is comparable across dialects and follows the prediction. As referential distance increases, the rates of null objects decrease in both Mexican and Peninsular Spanish. These results, however, only reflect a tendency in the data. The factor group referential distance[24] was included in the Varbrul analysis —collapsing the degrees of distance in two factors, one sentence and more than one sentence—and it was not selected as significant in any of the two dialects, but it would be interesting to observe the effect of this factor group in a larger sample.

Regarding the observation that higher-order entities[25] are available for remention only in the next sentence (Dahl & Hellman 1995; Schiffman, 1985), these data show that, even though reference to entities introduced two or more clauses before in the discourse is not impossible, most of the anaphoric uses of the *lo* and the null pronoun in both dialects refer to entities introduced in the immediately previous discourse, mainly in the previous sentence (85% in both dialects).

The second factor group included to measure accessibility was number of times the proposition is referred to in the discourse. According to some scholars, for a

https://doi.org/10.1017/S0954394509990111 Published online by Cambridge University Press

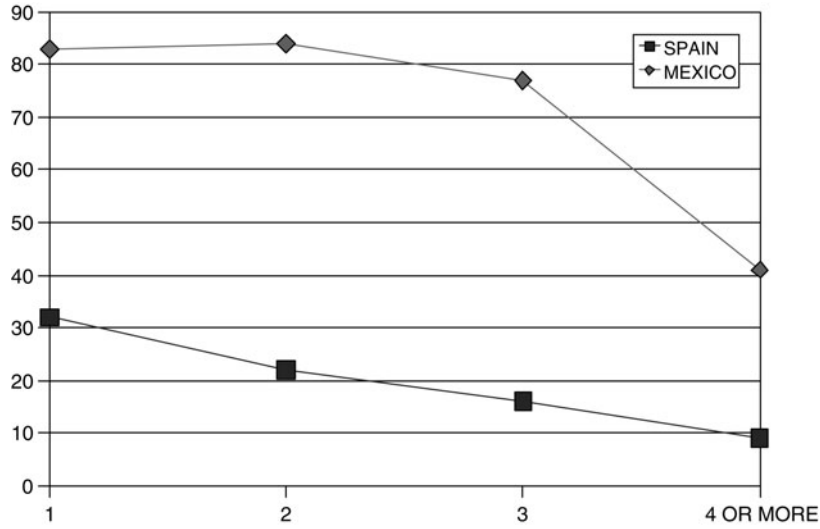


FIGURE 2. Percentages of null object by referential distance (number of sentences).

higher-order entity to be "in focus" in Gundel et al.'s (1993) scale and, thus, available for reference with a pronoun, it has to be referred to at least twice in the discourse (Borthen et al., 2003; Gundel et al., 2003; Hegarty, 2003). The explanation provided by these scholars for the distribution of English demonstratives and the pronoun "it" referring to higher-order entities presents the number of times an entity has been referred to as a measurement of the accessibility of the antecedent, suggesting that the more times an entity is referred to, the more accessible it should be for the speaker.

In order to incorporate and test this hypothesis, I coded for the number of times that the proposition is referred to in the previous discourse establishing an arbitrary limit of a maximum of 10 times. The coding of this factor group is exemplified in (24).

(24) A: *El lunes ya clase normal, ¿no?, según lo que decían ...*
B: *Bueno*, ***eso*** *dicen, pero yo no lo sé seguro;* (Mad Cult)
A: 'On Monday [we have] normal class, right? According to what they said …'
B: 'Well, so they say, but I don't know LO for sure;'

In (24), the target pronoun, *lo* in *no lo sé*, refers to the sentence *"el lunes ya [hay] clase normal,"* and this proposition is referred to twice in the discourse: by the whole sentence, *"el lunes ya [hay] clase normal"* and by the demonstrative pronoun *eso*.

The prediction that the null object would be linked to a higher number of mentions of the proposition than the overt *lo* is not met. The null pronoun is not more often used with entities repeatedly referred to in the discourse; on the

https://doi.org/10.1017/S0954394509990111 Published online by Cambridge University Press

contrary, the rate of null pronouns is higher in Spain for the first anaphoric mention (33%, 168 of 504) than it is when the referent has been previously mentioned twice (22%, 24 of 108) or more than twice (11%, 5 of 44). In Mexico, the difference between first and second anaphoric mention is very small regarding the rate of null pronoun (83% [424 of 510] for propositions mentioned once and 86% [94 of 109] for propositions referred to twice), but with entities mentioned three times or more, the rate of null pronoun is lower (76%, 38 of 50). The hypothesis that the distribution of the variants under study could be constrained by the number of times the proposition has been referred to is, thus, not supported by the data in either of the two dialects.

Finally, the factor group turn was included under the hypothesis that entities previously mentioned by the speaker would be more accessible and, accordingly, were expected to be more frequently coded as null. The data analyzed contradict this prediction. In both varieties, the null pronoun is more frequent when the last mention of the proposition was made by another speaker, in a different turn (36%, 141 of 388, in Spain and 93%, 334 of 360, in Mexico) than when it was made in the same turn (21%, 56 of 268, in Spain and 72%, 222 of 309, in Mexico) (Spain, chi square: 17.995, $p < .001$; Mexico, chi square: 51.904, $p < .001$). These results, and the possible interpretation that different turn favors the null pronoun in both dialects, should be taken cautiously because this factor group was not included in the Varbrul analysis due to interactions with kind of antecedent. Most of the interrogative antecedents (which favor the null object) were also in a different turn and that could explain the high rate of null objects in a different turn. It cannot be concluded, with the data analyzed, what the effect of turn might be on the variation under study here, but the analysis does not seem to support the hypothesis that the null object would be favored when the antecedent is last mentioned in the same turn, and, therefore, more accessible in the discourse than propositions introduced in different turns.

Maintaining the assumption that these were adequate measurements of accessibility, it should be concluded that the variation under study is not constrained by the accessibility of the antecedent. This finding is consistent with Maes (1996), who found that the theories of anaphora resolution (Ariel, 1988, 1990; Gundel et al., 1993) do not account for the distribution of the abstract personal pronoun *het* (it) and the demonstrative *dit/dat* (this/that) in Dutch. These results are, however, somehow unexpected because the idea of accessibility has been considered central in the distribution of anaphoric expressions and leave us with the question of what linguistic or discourse constraints explain this variation, and how would these be related, if at all, to the notion of accessibility.

Finally, it has been argued that the accessibility of an element is also linked to its semantic characteristics (Ariel, 1994:28) and, in this way, it is worth exploring whether the semantic feature proposed earlier, the completeness of the proposition, could be also linked to the accessibility of the antecedent in such a way that complete propositions could be seen as more accessible than incomplete propositions, in the same way that human entities are usually more

https://doi.org/10.1017/S0954394509990111 Published online by Cambridge University Press

accessible to the speaker and hearer than nonhuman entities.[26] Appealing as this hypothesis may be, if we assumed that completeness of the proposition is in fact a measure of accessibility in the way just explained, we would still have a disagreement with the main prediction of the anaphora resolution proposals. The analysis provided earlier shows that, in both Mexican and Peninsular Spanish, more accessible entities (complete propositions) are, in fact, not linked to the null pronoun but to the overt propositional *lo*, and less accessible discourse referents (here, incomplete propositions) favor the null pronoun instead of the overt clitic *lo*. That is to say, if the semantic content of the proposition (in this case, its completeness) indicates accessibility, the generalization that more accessible entities are coded with smaller or simpler anaphors does not hold for propositional anaphora, because more accessible, that is, complete, propositions have been shown to favor the overt pronoun *lo*, whereas incomplete, less accessible entities are more likely coded as zeros.

VARIATION AND LINGUISTIC CHANGE: SOCIO-DEMOGRAPHIC INFORMATION

A crucial aspect of the study of variation is the fact that synchronic variation has been proven to be, in many cases, a reflection of a change in progress. An appealing question about the variation between propositional *lo* and the null pronoun is whether it reflects a linguistic change in either or both dialects, that is, whether this variation is a stage in the movement from one linguistic state to another (Chambers, 2003:203). Although undeniable evidence that intervariety differences are actually the result of language change happening at different rates only comes from diachronic data, indication of a change in progress can be obtained from the analysis of sociodemographic constraints on the variation (Labov, 1994).

Some sociodemographic information was available in the corpora and was included in the statistical analysis. In the factor group age, speakers were distributed among three groups maintaining the division used in the *Habla Culta* project: $\leq$34 years old, 35–54 years old, and $\geq$55 years old. Regarding education, the information available allowed differentiating two educational groups, low education (informants in *Habla Popular* corpus and speakers from Monterrey who were illiterate, or who had primary education or unfinished secondary education) and high education (speakers from *Habla Culta* and speakers from the Monterrey corpus who had attended college).

Whereas none of the social factors were significant in the analysis of Peninsular Spanish, the results from Mexican Spanish suggest a change in progress (Table 4). First, women favor the null pronoun (.61) and men disfavor it (.34). These results, together with the fact that the null pronoun is not at all stigmatized in Mexican Spanish, nor are speakers conscious of the variation analyzed here, are of great interest, because it has repeatedly been shown that women are most often the innovators in linguistic changes from below, that is, changes that take place

https://doi.org/10.1017/S0954394509990111 Published online by Cambridge University Press

below the level of consciousness (Labov, 1990). The effect of women favoring the null pronoun is found across the data, independently of age and education of the speakers.

Turning to the factor group education, low education favors the null pronoun (.58) whereas high education disfavors it (.40). The cross-tabulation of sex and education shows a lower rate of null pronouns in highly educated speakers than in less educated speakers of both sexes (highly educated women: 81%, less educated women: 92%; chi square: 11.54, $p = .0006$; highly educated men: 73%, less educated men: 78%, chi square: .79, $p = .37$). The favoring effect of the null pronoun by less educated speakers can also be interpreted as indicating a linguistic change in progress, specifically a change from below: although the notion of "change from below" refers to the level of awareness of the speakers, it often correlates with a higher rate of the innovative variant in low educated speakers or speakers with a "low status in the social hierarchy" (Labov, 1966:128).

Finally, regarding age, the youngest generation favors the null pronoun (.62), the eldest generation slightly disfavors it (.45), and the generation between 35 and 55 years old clearly disfavors it (.37). These weights could be interpreted on first sight as suggesting a pattern of age grading (Chambers 1995:188; Chambers & Trudgill 1998:151), but it is necessary to note that the data are not well distributed according to age. We only have 123 tokens in the group of older speakers, and 85% of these tokens are also from low educated speakers. The cross-tabulation of the factor groups age and education shows a consistent pattern in low educated speakers in Mexico (first generation: 90%, second generation: 86%, third generation: 85%). Interestingly for the issue of language change, young speakers in both education groups show a high rate of null pronouns (90% for low educated speakers and 86% for high educated speakers), which can be interpreted as indicating that the null object is very widespread among young speakers, independently of their education, whereas the factor education is playing a clear role in the distribution of the two variants in older generations.

The sociolinguistic analysis just presented should be taken cautiously. The sociodemographic information that we possess is not very detailed, a greater number of speakers would be desirable, and information about other social indexes and style is not available in the corpora; nevertheless, these results suggest a change in progress in Mexican Spanish in the direction of losing the propositional *lo* in favor of the null pronoun. As for Peninsular Spanish, the social factors included in the analysis do not provide information that can lead us to infer that the same change in progress is taking place in this variety of Spanish.

The idea of a change in progress is very appealing because of the parallels with the change that took place in the DO system in Brazilian Portuguese. Null objects are now widespread in Brazilian Portuguese (Schwenter & Silva, 2003), and Cyrino (1997), analyzing diachronic data, shows that the first kind of DO that was coded as zero was the neuter *o* referring to propositional antecedents.

In order to test whether the direction of the change is, as it was in Brazilian Portuguese, toward the loss or limiting the use of *lo* in favor of the null pronoun,

https://doi.org/10.1017/S0954394509990111 Published online by Cambridge University Press

diachronic data would be needed. A preliminary search in the Mexican data in the *Corpus Diacrónico del Español* (CORDE)[27] of the most common verb forms that take the null pronoun or the clitic *lo* as its DO (*se/le _dije, te_dije, ya_sabía, _sabías*) shows only 3% (1/39) of null pronouns, and the results of the search of the same expressions in the 17th-century data of Spain in CORDE show no cases of null object (0 of 47), which suggests that the direction of the change is toward the progressive loss of the overt clitic pronoun *lo* in favor of the null object.

CONCLUSIONS

This study has established the widespread existence of null objects in monolingual varieties of Spanish, providing evidence that null objects with propositional antecedents are actually commonly found in two monolingual varieties of Spanish and that the rate of null objects in Mexican Spanish is actually strikingly higher than the use of the "canonical" form *lo*.

The comparison of the multivariate analyses shows that, in spite of the great overall difference in null pronoun rate in both dialects, the factors constraining the use of null objects in Peninsular Spanish also play a role in the distribution of the two variants in Mexico, and the relative strength of these factor groups and the constraint hierarchy within them are also parallel in both dialects, indicating that they share an underlying grammar in the use of propositional DO anaphors (cf. Tagliamonte, 2004:731).

Although the favoring effect on the null object of the construction *no (lo) sé* is stronger in the Peninsular results, and when these tokens are excluded the factor groups selected as significant slightly vary in this dialect, both variants are productive (within the group of verbs examined) in Mexican and Peninsular Spanish, and internal factors such as the mono- or ditransitivity of the verb and the presence or absence of a dative pronoun or the sentence type affect the variation. Other linguistic factor groups, namely type of antecedent and presence of a manner adverbial (and the adverb *ya*), selected only in Mexican Spanish but showing the same tendency in the Peninsular data, suggest that there might be semantic-pragmatic factors—the completeness of the proposition and information structure, respectively—playing a role in the choice of null objects with propositional antecedents.

An interesting case of divergence between the two varieties is that external conditioning is also affecting the variation in Mexican Spanish but not in the Peninsular corpora, and the sociodemographic data analyzed, although limited, suggest the existence of a change in progress. To verify this observation, a more detailed sociolinguistic investigation and a diachronic study of the null object and clitic *lo* with propositional antecedents would be of great interest.

Lastly, regarding theories of anaphora resolution, the data analyzed reveal that the notion of the accessibility of the antecedent, accepted as an accurate generalization able to explain the distribution of anaphors referring to first-order entities, does not account for the variation between the clitic *lo* and the null

https://doi.org/10.1017/S0954394509990111 Published online by Cambridge University Press

object referring to propositions. The phonetically zero pronoun under study does not correlate with more accessibility when this notion is operationalized as discourse measurements such as referential distance or turn, nor is it preferably used to refer to the most accessible entities if accessibility is determined by the semantic completeness of the entity.

These results, and the mere existence of propositional null objects in Mexican and Peninsular Spanish (and probably all Spanish varieties), suggest the importance of differentiating between first-order entities and higher-order entities in the study of phenomena that are normally analyzed using exclusively the distinction between animate and inanimate objects, such as the anaphoric representation of DOs but also anaphoric subjects and probably other pronoun-related phenomena, such as *leismo* or clitic doubling in Spanish.

NOTES

1. I will be using the generalizations Peninsular and Mexican Spanish throughout this article in order to differentiate these varieties from each other, but the data under Peninsular Spanish come from corpora from Madrid and Alicante, and the data under Mexican Spanish come from the speech of Mexico City and Monterrey (see Methodology section). No statistical difference was found when the data from Madrid were differentiated from the data from Alicante in the analysis of Peninsular Spanish nor when Mexico City and Monterrey were included as separate factors in the analysis of Mexican Spanish.

2. These quantitative results, however, should be taken cautiously because, as Meyerhoff explained, "since clausal complements were excluded from the corpus, it is not clear whether these anaphors should have been included as well" (Meyerhoff, 2002:333).

3. The characteristics that led me to consider them fixed or idiomatic expressions are their high frequency (especially in the speech of some speakers who repeatedly introduced certain expression in their discourse), the fact that they usually constitute a separate phonological or intonational phrase, the fixed morphosyntactic form, and their nonliteral meaning, which is usually a pragmatic value (epistemic markers, marks of reformulation, phatic, or connative expressions).

4. Interestingly, the same is true in the case of null complement anaphora. Some verbs accept this null anaphor and verbs that are semantically close do not (Brucart, 1999).

5. To further test whether these two verbs exhibit significantly different behavior, I conducted independent analyses with only the verbs *saber* and *decir* in each dialect. In Peninsular Spanish, type of antecedent was selected as significant in both data sets (*saber* and *decir* tokens) with the same direction of effect. Interrogatives favor null objects (*decir* .74; *saber* .60) and declaratives slightly disfavor it (*decir* .45; *saber* .40). Polarity of the sentence was selected for *saber* (negative polarity favoring null objects [.61] and positive polarity disfavoring [.24]), and the same direction of effect was found in the *decir* data set (.59 and .48, respectively), even though polarity was not selected as significant. Finally, a divergence was found in the factor group referential distance, selected as significant in both *saber* and *decir* data sets but with opposite effects: in *decir*, referential distance of one clause slightly disfavors null objects [.45] and referential distance greater than one clause favors it [.68], whereas in the *saber* data set, greater referential distance disfavors null objects [.29] and one-clause referential distance showed hardly no effect on the variation [.52]. In Mexican Spanish, the factor group kind of antecedent was selected for both verbs with the same direction of effect (interrogative antecedent favors the null object [.85 in *decir* and .62 in *saber*], but declarative antecedents disfavor null objects [.22 in *decir* and .19 in *saber*]). Also significant was the presence of a manner adverbial (.10 in *decir* and .15 in *saber*). The factor group sentence type showed the same direction of effect in both groups. Declarative sentences slightly disfavor the null object (.45 in both verbs) and nondeclaratives favor it (.79 in *decir* and .83 in *saber*), but it was only selected as significant in the *decir* data set. Finally, dative pronoun was selected as significant in the *decir* data set in Mexico and Spain but it could not be included in the *saber* analysis because all tokens of this verb are monotransitive (i.e., no dative pronoun).

6. Not selected as significant in the analyses were the following factor groups (factor weights from first stepping down). *Referential distance:* Spain: 1 clause: .51, more than 1 clause: .43; Mexico: 1 clause: .49; more than 1 clause: .53. *Verb tense:* Spain: present: .54; subjunctive: .50; preterit: .44; imperfect: .35; future/conditional: .21; Mexico: present: .45; subjunctive: .47; preterit: .61; imperfect:

https://doi.org/10.1017/S0954394509990111 Published online by Cambridge University Press

.72; future/conditional: .47. *Verb person*: Spain: first person: .52; second person: .54; other: .43; Mexico: first person: .55; second person: .49; other: .36. In Peninsular Spanish, none of the social factor groups were selected as significant. The weights are sex: women: .49; men: .50, and age (in years): ≤34: .51; 35–54: .50; ≥55: .45.

7. It is noted in the literature that an analysis including both social and linguistic factors will usually differ little from one in which both internal and external variables are considered separately (Poplack, 1997:304). In this case, when social and linguistic factors were analyzed separately, the results differ in the following way. In Mexico, the factor group polarity was no longer selected as significant and verb person was now selected; in Spain, the results of the linguistic factors are almost identical but age is also selected as significant. External and internal factor groups, in this case, are not completely orthogonal.

8. The manner expressions found in the corpora are: *con confianza* (with confidence), *bien* (well), *llorando* (crying), *así* (like that), *como usted* (like you), *en términos generales* (in general terms), *en serio* (seriously), *con naturalidad* (naturally), *textualmente* (textually), *con razón* (with reason), *justo* (fair), *difícil* (difficult), *en serio* (seriously), *de una forma* (in such a way), *en función de las ventas* (according to the sales), *delante de mi mujer* (in front of my wife), *por experiencia* (from experience), *cómo* (how), *en otro tono* (in a different tone), *mejor* (better), *detalladamente* (in detail), *de otra manera* (in a different way), *exactamente* (exactly), *muy bien* (very well), *bien a bien* (very well), *perfectamente* (perfectly well), *de memoria* (by heart), *de una manera especial* (in a special manner).

9. The tokens were originally coded as declarative main clause, declarative complement clause, main *wh*-interrogative clause, main yes/no interrogative clause, nominalized relative sentence (*lo que …*), indirect *wh*-question, and indirect if (*si*) interrogative, and the latter collapsed into interrogatives and declaratives.

10. Of the nondeclarative sentences in the Mexican corpora, 78% are interrogatives, but the favoring effect of nondeclarative sentences on the null pronoun is shared in the three type of sentences collapsed in this factor (90% of nulls in imperatives; 100% in exclamatory sentences; and 90.5% in interrogatives).

11. Not selected as significant were the following. Referential distance: Spain: 1 clause: .51, more than 1 clause: .43; Mexico: 1 clause: .49; more than 1 clause: .53. Verb tense: Spain: present: .54; subjunctive: .50; preterit: .44; imperfect: .35; future/conditional: .21; Mexico: present: .45; subjunctive: .47; preterit: .61; imperfect: .72; future/conditional: .47. Verb person: Spain: first person: .52; second person: .54; other: .43; Mexico: first person: .55; second person: .49; other: .36.

12. Also selected as significant, but not included in Table 9 are the three social factor groups: sex (women: .60; men: .35), age (≤34: .61; ≥55: .46; 35–54: .36), and education (low: .57; high: .41). The external factors were not excluded from this analysis because when these were excluded in the first, general analysis, polarity was no longer selected as significant.

13. In the Peninsular Spanish analysis (Table 8), due to interactions, the external factor groups and the internal factor groups that had never been selected as significant in previous runs were excluded: referential distance, verb tense, verb person, manner adverbial, age, and sex.

14. The weights of the factors under polarity were the following. Spain: positive sentence: .50; negative sentence: .49. Mexico: positive sentence: .47; negative sentence: .54.

15. A new multivariate analysis was conducted excluding the factor group polarity. In Peninsular Spanish (input: 0.239, $N = 656$, log likelihood [Ll] = −328.754), the same internal factor groups were selected as significant (dative pronoun: monotransitive, .67; pronoun, .44; ditransitive without dative pronoun, .14; type of antecedent: interrogative, .71; declarative, .38). In Mexican Spanish (input 0.91, $N = 669$, Ll = −205.575), the same factor groups plus the factor group verb person were selected as significant (dative pronoun: monotransitive, .57; pronoun, .51, ditransitive without dative pronoun, .15; manner adverbial: absent, .54; present, .12; type of antecedent: interrogative, .68; declarative, .29; sentence type: nondeclarative, .73, declarative, .47).

16. This distribution of the null pronoun cannot be explained by the opposition "primary vs. secondary objects" (Dryer, 1986). A primary object is an indirect object in a ditransitive clause or a direct object in a monotransitive clause, whereas a secondary object is a direct object in a transitive clause. If the object coded as null were the primary object, that would explain that monotransitive verbs accept the DO-null but it would predict that in ditransitive verbs, the dative pronoun would be more likely coded as zero, and this is not the case. When the verb semantically carries a dative object, the dative is usually explicit and the DO (secondary object) is coded as zero in our data. In other words, if this were a primary-secondary object system, in which the primary object were more likely coded as zero, the prediction would be that, together with *Tú sabes*, we would usually find *Tu lo dices* meaning *tú se lo dices*, with a null pronoun

https://doi.org/10.1017/S0954394509990111 Published online by Cambridge University Press

referring to the dative object, but this is not the case, because what we usually find is *Tú le dices*, with the null object referring to the propositional direct object.

17. In fact, we may find the same tendency with nonreferential null DO, as in (i); when a manner adverbial is included, the reading of the DO is hypothesized to be interpreted as more specific and the overt clitic pronoun preferred over the null object.

(i) *¿Compraste café?*
Sí, Ø compré
Sí, lo compré muy barato/de oferta
??Sí, Ø compré muy barato/de oferta
Did you buy (some) coffee?
Yes, I bought Ø (some)
Yes, I bought LO very cheap/on sale
Yes, I bought Ø (some) very cheap/on sale

18. In the Spanish data, besides this, only two sentences with the adverb *ya* had negative polarity.

19. The remaining factor groups selected as significant are the same factor groups contributing to the use of the null pronoun in the original analysis (Table 3). In this new analysis, the new factor group adverb has the greatest magnitude of effect, followed by dative pronoun (ditransitive with dative pronoun: .55; monotransitive .53; ditransitive without dative pronoun: .16), type of antecedent (interrogative: .67; declarative .29), sentence type (interrogative: .72; declarative: .47), and polarity (negative .61; positive .38).

20. This effect is comparable to the effect of the so-called completive or perfective particles in English, *up* and *out*, in phrasal verbs such as *drink up, use up, seek out*, or *work out*. These expressions are analyzed as serving to shift the focus of attention onto the result of the action, hence, onto the verbal object (Quirk et al., 1985:595), which explains why they are incompatible with object omission (García & Portero, 2002:12).

21. A deeper discussion of the notion of incompleteness of the proposition is beyond the limits of this article, but I would point out that interrogative sentences are only one of several linguistic constructions that introduce incomplete propositions in the discourse model; others are, for example, commands/petitions and the DO of verbs such as *contar* (to tell) (Reig Alamillo, 2008). An anonymous reviewer suggested using the notion of specificity to explain the results of the factor group type of antecedent. Several reasons led me to prefer the notion of "incompleteness of the proposition" over specificity. First, "specificity" has always been used, to the best of my knowledge, to describe a characteristic of NPs and, moreover, there is no agreement on exactly how specificity should be defined: a semantic feature based on "epistemic specificity" (Hellan, 1981; Ioup, 1977); a scope relation, that is, whether the NP has wide scope over an operator (Fodor & Sag, 1982); a derived pragmatic effect (Leonetti, 2003); or an epiphenomenon connected to a family of distinctions marked differently in languages (Farkas, 2001:85). I prefer not to add to the terminological/conceptual confusion by using the term specificity to describe a semantic feature of entities other than NPs.

Second, it was suggested that *epistemic specificity*, that is, whether the speaker has something in mind or is able to identify the referent (Hellan, 1981; Ioup, 1977), could be better suited to explain the effect of the factor group kind of antecedent. The idea of epistemic specificity is not the one that I want to convey in the linguistic analysis. Completeness or incompleteness of the proposition interacts with the knowledge that the speaker has about the content of the proposition in such a way that incomplete propositions do not always equate to a lack of knowledge on the part of the speaker. Compare, as an illustration, the pair *Me dijo qué vamos a comer* (she told me what we are eating) (incomplete proposition/knowledge about the content of the proposition) vs. *No sé qué vamos a comer* (I don't know what we are eating) (incomplete proposition/lack of knowledge) (Reig Alamillo, 2008). In fact, in many cases we cannot say whether the speaker knows or does not know the full content of an incomplete proposition. Thus, I prefer to avoid including epistemic specificity of the proposition in the analysis.

22. Although a declarative antecedent does not necessarily mean a complete, saturated proposition for the DO pronoun, this simplification would be enough for the purposes of this discussion.

23. This phrase was not excluded as a discourse marker because of the less fixed nature of the construction—different forms of the verb or presence or absence of the dative pronoun—and because it does not have a directly perceived discourse function but a literal meaning and a clear referent for the null object.

24. In both dialects, referential distance interacted with the factor group number of times the proposition was referred to—excluded from the Varbrul analysis—in such a way that very few tokens were mentioned more than one clause before and more than one time in the previous discourse, and

https://doi.org/10.1017/S0954394509990111 Published online by Cambridge University Press

most of the tokens were situated in the cell of "referential distance of one sentence" and "referred to once in the discourse." No interaction was observed between the factor groups referential distance and turn or referential distance and type of antecedent in Peninsular or Mexican Spanish.

25. The term "higher-order entities" is used in part of the literature dealing with abstract anaphora to differentiate it from "first-order entities" anaphora. In these studies, the term is used to refer not only to propositions but also to other semantic types of abstract entities (events, facts, etc.) (Borthen et al., 2003).

26. It seems intuitively correct to assume that the correlation between the semantic content of the proposition and accessibility would be such that propositions with (more) complete content are more accessible and lesser accessibility correlates with (more) incomplete content, but this is an assumption that I will not try to justify at this point.

27. No dates were entered in the search but the results are from 17th, 19th, and 20th centuries.

REFERENCES

Ariel, Mira. (1988). Referring and accessibility. *Journal of Linguistics* 24:65–87.

_____. (1990). *Accessing noun phrase antecedents*. New York: Routledge.

_____. (1994). Interpreting anaphoric expressions: A cognitive versus a pragmatic approach. *Journal of Linguistics* 30:3–42.

_____. (1996). Referring expressions and the +/– coreference distinction. In T. Fretheim & J. K. Gundel (eds.), *Reference and referent accessibility*. Amsterdam: Benjamins. 13–35.

Armary, Valérie. (1997). Vers une typologie des objets nuls en Français. *Verbum* 4:375–390.

Azorín, Dolores (ed.). (2002). *El proyecto Alcore: Alicante corpus oral de Español*. Alicante: Universitat d'Alacant/Departamento de Filología Española, Lingüística General y Teoría de la Literatura.

Bello, Andrés, & Cuervo, Rufino J. (1945). *Gramática de la lengua castellana*. Buenos Aires: Sopena.

Borthen, Kaja, Fretheim, Thorstein, & Gundel, Jeanette K. (2003). What brings a higher-order entity into focus of attention? Sentential pronouns in English and Norwegian. *Journal of Logic, Language and Information* 12:291–299.

Bosque, Ignacio, & Moreno, Juan Carlos. (1990). Las construcciones con *lo* y la denotación del neutro. *Lingüística* 2:5–50.

Brucart, José María. (1999). La elipsis. In I. Bosque & V. Demonte (eds.), *Gramática descriptiva de la lengua española*, 2. Madrid: Espasa. 2787–2866.

Campos, Héctor. (1986). Indefinite object drop. *Linguistic Inquiry* 17:354–359.

Chambers, Jack. (1995). *Sociolinguistic theory: Linguistic variation and its social significance*. Oxford: Blackwell.

_____. (2003). *Sociolinguistic theory*. Oxford: Blackwell.

Chambers, Jack, & Peter, Trudgill. (1998). *Dialectology*. Cambridge: Cambridge University Press.

Choi, Jinny K. (1998). *Languages in contact: A morphosyntactic analysis of Paraguayan Spanish from a historical and sociolinguistic perspective*. Ph.D. dissertation, Georgetown University.

_____. (2000). [-Person] direct object drop: the genetic cause of a syntactic feature in Paraguayan Spanish. *Hispania* 83:531–543.

Clements, J. Clancy. (1994). Notes on topicalization and object drop in Spanish. In M. Mazzola (ed.), *Issues in Romance Linguistics: Selected papers from LSRL XXIII*. Washington, DC: Georgetown University Press. 219–237.

_____. (2006). Null direct objects in Spanish. In C. Clements & J. Yoon (eds.), *Functional approaches to Spanish syntax*. New York: Palgrave. 134–150.

Company, Concepción. (2006). El objeto indirecto. In C. Company (ed.), *Sintaxis histórica de la lengua española*, I. Mexico: Universidad Nacional Autóma de México. 479–574.

Cornish, Francis. (2006). Compléments nuls vs pronoms objets manifestes en anglais en tant qu'anaphoriques: Syntaxe, sémantique ou pragmatique? *Cahiers de Grammaire* 30:89–101

Cyrino, Sonia. (1997). *O objeto nulo no português do Brasil*. Londrina: Editoral da Universidade Estadual de Londrina.

Dahl, Östen, & Hellman, Christina. (1995). What happens when we use an anaphor. Presentation at the XVth Scandinavian Conference of Linguistics. Oslo.

Depiante, Marcela. (2001). On null complement anaphora in Spanish and Italian. *Probus* 13:193–221.

Dryer, Matthew S. (1986). Primary objects, secondary objects, and antidative. *Language* 62:808–845.

Eckert, Miriam, & Strube, Michael. (2000). Dialogue acts, synchronising units and anaphora resolution. *Journal of Semantics* 17:51–89.

https://doi.org/10.1017/S0954394509990111 Published online by Cambridge University Press

Eguía, Myriam. (2002). The social meaning of Basque Spanish null object constructions. Paper presented at 31st New Ways of Analyzing Variation, Stanford University.
Esgueva, Manuel, & Cantarero, Margarita (eds.). (1981). *El habla culta de la ciudad de Madrid*. Madrid: Consejo Superior de Investigaciones Científicas.
Farkas, Donka. (2001). Specificity distinctions. *ZAS Papers in Linguistics* 23:85–101.
Fillmore, Charles J. (1986). Pragmatically controlled zero anaphora. In V. Nikiforidiu, M. VanClay, M. Niepokuj, & D. Feder (eds.), *Proceedings of the Twelfth Annual Meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society. 95–107.
Fodor, Janet D., & Sag, Ivan. (1982). Referential and quantificational indefinites. *Linguistics and Philosophy* 5:355–398.
Fónagy, Ivan. (1985). J'aime. Je connais. Verbes transitifs à l'objet latent. *Revue Romaine* 20:3–35.
Fox, Barbara A. (1987). *Discourse structure and anaphora*. Cambridge: Cambridge University Press.
García Velasco, Daniel, & Portero Muñoz, Carmen. (2002). Understood objects in functional grammar. *Working Papers in Functional Grammar* 76. Available at: home.hum.uva.nl/fg/publications_papers.html.
Gili y Gaya, Samuel. (1964). *Curso superior de sintaxis española*. Barcelona: Vox.
Givón, Talmy. (1983). Topic continuity in discourse: An introduction. In T. Givón (ed.), *Topic continuity in discourse: A quantitative cross-language study*. Philadelphia: Benjamins. 1–41.
Grevisse, Maurice. (1993). *Le Bon Usage: Grammaire Française*. Duculot: Louvain-la-Neuve.
Gundel, Jeanette, Hedberg, Nancy, & Zacharski, Ron. (1993). Cognitive status and the form of referring expressions in discourse. *Language* 69:274–307.
Gundel, Jeanette, Hegarty, Michael, & Borthen, Kaja. (2003). Cognitive status, information structure, and pronominal reference to clausally introduced entities. *Journal of Logic, Language and Information* 12:281–299.
Hamblin, Charles L. (1973). Questions in Montague English. *Foundations of Language* 10:41–53.
Hegarty, Michael. (2003). Semantic types of abstract entities. *Lingua* 113:891–927.
Hellan, Lars. (1981). On semantic scope. In F. Heny (ed.), *Ambiguities in Intensional Contexts*. Dordrech: Reidel. 47–81.
Iliescu, María. (1988). Reprise et non reprise pronominale d'un complément d'object direct à valeur neutre. *International Journal of Rumanian Studies* 6:69–79.
Ioup, Georgette. (1977). Specificity and the interpretation of quantifiers. *Linguistics and Philosophy* 1:233–245.
Kany, Charles. (1945). *American-Spanish syntax*. Chicago: University of Chicago Press.
Karttunen, Lauri. (1976). Discourse referents. *Syntax and Semantics* 7:361–385.
Labov, William. (1966). *The Social Stratification of English in New York City*. Washington DC: Center for Applied Linguistics.
_____. (1990). The intersection of sex and social class in the course of linguistic change. *Language Variation and Change* 2:205–254.
_____. (1994). *Principles of linguistic change*. Volume 1: Internal factors. Oxford: Blackwell Publishers.
Lambrecht, Knud, & Lemoine, Kevin. (1996). Vers une grammaire des compléments zéro en français parlé. In J. Chuquet & M. Frid (eds.), *Absence de marques et représentation de l'absence*. Rennes: Presses universitaires de Rennes. 279–309.
Landa, Miren Alazne. (1995). *Conditions on null objects in Basque Spanish and their relation to* leísmo *and clitic doubling*. Ph.D. dissertation, University of Southern California.
Larjavaara, Meri. (2000). *Présence ou absence de l'objet. Limite du possible en français contemporain*. Helsinki: Academia Scientiarum Fennica.
Leonetti, Manuel. (1999). El artículo. In I. Bosque & V. Demonte (eds.), *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa. 787–890.
_____. (2003). Specificity and object marking: The case of Spanish *a*. In K. Von Heusinger & G. Kaiser (eds.), *Arbeitspapier Nr. 113. Proceedings of the Workshop "Semantic and Syntactic Aspects of Specificity in Romance Languages."*; Fachbereich Sprachwissenschaft: Universität Konstanz. 67–101.
Lope Blanch, Juan M. (ed.). (1971). *El habla de la ciudad de México: Materiales para su estudio*. México: Universidad Nacional Autónoma de México.
Lope Blanch, Juan M. (ed.). (1976). *El habla popular de México: Materiales para su estudio*. México: Universidad Nacional Autónoma de México.
Maes, Alfons. (1996). The markedness of abstract-object in discourse. In W. De Mulder & L. Tasmowski (eds.), *Coherence and Anaphora. Belgian Journal of Linguistics* 10:161–184.
Marcos Marín, Francisco. (1992). *COREC: Corpus de Referencia de la Lengua Española Contemporánea: Corpus Oral Peninsular*. Available at: www.lllf.uam.es/~fmarcos/informes/corpus/corpusix.html.

https://doi.org/10.1017/S0954394509990111 Published online by Cambridge University Press

Masullo, Pascual J. (2003). Clitic-less definite object drop in River Plate Spanish. Paper presented at *Linguistic Symposium on Romance Languages (LSRL) XXXIII*, Indiana University.
Meyer-Lübke, Wilhelm. (1923). *Grammaire des langues romanes*, III. New York: Stechert.
Meyerhoff, Miriam. (2002). Formal and cultural constraints on optional objects in Bislama. *Language Variation and Change* 14:323–346.
Noailly, Michèle. (1997). Les mystères de la transitivité invisible. *Langages* 31:96–109.
Ojeda, Almerindo. (1984). A note on the Spanish neuter. *Linguistic Inquiry* 14:171–174.
Otheguy, Ricardo. (1978). A semantic analysis of the difference between el/la and lo. In M. Suñer (ed.), *Contemporary studies in romance languages*. Washington: Georgetown University Press. 241–257.
Palacios Alcaine, Azucena. (1998). Variación sintáctica en el sistema pronominal del español paraguayo: La elisión de pronombres objeto. *Anuario de lingüística hispánica* 14:431–454.
_____. (2000). El sistema pronominal del español paraguayo: Un caso de contacto de lenguas. In J. Calvo (ed.), *Teoría y práctica del contacto: El español de América en el candelero*. Madrid: Iberoamericana. 123–143.
Peterson, Philip. (1997). *Fact, proposition, event*. Dordrecht: Kluwer.
Poplack, Shana. (1997). The sociolinguistic dynamics of apparent convergence. In G. Guy, J. Baugh, & D. Schiffrin (eds.), *Towards a social science of language*. Amsterdam: Benjamins. 285–309.
Poplack, Shana, & Tagliamonte, Sali A. (2001). *African American English in the diaspora*. Oxford: Blackwell.
Quirk, Randolph, Crystal, David, Greenbaum, Sidney, Svartvik, Jan, & Leech, Geoffrey. (1985). *A comprehensive grammar of the English language*. New York: Longman.
Reig Alamillo, Asela. (2007). Variable propositional objects in Mexican and Peninsular Spanish. Paper presented at New Ways of Analyzing Variation 36, Philadelphia.
_____. (2008). C*ross-dialectal variability in propositional anaphora: A quantitative and pragmatic study of null objects in Mexican and Peninsular Spanish*. Ph.D. dissertation, The Ohio State University.
Reig Alamillo, Asela, & Schwenter, Scott A. (2007). Null objects and neuter *lo*: A cross-dialectal variationist analysis. In J. Holmquist, A. Lorenzino, & L. Sayahi (eds.), *Selected Proceedings of the Third Workshop on Spanish Sociolinguistics*. Somerville, MA: Cascadilla Proceedings Project. 113–121.
Rodríguez Alfano, Lidia. (2004). *El habla de Monterrey. Base de información para estudios en Ciencias del Lenguaje*. Monterrey: Universidad Autónoma de Nuevo León. Available at: http://www.filosofia.uanl.mx/hablamty/index.html.
Schegloff, Emanuel. (2007). *Sequence organization in interaction: A primer in conversation analysis*. I. Cambridge: Cambridge University Press.
Schiffman, Rebecca. (1985). *Discourse constraints on "it" and "that": A study of language use in career-counseling interviews*. Ph.D. thesis, University of Chicago.
Schwenter, Scott A. (2006). Null objects across South America. In T. L. Face & C. A. Klee (eds.), *Selected Proceedings of the 8th Hispanic Linguistics Symposium*. Somerville, MA: Cascadilla Proceedings Project. 23–36.
Schwenter, Scott A., & Silva, Gláucia. (2003). Anaphoric direct objects in spoken Brazilian Portuguese: Semantics and pragmatics. *Revista Internacional de Lingüística Iberoamericana* 2:99–123.
Solé, Carlos, & Solé, Yolanda. (1977). *Modern Spanish syntax: A study in contrast*. Lexington, MA: Heath.
Suñer, Margarita, & Yépez, María. (1988). Null definite objects in Quiteño. *Linguistic Inquiry* 14: 561–565.
Tagliamonte, Sali A. (2004). Comparative sociolinguistics. In J. Chambers, P. Trudgill, & N. Schilling-Estes (eds.), *Handbook of Language Variation and Change*. Malden, MA: Blackwell Publishers. 729–763.
Torres Cacoullos, Rena, & Schwenter, Scott. (2008). Constructions and pragmatics: Variable middle marking in Spanish *subir*(*se*) 'go up' and *bajar*(*se*) 'go down'. *Journal of Pragmatics* 40:1455–1477.
Torres Cacoullos, Rena, & Walker, James A. (2009). On the persistence of grammar in discourse formulas: A variationist study of *that*. *Linguistics* 47:1–43.
Toscano, Mateus. (1953). *El español en el Ecuador*. Madrid: Consejo Superior de Investigaciones Científicas.
Yaguello, Marina. (1998). La réalisation zéro des clitiques objet dans les constructions di-transitives du français parlé. In M. Bilger, K. van den Eynden, & F. Gadet (eds.), *Analyse linguistique et approaches de l'oral: recueil d'études offert en hommage à Claire Blanche-Benveniste*. Paris: Peeters. 267–274.
Yépez, María Victoria. (1986). *Direct object clitics in Quiteño Spanish*. M.A. thesis, Cornell University.

https://doi.org/10.1017/S0954394509990111 Published online by Cambridge University Press