

The synchrony and diachrony of differential object marking in Paraguayan Guaraní

CORY SHAIN

The Ohio State University

JUDITH TONHAUSER

The Ohio State University

ABSTRACT

This paper explores the synchrony and diachrony of differential object marking in Paraguayan Guaraní on the basis of a quantitative study of a corpus of naturally occurring data of the modern language and an investigation of object marking in a 17th-century catechism. We show that both animacy and topicality, but not definiteness, affect whether a direct object is marked in modern Guaraní, a finding that has implications for cross-linguistic theories of differential object marking, not all of which recognize topicality as a factor. We also find no categorical constraints on differential object marking in Guaraní, contrary to Bossong (1985b). Our study of the 17th-century catechism provides further support for Bossong's (1985b, 2009) claim that Guaraní did not have differential object marking when it came into contact with Spanish. The paper concludes with a discussion of the hypothesis that differential object marking in Guaraní resulted from contact with Spanish.

Differential object marking (DOM) refers to variation within a language as to whether or not a direct object is overtly marked; DOM is observed in a wide range of typologically diverse languages (cf., e.g., Aissen, 2003; Bossong, 1985a; Comrie, 1989; de Swart, 2007). The examples in (1) illustrate DOM in Paraguayan Guaraní, a Tupí-Guaraní language spoken by about four million people in Paraguay and surrounding countries. In (1), both (bold-faced) noun phrases are direct objects of the verb *o-hecha* (A3-see) 's/he/it saw': in (1a), the coordinated direct object noun phrase *Juán-chi ha Pirúlo* 'Juanito and Pirulo' is marked with the suffix *-pe*; in (1b), on the other hand, the direct object *ju'i* 'frog' is not marked with *-pe*.¹

For helpful discussions of the material, we thank Peter Culicover, Scott Schwenter, and Kathryn Campbell-Kibler, as well as the audience at the 2009 Congress on Hispanic and Lusophone Linguistics at The Ohio State University. We also thank the editor and four reviewers of the journal *Language Variation and Change* for their comments, which greatly helped improve the paper. We are grateful to our Guaraní language consultants for working with us on their language.

- (1) a. Context: While the frog was there,
 o-hecha **Juán-chi ha Pirúlo-pe** o-ñe-moĩ o-hupi i-po ichupe.
 A3-see Juan-DIM and Pirulo-PE A3-JE-put A3-raise B3-hand pronO.3
 ‘it saw Juanito and Pirulo getting ready to say good-bye.’
- b. Context: As the little boy was preparing a bath for him and his dog,
 o-hecha **ju’i** o-po-po o-hó-vo h-apykuéri-kuéra.
 A3-see frog A3-jump-jump A3-go-when B3-behind-PL
 ‘he saw the frog coming, jumping and jumping behind them.’

In the cross-linguistic literature on DOM, a number of factors have been observed to condition whether a direct object is marked or not, including animacy (e.g., Kannada [Dravidian: de Swart, 2007], Romanian [Indo-European: Mardale, 2008]), definiteness and specificity (e.g., Hebrew [Afro-Asiatic: Aissen 2003], Kiswahili [Niger-Congo: Morimoto, 2002]), topicality (e.g., Persian [Indo-European: Shokouhi & Kipka, 2003], Spanish [Indo-European: Leonetti, 2004]), telicity (Spanish: García García, 2005), and whether nonmarking would result in ambiguity (e.g., Yongren Lolo [Sino-Tibetan: Gerner, 2008], Malayalam [Dravidian: de Swart, 2007]). Languages with DOM also differ in whether the conditions on DOM are categorical or not (cf., e.g., von Heusinger & Kaiser, 2007; Tippetts & Schwenter, 2007). In Kannada, for example, animate direct objects, such as *sekretari* ‘secretary’ in (2a), must occur with the accusative marker *-yannu* (a categorical condition on marking) but inanimate objects, such as *pustaka* ‘book’ in (2b), may or may not be marked (a noncategorical condition), cf. de Swart (2007:178–9), citing Lidz (2006:11).²

- (2) Kannada (adapted from Lidz, 2006:11)
- a. Naanu **sekretari*(-yannu)** huDuk-utt-idd-eene.
 I.NOM secretary-ACC look.for-NPST-be-1sg
 ‘I am looking for a secretary.’
- b. Naanu **pustaka(-vannu)** huDuk-utt-idd-eene.
 I.NOM book-ACC look.for-NPST-be-1sg
 ‘I am looking for a book.’

This paper explores the synchrony and diachrony of DOM in Paraguayan Guaraní (henceforth Guaraní). The factors that condition DOM in the modern language are examined on the basis of a quantitative study of a corpus of naturally occurring data of modern Guaraní. Our approach thus differs from Bossong (1985b), the only other extensive discussion of DOM in Guaraní, which was based, as far as we can tell, on a nonquantitative study of texts and consultants’ judgments of isolated examples.³ The quantitative approach proves useful in identifying actual usage patterns and shows that topicality, an inherently discourse-based property not considered by Bossong, is a factor for DOM in Guaraní. We furthermore illustrate that conditions on DOM in this language are not categorical: Although Bossong claimed, for example, that definite direct objects that denote humans are categorically marked with *-pe*,

our corpus as well as our consultants' judgments show that such direct objects are acceptable both with and without the direct object marker *-pe*,⁴ as illustrated in (3).

- (3) Kuehe Juan o-hecha **pe pa'í(-pe)**.
 yesterday Juan A3-see that priest-PE
 'Yesterday Juan saw that priest.'

Our study shows that both animacy and topicality, but not definiteness, affect whether a direct object is marked in modern Guaraní. The study furthermore demonstrates that quantitative methods, which are underutilized in the exploration of DOM cross-linguistically (but see Aissen, 2003; von Heusinger & Kaiser, 2007; Tippets & Schwenter, 2007), can provide insight into, for example, subtle differences between noncategorical constraints on DOM.

The diachrony of DOM in Guaraní is explored on the basis of a study of a 17th-century Guaraní catechism. We find support for Bossong's (1985b, 2009) claim that Guaraní did not have differential object marking when it came into contact with Spanish. The paper concludes with a discussion of the hypothesis that DOM in Guaraní resulted from contact with Spanish.

ARGUMENT REALIZATION AND ARGUMENT MARKING IN GUARANÍ

This section introduces relevant aspects of argument realization and argument marking in Guaraní. Transitive and intransitive verbs in Guaraní are obligatorily inflected for person/number by either a set A or a set B cross-reference marker (cf. n. 1). Guaraní is a split-S language; the single argument of an intransitive verb, referred to as the S-argument following Dixon (1979), is cross-referenced either by a set A or a set B marker, depending on whether the verb is active or stative. With transitive verbs, either the A- or O-argument is cross-referenced on the verb, depending on which one is higher on the person hierarchy (1 > 2 > 3); see Mithun (1991) and Velázquez-Castillo (2002) for details on argument marking in Guaraní. Verbal arguments in Guaraní need not be realized by independent pronouns or full noun phrases, as illustrated by the examples in (4). In (4a), the first-person S-argument is realized only by the first-person singular set A cross-reference marker *a-* on the verbal stem *jeroky* 'dance'. In (4b), the first-person singular (direct object) O-argument of the transitive verb stem *hecha* 'see' (here realized as *recha*) is expressed with the first-person set B cross-reference marker *che-*; the (third-person) A argument is expressed by the overt noun phrase *Felipa*. In (4c), the direct object of the transitive verb stem *topa* 'find' is realized neither by a cross-reference marker (the third-person set A prefix *o-* cross-references the A-argument) nor by an overt noun phrase: The individual who is found is the (contextually salient) surgeon.

- (4) a. Context: What did you do at the party yesterday?
 A-*jeroky*.
 A1sg-dance
 'I danced.'
- b. Context: The speaker was hiding from a group of women including Felipa, but made a noise.
Felípa che-recha.
 Felípa B1sg-see
 'Felipa saw me.'
- c. Context: They went to the health station to see if the surgeon was there.
O-topa hikuái ha o-mbyvy chupe.
 A3-find pronS.3pl and A3-stitch pronO.3
 'They found him and he stitched her.'

As illustrated above, direct objects can be marked in Guaraní with the suffix *-pe*. The suffix is realized as *-pe* in oral environments, such as after *pa'i* 'priest' in (3), and as *-me* in nasal environments, such as after *kuña* 'woman':

- (5) *Kuehe a-hecha peteĩ kuñá-me*.
 yesterday A1sg-see one woman-PE
 'Yesterday I saw a woman.'

Similar to the marker used for DOM in other languages (e.g., Spanish *a*), this suffix also occurs with indirect objects and spatiotemporal dependents (where it expresses meanings such as 'at', 'in', 'on', and 'to'; Gregores & Suárez, 1967:236), as illustrated in (6a) and (6b), respectively. The suffix is glossed as PE regardless of its function, but the direct objects relevant to our study are bold-faced in the examples throughout the paper.

- (6) a. *Ña María, ña María, rei-kuaa-ko mba'é-pa o-jehu Rossáni-pe?*
 Doña Maria doña Maria A2sg-know-EMPH thing-QU A3-happen Rossani-PE
 'Doña Maria, doña Maria, do you know what happened to Rossani?'
- b. *Che-váje-pe o-ĩ va'ekue peteĩ karai*.
 B1sg-town-PE A3-be back.then one gentleman
 'In my town, there used to be a gentleman.'

The direct objects of interest for this paper are those that are contained in transitive or ditransitive clauses where the direct object is or could have been marked with *-pe*, as in (1) and (3). Thus, we exclude from consideration indirect objects and spatiotemporal dependents, because they are not direct objects, proposition-denoting and incorporated objects, because they are never marked with *-pe*, as well as direct object arguments of verbs that require those arguments to be marked with a different object marker, such as *-rehe* 'at, through, by, with, by means of' (Gregores & Suárez, 1967:240) or *-gui* 'ABL'.

ANNOTATION OF THE GUARANÍ CORPUS

The corpus we annotated consists of nine Guaraní texts;⁵ the total word count is about 6,500 words, which corresponds to about 20,000 English words as Guaraní is mildly polysynthetic. There were 233 clauses with (nonpronominal) direct objects in the corpus that conform to the inclusion criteria given in the previous section (pronominal direct objects are a special case, as discussed below). We annotated the direct objects of these clauses according to the four main factors that have been identified to play a role in DOM cross-linguistically, namely animacy, definiteness, specificity, and topicality.⁶ This section illustrates the factors in detail and identifies, for each factor, how we annotated the Guaraní corpus.

Animacy

In languages where animacy is a factor in DOM, direct objects higher on the animacy hierarchy in (7) are more likely to be marked than ones lower on the hierarchy (e.g., Aissen, 2003; Comrie, 1989; de Swart, 2007).

(7) **Animacy hierarchy:** Human > Animate > Inanimate

An example of a language where animacy plays a role in DOM is Kannada (de Swart, 2007). As illustrated in the introduction, animate direct objects in Kannada must take an accusative marker, whereas inanimate objects may or may not take a marker. Other languages for which animacy has been reported to play a role in DOM include Hindi (Indo-European: Comrie, 1989), Hup (Makú: Epps, 2008), Russian (Indo-European: Comrie, 1989), and Spanish (Indo-European: von Heusinger & Kaiser, 2003).

We determined the animacy of the 233 nonpronominal direct objects in our corpus on the basis of the denotation of the direct object. The examples in (8) illustrate a human direct object in (8a), a nonhuman animate direct object in (8b) and an inanimate direct object in (8c).

- (8) a. Context: After the woman gave birth, ...
I-memby o-heja.
 B3-child A3-leave
 'She left her child.'
- b. Ro-gueru la **ore-ñakyrã**.
 A1pl.excl-bring the B1pl.excl-cricket
 'We brought our cricket.'
- c. Context: When the boy and his dog came to the pond, ...
 O-mbo-guejy **i-mbayru-kuéra**.
 A3-CAUS-low B3-pack-PL
 'They lowered their packs.'

Because our corpus consists to a large extent of fables, where animals have humanlike abilities, anthropomorphized animals were categorized as human. Direct objects denoting body parts were annotated as animate.

Topicality

Topicality also has been shown to play a role in DOM cross-linguistically; more topical direct objects are more likely to be marked than less topical ones. Mardale (2008) and Leonetti (2004), for example, argued that DOM in Spanish is not only conditioned by animacy, but also by topicality. It is claimed that although the Spanish differential object marker *a* cannot be realized with nontopical inanimate direct objects, as in (9a), it is optional when such objects are topicalized, as in (9b).

(9) Spanish (Mardale, 2008:451, glosses added)

- a. Un buen sablazo de sol traspasaba (*a) la sacristia.
 a good cut of sun pierced the vestry
 ‘A good cut of sun pierced the vestry.’
- b. (A) la sacristia_i la_i traspasaba un buen sablazo de sol
 the vestry 3.fem pierced one good cut of sun
 ‘A good cut of sun pierced the vestry.’

Although postverbal direct objects are more topical in Guaraní than preverbal ones (Velázquez Castillo, 1995), we did not rely on the position of the object relative to the verb for the encoding of topicality, but instead annotated the direct objects in the corpus for topicality using Givón’s (1983) topicality measure, which has proven useful in previous research on Guaraní (e.g., Tonhauser & Colijn, 2010; Velázquez-Castillo, 1995). The topicality of a particular noun phrase is determined on the basis of two measures—referential distance and topic persistence. The first is a count of the number of clauses intervening between the noun phrase in question and the previous mention of its referent. The second is a count of consecutive subsequent clauses in which a mention of the referent of the noun phrase under consideration appears. Low referential distance is assumed to indicate high topicality, as is high topic persistence. We established the referential distance and topic persistence value for each of the 233 direct object noun phrases in the corpus. Following Givón (1983), we only considered the discourse referent of a noun phrase to be mentioned elsewhere if that referent was explicitly denoted by another noun phrase or was an implicit argument of another proposition, like the direct object in the first clause of (4c). Thus hypernym/hyponym pairs were not considered mentions of the same referent.

Definiteness and specificity

In languages where definiteness plays a role in DOM, direct objects that are more definite are more likely to be marked than those that are less definite (e.g., Aissen, 2003; de Swart, 2007). We assume the definiteness hierarchy of Aissen (2003, 437,444):

(10) **Definiteness hierarchy:**

- Personal pronoun
- > Proper name
- > Definite descriptive noun phrase
- > Indefinite specific noun phrase
- > Nonspecific noun phrase:

Definiteness plays a role in DOM in Hebrew, as illustrated in (11) with data from Aissen (2003:453). The preposition *'et-* (glossed 'ACC', for accusative case) is obligatorily used with definite direct objects (e.g., those marked with the definiteness marker *ha-*), as in (11a), but may not occur with indefinite ones, as in (11b).

(11) Hebrew

- a. Ha-seret her'a **'et-ha-milxama**
 the-movie showed ACC-the-war
 'The movie showed the war.'
- b. Ha-seret her'a **(*'et-)milxama**
 the-movie showed ACC-war
 'The movie showed a war.'

Our annotation of the definiteness of direct objects relies on the semantics of Guaraní noun phrase types, following Tonhauser and Colijn (2010). The first type of noun phrase in Aissen's hierarchy in (10) are pronouns. Whether direct object pronouns are part of the DOM system in Guaraní is an open question. As illustrated in Table 1, the direct object pronouns consist of the corresponding subject pronouns and the suffix *-ve*, not *-pe*, with two exceptions: the second-person plural direct object pronoun is marked with *-me*, the nasal allomorph of *-pe*, and the third-person object pronoun is the suppletive form (*i*)*chupe*.⁷ The suffix *-kuéra* on the third-person plural direct object pronoun is the (optional) plural marker.

Gregores and Suárez (1967:109,141) argued that the suffix *-ve* is an allomorph of *-pe* (cf. also Dietrich (1986:128) on Chiriguano [Tupí-Guaraní, Bolivia]). Support for this analysis is that the second-person plural pronoun is marked with (the nasal allomorph of) *-pe*. The analysis furthermore would allow for a parallel treatment of pronominal and nonpronominal direct objects in that they are both marked. One could, however, also take a position according to which *-ve* is not an allomorph of *-pe* (and the suffix *-me* of the second-person plural direct object pronoun is an allomorph of *-ve*). Support for this position is that

TABLE 1. *Subject and direct object pronouns*

Person/Number	Subject Pronoun	Direct Object Pronoun
1sg	<i>che</i>	<i>ché-ve</i>
1pl.incl	<i>ñande</i>	<i>ñandé-ve</i>
1pl.excl	<i>ore</i>	<i>oré-ve</i>
2sg	<i>nde</i>	<i>ndé-ve</i>
2pl	<i>peẽ</i>	<i>peẽ-me</i>
3sg	<i>ha'e</i>	(<i>i</i>) <i>chupe</i>
3pl	<i>hikuái</i>	(<i>i</i>) <i>chupe(-kuéra)</i>

the suffix *-ve* only occurs with object pronouns, hence its realization is not phonologically conditioned. Ultimately, the question of whether *-ve* is an allomorph of *-pe* cannot be resolved in this paper. Our quantitative analysis is restricted to nonpronominal direct objects, direct objects that exhibit variation in marking. If *-ve* is an allomorph of *-pe*, direct object pronouns are categorically marked; if *-ve* is not an allomorph, direct object pronouns are not part of the DOM system of Guaraní. We entertain these two options at the relevant places in the remainder of the paper.

The corpus included several proper names, next on Aissen's hierarchy, such as *Juan* and *Pirúlo*, which we coded as such. The next class of noun phrases on the definiteness hierarchy is definite noun phrases. Following Tonhauser and Colijn (2010), we assume that, in Guaraní, this class is defined as those noun phrases that refer to discourse- or hearer-old entities (Gundel, Hedberg, & Zacharski, 1993; Prince, 1981; Roberts, 2003). This class includes possessive noun phrases, as in (12a), demonstrative noun phrases, as in (12b), and noun phrases with the determiner *la* (borrowed from Spanish), as in (12c). The fourth type of Guaraní noun phrase that can be definite is bare noun phrases: noun phrases that only consist of nominal or adjectival predicates, but no other markers, as illustrated in (12d) and (12e). These noun phrases can refer to an entity previously introduced in discourse (discourse-old), as in the example in (12d), or to entities that had not been explicitly mentioned but are salient to the interlocutors in the discourse context (hearer-old), such as *múndo* 'the world' in the example in (12e).

- (12) a. O-ñami piko **ñande-vaka** Libória ra'e?
 A3-milk QU.EMPH B1pl.incl-cow Liboria yet
 'Did Liboria milk our cow yet?'
- b. Hasypevé-ko peina o-j-arreglá-ta **ko asúnto**.
 finally-EMPH here A3-JE-resolve-PROSP this matter
 'Finally, here this matter is getting resolved.'
- c. Context: A woman is talking about her grandmother's past chores.
 Ha upéi o-vende-pá-rire **la kamby** ...
 and then A3-sell-COMLETE-after the milk
 'And then, after she had sold all of the milk ...'
- d. Context: The boy arrives at a pond and climbs on a tree to see what he could catch.
 O-hasá-vo upéi o-hecha ju'í-pe o-guapy y mbyté-pe peteĩ yrupẽ-'ári.
 A3-pass-when then A3-see frog-PE A3-sit water middle-PE one sieve-one
 [...] Ha o-ñe-moĩ i-kuã chã'i-haguã-icha hese. [...] Juán-chi
 and A3-JE-put B3-finger bent-PURP-like at.pronO.3 Juan-DIM
 nd-o-jurá-i **ju'í-pe**.
 NEG-A3-grab-NEG frog-PE
 'In passing, he saw a frog sitting in the middle of the water on a water lily. [...]
 And he started to grab at the frog. [...] Juan didn't get the frog.'
- e. Ja-korre **múndo**.
 A1pl.incl-run world
 'We run the world.'

Guaraní has two types of indefinite noun phrases—noun phrases that denote entities that are discourse- and hearer-new: noun phrases marked with *peteĩ* ‘a/one’ and bare noun phrases. Both types of indefinite noun phrases can have a specific referent (indefinite specific noun phrase) or denote nonspecifically (nonspecific indefinite noun phrase). We follow Enç (1991) in assuming that proper names and definite noun phrases are necessarily specific. We annotated an indefinite noun phrase as specific if the speaker/writer could reasonably be assumed to have a particular individual in mind when uttering/writing the indefinite noun phrase (Donellan, 1966; Hintikka, 1973; Ioup, 1977; Partee, 1972). The examples in (13) illustrate the two types of indefinite specific noun phrases. The noun phrase in (13a) is marked with *peteĩ* ‘a/one’ and is considered specific in this discourse context because the boy had a particular dog. The bare noun phrase in (13b) is also indefinite, despite its lack of the overt indefinite determiner *peteĩ*, because it introduces the individual frog into the discourse (cf. the context of (12d)), and nothing about the context of utterance would render the frog familiar to the discourse participants. Example (13b) was furthermore classified as specific, because the boy sees a specific frog.

- (13) a. Context: There once was a man who had a son called Juan.
 O-guereko avei **peteĩ jagua**.
 A3-have also one dog
 He also had a dog.’
- b. Context: The boy arrives at a pond and climbs on a tree to see what he could catch.
 O-hasá-vo upéi o-hecha **ju’i-pe** o-guapy y mbyté-pe *peteĩ* yrupẽ-’ári.
 A3-pass-when then A3-see frog-PE A3-sit water middle-PE one sieve-on
 ‘In passing, he saw a frog sitting in the middle of the water on a water lily.’

The examples in (14) illustrate the two types of indefinite nonspecific noun phrases. The indefinite noun phrase *peteĩ mbatara* ‘a vagabond’ in the antecedent of the conditional in (14a) does not denote a specific vagabond. The indefinite noun phrase in (14b) is nonspecific because the speaker did not want to catch a particular cricket.

- (14) a. A-mono’õ-rire pe tapé-rupi *peteĩ mbatara*...
 A1sg-collect-after that path-through one fickle
 ‘After I take in a vagabond off the street ...’
- b. Context: When we were staying at my grandmother’s house, we often went into the forest.
 Ro-heka **ñakyrã** ha oi-me-raẽ-va vyra-’i ro-juga-haguã
 A1pl.excl-search cicada and A3-exist-still-RC bird-DIM A1pl.excl-play-PURP
 hese.
 with.pronO.3
 ‘We looked for cicadas and any birds there in order to play with them.’

RESULTS OF THE SYNCHRONIC STUDY OF DOM IN GUARANÍ

Of the 233 nonpronominal direct objects, 27 were marked with *-pe* or its allomorph *-me*. The results of the quantitative study suggest that animacy and topicality, but not definiteness, are factors relevant for DOM in Guaraní. After presenting our findings, we compare them to those reported in Bossong (1985b) and discuss their implications for cross-linguistic theories of DOM. In the following sections, statistical significance regarding binary feature values was determined via Fisher's exact test for independence. Statistical significance of the differences in scalar feature values, such as topicality measurements, was determined using the unpaired *t* test. All *p* values presented here are two-tailed; we assume a maximum *p* value for significance of .05.

Relevance of the individual factors

Animacy. Table 2 gives the distribution of direct objects that are marked with *-pe* ([+PE]) and those that are not marked with it ([-PE]) across the three animacy categories. Over one-half of the human direct objects were marked with *-pe* (56%) in contrast to only 5% of animate noun phrases and no inanimate noun phrases. Humans are more frequently marked than animates and inanimates to a statistically significant degree ($p < .001$ for both), and there is also a statistically significant difference in marking frequency between animates and inanimates ($p = .029$).

Because human direct objects are much more likely to be marked than animate and inanimate direct objects, animacy seems to be a factor relevant for DOM in Guaraní. Note also that of the 27 direct objects that are marked with *-pe* in the corpus, 24 are human.

The corpus data might also suggest that there is a categorical constraint against marking inanimate direct objects with *-pe* because no inanimate direct objects are marked. However, we have been able to identify examples of *-pe*-marked inanimate direct objects in texts not included in the corpus:

- (15) a. Pe tahachi n-oi-pysyrõ-i kuri **pe mohenda-há-pe.**
 that police.officer NEG-A3-save-NEG back.then that locate-NMLZ-PE
 'That police officer didn't save the computer.'

TABLE 2. *Animacy and DOM*

	[+PE]	[-PE]	% Marked
Human	24	19	56%
Animate	3	56	5%
Inanimate	0	131	0%
Total	27	206	12%

- b. Pe tahachi piko n-oi-nupā-i ra'e peteĩ mba'eyru-guatá-pe?
 that police.officer QU.EMPH NEG-A3-hit-NEG one cart-walk-PE
 'Didn't the police officer hit a car?'

Furthermore, Guaraní speakers we consulted accepted utterances where an inanimate direct object is *-pe*-marked:

- (16) a. Kuehe Aníta o-hupi **pe kysé-pe**.
 yesterday Anita A3-lift that knife-PE
 'Yesterday Anita lifted that knife.'
 b. E-ma'ẽ! Pe ita oi-nupa **ko kasõ-me**.
 A2sg.IMP-look that stone A3-hit this pants-PE
 'Look! That stone hit these pants.'

Although all four speakers we consulted preferred the variants of (16) where the direct objects were unmarked, the acceptability of the examples in (16), along with attestation of sentences like those in (15) in Guaraní texts, demonstrates that there is no categorical constraint in Guaraní against the marking of inanimate direct objects.⁸

Topicality. Table 3 gives the figures for the relation between topicality and DOM. The mean referential distance value for direct objects marked with *-pe* is significantly lower than that of direct objects not marked with *-pe* (9.37 versus 15.3) and the mean topic persistence value of direct objects marked with *-pe* is significantly higher than that of direct objects without *-pe* (3.26 versus .45).

According to these figures, direct objects marked with *-pe* are more topical than direct objects not marked with *-pe*, which suggests that topicality is a factor in DOM in Guaraní: The more topical a direct object is, the more likely it is to be marked with *-pe*.

Definiteness and specificity. Whether definiteness is relevant for DOM in Guaraní is less clear. If definiteness played a role in DOM in Guaraní, we would expect to observe a relationship between the rank of a direct object on the definiteness hierarchy and frequency of marking. This is not obviously the case, as shown in Table 4, because, for example, both proper names and indefinite specific noun phrases are more frequently marked than the definite ones.

Rather than dismiss definiteness as a relevant factor for DOM in Guaraní on the basis of the data in Table 4, we entertain the hypothesis that the definiteness hierarchy as given in (10) is not suitable for Guaraní because it makes too many subdivisions among noun phrases. An alternative classification of Guaraní noun phrases groups together proper names and definite noun phrases as definite noun phrases, and specific and nonspecific indefinite noun phrases as indefinite noun phrases:

- (17) Proper name > Definite NP > Indefinite specific NP > Nonspecific NP
 Definite > Indefinite

TABLE 5. (Revised) *Definiteness and DOM*

	[+ PE]	[–PE]	% Marked
Definite	21	139	13%
Indefinite	6	67	8%
Total	27	206	12%

TABLE 6. *Specificity and DOM*

	[+ PE]	[–PE]	% Marked
Specific	25	155	14%
Nonspecific	2	51	4%
Total	27	206	12%

the three factors as factor groups: [+/- human], [+/- topical], and [+/- specific]. With the factor group [+/- human], we distinguish direct objects that denote humans ([+ human]) from those that denote nonhuman animate or inanimate entities ([– human]), because the latter two form a natural group in that *-pe* marking is quite rare for them (cf. Table 2). Whereas topicality is a gradual notion for Givón (1983), we transformed our topicality measures into a binary opposition for the purpose of the Varbrul analysis. We considered direct objects with a referential distance of 15 or lower as topical ([+ topical]) and those with a referential distance greater than 15 as nontopical ([– topical]), with the cutoff point of 15 being the mean referential distance value of direct objects.⁹ The weight columns in Table 7 show the factor weight of the two binary factors for each factor group tested. A weight closer to 1 indicates that the factor favors marking, and a weight closer to 0 indicates that the factor disfavors marking; the range is the difference between the highest and lowest factor weights for each factor group, multiplied by 100. As shown in the table, both humanness and topicality (referential distance) contribute significantly to the model of variation, with humanness a stronger contributing factor than topicality (referential distance). Specificity did not contribute significantly to the model of variation. We, therefore, conclude that humanness and topicality (referential distance), but not specificity, condition DOM in Guaraní.

Humanness and topicality (referential distance) cannot be reduced to one or the other. As shown by the cross tabulation in Table 8, there is a statistically significant difference in marking frequency between human and nonhuman topics ($p < .001$), as well as a statistically significant difference in marking frequency between topical and nontopical humans ($p < .01$). Thus, stating the analysis strictly in terms of topicality (referential distance) or humanness alone would miss important generalizations.

TABLE 7. *Variable rule analysis results for humanness, specificity, and topicality*

Factor	Percentage of Marked Tokens	Number of Tokens	Factor Weight
[+ human]	56%	43	0.970
[- human]	2%	190	0.313
Range = 66			
[+ topical]	23%	75	0.691
[- topical]	6%	158	0.406
Range = 28			
[+ specific]	14%	180	0.566
[- specific]	4%	53	0.287

Corrected mean = .03; Total $N = 233$, Log likelihood = -42.639

TABLE 8. *Humanness, binary topicality (referential distance), and DOM*

		Topical		Nontopical		Total	
		[+ PE]	[-PE]	[+ PE]	[-PE]	[+ PE]	[-PE]
Human	No. of objects	17	5	7	14	24	19
	% marked	77%		33%		56%	
Nonhuman	No. of objects	0	53	3	134	3	187
	% marked	0%		2%		2%	
Total	No. of objects	17	58	10	148	27	206
	% marked	23%		6%		12%	

As none of the topical nonhuman direct objects are marked with *-pe*, our data also suggest that Guaraní does not provide support for Laca's (2006:431) finding that global constraints, such as topicality, can lead to a marked inanimate object, that is, can override the inherent unmarkedness of such objects that is due to local factors, such as animacy and definiteness.

Interim conclusions

The results of the preceding two sections motivate the following conclusions about the distribution of DOM in Guaraní:

(19) The distribution of DOM in Guaraní:

- a. If a direct object denotes an inanimate entity, it strongly disfavors marking.
- b. If a direct object denotes an animate entity, it disfavors marking, but it is more frequently marked than direct objects denoting inanimate entities.
- c. Whereas direct objects denoting humans are already marked with above-average frequency regardless of topicality, those direct objects that denote highly topical humans favor marking more strongly than those that denote less topical humans.

Comparison with Bossong's (1985b) analysis of DOM in Guaraní

A detailed discussion of DOM in Guaraní and related Tupí-Guaraní languages is found in Bossong (1985b). His analysis was based on data reported in grammars of Guaraní (e.g., Bottignoli, 1926; Dessaint, 1981; Gregores & Suárez, 1967; Guasch, 1976), texts contained in these grammars, a translation of the Bible, and native speaker judgments. Bossong assumed that the position of a particular direct object on the two hierarchies in (20), which he took to be universal, affects whether it is marked with *-pe* or not. The “Skala der Belebtheitsmerkmale” in (20a) is a version of the animacy hierarchy; it ranks pronouns (deix) above proper names (propr) above NPs denoting humans (hum) above NPs denoting personified animals (pers), above NPs denoting discrete entities (discr), above NPs denoting concrete entities (concr). The “Skala der Referenzmerkmale” in (20b) is a hierarchy of referential properties; it ranks objects whose referents are identifiable to both speaker and hearer above objects with referents identifiable only to the speaker, above objects with referents identifiable to neither.

(20) Bossong (1985b:3)

a. Skala der Belebtheitsmerkmale:

[+/- deix] < [+/- propr] < [+/- hum] < [+/- pers] < [+/- anim] < [+/- discr] < [+/- concr]

b. Skala der Referenzmerkmale:

[+ ident ego \wedge + ident tu] < [+ ident ego \wedge - ident tu] < [- ident ego \wedge - ident tu]

We have organized Bossong's (1985b:17–25) claims about the (non-)marking of different kinds of direct objects in Guaraní in (21) according to whether the direct object is obligatorily marked (21a), optionally marked (21b), or never marked (21c).

(21) a. **Obligatorily marked:**

- i. Definite direct objects denoting humans are always marked (p.20),¹⁰ as are personal pronouns and proper names (p.19f.).
- ii. Direct objects that denote animals that are identifiable individuals ([+ pers], ‘mit einer ‘Persönlichkeit’ (with a personality)) are always marked (p. 17) as are direct objects that denote anthropomorphized animals ([+ pers, - hum, + anim]) when the object is definite (p. 22).

b. **Optionally marked:**

- i. A direct object that is the theme and realized preverbally can be marked even when the object denotes an animate (nonhuman) entity (p. 18).
- ii. Indefinite direct objects can be marked if they denote a human ([+hum]) or an individually identifiable entity ([+ pers]) (p. 23f.).

- iii. Indefinite direct objects denoting humans can be marked but are usually not (p. 21).
 - iv. If the denotation of a direct object is not individually identifiable, the direct object is typically not marked even if it denotes a human (p. 17–18).
 - v. Direct objects that denote anthropomorphized animals ([+pers, –hum, +anim]) are usually unmarked when the object is indefinite (p. 22).
- c. **Obligatorily unmarked:**
- i. Direct objects that denote inanimate entities are never marked, regardless of whether they are definite or not (p. 24).
 - ii. Nondiscrete or nonconcrete entities are never marked (p. 24).

Bossong's analysis of DOM in Guaraní agrees with ours with respect to the types of direct object in (21b) for which the conditions on marking are noncategorical. However, our quantitative study provides counterexamples both for Bossong's claim that certain direct objects are obligatorily marked (21a) and his claim that certain direct objects are obligatorily unmarked (21c). The corpus data in (22) are empirical evidence against the first claim as both the definite direct object *pe mitã* 'that child' in (22a) and the proper name in (22b) are unmarked, contrary to (21ai).

- (22) a. Context: A child had been abandoned by its mother.
 O-mo-ngakuaa karai **pe mitã**.
 A3-CAUS-grow gentleman that child
 'The gentleman raised that child.'
- b. A-topa upépe **Libória** o-guapy.
 A1sg-find there Liboria A3-sit.down
 'There I found Liboria sitting down.'

Bossong's claim that direct objects that denote animals with a personality or anthropomorphized animals are obligatorily marked (21a_{iii}) is harder to assess because Bossong did not give principled criteria that would allow one to decide whether a particular animal has a personality or is anthropomorphized. Thus, one could take the position that the indefinite direct object *peteĩ jagua* 'a dog' in (23) is not anthropomorphized and does not have a personality, and hence is not marked with *-pe*. At the same time, one could argue that this particular dog has a personality (it has been given a name) and is anthropomorphized, as indicated by the fact that the dog later in the story chats with Juan and has independent thoughts, for example; it critiques Juan for not being able to catch the frog. If this is the case, we would expect the noun phrase to be marked on the basis of (21a_{iii})—whether it is unmarked because the dog does not have a personality or is anthropomorphized, or because marking would go against (21b_{iii}), cannot be resolved given the criteria Bossong provided.

- (23) Context: There once was a boy called Juan.
 O-guereko **peteĩ jagua** h-éra-va Pirúli.
 A3-have one dog B3-name-RC Piruli
 'He had a dog whose name was Piruli.'

Our consultants' judgments of such examples again varied, as illustrated with the discourse in (24), in which a toad, a frog, and a turtle are personified:

- (24) Mbohapy mymba o-jo-topa y rembe'ý-pe o-ñemongeta-haguã:
 three animal A3-RECIP-meet water beach-PE A3-chat-PURP
 peteĩ kururu, peteĩ ju'i ha peteĩ karumbe. Kururu i-pire-vai-étereĩ
 one toad one frog and one turtle toad B3-skin-bad-very
 o-karu heta-íterei-gui. O-hechá-vo **ju'í-pe**, oi-nupa chupe.
 A3-eat much-very-ABL A3-see-when frog-PE A3-hit pronO.3
 Karumbe o-ja'ó **kururú-pe**. Ju'i o-ñañaũ **karumbé-pe**.
 turtle A3-chastize toad-PE frog A3-hug turtle-PE
 'Three animals met at the side of a lake to chat: a toad, a frog, and a turtle. The
 toad was in a bad mood since he had eaten too much. When he saw the frog,
 he hit him. The turtle chastised the toad. The frog hugged the turtle.'

The (bold-faced) direct objects in these utterances are definite and denote either animals with a personality or anthropomorphized ones, but the three consultants we asked only judged the first of the three (*ju'i* 'frog') to be obligatorily marked, whereas marking was optional or even dispreferred for the other two (*kururu* 'toad' and *karumbe* 'turtle'). This again suggests that Bossong's claim in (21a_{ii}) may be too strong.

Bossong also argued that direct object pronouns are always *-pe*-marked. As discussed previously, this argument rests on the assumption that the suffix *-ve* is an allomorph of *-pe*. If this is the case, then direct object pronouns are obligatorily marked; in fact, they are the only kind of direct object that is obligatorily marked. If, on the other hand, direct object pronouns should not be considered *-pe*-marked versions of the subject pronoun series, then Guaraní has no direct objects that are obligatorily marked.

Evidence against Bossong's claim that the direct object types in (21c) are obligatorily unmarked comes both from our consultants' judgments and from the corpus. For example, three of our four consultants judged example (25) with the direct object *pe ita* 'that stone' to be acceptable both with the object marker *-pe* and without it.

- (25) Kuehe Juan o-nupā **pe itá(-pe)**.
 yesterday Juan A3-hit that stone(-PE)
 'Yesterday Juan hit that stone.'

Likewise, the definite and the indefinite direct object noun phrases in (15) denote inanimate entities but are marked with *-pe*. Thus, both consultants'

judgments and naturally occurring examples empirically falsify Bossong's claim in (21ci) that direct objects that denote inanimate entities are never marked.

Bossong also argued that direct objects that denote nonconcrete or nondiscrete entities are never marked with *-pe*, cf. (21cii). Nonconcrete referents are those that are not physical or tangible, but abstract (Hopper & Thompson, 1980), and include "things like air, voice, wind, and other intangibles," according to, for example, Zaenen, Carletta, Garretson, Bresnan, Koontz-Garboden, Nikitina, O'Connor, and Wasow (2004:121). Bossong (1983–1984) assumed that discrete entities are countable and that "discrete entities form a subclass of the concrete ones" (p. 9). Our corpus does not contain examples with such noun phrases, but the Guaraní speakers we consulted with about such examples gave somewhat mixed judgments. None of the three speakers considered (26a) with the nonconcrete direct object *mborayhu* 'love' or (27a) with the nondiscrete direct object *aramirõ* 'mandioka flour' acceptable with *-pe* (or its nasal allomorph *-me*). In contrast, only two of the three speakers did not consider (26b) with the nonconcrete direct object *yvytu* 'wind' or (27b) with the nondiscrete direct objects *aro* 'rice' acceptable with *-pe*, but one speaker judged both of these examples to be better with *-pe* than without.

- (26) a. Ai-pota **mborayhu**.
A1sg-need love
'I need love.'
- b. A-heka **yvytú(-pe)**.
A1sg-search wind-PE
'I search for wind.'
- (27) a. María oi-kotevẽ **aramirõ**.
María A3-need mandioka.flour
'Maria needs mandioka flour.'
- b. María o-hecha **aró(-pe)**.
María A3-see rice-PE
'Maria sees rice.'

The acceptability of (26b) and (27b) with *-pe*-marking to at least one of the three speakers suggests that a categorical rule such as Bossong's (21cii) may be too restrictive. However, a more systematic study using naturally occurring data might be better suited to shed light on this matter.

In conclusion, though we agree with Bossong that Guaraní has DOM, our study shows that the categorical constraints on DOM in Bossong's analysis are not empirically supported and that topicality plays a role in DOM in Guaraní. Our study also illustrates the value of using quantitative methods in the study of DOM (cf. also Aissen, 2003; von Heusinger & Kaiser, 2007; Tippetts & Schwenter, 2007). In contrast to the methodology used by Bossong, that is, elicitation of (binary) judgments of grammaticality (see also, e.g., Comrie, 1989; Leonetti, 2004; Morimoto, 2002; Næss, 2004; Schwenter & Silva, 2002; de Swart, 2007), our study, which combines both methods, reveals both that the

constraints are not categorical as well as subtle differences between noncategorical constraints on DOM in Guaraní (cf. Table 8 and the analysis in (19)).

Implications for cross-linguistic theories of DOM

Our quantitative analysis of DOM in Guaraní shows that topicality plays a role in determining whether a direct object is marked or not; cf. also, for example, Iemmolo (2010), Laca (1995), Leonetti (2004), and Shokouhi and Kipka (2003) for the relevance of topicality for DOM in a variety of other languages. Cross-linguistic theories of DOM, however, do not typically include topicality as a factor. The theory developed in Aissen (2003), for example, relies on animacy and definiteness: Prototypical direct objects are assumed to be inanimate and indefinite, with the result that atypical objects, those that are definite and human, are the most likely to be marked in a language with DOM. Another cross-linguistic theory of DOM is based on Hopper and Thompson's (1980) idea that, cross-linguistically, certain morphosyntactic and semantic features of a clause correspond to varying degrees of transitivity (cf., e.g., Næss, 2004, 2007; de Swart, 2007). Topicality also plays no role in this theory of DOM according to which clauses containing definite/animate objects are more prototypically transitive and thus more amenable to morphosyntactic realizations of transitivity, such as overt object marking. Though we are not the first to identify topicality as a relevant factor in DOM, our finding that topicality plays a role in DOM in Guaraní further emphasizes the need to include topicality in cross-linguistic theories of DOM. In the context of Aissen's theory, for example, the results of our study might suggest that typical direct objects denote nonhuman, nontopical entities, and that direct objects that denote human and/or topical entities are atypical and hence most likely to be marked; in terms of Hopper and Thompson (1980), the Guaraní results might suggest that clauses containing direct objects that denote human, topical entities are more prototypically transitive (see Shain, 2009, for further discussion).

THE DIACHRONY OF DOM IN GUARANÍ

This last section explores Bossong's hypothesis that Guaraní did not have DOM at the time it came into contact with Spanish in the 16th century,¹¹ but that "the formation of DOM with all its semantic properties, was actualized by language contact" with Spanish (Bossong, 2009:13). We focus here on the first part of the hypothesis, namely that Guaraní did not have DOM at the time it came into contact with Spanish. Bossong's evidence for this claim is his observation that several grammars of Guaraní written by Jesuits during the 16th and 17th centuries, such as de Anchieta (1595), de Aragona (ca. 1625), and de Montoya (1640a) did not mention the possibility of marking a direct object with *-pe*.¹² This was the case even for examples with proper name direct objects, such as (28), which de Montoya (1640a) claimed is ambiguous as to which noun phrase functions as the direct object.

- (28) Peru o-mbo'e Chua.
 Pedro A3-teach Juan
 'Pedro teaches Juan' or 'Juan teaches Pedro' (de Montoya, 1640a:35)

Bossong argued that if a direct object marker had been available to disambiguate examples such as (28), it is likely that one of these 16th- and 17th-century grammar writers would have mentioned it.

We agree with Bossong that this silence is suggestive, but we show here that his hypothesis can be strengthened by examining the marking of direct objects in early Guaraní texts. The text we exhaustively surveyed to identify whether there is any evidence for DOM is one of the earliest published in Guaraní, namely de Montoya's 17th-century catechism (de Montoya, 1640b).¹³ Our findings support Bossong's claim in two ways. First, we found no instances of direct objects marked with *-pe* or *upe*, even though the language already used *-pe* (and its allomorph *-me*) at the time to mark spatiotemporal dependents and *upe* to mark indirect objects. Second, we found several instances of direct objects that would be very likely be marked in modern Guaraní, such as human and topical direct objects, that were not marked in the 17th-century text. Some of these are given in (29) together with the referential distance (RD) and topic persistence (TP) value of the direct object referent; we added glosses as needed.

- (29) a. Ere-ayhu **Tupã**-ne mba'e pãvẽ asoce. [RD = 2, TP = 1]
 you-love God-will thing
 'Love God more than all other things.' (de Montoya, 1640b:10)
- b. A-rovia ave **Jesu Christo**. [RD = 2, TP = 9]
 I-believe Jesus Christ
 'I believe also in Jesus Christ.' (de Montoya, 1640b:5)
- c. Hũmãme Tupã **aipo Angeles** i-mõñãngi ra'e? [RD = 1, TP = 17]
 where God those angels he-create
 'Where did God create those angels?' (de Montoya, 1640b:72)
- d. Ha'e-nia o-hecha Tupã ne. [RD = 1, TP = 1]
 he he-see God will
 'He will see God.' (de Montoya, 1640b:33)

The direct objects in these examples denote God (28a,d), Jesus Christ (28b), or angels (28c). Their RD value is very low, indicating their high topicality. Despite denoting (super-)human entities and being highly topical, none of these direct objects are marked,¹⁴ thus strongly suggesting that 17th-century Guaraní did not have DOM. To conclude, our analysis of a 17th-century Guaraní text has provided further support for the hypothesis, originally formulated by Bossong, that Guaraní did not have DOM when it came into contact with Spanish.

Against this background, one can now explore the second part of Bossong's hypothesis, which is that the contact with Spanish led to DOM developing in Guaraní. This hypothesis is motivated especially because the two languages have been in contact for over 500 years, which has led to lexical and other grammatical borrowings in both languages (e.g., Choi 2000; Dietrich 1993,

1995; Giménez Caballero 1966; de Granda 1996; Usher de Herreros 1976). We cannot explore this hypothesis here but offer two points for consideration. The first point is that some but not all Tupí-Guaraní languages closely related to Guaraní also have DOM systems: Mbyá Guaraní does (Bossong, 2009, 1985b; Martins, 2003), but not, for example, Chiriguano (Dietrich, 1986:127f.). An investigation of the extent of contact between Spanish and these languages might provide insight into whether it is reasonable to conclude that Guaraní and Mbyá Guaraní, but not Chiriguano, developed DOM systems as a result of this contact. Relevant in this regard is the finding that Spanish already had DOM when it came into contact with Guaraní in the 16th century (cf. von Heusinger & Kaiser, 2005; Laca, 2006).

This also relates to our second point, which concerns the set of conditions on DOM, which differ in the three modern languages (Guaraní, Spanish, and Mbyá Guaraní). In both Guaraní and Spanish, animacy and topicality are relevant to DOM, but the two factors do not play the same roles (e.g., because human direct objects may be marked in Guaraní, and are more likely to be marked if topical, but must be marked in Spanish, regardless of their topicality), and constraints on DOM that are claimed to be categorical in Spanish (e.g., the marking of human direct objects) are not in Guaraní, although they may also not have been categorical in Spanish at the time the two languages came into contact, cf. Laca (2006). The conditions on DOM in Mbyá Guaraní seem to again be different: Martins (2003) argued that the object marker is to disambiguate transitive utterances where there is an indeterminacy as to which noun phrase is the subject and which is the object (cf. the 17th-century Guaraní example in (28)), whereas Bossong (1985b) argued that *-pe* is used in Mbyá Guaraní to mark only definite animate direct objects. Why are the conditions on DOM in Guaraní and Mbyá Guaraní different? One hypothesis is that DOM is borrowed for one particular type of direct object and then spreads to other types over time. According to this hypothesis, DOM in the two languages would be different because Guaraní has been in contact with Spanish longer than Mbyá has, see also Bossong (1985b:13), or because the two languages may have borrowed DOM for different kinds of noun phrase types. An alternative hypothesis is that perhaps the conditions on DOM in 16th-/17th-century Spanish, which Guaraní came in contact with, were different from those of the Spanish that Mbyá Guaraní came in contact with, such that DOM in Guaraní and Mbyá differs because the contact languages differ.

NOTES

1. The Guaraní examples in this paper are given in the standardized orthography of Guaraní—used in Paraguay (Ministerio de Educación y Cultura, 2004, Velázquez-Castillo, 2004:1421f.), except that all postpositions are attached to their host. Following this orthography, accents are not written for normally accented words (stress on the final syllable); stressed nasal syllables are marked with a tilde. The language has two sets of cross-reference markers (Mithun, 1991; Velázquez-Castillo, 2002): The set A prefixes (which mark transitive subjects and some intransitive ones) are *a(i)*– ‘A1sg’, *ja(i)*– ‘A1pl.incl’, *ro(i)*– ‘A1pl.excl’, *re(i)*– ‘A2sg’, *pe(i)*– ‘A2pl’, and *o(i)*– ‘A3’; the set B prefixes (which mark possessors, direct objects of transitive verbs and some intransitive subjects) are *che(r)*– ‘B1sg’, *ñande(r)*– ‘B1pl.incl’, *ore(r)*– ‘B1pl.excl’, *nde(r)*– ‘B2sg’, *pende(r)*– ‘B2pl’, and *i(ñ)*–/ñ– ‘B3’. The

two portmanteaux prefixes *ro(i)*– ‘12sg’ and *po(i)*– ‘12pl’ refer to a first-person subject and a second-person (singular/plural) object. The following glosses are used: pronO/S = object/subject pronoun, ABL = ablative case, CAUS = causative, COMPLETE = completive aspect, DIM = diminutive, EMPH = emphatic, excl = exclusive, IMP = imperative, incl = inclusive, JE = reflexive/passive, NEG = negation, NMLZ = nominalizer, PE = (in)direct object and spatiotemporal phrase marker, PL = plural, PROSP = prospective aspect/modal, PURP = purpose, QU = question, QU.EMPH = emphatic question, RC = relative clause, RECIP = reciprocal, SAY = reportative evidential, SG = singular.

2. The following glosses are used in the Kannada examples: 1sg = first-person singular, ACC = accusative case, NOM = nominative case, NPST = nonpast tense.

3. That Guaraní has DOM is also acknowledged in Gregores and Suárez (1967:156) and Velázquez-Castillo (2004:1426), who both mentioned that *-pe* occurs with direct objects that denote humans; Gregores and Suárez (1967:156) also recognized that it is sometimes used to mark direct objects that denote nonhuman animates.

4. We refer to *-pe* as a direct object marker because its morphosyntactic status is not agreed on in the Guaraní literature: whereas Adelaar (1994) and Tonhauser and Colijn (2010) assumed that *-pe* is a case suffix, Gregores and Suárez (1967) and Velázquez-Castillo (2004) considered it to be a postposition. See Shain (2009) for discussion.

5. The nine texts in the corpus were collected by the second author over the course of several years. Of the nine texts, four are fables, two are transcriptions of personal narratives of Guaraní consultants, two are transcriptions of Guaraní consultants telling a story on the basis of the picture book *A Boy, a Dog, and a Frog* (Mayer, 1967), and one is a theater play written in the 1970s. None of the texts were collected for the purposes of the current study, and all of the texts are judged natural by the Guaraní consultants we checked them with.

6. Because Bossong (1985b) as well as Shain (2009) argued that ambiguity avoidance is not a factor that conditions DOM in Guaraní, we did not include this factor here. Telicity or verbal semantics (von Heusinger & Kaiser, 2007) were also not included as factor groups in the study.

7. There are three pieces of evidence that *(i)chupe* should be considered a single, unanalyzable lexical item rather than consisting of *(i)chu* and the direct object marker *-pe*, although this may have been the appropriate analysis at an earlier stage of the language. The first is that *(i)chu* cannot appear by itself. The only other occurrence of *(i)chu* in the language is in *(i)chugui*, the third-person ablative argument, which may have been analyzable as a combination of *(i)chu* and the ablative marker *-gui* ‘ABL’ at an earlier stage of the language. As these two forms are the only ones in which *(i)chu* occurs, we argue that it is not a productive expression of the modern language. The second piece of evidence comes from the plural marker *-kuéra*, which generally intervenes between the noun and *-pe*, as illustrated in (ia). Evidence that *(i)chupe* is a fixed form comes from the fact that the third-person plural pronoun is not **(i)chukuéra-pe*, as would be expected if *(i)chupe* were analyzable as a complex form, but rather *(i)chukuéra*, as in (ib). (Cf. also *(i)chugui-kuéra*, not **(i)chu-kuéra-gui*.)

- (i) a. O-ĩ-ndaje raka'e peteĩ mitã tyre'ỹ o-hayhú-va mymba-kuéra-pe.
A3-be-SAY long.ago one child orphan A3-love-RC wild.animal-PL-PE
'There once was an orphan who loved animals.'
- b. Context: The frog followed the boy and his dog to their house.
O-puka o-hechá-vo chupe-kuéra.
A3-laugh A3-see-when pronO.3-PL
'It laughed when it saw them.'

The third argument that *(i)chupe* is not analyzable as consisting of *(i)chu* and the suffix *-pe* comes from the lexical stress system of the language. Lexical stress typically falls on the final syllable of a stem (when it does not, the stress position is orthographically marked, e.g., *apyka* ‘chair’ versus *óga* ‘house’). The suffix *-pe* does not attract stress, for example, *apyká-pe* ‘chair-PE’, not *apyka-pe* ‘chair-PE’. If *(i)chupe* consisted of the stem *(i)chu* and the suffix *-pe*, we would expect the lexical stress to fall on the final syllable of the stem—**(i)chú-pe*. The fact that *(i)chupe* is lexically stressed on the final syllable indicates that it is not a complex form.

8. Relative animacy—the animacy of the direct object in comparison to the animacy of the subject—has also been noted to play a role in DOM in some languages (cf., e.g., Tippetts & Schwenter, 2007, on Spanish). Because our corpus does not contain utterances where the subject is inanimate, and only one where the subject denotes a nonhuman animate, our quantitative study cannot identify whether this factor plays a role in DOM in Guaraní. Judgments elicited from our consultants, however, suggest that this factor should be included in future studies. First, all three speakers we consulted dispreferred the (inanimate) direct object to be marked in (ia), where the subject is more animate than the direct object, but two of the three speakers preferred *-pe*-marking in (ib), where the subject is less

animate than that of (ia) (the third speaker still preferred not to mark the direct object). Second, whereas two of the three speakers preferred the (animate) direct object not to be marked in (ia), where the subject denotes a human, the same speakers preferred the direct object to be marked in (iib), where the subject denotes an inanimate entity.

- (i) a. Peteĩ kuña o-hecha ko kasõ(-me).
 one woman A3-see this pants-PE
 'A woman sees/saw these pants.'
- b. Peteĩ mbarakaja o-hecha ko kasõ(-me).
 one cat A3-see this pants-PE
 'A cat sees/saw these pants.'
- (ii) a. María o-hecha ko vaká(-pe).
 Maria A3-see this cow-PE
 'Maria sees/saw this cow.'
- b. Pe ita oi-nupa ko vaká(-pe).
 that stone A3-hit this cow-PE
 'That stone hit this cow.'

9. Topic persistence, Givón's other topicality measure, did not turn out to be a significant factor group when included in the Varbrul analysis. It is beyond the scope of this paper to offer an explanation for this difference between referential distance and topic persistence.

10. Bossong's (1985b:20) example in (i) is a counterexample to this claim (we translated Bossong's German glosses and French translation into English; Ø indicates the absence of *-pe*-marking):

- (i) ore rendota o.me'ẽ.ta i.tajýra.Ø pe kuimba'e.pe
 our.excl chief A3.give.future his.daughter man.dative
 'Our chief will give his daughter to that man.'

Bossong posited two hypotheses to account for why this definite human direct object is not marked. The first is that the absence of marking might indicate that the daughter is not treated like a human by the father but as an (inanimate) thing; the second is phonological—the speaker avoids three uses of the sound *pe* in one utterance. Because (i) is not an isolated example (cf., e.g., (22)), we argue instead that such data constitute empirical evidence against a categorical rule of marked definite human direct objects.

11. The first Europeans arrived in the area that is now Paraguay in the early 1500s; Asunción was founded in 1537 by Juan de Salazar and Gonzalo de Mendoza. See, for example, Rubin (1985) and Gynan (2001) for the history of European settlement in the area and its effect on Guaraní.

12. These Jesuit grammars were published about 50–100 years after contact began and, therefore, do not necessarily represent precontact Guaraní. There is reason to believe, however, that at that time, Spanish influence on Guaraní had been rather limited. First, there was little Spanish immigration at the time as Paraguay was rather inaccessible and lacked natural resources (Rubin, 1985). Second, the Jesuit interest in Guaraní afforded the language a comparatively high social status; Guaraní became broadly used as the language of liturgy and education throughout the period of Jesuit involvement in the colony (Choi, 2003; Rubin, 1985).

13. Though this catechism cannot necessarily be considered naturally occurring data and was probably written in an elevated register, de Montoya seems to be regarded as a skilled linguist who produced an accurate representation of the language (see, e.g., Huonder, 1912:224, on Mulhall, 1881).

14. The suffix *-ne* on the direct object in (29a) is not a direct object marker but rather a future marker (de Montoya, 1640a:236), used for commands, as in (29a), or in statements about the future, as in (29d).

REFERENCES

- Adelaar, Willem F. H. (1994). The nasal/oral distinction in Paraguayan Guaraní suffixes. *Revista Latinoamericana de Estudios Etnolingüísticos* 8:125–133.
- Aissen, Judith. (2003). Differential object marking. *Natural Language and Linguistic Theory* 21 (3):435–483.
- de Anchieta, Joseph. (1595). *Arte de grammatica da lingoa mais vsada na costa do Brasil*. Coimbra: Antonio de Mariz.
- de Aragona, Alonzo. (ca. 1625). Breve introducción para aprender la lengua guaraní. New edition edited by Bartomeu Meliá and published in 2003 in *Amerindia* 4:23–61.
- Bossong, Georg. (1983–1984). Animacy and markedness in universal grammar. *Glossologia* 39(2–3): 7–20.

- (1985a). *Empirische Universalienforschung*. Tübingen: Narr.
- (1985b). Markierung von Aktantenfunktionen im Guaraní. In F. Plank (ed.), *Relational typology*. Berlin: Mouton de Gruyter. 1–29.
- (2009). The typology of Tupi-Guarani as reflected in the grammars of Jesuit missionaries: Anchieta (1595), Aragona (c. 1625), Montoya (1640) and Restivo (1729). *Historiographica Linguistica* 36(2/3):225–258.
- Bottignoli, Justo. (1926). *Gramática Razonada de la Lengua Guaraní*. Turin, Italy: Sociedad Editora Internacional.
- Cedergren, Henrietta, & Sankoff, David. (1974). Variable rules. *Language* 50:333–355.
- Choi, Jinny K. (2000). [–person] direct object drop. *Hispania* 83:531–543.
- (2003). Language attitudes and the future of bilingualism. *International Journal of Bilingual Education and Bilingualism* 6(2):81–94.
- Comrie, Bernard. (1989). *Language universals and linguistic typology*. 2nd ed. Chicago: University of Chicago Press.
- Dessaint, Michel. (1981). *La langue Guaraní. Esquisse d'une typologie interne du Guaraní paraguayen contemporain*. Ph.D. dissertation, Practical School of High Studies.
- Dietrich, Wolf. (1986). *El idioma chiriguano*. Madrid: Ediciones Cultura Hispánica. Instituto de Cooperación Iberoamericana.
- (1993). Mbyá, guaraní, criollo y castellano. *Signo & Seña* 3:56–71.
- (1995). El español del Paraguay en contacto con el guaraní. In K. Zimmerman (ed.), *Lenguas en contacto en Hispanoamérica*. Frankfurt: Vervuert. 203–216.
- Dixon, R. M. W. (1979). Ergativity. *Language* 55(1):59–138.
- Donellan, Keith S. (1966). Reference and definite descriptions. *Philosophical Review* 75(3):281–304.
- Enç, Mürvet. (1991). The semantics of specificity. *Linguistic Inquiry* 22:1–26.
- Epps, Patience. (2008). Hup's typological treasures. *Linguistic Typology* 12:169–193.
- García García, Marco. (2005). Differential object marking and informativeness. In K. von Heusinger, G. A. Kaiser, & E. Stark (eds.), *Proceedings from the workshop "Specificity and the Evolution/Emergence of Nominal Determination in Romance"*. Konstanz: Universität Konstanz. 17–31.
- Gerner, Matthias. (2008). Ambiguity-driven differential object marking in Yongren Lolo. *Lingua* 118:296–331.
- Giménez Caballero, Ernesto. (1966). El yopará in Paraguay. In *Congreso de Academias de la lengua española 4. Actas y labores (Buenos Aires, de 30 de nov. al 10 de dic. de 1964)*. 124–136.
- Givón, Talmy. (1983). Topic continuity in discourse. In T. Givón (ed.), *Topic continuity in discourse: A quantitative cross-language study*. Amsterdam: John Benjamins Publishing Company. 1–41.
- de Granda, Germán. (1996). Interferencia y convergencia sintáctica e isogramatismo amplio en el español paraguayo. *International Journal of the Sociology of Language* 117:63–80.
- Gregores, Emma, & Suárez, Jorge A. (1967). *A Description of Colloquial Guaraní*. The Hague: Mouton de Gruyter.
- Guasch, Antonio. (1976). *El idioma guaraní. Gramática y antología de prosa y verso*. 4th ed. Asunción: Loyola.
- Gundel, Jeannette K., Hedberg, Nancy, & Zacharski, Ron. (1993). Cognitive status and the form of referring expressions in discourse. *Language* 69(2):274–307.
- Gynan, Shaw N. (2001). Language planning and policy in Paraguay. *Current Issues in Language Planning* 2(1):53–118.
- von Heusinger, Klaus, & Kaiser, Georg A. (2003). The interaction of animacy, specificity, and definiteness in Spanish. In K. von Heusinger, G. A. Kaiser, & E. Stark (eds.), *Proceedings of the workshop "Semantic and Syntactic Aspects of Specificity in Romance Languages"*. Konstanz: Universität Konstanz. 41–65.
- (2005). The evolution of differential object marking in Spanish. In K. von Heusinger, G. A. Kaiser, & E. Stark (eds.), *Proceedings of the workshop "Specificity And The Evolution / Emergence of Nominal Determination Systems in Romance"*. Universität Konstanz, Konstanz. 33–70.
- (2007). Differential object marking and the lexical semantics of verbs in Spanish. In G. A. Kaiser & M. Leonetti (eds.), *Proceedings of the workshop "Definiteness, Specificity and Animacy in Ibero-Romance Languages"*. Konstanz: Universität Konstanz. 85–110.
- Hintikka, Jaako. (1973). Grammar and logic. In J. Hintikka, P. Suppes, & J. Moravcsik (eds.), *Approaches to Natural Languages*. Dordrecht: Reidel. 197–214.
- Hopper, Paul J., & Thompson, Sandra A. (1980). Transitivity in grammar and discourse. *Language* 56(2):251–299.

- Huonder, Anthony. (1912). Antonio Ruiz de Montoya. In *The Catholic encyclopedia*. New York: Robert Appleton Company.
- Iemmolo, Giorgio. (2010). Topicality and differential object marking: Evidence from Romance and beyond. *Studies in Language* 34(2):239–272.
- Ioup, Georgette. (1977). Specificity and the interpretation of quantifiers. *Linguistics and Philosophy* 1(2):233–245.
- Laca, Brenda. (1995). Sobre el uso del acusativo preposicional en español. In C. Pensado (ed.), *El complemento directo preposicional*. Madrid: Visor Libros. 61–91.
- (2006). El objeto directo. La marcación preposicional. In C. Company (ed.), *Sintaxis histórica de la lengua española*. Mexico City, Mexico: UNAM. 421–475.
- Leonetti, Manuel. (2004). Specificity and differential object marking in Spanish. *Catalan Journal of Linguistics* 3:75–114.
- Lidz, Jeffrey. (2006). The grammar of accusative case in Kannada. *Language* 82(1):10–32.
- Mardale, Alexandru. (2008). Microvariation within differential object marking. *Revue Roumaine de Linguistique (RRL)* 53(4):449–467.
- Martins, Marci Fileti. (2003). *Descrição e análise de aspectos da gramática do Guarani Mbyá*. Ph.D. dissertation, Universidade Estadual de Campinas.
- Mayer, Mercer. (1967). *A Boy, a Dog and a Frog*. New York: Penguin Books.
- Ministerio de Educación y Cultura. (2004). La educación bilingüe en la reforma educativa Paraguaya. Available at: www.esuelaviva-mec.edu.py/investigaciones/60.pdf, accessed August 30, 2008.
- Mithun, Marianne. (1991). Active/agentive case marking and its motivations. *Language* 67:510–546.
- de Montoya, Antonio Ruiz. (1640a). *Arte, y vocabulario de la lengua guarani*. Reprinted in 1878 by Julio Platzman, Leipzig.
- (1640b). *Catecismo de la lengua guarani*. Reprinted in 1878 by Julio Platzman, Leipzig.
- Morimoto, Yukiko. (2002). Prominence mismatches and differential object marking in Bantu. In M. Butt & T. H. King (eds.), *Proceedings of the LFG02 Conference, National Technical University of Athens, Athens*. Stanford: CSLI Publications. 292–314.
- Mulhall, Marion. (1881). *Between the Amazon and the Andes*. London: Edward Stanford.
- Næss, Åshild. (2004). What markedness marks. *Lingua* 114:1186–1212.
- (2007). *Prototypical transitivity*. Amsterdam: John Benjamins Publishing Company.
- Partee, Barbara H. (1972). Opacity, coreference and pronouns. In D. Davidson & G. Harman (eds.), *Semantics of natural language*. Dordrecht: Reidel. 415–441.
- Prince, Ellen F. (1981). Toward a taxonomy of given/new information. In *Radical pragmatics*. New York: Academic Press. 223–255.
- Roberts, Craige. (2003). Uniqueness in definite noun phrases. *Linguistics and Philosophy* 26(3):287–350.
- Rubin, Joan. (1985). The special relation of Guarani and Spanish in Paraguay. In N. Wolfson & J. Manes (eds.), *Language of inequality*. Berlin: Mouton de Gruyter. 111–120.
- Schwenter, Scott A., & Silva, Gláucia. (2002). Overt vs. null direct objects in spoken Brazilian Portuguese. *Hispania* 85(3):577–586.
- Shain, Cory. (2009). *The distribution of differential object marking in Paraguayan Guarani*. M.A. thesis, The Ohio State University.
- Shokouhi, Hussein, & Kipka, Peter. (2003). A discourse study of Persian *râ*. *Lingua* 113:953–966.
- de Swart, Peter. (2007). *Cross-linguistic variation in object marking*. Ph.D. dissertation, Radboud Universiteit Nijmegen.
- Tippets, Ian, & Schwenter, Scott. (2007). Relative animacy and differential object marking in Spanish. Presented at NWAV 36, Philadelphia.
- Tonhauser, Judith, & Colijn, Erika. (2010). Word order in Paraguayan Guarani. *International Journal of American Linguistics* 76:255–288.
- Usher de Herreros, Beatriz. (1976). *Castellano-Paraguayo. Notas para una gramática contrastiva Castellano-Guaraní*. Asunción: Centro de Estudios Antropológicos.
- Velázquez-Castillo, Maura. (1995). Noun incorporation and object placement in discourse: The case of Guarani. In P. Downing & M. Noonan (eds.), *Word order in discourse*. Amsterdam: John Benjamins. 555–579.
- (2002). Grammatical relations in active systems: The case of Guarani. *Functions of Language* 9(2):133–167.
- (2004). Guarani (Tupí-Guaraní). In G. E. Booij, C. Lehmann, J. Mugdan, & S. Skopeteas (eds.), *Morphologie/Morphology: Ein internationales Handbuch zur Flexion und Wortbildung/An*

international handbook on inflection and word-formation. Vol. 2. Berlin: Walter de Gruyter. 1421–1432.

Zaenen, Annie, Carletta, Jean, Garretson, Gregory, Bresnan, Joan, Koontz-Garboden, Andrew, Nikitina, Tatiana, O'Connor, M. Catherine, & Wasow, Tom. (2004). Animacy encoding in English. In B. Webber & D. Byron (eds.), *Proceedings of the ACL 2004 Workshop on Discourse Annotation*. 118–125.