UC Santa Cruz

Title

A variationist account of differential object marking as a contact feature in Paraguayan Guarani

Permalink

https://escholarship.org/uc/item/8pd991rh

Author

Bittar, Josefina

Publication Date

2023-12-14

Copyright Information

Peer reviewed

8 A variationist account of differential object marking as a contact feature in Paraguayan Guarani

Josefina Bittar

Abstract

This chapter explores a corpus of present-day spoken Paraguayan Guarani to assess the claim that its differential object marking (DOM) originated from contact with Spanish. Data come from a 40-hour corpus of sociolinguistic interviews conducted in Paraguay. It explores internal and external factors that favor the occurrence of locative/dative -*pe* as a human direct object marker. Overall, it was found that -*pe* suffixes to loan nouns more frequently than to native-origin nouns, and that younger – more Spanish-dominant – speakers use the marker more frequently than older speakers. Thus, it is proposed that the Guarani locative/dative marker -*pe* underwent contact-induced grammaticalization (Heine and Kuteva 2005) to become an object marker (LOC/DAT > DOM), a process triggered by loans, which entered the language as parts of larger constructions. In turn, these borrowed constructions, whose slots were initially favored by loans, became more productive over time and started recruiting native-origin elements.

Keywords: differential object marking, Paraguayan Guarani, Spanish, grammatical borrowing, grammatical variation

8.1 Introduction

This study explores data from present-day spoken Paraguayan Guarani (henceforth, Guarani) to assess the claim that the language's differential object marking (henceforth, DOM) originated from contact with Spanish. DOM is understood as the different treatment that direct objects get within a single language (Döhla 2014), which can be motivated by a wide range of factors cross-linguistically (most famously, animacy and specificity of the direct object) (Sinnemäki 2014). The two languages studied here, Guarani and Spanish, are typologically different and are widely spoken in Paraguay, a South American country where more than half of its seven million inhabitants are bilingual. Consider the following examples.[1]

(8.1) *A-mongakuaa-pa   la   che   familia-kuéra.*
       1-raise-COMPL   the   my   child-PL
       "I raised all my children".

(8.2) *O-mongakuaa-pa-ite   la   i-familia-kuéra-pe.*
3-raise-COMPL-AUG   the   3-child-PL-OBJ
"He raised all his children".

While in example (8.1) the direct object *familiakúera* ("children") does not get the object marker, in example (8.2) it is marked with the suffix *-pe*. This Ø/-*pe* alternation is common in present-day Guarani (Shain and Tonhauser 2010). Nonetheless, studies suggest that the DOM is a new feature in the language, as it is absent in 16th- and 17th-century texts (Bossong 2009). Pointing at the recent diffusion of the feature, Bossong (2009) and Shain and Tonhauser (2010) claim that the DOM originated from contact with Spanish, which generally marks human direct objects with the preposition *a* and nonhuman direct objects with Ø.

Going beyond the discussion on *when* Guarani DOM emerged, this study is concerned with *how* the feature originated: It proposes that DOM arose from a contact-induced grammaticalization process, the replica type in Heine and Kuteva's (2005) typology. In this process, speakers of the recipient language recruit internal items to replicate a grammatical category in the donor language. In the case of Guarani, following the model of Spanish, speakers recruited the locative postposition and dative marker *-pe* to form a human direct object marker.

8.2 Theoretical background

The concept of typological distance is key in contact studies because typologically different systems are expected to evolve in different ways (Aikhenvald 2002; Thomason 2008). Thus, if the unrelated contact languages show similar patterns – in particular with respect to the linguistic structures that differentiate them – a contact hypothesis can be proposed. Thus, determining whether contact with Spanish has resulted in DOM in present-day Guarani requires an examination of universal tendencies of DOM, the presence of DOM in Tupi-Guarani languages, and shared characteristics of DOM in Spanish and Guarani. This study takes a multiple causation approach to change, as internal and external factors can conspire together in the appearance and spread of a new feature (Thomason 2008).

8.2.1 Grammatical relations in Guarani

Guarani is an agglutinative language (Estigarribia 2017) with an SVO word order (Gregores and Suarez 1967). However, the language exhibits typological characteristics of SOV languages: It has postpositions, possessor-possessed word order, and a wider variety of suffixes than prefixes.

In terms of predicate argument structure, Guarani and Spanish are quite distant. One of the most interesting characteristics of Guarani is its cross-referentiality in transitive constructions: the first person is usually marked in the transitive sentence regardless of whether it functions as a subject or an object. Another key Guarani feature is its lack of nonhuman third-person direct object pronouns. Spanish does not allow for cross-referentiality, and it uses nonhuman third-person object pronouns.

8.2.2 *Differential object marking: Definition and universal tendencies*

Differential object marking is the disparity in overt marking in a language's direct objects (Döhla 2014). Many scholars have tried to explain which criteria underlie the marking of objects in these cases. Iemmolo and Klumpp (2014), for example, claim that semantic or pragmatic properties of the object – which include animacy, definiteness, and specificity – are usually the determining factors. In his typological study of DOM, however, Sinnemäki (2014) explored 744 languages and found no universal correlation between animacy or definiteness and DOM. He found that while, cross-linguistically, there is a statistically significant preference for restricted case marking (such as DOM) versus non-restricted case marking, languages rely on diverse factors to draw the lines between which elements in the language get marked and which ones do not. These factors are not limited to animacy and definiteness, but also include "common/proper [nouns], kin terms, tense/aspect, information structure, and disambiguation" (Sinnemäki 2014: 300).

Going beyond the tradition of examining semantic properties of the direct object as a motivation for DOM, some scholars have explored verb semantics and constructions. In the case of Spanish, Heusinger and Kaiser (2011) have correlated affectedness with overt marking: Verbs like "kill" and "hit" highly affect the object, therefore, the object receives the marking. Delbecque (2002), on the other hand, has proposed conflict in force dynamics between the two event actors as a determining factor in the emergence of DOM in Spanish: The more agent-like role the object takes in relationship to the subject, the more likely it is that the marked construal will be used.

In sum, contrary to traditional accounts of DOM, which put animacy and definiteness as the main influential factors of overt DO marking, languages appear to treat DOM in diverse ways (Sinnemäki 2014). Also, even when factors like animacy have been proposed as essential to DOM in certain languages like Spanish (Tippets 2011), a single factor does not account for the entire paradigm within a language. This cross-linguistic variation of DOM and the language-specific complexity of the feature point to the fact that when two typologically distant languages, like Spanish and Guarani, share similar DOM patterns, universal tendencies are an unlikely explanation.

With respect to DOM in contact, a recent study on Basque showed that Spanish loan verbs favored the use of the marker among bilingual speakers (Rodríguez-Ordóñez 2020). Rodríguez-Ordóñez followed Heine and Kuteva (2010), who claim that grammatical structure enters the language through the semantics of the loan verb.

8.2.3 *DOM in Spanish*

In broad terms, Spanish marks the direct object with *a* when its referent is human (Laca 2006), as in (8.3): *su amigo* ("her friend") takes the object marker. When the object is inanimate, it does not take a marker, as in (8.4), where *su abrigo* ("her coat") is not preceded by *a*.

(8.3)   *Llev-ó*   *a   su   amigo   a   la   escuela.*
    Take-3.SG.PST   OBJ   her   friend   to   the   school
    "She took her friend to school". (ED)

(8.4)   *Llev-ó*   *su   abrigo   a   la   escuela.*
    Take-3.SG.PST   her   coat   to   the   school
    "She took her coat to school". (ED)

Naturally, there are many exceptions to the animacy-motivated account of *a*-marked objects, as inanimate objects and nonhuman animates (like pets) sometimes receive the marking, and animate objects sometimes lack the marking. In addition, there are other factors that favor *a*-marked objects in Spanish: Specificity of the referent, position of the object in the sentence, and presence of other verbal arguments (Fábregas 2013 provides a full account of these factors).

In Spanish, DOM is present in the language as early as the 12th century (Döhla 2014). Scholars have explored the role of referentiality, verb semantics, and constructions in the usage rate of the human object marker across centuries. Overall, as seen by Company Company (2003), the use of DOM with common human nouns went from 35% in the 15th century to 57% in the 21st century. However, when classifying these nouns into human definite and human indefinite, the advancement of DOM is even more visible. According to Laca's (2006) corpus study, while in the 15th century 58% of human definite noun tokens were marked, in the 19th century, the percentage increased to 96%. Similarly, with indefinite human objects, the markedness rate went from 0% in the 15th century to 41% in the 19th century (Laca 2006). Thus, Laca proposes that DOM in Spanish has advanced following a referentiality path:

Human Proper Name > Human Def NP > Human Indef NP > Human Bare Nouns

(Laca 2006)

With respect to verb semantics, Heusinger and Kaiser (2011) proposed that the advancement of DOM in Spanish followed these paths:

definite noun phrases: PERCEPTION, FEELING, ACTION >> PURSUIT, KNOWLEDGE
indefinite noun phrases: PERCEPTION > FEELING, ACTION > KNOWLEDGE > PURSUIT

(Heusinger and Kaiser 2011)

As for the source of Spanish DOM, some scholars have argued that it is an extension of directional locative *a* (LOC > DOM), and some have proposed it originates from the indirect object marker *a* (DAT > DOM) (See Fábregas 2013: 5–10 for a full review). This study adopts a neutral position in this discussion, and it is anchored in the assumption that DOM in Spanish followed this grammaticalization path: LOC / DAT > DOM.

Finally, speaking from a more synchronic point of view, Delbecque (2002) has claimed that the *a/Ø* alternation depends on how the event is construed. If the patient-like participant is perceived as actively responsible for the agent's experience, then the *a*-construal will be used. This explains why a sentence like *Dejé a Madrid* ("I left Madrid") is possible even when *Madrid* is a nonhuman object.

8.2.4 *DOM in Paraguayan Guarani and other Tupi-Guarani languages*

In Guarani, the human direct object is marked with *-pe* (Gregores and Suárez 1967) (not to be confused with *pe*, the distal demonstrative), but it has been described as optional (Shain and Tonhauser 2010). Thus, both sentences in (8.5) and (8.6) are acceptable in the language:

(8.5)   *Ai-kuaa   nde   sy.*
         1-know   your   mother
         "I know your mother". (ED)

(8.6)   *Ai-kuaa   nde   sý-pe.*
         1-know   your   mother-OBJ
         "I know your mother". (ED)

The Guarani DOM has been attested since the 20th century (Shain and Tonhauser 2010). As Shain and Tonhauser (2010) noted, the available diachronic data points to the fact that when contact between Spanish and Guarani began, in the 16th century, Spanish exhibited DOM but Guarani did not. This discrepancy is essential to argue for any contact-induced change (Thomason 2001). While it is known that the frequency of DOM in Spanish increased with the centuries, there are no studies on Guarani that explore a diachronic frequency change of this feature.

With respect to Guarani, Tonhauser and Shain's (2010) study remains, to my knowledge, the only quantitative exploration of DOM in Paraguayan Guarani or any other Tupi-Guarani language. Of the data they analyzed (43 tokens), 56% of human direct objects were marked. Their findings also show that, unlike Spanish, definiteness does not appear to play a role in the development of DOM. Instead, they propose animacy, and to a lesser extent, topicality, as the determining factors in the expression of the object marker. Bossong (2009), however, claims that it is relative animacy that determines the occurrence of the marker: When subject and object nouns are equal in strength on the hierarchy, the marker is obligatory (243).

Regarding other Tupi-Guarani languages, Bossong (2009) states that Mbyá or Apopokuva show patterns of DOM but its usage is not as advanced as in Paraguayan Guarani. Roessler (2019), however, finds DOM in four Tupi-Guarani languages, which she groups in a cluster: Paraguayan Guarani, Mbyá, Avá/Chiripá, and Paĩ-Tavỹterã/Kaiowá. She claims that DOM in the four languages can be accounted for by animacy and specificity, but that the patterns are understudied: Paraguayan Guarani is the only of the four languages whose DOM has

been discussed in the literature. Roessler, Gasparani, and Danielsen (2014) are skeptical about DOM being a contact feature because it is present in all languages of the Tupi-Guarani sub-group 1 (which includes the aforementioned four languages and Aché, Chiriguano, and Ñandeva), and in Guarayo (sub-group 2).

Finally, indirect objects in Guarani are obligatorily marked with *-pe*. As early Jesuit grammars of Guarani give examples of *-(u)pe* as a dative marker, it could be assumed that this use precedes colonial times, as seen in the following examples:

(8.7)   *A-há   Perú   upê.*
        1-go   Pedro   LOC
        Spa: voy à Pedro
        "I go towards Pedro". (Ruiz de Montoya 1639: 76)[2]

(8.8)   *A-rahá   Perú   upê.*
        1-take   Pedro   DAT
        Spa: Llevólo à Pedro
        "He took it to Pedro".(Ruiz de Montoya 1639: 76)

(8.9)   *Ai-meê   xe-rúba-pé.*
        1-give   my-father-DAT
        Por: deyo á meu pay
        "I give it to my father".
        (Anchieta 1595: 90)

Ruiz de Montoya (1639) and Anchieta (1595) associate this dative marker with the indirect object pronoun *(i)chupe*. Ruiz de Montoya (1639) also relates it to the locative *-(u)pe*. Thus, as other researchers have proposed (e.g., Gimeno 2012), based on common paths of grammaticalization (Croft 2002), locative *-pe* grammaticalized into a dative marker. As *-pe* had these two functions (locative and dative) before colonial times, it is difficult to distinguish which one was the source of the DOM. Thus, we propose the following grammaticalization process: LOC / DAT > DOM. Thus, as could be seen, Spanish *a* and Guarani *-pe* show similar behaviors across three functions: Locative, indirect object marking, and differential object marking.

8.2.5 *Replica contact-induced grammaticalization*

According to Heine and Kuteva (2005), there are two types of contact-induced grammaticalization: Ordinary grammaticalization and replica grammaticalization. In both types, speakers of one language identify a grammaticalization process in the donor language, and replicate it in the recipient language, by forming an analogy. The difference between the two types is that the ordinary type entails common grammaticalization paths documented cross-linguistically. The replica type refers to rare processes identified in two systems in contact; the only explanation for why these processes took place in both systems is language contact, that is, speakers of the recipient language assumed this process happened in the donor language and replicated it. Thus, the replica type requires more agency, noticing,

and awareness on the part of the speakers. In the Guarani-Spanish scenario, Guarani speakers seem to have "replicated" the Spanish process of a locative adposition/dative marker grammaticalizing to a human direct object marker, but following a common path. In fact, as noted earlier, parallel diachronic processes have been identified in many Tupi-Guarani languages, as well as several others which appear to have DOM as a result of contact: Italian Romanian in contact with Standard Italian (Cohal 2014), Greek in contact with Turkish (Karatsareas 2020), and Basque in contact with Spanish (Rodríguez-Ordóñez 2020). With the terminology of Matras and Sakel (2007), Guarani DOM is a case of borrowing of pattern without matter, that is, the case-marking function (pattern) of Spanish *a* gets borrowed, but not its form (matter).

Finally, conversations being the most frequent type of discourse, it could be argued that they are also the locus of contact-induced change. This is not a foreign idea to usage-based approaches, which have placed daily linguistic interaction at the core of language change in general (e.g., Bybee 2015). In addition, variationist approaches also contribute to the exploration of contact-induced change, as they stem from the understanding that change is gradual and incremental, and emerges from vernacular, non-standard styles (Poplack 2020).

8.3 Research questions and hypotheses

This study aims at exploring two questions: One specifically regarding the origin of DOM in Paraguayan Guarani, and the other concerning the mechanisms of language borrowing in general. Thus, the research questions are the following:

1. Is DOM in Guarani a contact feature?
2. If DOM is a contact feature, how was it borrowed?

As for the first question, following Bossong (2009) and Shain and Tonhauser (2010), this study supports the hypothesis that DOM in Guarani is a contact feature. It originated by replicating a Spanish pattern. Regarding the second question, while Shain and Tonhauser (2010) provide evidence of the absence of DOM in the early stages of contact with Spanish, the present study provides an account of how this pattern emerged and spread in Guarani. It attempts to demonstrate that the clues of contact can be found in natural speech, as suggested by both usage-based approaches and variationist approaches to language change.

With respect to the mechanisms of borrowing, this study hypothesizes that the Guarani DOM was borrowed as an element of a Spanish construction and not as an independent feature. This borrowed Spanish construction included a transitive verb, a human direct object, and the DOM: [V DOM NPhuman]. It is hypothesized that the borrowing of the DOM Spanish construction can be attested synchronically in Guarani, by showing that the presence of loan verbs or loan nouns will predict the occurrence of the DOM. Loanwords are of special interest in determining the contact origin because they indicate, somewhat irrefutably, that contact-induced change took place. Thus, if the grammatical feature under study

appears more often with loanwords, it can be argued that they share origins, that is, that both the loanword and the grammatical loan are products of contact.

The sentences (8.10) and (8.11), produced by the same speaker (a 31-year-old man from a rural area), exemplify the predictions. Both utterances include the phrase "to help our farmer(s)" which, in turn, includes the loan *agrikultor*, from the Spanish *agricultor* ("farmer"). However, in (8.10), the speaker uses the loan *ajuda*, from the Spanish *ayudar* ("to help"), while in (8.11) he uses the native-origin counterpart *pytyvõ*. The object marker *-pe* occurs in (8.10), the sentence that includes the loan verb, but not in (8.11), the sentence with the native-origin verb.

(8.10) *Nd-aipó-ri*   *peteĩ*   *política*   *en*   *sí*
NEG-there.is-NEG   one   policy   in   itself
"There isn't a policy in itself …
o-ajuda-haguã-icha   *ñande*   *agrikultor-es-kuéra-pe*
3-help-to-as   our   farmer-PL-PL-OBJ
as to help our farmers".

(8.11) *Sa'i [...]*   *o-me'ẽ-a*   *la*   *goviérno*   *la*
Little   3-give-NMLZ   the   government   NMLZ
"It is little what the government gives …
oi-pytyvõ-haguã   *ñande*   *rapicha*   *agrikultor*
3-help-to   our   fellow   farmer
to help our fellow farmer".

Likewise, examples (8.12) and (8.13) refer to the action of raising children, both including the native-origin verb *mongakuaa* ("raise"). However, (8.12), which was produced by a 68-year-old woman from the city, refers to children using the native-origin noun *membykuéra*, while (8.13), which was produced by a 33-year-old man from the same neighborhood, uses the loan *família*, from the Spanish *familia* ("child"). Of the two sentences, only (8.13) includes the marker *-pe*, which suffixes the loan.

(8.12) *E-mongakuaa-porã*   *la*   *ne*   *memby-kuéra.*
2-raise-well   the   your   child-PL
"You raise your children well".

(8.13) *O-mongakuaa-pa-ite*   *la*   *i-família-kuéra-pe.*
3-raise-COMPL-AUG   the   3-child-PL-OBJ
"He raised all his children".

Thus, following examples (8.10) through (8.13), I predict a loan verb or an object noun, or the combination of both, will favor *-pe*. In summary, I hypothesize that:

1. DOM in Guarani is a contact feature, that is, a feature that emerged from contact with Spanish.
2. Guarani borrowed the Spanish DOM construction [V DOM OBJhuman], which led to the grammaticalization of the locative/dative marker into a human direct object marker.

8.4 Methods

This study explores DOM in a corpus of spoken Guarani comprised of 40 sixty-minute-long interviews. The interviews were recorded by me in two sets of sessions: the first one, in June 2015, in the Asunción metropolitan area (Paraguay's capital); the second one, in October 2019 and January 2020, in the countryside (San Juan Nepomuceno and surrounding towns). All the speakers were recruited with the snowball sampling method and were self-assessed native speakers of Guarani.

Most of the Asunción interviews were recorded in the Bañado Sur area, a group of neighborhoods by the river. The main interviewer was Israel Pedrozo Candia, a well-known member of the community and a self-assessed native Guarani speaker. I, a heritage-language speaker of Guarani, was present in the interviews, handling equipment at the beginning of the interview and asking questions toward the end. The same logistics were applied to the countryside interviews, which were recorded in San Juan Nepomuceno, a town located 200 kilometers from Asunción, which has an urban downtown but is comprised mainly of rural communities. The main interviewer was Antonio Adrián Zena Mereles, a San Juan native, who is a self-assessed native Guarani speaker and a well-known person from a well-known family in the town.

The interviews aimed at eliciting the vernacular style of the speaker and thus followed the interviewing practices proposed for sociolinguistic fieldworkers (e.g., Labov 2001; Shilling 2013: 92–112). The interviewee was asked basic demographic questions at the beginning but then they spoke about a topic of their interest with little to no intervention from the interviewers. In some interviews, family or community members were present, mainly listening and sometimes intervening from time to time: This aided in the prompting of the interviewee's colloquial style.

The corpus for this study has 20 Asunción interviews and 20 San Juan interviews. The Asunción interviews were selected from a larger corpus of 35 interviews. The main criterion in the selection was interviewer/interviewee speaking time ratio (the longer the interviewee speaking time, the more likely to be selected). Demographic balance was the other criterion: The selection includes ten female and ten male speakers, from a wide range of ages: 18 to 75. All the 20 San Juan interviews were recorded for the purpose of this study: ten speakers were female, and ten were male, from an age range of 18 to 85.

The 40 interviews were fully transcribed in the software Transcriber by native Guarani speaker Antonio Adrián Zena Mereles and me. The interviews transcribed by me were verified by Mr. Zena. From there, transcriptions were combined into a single Excel spreadsheet, which included demographic information of the speaker: Name, age, sex, and location. All tokens with full noun human direct objects were manually selected and copied onto a new spreadsheet.

Previous studies (Bossong 2009; Shain and Tonhauser 2010) and preliminary exploration of this study's data showed that *-pe* is not used with inanimate objects and, thus, there is no variation with respect to its use. This was verified by the data: Inanimate direct objects were categorically unmarked. Nonhuman animal

direct objects were marked with *-pe* in four of the 71 extracted tokens, which shows a strong preference for zero-marking. Human direct objects, however, show more variation and represent a much larger portion of the animate direct object tokens. Thus, only full human object noun phrases were included in the statistical analysis. Also, as the purpose of this study is to test the influence of Spanish on Guarani, in addition to having a direct human object, the selected tokens had to meet the following criteria: Their translation to Spanish had to include *a*, and the use of *-pe* had to be variable. To my knowledge, there are no comprehensive accounts of DOM in Paraguayan Spanish, therefore, for the first criterion, I relied on my own use of DOM in this variety (native to me) to decide whether the translation required a DOM *a* or not. For example, sentence (8.14a) includes a human direct object without the marker. However, its translation to Spanish (8.14b) would not include the marker either, because the woman (60) is talking about wanting a generic non-specific partner. Tokens like these ones were not included. For the second criterion, I consulted with native speaker Mr. Zena to verify that both *-pe* and lack of *-pe* were allowed in each token.

(8.14a)   *Maerã-piko   che   ai-pota   kompañéro?*
          Why-INT   I   1-want   partner
          "Why would I want a partner?"

(8.14b)   ¿Por qué querría yo (*a) un compañero?

Interestingly, the marker *-pe* never occurred in a sentence that would not include the marker *a* in Spanish. The opposite, however, was true for some sentences, that is, *-pe* was not allowed in sentences whose Spanish translation would require *a*, per my consultant Mr. Zena. These tokens comprised sentences with nominalized transitive verbs and sentences with three human participants. Both types of sentences did not allow the DOM marker, and thus, they were not included in the statistical analysis, as they are not variable.

More specifically, data show that *-pe* is not used with nominalized verbs (*-a* is the nominalizer, careful pronunciation: *-va*) which take complements, like in (8.15), where *grúpo de hóvenes* ("youth groups") is complementing the nominalized verb *omba'apóa* ("who work"). However, *-pe* occurs, optionally, with nominalized verbs that do not take complements, like in (8.16) and (8.17). In (8.16), the marker suffixes the nominalized verb *oka'úa* ("those who drink") but in (8.17), the marker does not suffix the nominalized verb *ouramóa* ("those who just came").

(8.15) *Che nd-ai-kuaa-i   la hénte o-mba'apó-a   la   grúpo de hóvenes.*
         I   NEG-1-know   the people 3-work-NMLZ   the group of young
             -NEG   people
        "I don't know the people that work with youth groups".

(8.16) *Nd-o-hecha-sé-i-oi   la   o-ka'ú-a-pe.*
         NEG-3-see-VOL-NEG-EMPH   the 3-drink-NMLZ-OBJ
        "He really doesn't want to see those who drink".

(8.17) *Ro-hecharamo-iterei   la   o-u-ramó-a.*
        1.EXCL-appreciate-AUG   the 3-come-REC.PST-NMLZ
        "We truly appreciate the newcomers".

The variable use of the marker in intransitive subordinate clauses and the categorical unmarkedness of transitive subordinate clauses were confirmed by my consultant, with whom I elicited sentences like (8.18) (with an intransitive subordinate clause) and (8.19a) (with a transitive subordinate clause), followed by its ungrammatical versions (8.19b) and (8.19c).

(8.18) *Ei-kuaa   pe   mitarusu   o-pitá-a(-pe).*
        2-know   that young.man   3-smoke-NMLZ(-OBJ)
        "You know the young man who smokes". (ED)

(8.19a) *Ei-kuaa   pe   mitarusu   o-pitá-a   mariuána.*
        2-know   that young.man   3-smoke-NMLZ   marihuana
        "You know the young man who smokes marihuana". (ED)

(8.19b) * *Ei-kuaa   pe   mitarusu   o-pitá-a-pe   mariuána.*
        2-know   that young.man 3-smoke-NMLZ-OBJ marihuana

(8.19c) * *Ei-kuaa   pe   mitarusu   o-pitá-a   mariuána-pe.*
        2-know   that young.man 3-smoke-NMLZ   marihuana-OBJ

This Guarani pattern differs from Spanish in that the Spanish DOM *a* is not discouraged when the DO is a transitive subordinate clause. The marker is used with both intransitive and transitive subordinate clauses, as in (8.20a) and (8.20b), respectively. That is, the type of subordinate construction (intransitive versus transitive) does not predict DOM in Spanish. However, it appears to predict it in Guarani.

(8.20a) *Conoc-es   a   ese   joven   que   fum-a.*
        know-2.PRES   OBJ that young.man   who smoke-3.PRES
        "You know the young man who smokes". (ED)

(8.20b) *Conoc-es   a   ese   joven   que   fum-a   marihuana.*
        know-2.PRES OBJ that young.man who smoke-3.PRES   marihuana.
        "You know the young man who smokes marihuana". (ED)

Thus, in general, Guarani DOM is not used with transitive subordinate clauses. However, there is an exception to this pattern. When the DO in the subordinate clause has a human referent, the marker can be used as a disambiguation device. In examples (8.21) and (8.22) *-pe* indicates the object of the dependent clause: In (8.21) it is *tia* ("aunt"); in (8.22) it is *mitarusu* ("young man"). The disambiguation function of the DOM is further evidenced by the ungrammatical example (8.23), in which both nouns take the object marker.

(8.21) *Ai-kuaa   pe   mitarusu   o-hayhú-a   che   tiá-pe.*
        1-know   that young.man   3-love-NMLZ   my   aunt-OBJ
        "I know the young man who loves my aunt". (ED)

(8.22)   *Ai-kuaa   pe   mitarusu-pe   o-hayhú-a   che   tia.*
        1-know   that   guy-OBJ   3-love-NMLZ   my   aunt
        "I know the young man whom my aunt loves". (ED)

(8.23)   ** Ai-kuaa   pe   mitarusu-pe   o-hayhú-a   che   tiá-pe.*
        1-know   the   young.man-OBJ   3-love-NMLZ   my   aunt-OBJ

The aforementioned Guarani pattern differs from Spanish in that a translation of (8.21) would require the use of the marker before both human direct objects (8.21b).

(8.21b)   *Conozc-o   a   ese   muchach-o   que am-a   a   mi tía.*
        know-1.PRES   OBJ   that   youngster-MASC   that   love-3.PRES   OBJ   my   aunt
        "I know that guy who loves my aunt". (ED)

Another interesting event type where the marker is not used in Guarani is in simple transitive constructions with three human participants. Likewise, in many dialects of Spanish, when human DO and human IO are present in a sentence, marking the human DO is discouraged (Moreno-Fernández, Penadés-Martínez, and Ureña-Tormo 2019), as seen in example (8.24a) and its ungrammatical version (8.24b). The Guarani translations of these sentences also discourage the direct object marker, as seen in (8.25a) and (8.25b).

(8.24a)   *Le   d-i   mi   hij-o   a   mi   tía.*
        IOBJ   give-1.PST   my   child-MASC   IOBJ   my   aunt
        "I gave my son to my aunt". (ED)

(8.24b)   ** Le   d-i   a   mi   hij-o   a   mi   tía.*
        IOBJ   give-1. PST   OBJ   my   child-MASC   IOBJ   my   aunt

(8.25a)   *A-me'ẽ   che   memby   che   tiá-pe*
        1-give   my   child   my   aunt-IOBJ (ED)

(8.25b)   ** A-me'ẽ   che   memby-pe   che   tiá-pe*
        1-give   my   child-OBJ   my   aunt-IOBJ

Corpus data also evidences the absence of the DOM in events with three human participants (examples (8.26), (8.27), and (8.28)).

(8.26)   *O-u   o-raha-pa   ndehegui   la   mitã.*
        3-come   3-take-COMPL   from.you   the   child
        "They come and take all your children from you".

(8.27)   *O-heja   la   mitã   pe   kuñatai-me.*
        3-leave   the   child   that   lady-IOBJ
        "She leaves the child to that lady".

(8.28)   *A-guerú-ta   peẽ-me   pende   arpista-rã,   peteĩ   mitã-'i.*
        1-bring-FUT   3.PL-IOBJ   your.PL   harpist-FUT   a   child-DIM
        "I will bring you your future harpist, a kid".

Interestingly, the DOM is allowed in constructions where the third participant, the receiver, takes the *-ndi* postposition, equivalent to Spanish *con*, as seen in example (8.29).

(8.29) *E-heja   la   nde   memby   michĩ-a-pe   la   iñ-erman-o   majór-ndi*
       2-leave the your   child   little-NMLZ-OBJ the 3-sibling-MASC   older-with
       "You leave your little child with their older brother …
       o   iñ-erman-a   majór-ndi.
       or   3-sibling-FEM   older-with
       or with their older sister".

Finally, the variable tokens (those where the DOM was optional) were coded. The dependent variable was the presence or absence of *-pe*. The linguistic independent variables included were: Linguistic origin of the object noun (Guarani or Spanish) and linguistic origin of the verb (Guarani or Spanish). In addition to the mentioned linguistic factors, demographic information of the speaker – age (continuous), sex (female or male), and location (countryside or capital) – were included. With respect to location, the variable *countryside or capital* was included because, traditionally, more rural-like areas have been more Guarani-dominant while more urban areas have been more Spanish-dominant (Rubin 1968). Thus, location could have an effect on the production of any presumed contact-induced change: If the feature is more used in urban areas, where Spanish is more widely used, a contact effect argument could be made.

In summary, the variables that were accounted for in the study were:

The dependent variable: Absence vs. presence of DO marker *-pe*

Social independent variables: Age (18–85), sex (male or female), location (rural or urban)

Linguistic independent variables: Origin of the predicate's object noun (Guarani or Spanish), linguistic origin of the predicate's verb (Guarani or Spanish)

Finally, all interviewee-produced tokens (N = 613) were correlated with the aforementioned variables through a logistic regression in R.

8.5 Results

Of the 613 tokens produced by the interviewees, 44% were marked with *-pe*. The logistic regression showed the age of the speaker ($p < 0.001$) and the language origin of the object noun ($p = 0.01$) as the statistically significant factors favoring the overt marking of the direct object. Details from the regression are shown in Table 8.1. The younger the speaker, the higher their usage rate of *-pe*, as seen in Figure 8.1.

In addition, when the object noun is a loan from Spanish, instead of a native-origin noun, the marker is more likely to occur, as summarized in Table 8.2.
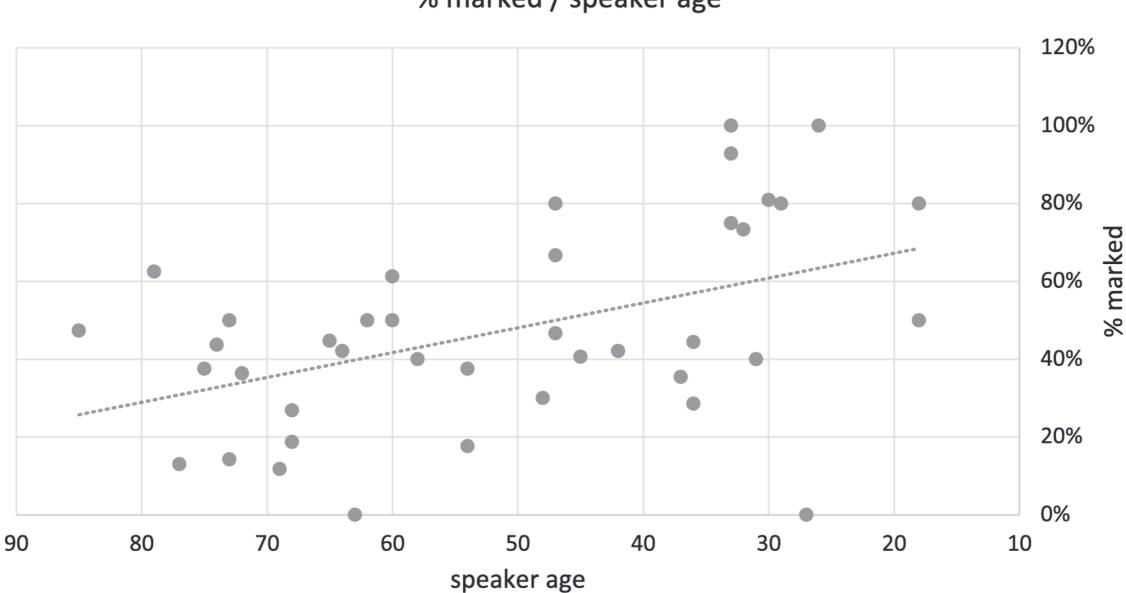
Overall, the other social factors which were not determined to be significant (location and gender) did not have an effect on the dependent variable. The

Table 8.1 DOM logistic regression results

	Estimate	Std. error	z value	Pr(>|z|)
(Intercept)	0.966050	0.354232	2.727	0.00639**
sexMale	0.239504	0.180809	1.325	0.18530
locationU	0.030445	0.174068	0.175	0.86115
age	−0.028253	0.005376	−5.255	1.48e-07***
ObjetoS	0.435645	0.172585	2.524	0.01160*
VerboS	0.106409	0.185932	0.572	0.56712

Figure 8.1 Rate of DOM usage per speaker.

Table 8.2 Marking across Guarani-origin nouns vs. Spanish-origin nouns

Linguistic origin of the noun	Not marked	pe-marked	Total
Guarani	192 (61%)	124 (39%)	316
Spanish	151 (51%)	146 (49%)	297
Total	**343**	**270**	**613**

correlation of age with the presence of the DOM was unexpected. However, this finding is compatible with our predictions, as will be explained in the analysis section.

With respect to internal factors, it was predicted that the presence of loan verbs and Spanish-origin object nouns would favor the occurrence of -*pe*. This was true only for the object nouns, as the presence of loan verbs did not have an effect on the marker (at least, not synchronically).

A further exploration of the data showed that certain verbs, regardless of origin, correlate with a higher presence or absence of the marker. For example, "help", "respect", and "listen" showed a very high rate of DOM presence (80–100%) while "take", "correct", and "put" show a low presence of DOM (0–7%). Likewise, it was found that direct objects whose referent was young people (e.g., "my children") were marked less often than objects that referred to adults, 25% compared to 55%, respectively.

8.6 Analysis

The logistic regression showed that *-pe* was more likely to occur in the presence of borrowed object nouns from Spanish, but not in the presence of loan verbs. However, the most significant predictor of the occurrence of the marker was the age of the speaker. Both of these results provide further evidence for *-pe* as a contact feature in Guarani and support the contact hypothesis.

First, the fact that a loan object noun favors the occurrence of the Guarani human object suffix suggests that the DOM is a contact feature. The borrowed lexical item proves the existence of a contact phenomenon in the sentence; that is, in an utterance like *oheja iprofesor* ("leave his teacher"), which includes the loan *profesor* ("teacher"), contact is self-evident. Results indicate that the presumed borrowed grammatical feature is used more frequently in sentences like these, where contact is evident, and thus, I argue the DOM emerges in this type of context. Put differently, if *-pe* were not a contact feature, it would not matter whether the suffixed noun was of native or Spanish origin. This pattern is supported by usage-based approaches to contact-induced change, which posit that items that are frequently used together may be borrowed together (Backus 2013) regardless of whether the items are grammatical or lexical. Furthermore, it was the object loan and not the verbal loan that had an effect on the object marker, a pattern that can also be accounted for by co-occurrence. In Guarani, thus, when borrowing a noun that occurs with a case marker, the speaker might borrow the schematic unit [DOM OBJ]. Another explanation is that verbs played a role in the emergence of the pattern, as we hypothesized the DOM was part of a "larger" loan, the DOM Spanish construction: [V DOM OBJhuman]. However, over time, the DOM extended to verbs of native origin, and thus, the verb effect was lost.

Second, in Paraguay, age correlates with the degree of bilingualism. Therefore, age being a highly significant predictor of the DOM's presence is another piece of evidence that the feature is replicated from Spanish. The younger the person, the more likely they are a self-assessed bilingual. In addition, according to census data, the percentage of monolingual Spanish speakers increased by 5% (from 10% to 15%) from 2002 to 2012 (Paraguay 2003; Paraguay 2016). During fieldwork, this generational difference with respect to language preferences was highly evident. In rural areas, while older speakers would not speak Spanish to me, even knowing it is my dominant language, younger speakers would address me in Spanish when meeting me for the first time. In urban areas, it was the younger speakers who would code-switch the most before, during, and after the interviews.

As for the second question in our study – how does the DOM enter the Guarani grammar? – results suggest that Guarani speakers did not borrow *-pe* alone, but as an item in the Spanish DOM construction [V DOM OBJ]. At the beginning, this Spanish construction included loanwords. In turn, the borrowing of this construction led to the replica grammaticalization of the Guarani locative/dative marker to a human direct object marker. As part of its ongoing conventionalization (and as bilingualism increases in Paraguay), the Guarani DOM construction "accepts" more native-origin verbs and object nouns. This pattern is expanded by younger speakers as they not only include more Spanish loanwords into Guarani but also extend its use to native-origin nouns.

Another finding in this study is the effect of verb meanings and relative animacy in the occurrence of the DOM. These factors were not included in the logistic regression, as the token numbers vary widely across verbs. Nonetheless, the patterns found provide further explanations on the DOM in Paraguayan Guarani but also give insight for future research. With respect to verb meanings, it was found that across verbs with five tokens or more, there was a wide range of DOM usage. On one end, three verbs categorically correlated with the absence of the marker: "correct", "put", and "gather". On the other end, one verb categorically correlated with the presence of the marker: "help". With respect to relative animacy, object nouns were more likely to occur with the DOM when they have adult referents than child referents, as summarized in Table 8.3.

These two effects (verb meanings and relative animacy) are better accounted for when explored together. For example, there are ten tokens of the loan *korrehi* "correct", all of which have objects that refer to children: *mitã* ("child"), *família* ("offspring"), *alúmno* ("student"); none of them gets the case marker. Thus, it could be argued that the specific collocation (and not the verb or the noun alone) "correct + child" discourages the marker. In this type of event, it would be unexpected to have a child correct an adult. As proposed by Comrie (1989) for universal patterns, Delbecque (2002) for Spanish, and Bossong (2009) for Guarani, a marker would be expected when the object is higher in animacy than the subject. In this case, it could be argued that while children and adults are both equally animate, the prototypical correcting event would have an adult subject and a child object. If these roles are reversed, a less prototypical relation is construed and therefore the marker is needed, but tokens like "children correct their parents" are not available in the data. However, other verbs can illustrate the occurrence of the marker when both subject and object can perform the action. The sentences below include feeding events (the corpus has five tokens which include the verb "feed"). When an adult feeds another adult, as in example (8.30), the marker is used in the

Table 8.3 Percentage of DOM usage according to type of object referent

Object referent	Ø	-pe	total	% marked
Adult	170	211	381	55%
Child	173	59	232	25%

Table 8.4 Percentage of DOM usage with child referents

Object noun type	Ø	-pe	total	% marked
Native-origin	94	10	104	10%
Loan	28	17	45	38%

two tokens, but when an adult feeds a child, like in example (8.31), the marker is only used in one out of the three tokens.

(8.30) *Ro-mongaru-haguã   mamá-pe   umía-pe   merkádo-pe   ro-mba'apo.*
        1.EXCL-feed-to   mom-OBJ   those-OBJ   market-in   1.EXCL-work
        "We work at the market to feed mom and others".

(8.31) *A-mongaru   che   família.*
        1-feed   my   child
        "I feed my child".

These patterns provide further evidence that DOM is favored in the presence of a loan. Even when the referent of the object is a child (or children) and, thus, the marker is dispreferred, a loan with a child human referent increases the use of the marker. While only 10% of native-origin object nouns with child referents are used with the marker, loans with child referents are marked 38% of the time, as summarized in Table 8.4.

Even though relative animacy has been described as significant in DOM patterns across the world's languages (Comrie 1989) and also in Spanish specifically (Delbecque 2002; Tippets 2011), the child vs. adult distinction with respect to human object markers is not described in the Spanish DOM literature. However, this pattern warrants further study, particularly, on Spanish varieties in contact with Guarani.

8.7 Discussion

This study has shown that the usage of DOM in Paraguayan Guarani is favored by social and linguistic factors. With respect to social factors, the higher usage rate of *-pe* with human objects among younger speakers compared to older speakers shows a change in progress in Guarani. This convergence of Guarani and Spanish is not only accounted for by the increasing bilingual population, but also by positive attitudes toward Spanish. Studies have shown that positive attitudes favor linguistic convergence. Döhla (2014), for example, showed that the rate of DOM usage in Portuguese increased from the 13th until the 17th century but later declined. Döhla attributes this change in language structure to a shift in attitudes toward Spanish. Linguistic divergence between Spanish and Portuguese coincided with the period of national identity formation in Portugal, which included negative attitudes toward Spain. Thus, in Paraguay, the high value of Spanish, as

it is the language for social mobility (Choi 2003), creates a favorable scenario for convergence between Spanish and Guarani.

Regarding linguistic factors, overall, the marker is favored when it co-occurs with a Spanish-origin loan object noun and appears to be disfavored when the events show the prototypical arrangement of the subject having more agency than the object. This last factor explains why objects with child referents, who have less agency than adults, are marked less frequently than nouns with adult referents. The effect that the presence of loans has on the occurrence of the DOM points to contact with Spanish as the origin of the feature.

For this study, 40 hours of recordings were analyzed. Some of these patterns, however, emerged from very few tokens or had many confounding factors. For example, word order could have an effect on the realization of DOM. However, to better study the effect of word order, other factors would need to be somewhat controlled, like the type of verb, the type of object, etc. Thus, while naturally occurring data provide essential instances of spontaneous linguistic usages that are at the core of language change, their variability can limit the analysis. Elicitation or experimental methods can be used in future research to verify the patterns that were learned from the data explored in this study.

8.8 Conclusion

This study aimed at exploring DOM as a contact feature in Paraguayan Guarani in a corpus of natural speech. It was hypothesized that: (1) DOM is a feature that emerged from contact with Spanish, and (2) that as a contact feature, the direct object marker would occur more often in the presence of Spanish loanwords. A logistic regression showed that -*pe* was more likely to occur in the presence of borrowed object nouns, but not in the presence of loan verbs. However, the most significant predictor of the occurrence of the marker was the age of the speaker. Both of these results provide evidence for -*pe* as a contact feature in Guarani.

The data explored here indicate that in Guarani, the object marker -*pe* grammaticalized from the Guarani locative/dative marker after the Spanish construction [V DOM OBJ] was borrowed. This Spanish construction included a loanword (the object: OBJ), which explains why DOM is (still) favored by a loan object noun. Thus, the findings of this study point to the need to explore borrowing from a lexical-grammatical continuum approach, instead of a lexical vs. grammatical dichotomy, as words and particles that occur together (like nouns and case markers) might be borrowed together.

Notes

1 Examples come from the spoken corpus unless they are marked with "ED" (Elicited Data) or are followed by a citation.
2 Examples (8.7) to (8.9) are extracted with their original orthography and translations to Spanish (Spa) and Portuguese (Por). I added the interlinear gloss and the translation to English.

References

Aikhenvald, A. Y. 2002. *Language contact in Amazonia*. Oxford: Oxford University Press.

Backus, A. 2013. Towards a usage-based account of language change: Implications of contact linguistics for linguistic theory. In *Questioning language contact: Limits of contact, contact at its Limits*, ed. R. Nicolaï, 91–118. Leiden: Brill.

Bossong, G. 2009. The typology of Tupi-Guarani as reflected in the grammars of four Jesuit missionaries: Anchieta (1595), Aragona (c. 1625), Montoya (1640) and Restivo (1729). *Historiographia Linguistica* 36.2–3:225–258.

Bybee, J. L. 2015. *Language change*. Cambridge: Cambridge University Press.

Choi, J. K. 2003. Language attitudes and the future of bilingualism: The case of Paraguay. *International Journal of Bilingual Education and Bilingualism* 6.2:81–94.

Cohal, A. L. 2014. *Mutamenti nel romeno di immigrati in Italia*. Milan: FrancoAngeli.

Company Company, C. 2003. La gramaticalización en la historia del español. *Medievalia* 35:3–61.

Comrie, B. 1989. *Language universals and linguistic typology: Syntax and morphology*. Chicago, IL: University of Chicago Press.

Delbecque, N. 2002. A construction grammar approach to transitivity in Spanish. In *The Nominative and Accusative and Their Counterparts*, ed. K. Davidse, and B. Lamiroy, 81–130. Amsterdam: John Benjamins.

Döhla, H. J. 2014. Diachronic convergence and divergence in differential object marking between Spanish and Portuguese. In *Stability and divergence in language contact. Factors and mechanisms*, ed. K. Braunmüller, S. Höder, and K. Kühl, 265–289. Amsterdam: John Benjamins.

Estigarribia, B. 2017. A grammar sketch of Paraguayan Guarani. In *Guarani linguistics in the 21ˢᵗ century*, ed. B. Estigarribia, and J. Pinta, 7–85. Leiden, Netherlands: Brill.

Fábregas, A. 2013. Differential object marking in Spanish: State of the art. *Borealis* 2.2:1–80.

Gregores, E., and J. A. Suárez. 1967. *A description of colloquial Guarani*. The Hague: Mouton.

Heine, B., and T. Kuteva. 2005. *Language contact and grammatical change*. Cambridge: Cambridge University Press.

Heine, B., and T. Kuteva. 2010. Contact and grammaticalization. In *The Handbook of Language Contact*, ed. R. Hickey, 86–105. Oxford: Wiley-Blackwell.

Iemmolo, G., and G. Klumpp. 2014. Introduction. *Linguistics* 52.2:271–279.

Karatsareas, P. 2020. The development, preservation and loss of differential case marking in Inner Asia Minor Greek. *Journal of Language Contact* 13.1:177–226.

Labov, W. 2001. The anatomy of style-shifting. In *Style and sociolinguistic variation*, ed. P. Eckert, and J. R. Rickford, 85–108. Cambridge: Cambridge University Press.

Laca, B. 2006. El objeto directo. La marcación preposicional. In *Sintaxis histórica del español*. Primera parte: La frase verbal, ed. C. Company Company, 423–475. México: Universidad Nacional Autónoma de México.

Matras, Y., and J. Sakel. 2007. *Grammatical borrowing in cross-linguistic perspective*. New York: Mouton de Gruyter.

Moreno-Fernández, F., I. Penadés-Martínez, and C. Ureña-Tormo. 2019. *Gramática fundamental del español*. London: Routledge.

Paraguay, Dirección General de Estadísticas, Encuestas y Censo. 2003. *Censo 2002: Resultados nvestigation*. Asunción, Paraguay.

Paraguay, Dirección General de Estadísticas, Encuestas y Censo. 2016. *Censo 2012: Resultados nvestigation.* Asunción, Paraguay.

Poplack, S. 2020. A variationist perspective on language contact. In *The Routledge handbook of language contact*, ed. E. Adamou, and Y. Matras, 46–62. London: Routledge.

Rodríguez-Ordóñez, I. 2020. Basque differential object marking as a contact-induced phenomenon: How and why?. *Journal of Language Contact* 13.1:227–270.

Roessler, E. M. 2019. Differential object marking and object scrambling in the Guarani Language cluster. *Lingvisticæ nvestigations* 42.1:31–55.

Roessler, E. M., N. Gasparani, and S. Danielsen. 2014. The southern clade of Tupi-Guarani languages: Adposition issues. Paper presented at Amazonicas *V*, Belém do Pará, Brazil.

Rubin, J. 1968. Bilingual usage in Paraguay. *Readings in the Sociology of Language* 2:512–530.

Ruiz de Montoya, A. 1876 [1639]. *Arte bocabulario, tesoro y catecismo de la lengua guarani.* Leipzig: Julio Platzmann.

Schilling, N. 2013. *Sociolinguistic fieldwork.* Cambridge: Cambridge University Press.

Shain, C., and J. Tonhauser. 2010. The synchrony and diachrony of differential object marking in Paraguayan Guarani. *Language Variation and Change* 22.3:321–346.

Sinnemäki, K. 2014. A typological perspective on differential object marking. *Linguistics* 52.2:281–313.

Thomason, S. 2001. *Language contact.* Washington, DC: Georgetown University Press.

Thomason, S. 2008. Social and linguistic factors as predictors of contact-induced change. *Journal of Language Contact* 2.1:42–56.

Tippets, I. 2011. Differential object marking: Quantitative evidence for underlying hierarchical constraints across Spanish dialects. In *Selected proceedings of the 13th Hispanic linguistics symposium*, ed. L. A. Ortiz-López, 107–117. Somerville, MA: Cascadilla.

von Heusinger, K., and G. A. Kaiser. 2011. Affectedness and differential object marking in Spanish. *Morphology* 21.3:593–617.