Null Objects across South America

Scott A. Schwenter
The Ohio State University

1. Introduction

The phenomenon of null direct objects (a.k.a. “object drop”) is one that has received a growing amount of interest from researchers in Hispanic linguistics during the past 10-15 years. In the majority of cases, investigators have been interested in one of two perspectives on the phenomenon: either (a) the role of language contact in promoting the occurrence of null objects (e.g., Choi 2000, Palacios Alcaine 2000), or (b) the correct syntactic characterization of the null object, typically couched in a generative syntactic framework (e.g., Landa 1995, Suñer & Yépez 1988).

Among the agreed-upon empirical findings of this research is the observation that null direct objects occur only with ANAPHORIC D[irect] O[bject], that is, those whose referents are recoverable from the ongoing discourse context as well as sufficiently salient in that context to be encoded pronominally. This does not mean, necessarily, that the DO referent must have been previously referred to linguistically in the discourse; it could instead be accessible due to its perceptual salience to the interlocutors in the extralinguistic situation. In addition, all researchers are in agreement that null DOs are typically, if not invariably, considered to be “anomalous” from the perspective of “standard” or “pan-Spanish” norms (whether Latin American or Peninsular). Indeed, one might say that it is this anomaly that makes the phenomenon so interesting to examine in the first place.

Another observation that is repeated throughout the literature on null objects is that typically they do not occur with DOs that have human referents. Thus, a sentence like Yo vi Ø ayer ‘I saw (it) yesterday’ with a null object will tend strongly to be interpreted as having a non-human, or even inanimate referent (e.g., el programa ‘the program’) instead of a human one (e.g., Juan); the latter will tend to be realized pronominally as either lo (the etymologically accusative 3sg. masculine pronoun, i.e., ‘him’) or (much more likely; see below) le (the etymologically dative 3sg. pronoun). However, a fact which has escaped the attention of nearly all linguists who have dealt with the issue of null objects in Spanish (or Portuguese) is that those DOs “coded” as null objects (i.e., as zeroes) are actually those which display the prototypical properties of DOs cross-linguistically (Company 2002a, Comrie 1979, 1989, Dahl 2000, Hopper & Thompson 1980). The prototype of the DO has been characterized as follows by Company (2002a:206) in a recent study of the ongoing spread of accusative a in Spanish.

[A] DO refers very frequently to inanimate beings, things or abstract concepts, fully affected by the action of the verb; a DO usually has neither energy nor volition and it usually undergoes some change of state, provoked by the energy of the agent via the transitivity of the verb. A thing is easier to change than a human being, therefore a non-human entity seems to represent the prototype of a DO better than a human…

* I am grateful to Tim Face and Carol Klee for extending an invitation to me to give this paper as a plenary lecture at HLS 8. Versions of the paper were also presented at Indiana University, UNAM, and the VIII Encuentro Internacional de Lingüística in Hermosillo. I would also like to thank the audiences at all these venues for probing comments. I am especially indebted to Bruno Estigarribia and Pascual Masullo, who have both spent a lot of time talking with me about direct objects and related phenomena. My OSU colleagues Terrell Morgan and Gláucia Silva have been extremely helpful on many points. Unfortunately, I alone am ultimately responsible for any misinterpretations.

1 It seems plausible to trace this attention, in great measure, to Raposo’s (1986) syntactic analysis of null direct objects in European Portuguese.
Thus, putting it slightly differently, the research focus on NULL objects (or “object drop”) has tended to overlook an empirical reality, despite the fact that most scholars seem to have recognized it in passing: in so-called null object dialects, the coding of anaphoric DOs is actually VARIABLE—some DOs are coded as pronouns, while others (the majority) are coded as zeroes. Why should this be the case?

The primary goal of the present paper is to attempt to provide an answer to this question. In order to do so, three steps will need to guide the discussion. First, I will present a survey of the factors that determine the choice of form for anaphoric DOs in both Brazilian Portuguese and null-object dialects of South American Spanish, principally Paraguayan (Asunción) and Quiteño Spanish. After laying out the empirical facts about the coding of anaphoric DOs in these dialects, I will demonstrate that a clear parallelism exists between Brazilian Portuguese and these dialects of South American Spanish with respect to this coding, i.e., the characteristics of those anaphoric DOs realized as null objects and those realized overtly (typically as pronouns) across the varieties in question are essentially the same. Finally, I will show that there is likewise a strong parallelism—in both form and function—between the coding of anaphoric DOs in the aforementioned varieties and the much more well-known phenomenon of differential object marking (DOM), which in Spanish is manifested by accusative a.

The desired result of this study is to provide a broader explanation for null vs. overt anaphoric DO variation, and also to begin to situate this variation with respect to other DO phenomena which, it appears, are regulated by similar factors or constraints.

2. The form of anaphoric direct objects in spoken Brazilian Portuguese

The frequent occurrence of null direct objects in Brazilian Portuguese is a well-known phenomenon. Although the relationship between null objects in Spanish and null objects in Brazilian Portuguese is made briefly by a number of scholars looking at the phenomenon in the former language, there have been no studies to date that carry out a close comparison of what factors promote or inhibit the occurrence of null vs. overt forms in the two languages. Rather, what is normally noted is that, among the Romance languages, Brazilian Portuguese (BP) is an exceptional variety in that it allows null anaphoric DOs with definite/specific referents. Moreover, and unlike the case of Spanish, there does not seem to be great variation in the frequency of null objects across the different dialects of BP. Note the well-known contrasts between so-called “standard” Spanish in (1a), spoken BP in (1b), and spoken European Portuguese (EP) in (1c).

(1a) Juan compró un libro nuevo. Ayer lo/*/Ø trajo a clase.    (“standard” Spanish)

‘Juan bought a new book. Yesterday he brought it to class.’

(1b) O João comprou um livro novo. Ontem ele trouxe Ø à aula.    (BP)

(1c) O João comprou um livro novo. Ontem trouxe-o à aula.    (EP)

The presence of the clitic pronoun lo is obligatory in “standard” Spanish (1a), but the overwhelming preference (see further below) in BP is for a null object as in (1b). While EP also allows null objects to some extent (Raposo 1986, inter alia), my EP informants are in agreement that a null object in (1c) would be odd; the enclitic –o therefore is the default option.2

Much as has been the case in the study of Spanish null objects, the debate over null objects in both BP and EP has been mainly limited to the syntactic sphere, and specifically to the issues surrounding the correct syntactic characterization of the null form itself (Cyrino 1997, Farrell 1990, Kato 1993, Raposo 1986, inter alia). This focus on syntax has led to a clear gap in the explanation of what kinds of DOs tend to appear as null objects, and which do not.

2 Interestingly, several EP speakers I have queried are in general disagreement with many of the acceptability judgments in Raposo (1986).
Of the full Portuguese DO clitic system that is still vibrant in present-day EP, only 2sg. *te* and 1sg. *me* remain current in spoken BP, a fact whose importance will become clearer in the discussion below. The other clitic forms are heavily restricted to more formal styles, as exemplified in (2) and (3).

(2) Eu o/a vi semana passada. (educated/careful speech only)
   ‘I saw him/her last week.’

(3) Eles querem convidá-las para a festa. (educated/careful speech only)
   ‘They want to invite them to the party.’

In spoken BP, the 3rd person tonic pronouns *ele(s)/ela(s)* are used instead of the clitic forms, as in (2’) and (3’) below. This use of the etymologically nominative pronouns for accusative DO functions is virtually unknown in EP.

(2’) Eu vi ele/ela semana passada
   ‘I saw him/her last week.’

(3’) Eles querem convidar elas para a festa.
   ‘They want to invite them to the party.’

In Schwenter and Silva (2002), we argued that there exists a clear continuum which orders the degree of acceptability of null direct objects in BP. This continuum is highly sensitive to two distinct characteristics of the referent of the third person direct object: animacy and specificity. The first of these characteristics is an inherent property of the referent (whether real or fictitious), while the second is determined in the discourse itself, and is dependent on whether the speaker believes her interlocutors can identify a particular token referent with the DO expression, or whether they can merely identify an exemplar of the class in question. Thus for us, as noted also by Lyons (1999), definiteness and specificity are orthogonal dimensions, such that both definites and indefinites may be either specific or non-specific. Indeed, our analysis revealed that definiteness, per se, does not have an effect on the choice of coding for anaphoric DOs in BP.

The ensuing quantitative analysis of naturally-occurring BP spoken data realized in Schwenter and Silva (2003) confirmed our earlier qualitative analysis (2002). Anaphoric DOs in BP were indeed most likely to be realized overtly (i.e., as a tonic pronoun) when their referents were both animate and specific. In stark contrast, third-person anaphoric DOs were found overwhelmingly to be encoded as null objects. The global results in Table 1 (Schwenter & Silva 2003:106) stem from our analysis of 12 speakers and 1250 total third-person anaphoric DOs extracted from the PEUL corpus of spoken Rio de Janeiro Portuguese (compiled in the early 1980s).

<table>
<thead>
<tr>
<th>Form</th>
<th>Count</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Null Object</td>
<td>907</td>
<td>72.6%</td>
</tr>
<tr>
<td>Clitic</td>
<td>4</td>
<td>0.3%</td>
</tr>
<tr>
<td>Tonic Pronoun</td>
<td>151</td>
<td>12.1%</td>
</tr>
<tr>
<td>Lexical NP</td>
<td>188</td>
<td>15.0%</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>1250</td>
<td><strong>100%</strong></td>
</tr>
</tbody>
</table>

Table 1: Global results for third-person anaphoric DO form in PEUL corpus

As the data in Table 1 make clear, null objects are by far the most common coding option for third-person anaphoric DOs overall, constituting nearly three-fourths of all the tokens (cf. the nearly identical results reported in Duarte 1989 and Tarallo 1996). However, once these global results are broken down in terms of animacy and specificity, it is the [+animate, +specific] anaphoric DOs that

---

3 Note also that the third-person clitics *o* and *a*, which would be enclitics in a sentence like (2) if rendered in EP, appear as proclitics when they are employed in BP.
are more frequently encoded by tonic pronouns than by either null objects or lexical NPs. The overall results for the coding of third-person anaphoric DOs in the PEUL corpus appear in Table 2.

<table>
<thead>
<tr>
<th></th>
<th>+an/+sp</th>
<th>+an/-sp</th>
<th>-an/+sp</th>
<th>-an/-sp</th>
</tr>
</thead>
<tbody>
<tr>
<td>Null</td>
<td>50 (24.4%)</td>
<td>102 (75.5%)</td>
<td>151 (87.3%)</td>
<td>604 (81.9%)</td>
</tr>
<tr>
<td>Pron.</td>
<td>126 (61.5%)</td>
<td>12 (8.9%)</td>
<td>0 (0%)</td>
<td>13 (1.8%)</td>
</tr>
<tr>
<td>Lex. NP</td>
<td>29 (14.1%)</td>
<td>21 (15.6%)</td>
<td>22 (12.7%)</td>
<td>120 (13.3%)</td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td><strong>205 (100%)</strong></td>
<td><strong>135 (100%)</strong></td>
<td><strong>173 (100%)</strong></td>
<td><strong>737 (100%)</strong></td>
</tr>
</tbody>
</table>

Table 2: Anaphoric 3rd person DOs in PEUL, by animacy/specificity

As these results make clear, only the [+animate, +specific] class of DOs shows frequent coding via a pronoun (61.5%), while the other classes of DOs show no more than 8.9% pronominal coding. This same class, however, shows very little use of null objects relative to the other classes, 24.4% vs. at least 75.5% for the other three DO-types. Thus, it is precisely the [+animate, +specific] DOs which are anomalous among the DOs when considered in their totality. Statistical analysis confirms this conclusion: the overall distribution in Table 2 is statistically significant ($p < 0.001$). There is likewise a statistically significant difference between the specific and non-specific animate objects ($p < 0.001$), validating the regulating effect of specificity on [+animate] DOs. However, no evidence was found for specificity affecting the distribution of the inanimate DOs: the difference between the two categories of inanimate DOs is NOT statistically significant ($p < .20$). Thus, it seems reasonable to claim that animacy serves as the primary filter for the form of anaphoric DOs, and specificity as the secondary filter, insofar as its effects are felt on [+animate] DOs only.

The rise of null objects in BP has resulted, most drastically, from the loss of the third-person DO clitics (unlike European Portuguese, as noted above, where the clitics remain). As Table 1 shows, we found only 0.3% third-person clitics, versus over 96% for both first- and second-person (singular)\(^4\) clitics (me and te). Why should first- and second-person clitics be resistant to this change? An answer that suggests itself is again related to animacy and specificity: the first- and second-person clitics me and te refer, unless used generically, to inherently animate (normally human) and specific referents, the speaker and his/her interlocutor. These are the properties, as noted in Section 1 above, that are highly marked for DOs. And, as noted by Dahl (2000), “egophoric” referents, i.e., speaker and addressee, are very rarely found in the grammatical function of DO, as opposed to that of subject, which very often tends to be egophoric.

The main conclusion of our prior research on BP is that it is a variety that displays a “split” system of marking anaphoric DOs. All animate (mainly human) and specific anaphoric DOs are preferentially encoded overtly, while all others are preferentially encoded as null objects. This split system is represented in its entirety in Table 3.

<table>
<thead>
<tr>
<th></th>
<th>Preferably overt</th>
<th>Preferably null</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st person</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>2nd person</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>3rd person [+an/+sp]</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>All other 3rd person</td>
<td>no</td>
<td>yes</td>
</tr>
</tbody>
</table>

Table 3: Anaphoric DOs in BP, by preferred overt/null realization

What is perhaps most striking about this configuration is that it is the “other” 3rd person category that most closely resembles the prototype for direct objects described in the quote from Company above,

\(^4\) The first-person plural clitic pronoun nos (not to be confused with subject nós) has also been virtually lost in spoken BP, but it has been replaced by the innovative 1st plural pronoun a gente (lit. ‘the people’). Null objects are not found for 1st plural referents.
and it is precisely this class of anaphoric DOs that tends to appear (in conversational discourse, at least) as null objects. In the next section, we will see that a similar system of anaphoric DO coding appears to be prevalent in at least two dialects of South American SPANISH.

3. Null objects in South American Spanish dialects

As is well-known, most dialects of Spanish allow null direct objects only when the referent of the DO is non-specific (but not necessarily indefinite; cf. Campos 1986; also Clements 1994). Thus, a null object is permissible in (4) below, where the noun café is being used generically, but not in (5), where the NP el periódico designates a specific newspaper. In (6), where the indefinite NP una revista is understood as referring to a specific magazine which the speaker believes can be uniquely identified by her interlocutor, a null object is likewise not permissible. However, on the alternate, non-specific, interpretation of the NP in (6) as in (6’), i.e., as referring to any magazine, a null object would be allowed.

(4) Fui a la tienda a comprar café pero no tenían Ø.
   ‘I went to the store to buy coffee but they didn’t have (any).’

(5) Fui a la tienda a comprar el periódico pero no lo/*Ø tenían.
   ‘I went to the store to buy the newspaper but they didn’t have it.’

(6) Fui a la tienda a comprar una revista (específica) pero no la/*Ø tenían.
   ‘I went to the store to buy a (specific) magazine but they didn’t have it.’

(6’) Fui a la tienda a comprar una revista (cualquiera) pero no *la/Ø tenían.
   ‘I went to the store to buy a (=any) magazine but they didn’t have (one).’

Despite the patterns of (in)acceptability given in (4) through (6), however, it has also been known at least since Kany (1945) that several dialects of South American Spanish, many but not all of which are or have been in contact with indigenous languages, also allow anaphoric DOs with definite/specific reference to occur as null objects. Thus, Kany pointed out 60 years ago that, “lo ... is very frequently omitted in American Spanish, particularly in conversation and nearly always so when an indirect-object pronoun is expressed ... It is especially diffused throughout South America. We find such omissions characteristic also of Brazilian Portuguese as opposed to standard [sic] Portuguese” (1945:114).

Although Kany’s quote states that null objects occur “nearly always” when an indirect-object pronoun is overtly expressed, it is crucial to note that in fact ALL of the examples provided by Kany (1945) come from ditransitive constructions. His examples are culled from literary texts written by authors from the following South American countries: Argentina, Bolivia, Chile, Colombia, Ecuador, and Peru. In each case, the process that is being assumed is that of a null object arising from a ditransitive pronominal construction, i.e., where the clitic sequence se lo(s)/la(s) is realized alternatively as le Ø via the “dropping” of the DO clitic pronoun. Thus, this process is yet another example where the dative (and its corresponding clitic pronoun) is “seizing” the accusative. Indeed, as noted by a number of authors, the pattern whereby DAT > ACC is one that permeates Spanish more generally in both diachrony and synchrony (Company 2002b).

Unfortunately, very little else has been done on the phenomenon of ditransitive object drop in South American Spanish dialects since Kany’s original observation. My own personal observation, however, is that the null object found in ditransitive constructions is not exclusive to the dialects that Kany mentions nor is it restricted to SOUTH America. I have for instance heard instances of le Ø instead of se lo(s)/la(s) in non-South American countries such as Mexico. The process of “dropping” such DO clitics appears to be especially frequent, but by no means obligatory, when the referent of the direct object pronoun is a proposition instead of an entity; thus, both ya te Ø dije and ya te lo dije are possible for some speakers. By contrast, in Spain I have regularly observed over the past three years that the null object version ya te Ø dije is only marginally acceptable in Peninsular Spanish (with the
exception of Basque Spanish), and the version with the overt DO pronoun is highly preferred. Obviously, however, these kinds of anecdotal observations need to be investigated much more thoroughly before any conclusions can be drawn.

But while null objects in ditransitive constructions appear to be fairly widespread (if nearly wholly unstudied, despite Kany’s observations), cases of specific null objects in monotransitive constructions are much more difficult to come across (but, paradoxically, seemingly more well-studied than their ditransitive counterparts). Null objects in monotransitives in South America\(^5\) have been attested in (at least) northwest Argentina, Bolivia, Ecuador (Quito), Peru, Uruguay (on the Brazilian border; Ana Maria Carvalho, personal communication), and especially Paraguay. In addition, Masullo (2003, personal communication) has recently noted that null DOs (what he calls “definite object drop”) are possible in non-contact River Plate Spanish, but only in cases where the referent can be anaphorically recovered from the immediate context of utterance. Again, this phenomenon is not limited to ditransitives, but can also be found in monotransitive sentences such as (7) and (8) (naturally-occurring River Plate Spanish examples from Masullo [personal communication]).\(^6\)

(7) A: Queremos el postre.  
‘We want the dessert.’
B: Ya Ø traigo.  
‘I’ll bring (it).’

(8) A: Tengo un calmante para dormir.  
‘I’ve got a sedative in order to sleep.’
B: No Ø tomes. Te va a hacer mal.  
‘Don’t take (it). It’ll do you harm.’

One variety where the alternation between null and overt anaphoric DOs has been studied in considerably greater depth, including the analysis of naturally-occurring conversational data (Yépez 1986), is colloquial spoken Quiteño Spanish. Suñer and Yépez (1988:513) claim that in this dialect the “referent of the omitted DO is obligatorily understood as inanimate,” e.g., \(Vi \Theta en la televisión\) must be understood as ‘I saw IT on TV’. Yépez (1986) provides the following quantitative data from monolingual Quiteños, with respect to the choice between \(le(s)\), \(lo(s)/la(s)\), or \(\Theta\) for the coding of anaphoric DOs.

<table>
<thead>
<tr>
<th></th>
<th>(LE)</th>
<th>(LO, LA)</th>
<th>(\Theta)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[+animate]</td>
<td>90 (97.8%)</td>
<td>2 (2.2%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>[-animate]</td>
<td>28 (26.7%)</td>
<td>11 (10.5%)</td>
<td>66 (62.8%)</td>
</tr>
</tbody>
</table>

Table 4: Choice of 3p DO pronoun in conversational Quiteño (adapted from Yépez 1986:33-4); \(\chi^2 = 104.4, df = 2, p < .001\)

\(^5\) Beyond the frontiers of South America, there has also been considerable research done on null objects in the Spanish of the Basque Country (e.g., Franco & Landa 2003, Landa 1995, Urrutia & Fernández Ulloa 1997). As shown by Eguía (2002), place of origin of the speaker is the most accurate predictor of null DO usage in this region, and the frequency of null DOs is heavily influenced by bilingualism (contra Landa 1995). Thus, Eguía found that although null DOs were present to some degree in the Spanish of monolinguals in Bilbao (15%), the phenomenon was much more prevalent in the Spanish of speakers in the Basque-dominant town of Germika-Bermeo (49%). In this sense, the Basque Spanish situation is distinct from that found in Paraguayan or Quiteño Spanish (see more below).

\(^6\) While the examples in (7) and (8) do not appear to be generally accepted by speakers of other dialects of Spanish, at least some “immediate context” null DO examples may be more widespread than initially expected. For instance, speakers of all dialects seem to accept null objects in imperative examples like \(apaga \Theta ‘turn off \Theta’\), when uttered in the presence of a TV or light that is currently turned on. It appears that if the joint salience of the DO referent is such that the speaker believes her interlocutor can identify that referent without an overt pronoun, then a null object construction is possible.
As can be seen, null objects in this variety are wholly absent when the DO referent is [+animate]. In this case, the coding is always overt, and nearly always (almost 98%) realized via le(s). In stark contrast, [-animate] anaphoric DOs are preferentially encoded by zeroes (nearly 63%), while overt coding of [-animate] referents is much less frequent, but in any case more frequently found with le(s) as opposed to los/la(s). Although Yépez does not provide information about the role of definiteness or specificity in Quiteño, it seems clear that this variety resembles quite closely the overall pattern of anaphoric DO coding detailed for BP in Section 2 above.

Another variety of South American Spanish that follows the same overall pattern as that found in Quiteño is Paraguayan Spanish, specifically as spoken in the capital, Asunción. A typical example of the null objects found in Paraguayan Spanish can be seen in the following excerpt, which comes from a folk tale entitled “El ñandutí”, as told by Miguel, a native speaker from Asunción (courtesy of Jessica A. Davis and Terrell A. Morgan).

(9) ...Y uno de los indígenas, que era el más humilde de todos los pretendientes, encontró en un árbol una araña tejiendo, haciendo su telaraña, entonces dijo, “Ésta es para mi doncella.” Bueno, entonces, se fue, se subió el árbol y quiso agarrar θ y la telaraña obviamente se despedazó en el momento. Esto (θ) vio la abuela, θ vio la abuela entonces muy apenada por lo que acaba de pasar al nieto, le dijo, “ve a dormir, que mañana en la mañana, vas a tener lo que quieres.” La abuela se pasó toda la noche tejiendo para el nieto, entonces, y la abuela hizo la misma, una réplica digamos, de la telaraña de su cabello. Y al día siguiente le entregó θ ...le entregó θ a su nieto le dijo, “esto era lo que querías y puedes llevarle θ a tu doncella.” Nuestro indio guaraní le llevó θ al día siguiente a la doncella y obviamente entre todos los obsequios que la doncella tenía, fue lo que más le cautivó ...

As the null objects in this passage nicely illustrate, the phenomenon is found with both ditransitive verbs, such as entregar ‘to deliver,’ as well as monotransitive verbs such as agarrar ‘to grab,’ in Paraguayan Spanish.

In the comprehensive research conducted by Choi (1998, 2000), it was shown clearly that in Asunción, “the omission of [-person] direct object clitics is prevalent in the speech of ALL SOCIAL CLASSES, AMONG BOTH BILINGUALS AND MONOLINGUAL SPANISH SPEAKERS” (2000:536; emphasis added). Choi’s results reveal that null forms are exceedingly common for [-person] direct objects in this variety, no matter what speakers’ level of bilingualism: the 5 monolingual speakers in her sample used null forms 90% of the time; the 7 least-educated bilinguals used null forms 98% of the time. The other two groups included in her study—bilinguals with two intermediate levels of education—had results falling between the two percentages of the end groups. The full results from Choi’s study in Table 5 reveal no statistically significant difference between the groups, whether considered in pairwise fashion or collectively.

7 ‘...And one of the indigenous men, who was the most humble of all the admirers, found a spider in a tree weaving, making its spiderweb, then he said, “This [the spiderweb] is for my maiden.” Well, then, he went, he climbed the tree and he wanted to grab [it] and the spiderweb obviously immediately fell apart. This the grandmother saw, the grandmother saw then very embarrassed by what had just happened to her grandson, [and] she told him, “go off and sleep, tomorrow morning, you’ll have what you want.” The grandmother spent all night weaving for her grandson, so, and the grandmother made the same one, a copy let’s say, of the spiderweb from her hair. And the next day she delivered [it] ...she delivered [it] to her grandson [and] she said to him, “this was what you wanted and you can take [it] to your maiden.” Our guarani took [it] the next day to the maiden and obviously among all the gifts that she had, it was the one that captivated her the most...’
Table 5: [-person] DO realization in Paraguayan Spanish (Choi 2000:541-3); \(\chi^2 = 4.45, \text{df} = 3, p \leq 1.0\)

Further corroboration of Choi’s findings has recently been provided by Morgan (2004), who asserts that, based on his own fieldwork in Asunción, “Object clitics lo, la, los, and las are vanishingly rare at the vernacular level [in Paraguay].” The naturally-occurring examples in (10) and (11) are from Morgan’s corpus.

(10)  A: ¿Dónde encontraste esa blusa?
‘Where did you find that blouse?’
B: Ø Compré en el mall.
‘I bought (it) at the mall.’

(11)  A: ¿Viste Spiderman 2?
‘Did you see Spiderman 2?’
B: Sí, fui a ver Ø con Julia.
‘Yes, I went to see (it) with Julia.’

An additional connection of importance between Quito and Asunción which makes them even more similar is that both are generalized leísta dialects. Thus, Suñer and Yépez (1988:511) had already noted that “it can be claimed that colloquial Q[uiteño]S[panish] has carried leísmo to conclusion.” The Quiteño data from Yépez (1986) in Table 4 above provide strong support for this assertion. And in fact Lipski (1994:310) had already observed the affinity for leísmo shared by the two dialects: “Paraguayans share with Ecuadorians the use of le and les as (masculine) direct object clitics, instead of lo/los, found elsewhere in Latin America. At the vernacular level, le and les are used even with feminine referents” (see also de Granda 1982, 1994). Three representative naturally-occurring examples of leísmo from Paraguayan Spanish are shown in (12) through (14) (from Morgan 2004).

(12)  ¡…venga para acá ahora mismo!—gritaba su mamá tan alto que le oían hasta en la otra cuadra.
‘… come here right now!—his mother yelled so loud that they could hear her even on the other block.’
(from a short story; Mereles 2003)

(13)  Michi le come.  [Michi = el gato; le = el pájaro]
‘Michi eats him’ [Where Michi is the cat and the pronoun le ‘him’ refers to the bird]

(14)  [Older gentleman talking about foreigners’ unfamiliarity with Paraguay]
No saben encontrarle en el mapa.
‘They don’t know how to find it on the map.’

Choi (1998) analyzed recorded interviews (30-35 mins. in length) with 24 Paraguayan speakers, grouped more or less evenly as in Table 5 above. She found no use of le(s) with [-animate] referents, while [+animate] referents were coded ONLY by le(s) (N=218). Likewise, Choi found NO spontaneous use of lo/la in conversations for [+animate] referents; recall that the overall usage of lo/la in Asunción for DOs was very low, as seen already in Table 5 above.

What these collective results suggest is that there is a new system of anaphoric DO coding in Quiteño and Paraguayan Spanish (albeit apparently more established in the latter variety): animates

<table>
<thead>
<tr>
<th></th>
<th>Null Object</th>
<th>Clitic (lo, la)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bilinguals, 3-6 yrs. school</td>
<td>50 (98%)</td>
<td>1 (2%)</td>
</tr>
<tr>
<td>Bilinguals, 9-12 yrs. school</td>
<td>56 (95%)</td>
<td>3 (5%)</td>
</tr>
<tr>
<td>Bilinguals, 12+ yrs. school</td>
<td>71 (90%)</td>
<td>8 (10%)</td>
</tr>
<tr>
<td>Monolinguals</td>
<td>69 (90%)</td>
<td>8 (10%)</td>
</tr>
<tr>
<td><strong>TOTALS</strong></td>
<td><strong>246 (92%)</strong></td>
<td><strong>20 (8%)</strong></td>
</tr>
</tbody>
</table>
(again, primarily humans) are coded preferentially by le(s), while inanimates are coded preferentially by Ø. A very similar distinction in coding between different types of anaphoric DO has also been noted by Landa (1995) for Basque Spanish. The result of this systemic configuration is that there is now maximal differentiation of marked and unmarked DOs, where “markedness” is taken to be the correspondence between the features of the DO referent (e.g., whether +animate or -animate), and the method of coding a given anaphoric DO token (e.g., either as a pronoun or as a zero). Just as in BP, the more marked DO referents are most often encoded pronominally, while the less marked DO referents are encoded as zeroes.

To conclude this section, it should be pointed out that the situations found in both Quito and Asunción differ considerably from another null-object system described in the speech of four levels of Quechua-Spanish bilinguals in Peru by Paredes (1996). Paredes’ results show that degree of bilingualism, and specifically one’s greater or lesser proficiency in Spanish, are clear predictors of the amount of null object use. However, consistent with the varieties in Quito and Asunción, notice that [+human] DOs are once again much more rarely realized as null objects when compared to [-human] DOs, no matter what the speaker’s level of proficiency, as shown in Table 6.

<table>
<thead>
<tr>
<th></th>
<th>[-human]</th>
<th>[+human]</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very Low Proficiency</td>
<td>75/137 (55%)</td>
<td>6/24 (25%)</td>
<td>81/161 (50%)</td>
</tr>
<tr>
<td>Low Proficiency</td>
<td>68/204 (33%)</td>
<td>13/81 (16%)</td>
<td>81/285 (28%)</td>
</tr>
<tr>
<td>Moderate Proficiency</td>
<td>45/357 (13%)</td>
<td>1/45 (2%)</td>
<td>46/402 (11%)</td>
</tr>
<tr>
<td>High Proficiency</td>
<td>50/571 (9%)</td>
<td>8/206 (4%)</td>
<td>58/777 (7%)</td>
</tr>
<tr>
<td><strong>TOTALS</strong></td>
<td>238/1269 (19%)</td>
<td>28/356 (8%)</td>
<td>266/1625 (16%)</td>
</tr>
</tbody>
</table>

As Paredes makes clear (1996), null objects in Peruvian Spanish exemplify a “bilingual continuum”—the more proficient one’s Spanish, the fewer null objects in one’s speech. Indeed, Lipski’s generalization that “Null direct objects, lacking any clitic, are very frequent in highland Peru, but not in coastal dialects” (1994:326) could be reconstrued as bilinguals versus monolinguals in light of Paredes’ data. There are, therefore, considerable differences between Paraguayan/Quiteño, on the one hand, and Peruvian Spanish, on the other, with respect to the degree to which null objects form part of the monolingual variety (or varieties) of Spanish spoken in each locale. This is not to say, of course, that the Paraguayan/Quiteño null objects did not arise initially as the result or influence of language contact, but rather simply that the monolingual conversational norm in these varieties has now incorporated null objects as an integral part of the anaphoric DO system, something that has apparently not yet occurred in monolingual Peruvian Spanish.

### 4. Anaphoric direct objects and “Differential Object Marking”

The patterns of anaphoric DO coding detailed above in both Portuguese and Spanish varieties of South America have potentially greater significance because of the clear parallels they provide with other DO-related phenomena across languages, and most especially with what is known widely as “Differential Object Marking” or DOM (cf. Aissen 2003, Bossong 1991, Comrie 1979). DOM is the label conventionally used to denote the special accusative case marking of certain direct objects in a language. DOM is manifested on these DOs by additional morphological marking that non-DOM DOs do not receive, e.g., by a in Spanish, pe in Rumanian, -ko in Hindi, -rā in Persian, which contrast with the lack of DOM on other DOs. As noted by many scholars, the kinds of features that are normally sensitive to DOM are animacy, definiteness, and specificity. Some no doubt familiar contrasts from DOM in Spanish—regulated by animacy and specificity—are provided in (15) and (16) (see Pensado 1995a for a comprehensive overview of Spanish DOM from multiple perspectives).

(15a) Ayer vi Ø/*a tu libro.   
‘Yesterday I saw your book.’

[-anim, +spec]
(15b) Ayer vi *Ø/a tu hermana.
    ‘Yesterday I saw your sister.’

(16a) Quiero entrevistar Ø/?a una persona que sepa catalán.
    ‘I want to interview a person (=any person) who knows Catalan.’

(16b) Quiero entrevistar *Ø/a una persona que sabe catalán.
    ‘I want to interview a (specific) person who knows Catalan.’

While the preceding four paradigmatic examples are useful in illustrating the similarities between
null/overt marking on the one hand, and accusative marking, on the other, the examples also obviously
do not present the whole picture of complex variation that obtains in the accusative-marking system of
Spanish. As pointed out by Laca (2002), among others, it is necessary to distinguish between inherent
referential properties of DOs and contextual factors, since, as Laca shows, both kinds of factors
combine to determine the overall pattern of accusative a marking in Spanish. Schwenter (2004) argues
that a similar distinction is important to account for the distribution of overt vs. null anaphoric DOs in
BP.

Inherent properties such as animacy are clearly primary determinants of DOM in a wide variety of
languages, and many authors (e.g., Aissen 2003, Comrie 1979) have proposed correlations between
DOM and animacy and definiteness/specificity scales, e.g., human > animate > inanimate and definite
> specific indefinite > non-specific indefinite. Of most relevance to the topic of this paper is the fact
that DOs marked by accusative a in (pan-) Spanish typically share the same inherent features as those
anaphoric DOs that tend to be coded overtly in BP and Paraguayan/Quiteño Spanish. On the other
hand, those Spanish DOs not regularly a-marked display features that are similar to those direct objects
in BP and Paraguayan/Quiteño Spanish that may, and overwhelmingly do, occur as NULL OBJECTS.

The question that remains, however, is whether these inherent properties are the main
determinants of DO coding. According to Laca (1995 [1987], 2002), and also to Leonetti (2004), the
true discourse function of accusative a in Spanish is not as a marker of animacy/humanness, but rather
as a marker of HIGH TOPICALITY (where “topicality” = “aboutness”).

…aunque no hay duda de que el uso de la preposición [a] está parcialmente gramaticalizado
en torno a un foco categorial, el de “persona”, este proceso de generalización no va
acompañado de una pérdida de la función originaria, la de indicar la alta “topicidad” del
objeto y su independencia dentro del predicado (Laca 1995 [1987]:89).

As implied by Laca’s quote, from a diachronic perspective this view of the functional origins of a as a
topic marker is amply supported by studies such as those of Melis (1995), Pensado (1995b) and Detges
(2005), among others.

Indirect quantitative support for this position in synchrony comes from Bentivoglio’s (1983) study
of topic continuity in Latin American Spanish. Bentivoglio operationalized the notion of topicality,
using Givón’s now well-known measures of referential distance (RD) and topic persistence (TP). RD
and TP measure, respectively, the distance (in clauses) from a given mention of a referent to its
previous mention; and the number of times that the same referent is referred to in the subsequent 10
clauses following any given mention of that referent. Bentivoglio’s overall results revealed that
[+human] NPs displayed much greater topicality (in her operationalized terms) than [-human] NPs, as
shown by the figures in Table 7.

---

8 Indeed, some of this complexity is already apparent in the example in (16a), where the presence/absence of
accusative a is highly variable and often subject to factors besides animacy and specificity.
9 ‘…although there is no doubt that the use of the preposition [a] is partially grammaticalized around a categorial
center, that of “persons,” this process of generalization is not accompanied by a loss of the original function, that
of indicating the high “topicality” of the object and its independence within the predicate.’
Referential Distance | Topic Persistence
--- | ---
+Human (N=149) | 8.89 | 1.60
-Human (N=272) | 13.81 | 0.30

Table 7: Topicality of Spanish [+definite] NPs, as measured by RD and TP (Bentivoglio 1983)

The lower RD score for [+human] referents (8.89), relative to the score for [-human] referents (13.81), coupled with the higher TP score (1.60 vs. .30) for the former class, are clear indicators of the relatively higher topicality of the [+human] class of definite NPs. The [-human] class, in contrast, shows a higher RD score and a lower TP score, and therefore exhibits relatively lower topicality. Since DOs tend strongly to have referents that are [-human], and therefore to display relatively low topicality overall (compared to, e.g., subjects and IOs), accusative a can be seen as a strategy to overtly signal the heightened topicality of some DOs, overwhelmingly but not solely those that are [+human]: “the contribution of a is the encoding of an instruction to process the [direct] object ... as an internal topic, that is, as a prominent and referentially autonomous argument” (Leonetti 2004:94).

Bentivoglio’s results in Table 7 mesh well with those of Schwenter and Silva (2003) for spoken BP, who found a close correlation between the topicality of anaphoric DOs (as measured by RD and TP) and their encoding. Thus, the greater the topicality of the anaphoric DO in BP, the more likely its referent was to be encoded overtly; the lesser the topicality, the more likely its referent was to occur encoded as a null object. Unfortunately, similar studies of Paraguayan or Quiteño Spanish do not exist. However, it is highly probable that anaphoric DOs in these varieties pattern the same way, especially in light of Bentivoglio’s findings in Table 7 above regarding the relative topicality of Spanish NP types overall.

The parallels between anaphoric DOs in BP and Paraguayan/Quiteño Spanish, on the one hand, and pan-Spanish DOs, on the other, are summarized in Table 8 (adapted from Schwenter & Silva 2002:584).

<table>
<thead>
<tr>
<th>BP/PQSp</th>
<th>PAN-SPANISH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower Topicality DOs</td>
<td>Null Pronominal (Ø)</td>
</tr>
<tr>
<td>(Less Overt Coding)</td>
<td></td>
</tr>
<tr>
<td>Higher Topicality DOs</td>
<td>Overt Pronominal (ele/le)</td>
</tr>
<tr>
<td>(More Overt Coding)</td>
<td></td>
</tr>
</tbody>
</table>

Table 8: Parallels between BP/PQSp and Spanish accusative a

The most important feature of this table is the correlation that holds between degree of topicality (lower vs. higher) and degree of overt coding (less vs. more). Lower topicality, both in the case of anaphoric DOs and in the case of the DOM system, leads to less overt coding, i.e., to null objects or to the lack of accusative a. Higher topicality, by contrast, is reflected by more overt coding in each case, i.e., to an overt pronoun or to the presence of the accusative a. Notice that this pattern is essentially the opposite of what is normally found for subjects in Spanish, which have a strong tendency to occur as zeroes when they display higher topicality, but tend to occur either as pronouns or lexical NPs when their topicality is lower (Bentivoglio 1983; Schwenter & Silva 2003).

The upshot of this discussion, then, is that the variable nature of both null object and DOM systems is regulated by the same factors, including both factors inherent to the DO referent in question and also discourse-contextual factors such as topicality. It does not appear, however, that the correlation between distinct DO-related phenomena stops there, since there also seems to be a clear connection between null/overt objects and the phenomenon of clitic doubling (or lack thereof). Recent work by Colantoni (2002:335) has noted that null objects and clitic doubling “represent two opposite patterns ... Taking into account the position of the referent in the animacy scale C[litic] D[oubling] and N[ull] O[bjects] show a complementary distribution. CD is more frequent with animate referents,  

10 Indeed, there do not seem to be widely-available corpora for Paraguayan (Asunción) or Quiteño Spanish. Neither of these varieties was included in the Norma Culta project, which collected samples of Spanish spoken in the major cities of Spain and Latin America.
while NOs are found only with inanimate referents.” A similar point had already been made for Basque Spanish by Landa (1995).

The observation that clitic doubling and DOM are also linked in a similar fashion has been made recently by Estigarribia (2003), who demonstrates a strong parallelism between those DOs available for clitic doubling (but not necessarily doubled) and those which receive DOM. He states that, “The distribution of both phenomena seems to be determined by the same factors: definiteness, animacy and perhaps topicality” (2003:26), and also notes that clitic doubling is available with what are considered typologically to be the most marked DOs, i.e., those that are [+human]. The corpus-based research of Dumitrescu (1997) further sustains the claim, tentatively suggested in Estigarribia’s quote, that both accusative -a-marking and clitic doubling are strategies for enhancing the topicality of a DO.

Dumitrescu (1997:332) also argues that the absence of both of these strategies has a clear motivation: “la falta tanto de a, como de duplicación, marca el objeto directo en cuestión como carente de topicalidad y también de prominencia de ningún tipo” (‘the lack of a, as well as of duplication, marks the direct object in question as lacking in topicality or prominence of any kind’). Of course, it is overwhelmingly the case that most DOs do not occur with either accusative a or clitic doubling, i.e., they are characterized by their LACK of extra coding, just as most anaphoric DOs in null-object varieties appear as zeroes (i.e., lack of coding par excellence). Thus, clitic doubling is another DO-related phenomenon that fits the overall pattern revealed in Table 8 above: extra coding (i.e., the linguistic material to realize doubling) is added to a DO referent in order to formally mark the heightened topicality of the DO referent vis-à-vis what is normally the case for DOs. Although future research is needed to further elucidate the connections between these phenomena, it does not appear coincidental that the same factors (whether inherent or contextual in nature) recur in each case.

5. Conclusion

The distribution of null and overt anaphoric DOs in BP is affected by animacy, and secondarily by specificity. Among South American dialects of Spanish that permit null DOs, Paraguayan and Quiteño Spanish are the closest to BP with regard to the overall system of anaphoric DO marking. In these latter varieties, a combination of internal systemic pressures (e.g., the dative “usurping” the functions of the accusative) and substrate influence appears to be in operation, although the overall system of anaphoric DO marking appears now to be a solid feature of monolingual Spanish in these regions. Null objects or “object drop,” even by monolinguals, is much more frequent than normally assumed in Spanish, and provides a clear parallel to Brazilian Portuguese, i.e., it seems to be an emerging areal phenomenon in South America, despite the fact that there is no obvious evidence that any of these three varieties are in close language contact with each other.

A more important point from a theoretical perspective is the clear parallelism between the system of variable anaphoric DO marking in the varieties examined here and the well-known system of DOM in Spanish. It appears beyond doubt that the motivations for the anaphoric DO system are motivated by the same factors that motivate the system of DOM: to code “atypical” or marked DOs, due to both inherent and contextual features, in a way that signals their special status, both in the ongoing discourse and vis-à-vis DOs with more typical or unmarked referents. More research is needed to see how far these factors actually reach to influence the realization of DOs in other constructions, for instance in clitic doubling.

To close, I would like to make a plea for overcoming what I would characterize as a particular kind of “deer in the headlights” phenomenon that appears to plague Hispanic linguistics, and which, in my opinion, has kept the kind of analysis I have presented here from being presented until now. In the particular case of null objects, I believe this phenomenon has occurred in two ways. First, because null objects are clearly not a typical feature of most varieties of Spanish, they have been dealt with solely as dialectal anomalies that are (somehow) contact-induced, something which is no doubt true, but also probably not in dispute. However, despite the widespread availability of the prior analyses that I have summarized in this paper, the fact that null objects are now clearly features of some MONOLINGUAL varieties of Spanish has been almost completely overlooked. Second, also owing to the “anomalous” status of null objects dialectally, there has been very little analysis of the SYSTEM of anaphoric DO marking of which they are a part. What is abundantly clear from the quantitative data presented above,
from both BP and the Spanish varieties, is that null objects are the UNMARKED coding option for anaphoric DOs in these varieties. Therefore, instead of strictly debating the question of the genetic cause or syntactic status of NULL objects exclusively, what is at least as worthy of our attention is the question of why it is the case that some anaphoric DOs in these varieties do NOT appear as null objects. I would hope that future researchers of such seemingly “anomalous” phenomena are not blinded by the bright lights of dialectal oddities, and that they will be able to analyze these phenomena on their own terms, as part of a larger system of choices that the language presents to its speakers.

References


