

Variable plural marking in Palenquero Creole

ESTILITA MARÍA CASSIANI OBESO

The Pennsylvania State University

HIRAM L. SMITH

Bucknell University

ABSTRACT

One of the most salient putative African features of Palenquero, an Afro-Hispanic creole spoken in northern Colombia, is the prenominal plural marker *ma*. However, plural number is not categorically marked with *ma*, which alternates with bare forms in plural contexts and also occurs in singular contexts. In a principled sample of noun phrases ($n = 1,186$) from the spontaneous speech of twenty-seven Palenquero-Spanish bilinguals, the rate of *ma* (versus zero) is 51% in plural and 13% in singular contexts. Singular *ma* is favored with subjects and specific objects, consistent with an association with definiteness. In plural contexts, where it is robust, selection of *ma* is favored with specific and generic referents in subject role. This conditioning indicates that plural marking is favored for discourse referential nouns, in accordance with the cross-linguistic generalization that morphological marking tends to appear on instances that approach the prototypical function of a category (Hopper & Thompson, 1984).

Typological approaches highlight multiple patterns of number marking cross-linguistically. For example, there are languages that exhibit singular and plural forms, while others have additional dual, trial, quadral, and paucal number values (e.g., Corbett & Mithun, 1996:2). There are languages where the category of number is obligatory, while, for others, there is a ‘transnumeral,’ or ‘general number’ system, where the meaning of the noun may be expressed without reference to number (Corbett, 2000:10). In spite of this, there are constraints on plural marking in the languages of the world (Croft, 2003:134). That is, the marking of plurality on nouns can be determined by animacy and referentiality hierarchies, or topicality constraints (Corbett, 2000; Corbett & Mithun, 1996; Croft, 2003:134).

Creoles and contact-induced languages provide a rich environment to study plurality. As an example, we have the Afro-Hispanic creole, Palenquero (PL), or *Lengua* (as it is called by the locals), which is spoken in northern Colombia (see map in Figure 1). PL is a Spanish-lexified language with contributions from the Kikongo Language Cluster (Lipski, 2005; Moñino, 2016; Schwegler, 2016, 2017). It has been noted that the putative plural marker in PL differs substantially from Spanish. While the latter generally marks plurality with an



FIGURE 1. Location of Palenque in Colombia.

Source: <https://royalsocietypublishing.org/doi/10.1098/rspb.2015.2980>.

-s suffix ([vowel + -s] and [consonant + -es]), PL employs the prenominal plural morpheme *ma* (Lipski, 2012; Moñino, 2007; Schwegler, 2007), which is most likely derived from the pluralizing prefix attested in Kikongo languages (Guthrie, 1970; though see Megenney, 1986:91; cf., Moñino, 2007:51). Example (1) illustrates the use of plural *ma*. However, *ma* does not categorically “mark” plurality, as seen in (2), where a bare form occurs in a plural context, or in (3), where *ma* occurs in a singular context. (In the examples, we gloss *ma* as PL and use \emptyset to indicate the absence of *ma*, for ease of identification.)

- (1) *Nunca, ma mailo suto trababá nu.*
 Never PL husband POSS work NEG
 ‘Never, our **husbands** do not work.’ (Female, 61, Recording 2, 06:52)¹
- (2) *Ma jende asé Øndulse en Semana Santa.*
 PL people make candies PREP Holy Week
 ‘People make **candies** during Holy Week.’ (Male, 62, Recording 32, 07:44)
- (3) *Ma kombilesa mi asé aprendé e ma kultura ke suto ten ayá.*
 PL friend POSS HAB learn DET PL culture that we have there

'My friends learn about **the culture** that we have there.' (Male, 27, Recording 65, 04:15)

Traditional views provided isomorphic schemas for the PL number system, whereby there is a one-to-one relation between form and meaning (e.g., Faingold, 1994; Friedemann & Patiño Roselli, 1983:138–49; Megenney, 1986:149–50). Later PL scholars have acknowledged variation as to the potential meanings of both bare and overtly coded forms, though they provide different explanations for some features of the system. The call has gone out for finer-grained inquiries into the variable expression of number marking in this language (Schwegler, 2007:208, 220). Answering the call for advancing research in this area, and for variationist research in creole studies more generally (Meyerhoff, 2009; Sankoff, 1990:296), the current study seeks to take advantage of the variation illustrated in (1)–(3) to ascertain to what degree *ma* is a marker of plurality. Building on prior work, the focus of this study is the discovery and interpretation of the conditioning factors on speakers' choice of prenominal *ma* over zero-coded forms, with the aim of providing a fresh perspective on creole marking.

Results show that, as expected from the prior literature, *ma* is indeed favored in plural contexts. However, variation is still robust within plural contexts, where *ma* is favored in subject role and with specific reference. These results point to the role of discourse referentiality and prototypicality of nouns in plural marking. Such a role is consistent with typological findings regarding the intersectionality of how a noun is used in discourse and the degree to which it is marked for nominal categories (Hopper & Thompson, 1984). Thus, while exhibiting widely cited creole characteristics, such as a substrate-derived plural morpheme and zero-coded forms, we find that variation in marker presence and absence in this creole is consistent with cross-linguistic marking patterns.

We begin with an overview of the literature concerning bare nouns in PL, highlighting the general consensus on plural marking in this creole. Second, we describe the Palenquero bilingual speech community and the corpus used in the present study. Focusing on forms with the prenominal determiner *ma* and bare nouns, we then present an analysis of PL noun phrase patterns based on syntactic and discursive factors. In the final section, we discuss the results and how they contribute to our understanding of plural marking in this Afro-Hispanic language.

PLURAL MARKING IN PALENQUERO

Table 1 is a depiction of Palenquero's article system as described under the traditional view (e.g., Faingold, 1994; Friedemann & Patiño Roselli, 1983:138–49; Megenney, 1986:149–50; see Schwegler, 2007:209 for discussion). The traditional view holds that bare (or zero-coded) nouns in PL categorically signal

TABLE 1. *Traditional view of Palenquero's article system*

	Definite, generic, and mass	Indefinite	Examples
Singular	∅	<i>un</i>	<i>∅ pelo asé ndrú mí mucho.</i> 'The dog sleeps a lot.' <i>Un pelo asé ndrú mí mucho.</i> 'A dog sleeps a lot.'
Plural	<i>ma</i>	<i>un ma</i>	<i>Ma pelo asé ndrú mí mucho.</i> '(The) dogs sleep a lot.' <i>Un ma pelo asé ndrú mí mucho.</i> 'Some dogs sleep a lot.'

singular definite meaning, as well as plural generic and mass interpretations, while overt expression with *ma* and *un ma* (definite and indefinite plural, respectively) are commonly interpreted as marking plurality.

More recent analyses of PL number marking, on the other hand, provide examples that clearly demonstrate that plural referents are not always marked by *ma*; rather, bare nouns may allow for both singular and plural interpretations (e.g., Lipski, 2012; Moñino, 2007). In particular, although *ma* has been traditionally associated with definiteness and plurality, it is claimed that “[the] definite plural marker *ma* is never obligatory, and therefore never predictable, and may (upon closer analysis) not signal ± definite. At the same time, it is true that when *ma* is present, the noun is always marked for plurality (but not necessarily for definiteness)” (Schwegler, 2013).

Against the traditional account, PL bare nouns are said to be transnumeral, that is, neutral as to number, with their interpretation driven entirely by context. Thus, as Schwegler (2007:211) illustrated, a decontextualized phrase containing a bare noun such as *puetta ri kasa* (lit. ‘door of house’) can have several readings: singular, plural, definite, or indefinite. Under this analysis, *ma* marks plural meaning (of various types) only, but plural meaning (and other meanings) can be expressed by zero.

According to a context-driven interpretation for PL number, then, overt marking is not needed to determine plurality. Indeed, it has been predicted that plural marking should be relatively infrequent:

If the need to disambiguate number were the primary reason for the use of *ma*, then its relative frequency should predictably become reduced whenever preceding discourse has made explicit the value of a bare noun (Schwegler, 2007:213).

Plural marking in noun phrases, *when it occurs at all*, is affected by the preposed plural marker *ma* (Lipski, 2012:30, our emphasis).

Nevertheless, scholars have noted that their samples turned up examples where *ma* was used redundantly and often in unambiguously plural contexts (Moñino, 2007;

Schwegler, 2007). As to *why* redundant *ma* appears when the context is unambiguously plural, Moñino (2007) proposes that, in the early phases of the creole, PL speakers used *ma* to make their language less intelligible to outsiders and to signal that speakers were using *Lengua* (PL) and not Spanish (as summarized in Schwegler, 2007:216). Such a view suggests that *ma* would be used as a symbol of social and cultural identity. Schwegler disagrees, citing, among other reasons, that *ma*'s structural transparency would be one of the earliest features learned by L2 speakers (Schwegler, 2007:216–7; cf., Lipski, 2012:30–1).

As we mentioned earlier, noun reference and semantics are also implicated in accounts of PL number. There is broad consensus that both *ma* and bare nouns may have definite, generic, and mass meanings. It has been claimed, however, that *ma* occurs more frequently with definite reference than with generic plural meaning (Lipski, 2012:31). Further, it has also been claimed that *ma* cannot occur at all with mass nouns (Moñino, 2013; Pérez Tejedor, 2006:43).

In sum, all of this research provides substantial insight into plural marking in PL. However, there have been calls for “more ambitious investigations into the grammar of Palenquero,” as “it remains altogether unclear what primary factor (or factors) motivate [*ma*'s] use in noun phrases” (Schwegler, 2007:208, 220). In order to dig deep into the grammar, here we undertake systematic quantitative analysis of the factors that influence speakers' selection of *ma*. We turn now to the speech community, the corpus, and the methods for analysis.

DATA AND METHODS

Palenquero community and corpus

PL is the only Spanish-lexified creole in mainland Latin America (Schwegler, 2011:446; cf., Jacobs & Parkvall, 2020) and its only demonstrable substrate has been identified as Kikongo (Granda, 1971; Schwegler, 2016, 2017). PL is spoken in San Basilio de Palenque, a small Afro-Hispanic village of approximately four thousand residents, located less than eighty kilometers from Cartagena de Indias, Colombia (Figure 1). For two centuries, Cartagena and its surrounding territories were the epicenter for the Spanish slave trade in the Americas (Wheat, 2011). San Basilio was founded between 1655–1674 (Navarrete, 2007:19), and is the only surviving palenque ‘palisade fort’ (among dozens extant during the colonial period) formed by Africans who had escaped bondage and became maroons in the numerous mountain slopes, swamps, and hinterlands of the jurisdictional district of Cartagena (Navarrete, 2008).

Currently, local members of the community still speak PL. Although with the stigmatization of this language there was a break in intergenerational transmission, in spite of this, in the past two decades, PL has undergone a remarkable resurgence of use due to language revitalization programs, activism, and education at the community level.

The data for this study were taken from *The Bilingual Corpus of Palenquero Creole: San Basilio de Palenque, Cartagena, and Barranquilla* (Smith, 2011–2014),² a collection of sociolinguistic interviews conducted by the second author throughout extended visits to the community from 2011 to 2014. The goal of the sociolinguistic interview is to tap into the *vernacular*, the speech register in which minimal attention is paid to speech. It is inherently variable yet observed to be the most regular and systematic of an individual's total set of speech registers (Labov, 1984:29).

A primary approach to conducting the interviews was through participant observation, whereby “the analyst integrates themselves within the community under investigation, either by engagement in local affairs and/or developing personal associations with members” (Tagliamonte, 2006:20). Community members also were enlisted to help find participants, which, over time, allowed the researcher to be accepted in, gain access to, and befriend the community in a shorter period of time and in a more thorough manner than “going it alone” (Milroy, 1980:47). Here we extend thanks to Florentino “Niño” Estrada, Rosalio Salgado, Ángel Valdez Herazos, Walberto Torres, and Luís Hénder Martínez for their aid in finding participants, and, in many cases, playing a principal role in facilitating, and even conducting, the interview. The interviews were conducted in settings familiar to the participants, usually in private homes, but also in plazas, streets, and other public spaces.

The interviews were transcribed to create a corpus of widespread usability, one that is not tied to particular linguistic features or research questions, and which could serve as a language archive (Torres Cacoullous & Travis, 2018:39–49). Interviews have been transcribed using the transcription software *Elan* (Lausberg & Sloetjes, 2009) and revised by a team of research assistants who were either native speakers or heritage speakers of PL. In addition to the first author, who is a member of the community, special thanks are also due to Basilia Pérez Márquez, Juana Paula “Pavi” Herazo Tejedor, and Cristina de la Hoz Márquez for their tireless work on the transcriptions.

For this study, we draw on twenty-seven interviews with mostly older speakers, born and raised in Palenque, who regularly use PL and are generally acknowledged to be good speakers of the language. We restrict this study to fluent speakers to provide a baseline in order to, as Poplack states, first “[establish] the nature of the system, against which we can subsequently assess what may be characterized as deviant with regard to it” (Poplack, 1993:252).³ The participants were males ($n = 13$) and females ($n = 14$) whose ages ranged from twenty-seven to eighty-eight years old (mean age = 63).

The tokens for this study ($n = 1,186$) were extracted from uninterrupted twenty-minute blocks of speech, comprising a total of seven hours and twenty minutes. All full noun phrases (including bare nouns), regardless of number readings (singular or plural), as well as any prenominal particles, were exhaustively extracted from this subcorpus. A reliability check for the extraction and coding was performed by the authors.⁴

TABLE 2. *Prenominal noun phrase particles in PL*

Labels	Forms	Examples
Bare	∅	<i>Aseba trompiá ku kuagro ri Abajo o ri Arriba.</i> '[We] used to fight with groups from [Barrio] Arriba and [Barrio] Abajo.' (Male, 27, Recording 76, 2:41)
Plural	<i>ma</i>	<i>Ma abispa aseba piká ma kompañero mi.</i> ' The wasps would sting my friends.' (Male, 60, Recording 5, 18:19)
Definite	<i>e</i> <i>lo, la</i>	<i>E moná a ta malo.</i> ' The kid is sick.' (Female, 80+, Recording 1, 5:41) <i>Ke lo tratte ma chisme limpiá lendro kasa.</i> ' The dishes , the thing-a-ma-jig washes the dishes in the house.' (Male, 62, Recording 32, 2:02)
Demonstrative	<i>ese,</i> <i>eta/ete</i>	<i>Ahora awe suto ta enfermá ku ese ma kumina.</i> 'Nowadays, we are getting sick from that food .' (Female, 53, Recording 42, 19:34) <i>Ma jende a deklará ma ete año andí ri afrocolombianía.</i> 'People have declared this year [The Year] of Afro-Colombianism.' (Male, 27, Recording 65, 08:51)
Indefinite	<i>un</i>	<i>Bo a tené ke kogé un oja.</i> 'You have to take a leaf .' (Male, 70+, Recording 9, 12:00)
Numeral/Quantifier		<i>Akí ma jende, ma ombre, poleba tené ata tre changáina.</i> 'Here people, men, could have up to three women .' (Male, 30+, Recording 29, 13:41)

Distribution of determiners/prenominal particles in PL corpus

The prenominal forms found, besides *ma*, were demonstratives, definite articles, indefinite articles, numerals, and other quantifiers (Table 2). Note that the traditional labels given to the particles are not meant to characterize their functions but are used for convenience.

The overall distribution of all prenominal particles is represented in Figure 2. The most notable finding is the relative frequency of bare nouns compared with noun phrases marked with prenominal *ma*. From a total of 1,186 noun phrases, bare forms constitute approximately half of the tokens (48%, $n = 564$); of prenominal elements, the most frequently occurring is *ma* (38%, $n = 499$); the remainder is comprised of all other co-occurring prenominal elements, which, combined, make up 14% ($n = 173$) of the full NP sample.⁵ The fact that there are more bare nouns overall is not surprising given Palenquero's status as a creole language and observations, which suggest that number may be driven by context rather than overt morphology. Still, these distributions raise important questions: Why is there such a high rate of *ma*? To what degree is it indexing plural marking?

Given the high proportion of noun phrases co-occurring with prenominal *ma* and bare forms, in the remainder of this paper we limit our discussion to the distributions for *ma* and zero. The next two sections consider the extraction and coding of the tokens.

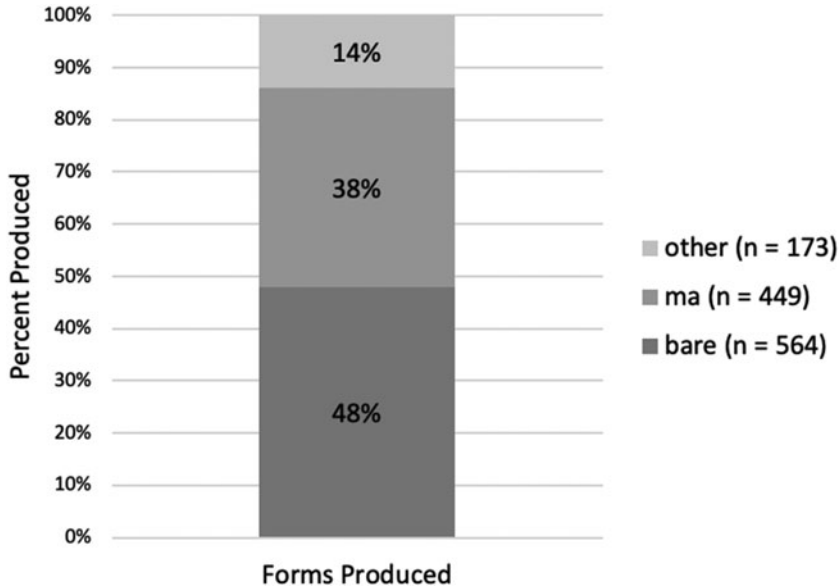


FIGURE 2. General distribution of forms produced ($n = 1,186$).

VARIATIONIST METHOD

Data extraction

An overarching goal of this study is to discover “what factors currently trigger (or constrain) the inclusion or exclusion of plural *ma*” (Schwegler, 2007:213). To do this, we use the variationist method (Labov, 1969, 2008), which combines a consideration of frequency and patterns of use. While there are previous studies of *ma* employing quantitative approaches (Moñino, 2007; Schwegler, 2007), we chose the variationist method because of the principle of accountability (Labov, 1972:72), which requires that we count all occurrences and nonoccurrences of the form of interest. The variable context is initially broadly defined as full (lexical) noun phrases regardless of number (e.g., singular versus plural) appearing with zero (a bare noun) or with pronominal *ma*.

Not included in the present study were mass nouns such as *maní* ‘peanuts,’ *kumina* ‘food,’ *tiempo* ‘time,’ and *ropita* ‘clothes’ ($n = 78$), only eight tokens of which (10%) co-occurred with *ma* (cf., Moñino, 2013; Pérez Tejedor, 2006:43). The noun *jende* ‘people, folks’ was excluded from this analysis because of its high frequency ($n = 213$).⁶ Additionally, conservatively left out of the analysis were proper nouns, in particular geographic names that did not occur with *ma* in these data (*Palenque*, *Lengua Palenquera*, *Cartagena*, *África*, *Venezuela*, etc.) ($n = 73$). Finally, we excluded the few instances of a noun with a Spanish determiner and gender agreement⁷ ($n = 4$).

Data coding

Number. Regarding the category number, we test to what extent the morpheme *ma* marks plurality. To avoid circularity, we rely on the context to establish plurality, instead of the presence of grammatical marking. It has been claimed that older, fluent Palenquero speakers never use *ma* in unambiguously singular contexts, but that singular uses are an innovation of younger L2 learners (e.g., Lipski, 2012:31, 2020:12).

In example (4), *trabajo* ‘job’ is coded as singular even if it has the pronominal *ma*, because the context tells us that the noun is singular. The participant-observer/guide is responding to a woman’s question: Is this [sitting down for an interview] voluntary or is [the interviewer] going to pay me? The guide responds that “The job that he gives you—he has to pay you for it.” The ‘work’ to do, or ‘job,’ or ‘task’ in this case, is the one act of participating in the sociolinguistic interview. Similarly, in example (3) above, the noun is preceded by *e ma*, yet *kuttura* ‘culture’ is a singular noun, which, in this context, refers to the Palenquero culture. The second author asked the participant, who is a school teacher, if he teaches *kuttura africana* ‘African [Palenquero] culture’ to his friends, or only to outsiders. After responding that all of his friends learn about “the culture we have here” from him, he adds: *I asé enseñá e kusa* ‘I teach about it [lit. ‘the thing’].’ In example (5), *e ma lengua ete* ‘this language’ is an unmistakably singular noun that refers to the Palenquero language. In the immediately preceding context, the woman uses several expressions that indicate she is talking about something one of a kind, such as, “This language is from here. It’s not from anywhere else. It is spoken by the old women who were born right here.” We also note the surrounding linguistic elements, such as the co-occurring singular proximal demonstrative *ete* ‘this’ in (5).⁸

On the other hand, *monasito* and *majanasito* ‘kid’ in (6) and (7) are coded as plural, based on the immediate linguistic context. In both cases, the speakers are making a general statement about kids. Consider now example (2) above, where the bare noun *ndulse* ‘candies’ has plural meaning, despite the absence of overt marking or anything in the immediate discourse context to establish plurality. Nevertheless, we can confidently code this token as plural based on the authors’ general knowledge of the cultural traditions in Palenque, which is home to several annual festivals (such as *Festival de Tambores* ‘Festival of Drums,’ *Semana Santa* ‘Holy Week’). Palenquero women typically make various types of homemade sweet treats (*alegrías*, *dulce de coco*, *dulce de papaya*) to sell at these festivals and as a daily means of subsistence. The small number of cases where we could not determine the number were coded as “ambiguous” and were excluded ($n = 27$ such cases).

- (4) <A su trabajo>⁹ *ma trabajo* si lo ke te ponélo a ten ke
 For POSS work PL job POSS 3sg that IO put-DO PRES have to
 pagálo.
 pay-DO

- ‘Your work—the **job that** he gives you, [he] has to pay you.’ (Male, 62, Recording 2, 00:22)
- (5) *E ma lengua ete ke suto tené aká es ri por ayá nu, sino ri propio.*
 DET PL language DEM that we have here is from over there NEG rather from us.
 ‘**This language** that we have here is not from anywhere else, but is [exclusively] ours.’
 (Female 61, Recording 2, 10:12)
- (6) *Ma monasito ahora a ta degenerao.*
 PL kids today PRES COP deprived
 ‘Nowadays, **kids** are deprived.’ (Male, 58, Recording 34, 12:18)
- (7) *Ø Majanasito asé chitiá lengua palenquera. <Ellos>—ke ané a kriá akí.*
 Kids HAB speak Lengua Palenquera they that they PRET raise here
 <Trajeron uno con dos años>. <El> uto a miní ku año y
 brought one with two years. The other PRET come with one and
 medio.
 half
 ‘**Kids** speak Lengua Palenquera. Those that were raised here. [They] brought
 one that was two years old. The other one came when [he was] one and a half
 years old.’ (Female, 60+, Recording 71, 03:44)

Specificity. Much of the previous literature on Palenquero *ma* has focused on definiteness, because, as we discussed previously (Table 1), *ma* is regarded as a marker, not only of plurality, but also of definiteness. Regarding definiteness in PL, Lipski (2012:30) observed that *ma* is more frequently used with definite reference, but also with generic plurals. On the other hand, Moñino (2007) argued that PL bare nouns also can be plural, generic, and definite. Schwegler (2007:211) stated: “Moñino and I believe that the motivation for the behavior of *ma* (or *un*, *un ma*) is unrelated to functional (i.e., strictly grammatical) considerations. As a result, there exists no predictable correlation between the co-occurrence of bare nouns and reference type.”

Here, instead of coding NPs based on definiteness, which is variously understood¹⁰ (and often difficult to code independently of the presence of a language-particular marker of definiteness, such as the definite article in English), we coded based on the construct of specificity, which we may code for based on the context. Specificity is a category of information flow, which refers to how speakers package ideas as they are participating in the discourse. Cross-cutting information flow categories are identifiability, activation state, discourse function, and specificity-generality (Dubois, 1980; Thompson, 1997; Travis & Torres Cacoullos, 2020:130–4).

We focus on specificity to test here whether overt marking (*ma*) correlates with generics, as has been proposed, and with specific over nonspecific NPs. Following Torres Cacoullos and Aaron’s (2003:307–8) study on the distribution of bare noun phrases in Spanish and English, there are three factors in this group. Specific nouns refer to specific people or things that are not considered to be interchangeable, as in (8), while a nonspecific NP involves the notion of nonunique existence, that is, a noun labeled nonspecific is any member of a class, as in (9) (see Ashby & Bentivoglio, 1993:69). Finally, generic NPs refer to an entire class of entities, as

illustrated in (10). Torres Cacoullós and Aaron (2003:310) found that, for both Spanish and English, nonspecific uses of noun phrases favor zero (bare nouns), whereas specific uses disfavor bare forms. However, while generic uses disfavor bare forms in Spanish, in English this is a favorable context for bare forms (as with ‘mothers’ in the English translation of [10]). If *ma* is associated with definiteness, it should be favored more by specific than nonspecific uses of NPs (cf., Torres Cacoullós & Vigil, 2003).

- (8) *Ane asé ablálo tambié “mango puerko.” Ma chikito ke bo kumé anoche*
 They HAB call-it also mango pig. PL little ones that you eat last night
é ri puerko.
 is from pig
 ‘They also call it “pig mango.” **The little ones** that you ate last night are the
 “pig” kind.’ (Male, 62, Recording 9, 18:44)
- (9) *Ayá, en posá ri ele, konbesá ma monasito andi lengua suto.*
 There in house of him, speak PL kid in language 1PL
 ‘Over there, in his house, **the kids** speak Palenquero.’ (Male, 30+, Recording
 29, 08:31)
- (10) *Ma mamá ri ahora, asé salí a peleá po majaná ri ahora.*
 PL mother of today, HAB leave to fight for children of today
 ‘**Mothers** nowadays, go out and fight for these **kids** today.’ (Male, 53,
 Recording 21, 08:05)

In example (8), the noun *chikito* ‘the little ones’ is specific, because it is referring to specific mangoes from the speaker’s refrigerator that he had given to the second author the night before. In example (9), the word *monasito* has a nonspecific reference because the speaker is referring to any of the kids, not one(s) in particular, in the person’s house. No specific child is identified. In example (10), the speaker uses the word *majaná* to refer to children in general, as indicated by *ri ahora*, hence, ‘kids today/nowadays.’ Similarly, ‘mothers’ refers to the entire class. Thus, we coded tokens like these as generic.

However, there were also cases in which the nouns did not have any local linguistic cues as to specificity; yet, we were still able to code for specificity according to the broader discourse context, as is shown in the following example (11). Notice that the noun *kabesa* ‘head’ does not have a determiner, yet, in this context, it is obvious that the ‘head’ spoken of is the speaker’s head. Here, the participant describes a boxing match between him and another man. The result was, the other man came out *sano* ‘unharméd,’ but “I left with [my] head split open.” The speaker’s head is not interchangeable; thus, we coded this token as specific.

- (11) *Pogke kuando a miní, lo ke ta peleá, a kelá sano y uto—*
 Because when PRET come, ART that PROG fight PRET remain safe and other
I a salí yo ku kabesa partilo.
 1st sg PRET leave 1st sg with head split.
 ‘Because when he came out, the guy that was fighting, he was fine, but the other
 one—I left with [my] **head** split open.’ (Male, 62, Recording 32, 04:31)

Syntactic role. Subject position has been found to be the least favorable syntactic role for bare nouns, since subjects tend to have identifiable referents, whereas objects are more favorable to bare nouns, in Spanish and English (Torres Cacoullós & Aaron, 2003:309). In coding for syntactic role, we seek to determine where in a sentence *ma*—contrasting with zero—tends to occur, for example, in subject (12), object (13), or oblique positions (14).

- (12) *Ma mamá suto¹¹ a ta muetto a tiembo.*
 PL mother 1PL PRET COP dead long time
 ‘Our mothers have been dead for a long time.’ (Female, 70+, Recording 18, 10:56)
- (13) *Ma changaína ke asé bae pa Cartagena a mbendé ma ndulse y fruta.*
 PL women that HAB go to Cartagena to sell PL candies and fruit
 ‘Women that go to Cartagena to sell candies and fruit.’ (Male, 27, Recording 65, 08:24)
- (14) *Suto konserbá ese lengua pa majaná suto má tadde.*
 1PL conserve that language for kids 1PL more later
 ‘We will conserve the language for our kids in later generations.’ (Female, 53, Recording 42, 12:35)

Animacy. Several hierarchies constrain number marking: a person hierarchy, where first and second person outrank third person; a referentiality hierarchy, where pronouns outrank common nouns; and an animacy hierarchy, where humans outrank nonhuman animates, which, in turn, outrank inanimates (Corbett, 2000; Croft, 2003:128–32). This set of universals predicts that, despite the parochial facts of number marking in a particular language, the more animate the (pro)noun is, the more likely it is to co-occur with markers for nominal categories such as number. Thus, the person-referential form-animacy hierarchy is often depicted in diagram form as showing rankings descending as we read from left to right: first/second person pronouns < third person pronouns < kin < proper names < humans < human common nouns < nonhuman animates < inanimate common nouns. This constraint-ranking also suggests that plural marking begins earlier rather than later along this constraint hierarchy. Thus, we coded tokens according to animacy, distinguishing between animate and inanimate nouns.

Locative. It has been suggested that PL noun phrases preceded by particular locative items are inherently definite (e.g., *akí* ‘here,’ *at* ‘there’ [proximal], *ayá* ‘there’ [distal]) (Schwegler, 2007:210), whereas others (those preceded by *andi* ‘in, at, over there’) can be definite or indefinite [examples [15] and [16]]. We test if the presence or absence of a locative will influence the occurrence of *ma* as a way to examine the relation between *ma* and definiteness. If *ma* tends to have definite reference, we might expect it to be favored in the presence of pronominal locatives that are consistent with this function. While Schwegler commented only on particular locatives, we included all those that appeared in the corpus (e.g., *en* ‘in,’ *lentro* ‘inside,’ etc.).

- (15) *I ten ma baka. Bo sabeba traé bo solo, pasando trabajo, bo*
 I have PL cow you HAB-PASTIMP bring you alone pass-PROG work you
solo aí monte.
 alone there mountain
 ‘I had the cows. You used to bring [them] all by yourself, working, you by
 yourself [up] **there in [the] mountain.**’ (Male, 60, Recording 5, 10:28).
- (16) *Suto sabeba ta andi grupo.*
 1PL HAB-PASTIMP COP in group
 ‘We used to be **in groups** [kuadros].’ (Male, 53, Recording 21, 04:03)

RESULTS

Figure 3 presents the rates of *ma* versus bare nouns according to noun number. In singular contexts, *ma* occurs 13% ($n = 34/252$) of the time as opposed to bare forms. While it was expected that bare nouns would express more meanings than just singular number (e.g., Moñino, 2007; Schwegler, 2007), the fact that *ma* does occur with singular nouns (in fluent adult speakers), albeit less frequently than with plural nouns (13%, $n = 34/252$), was not (cf., Lipski, 2012, 2020:5).

The more striking result is for the plural context, in which we see robust variation in the use of the prenominal *ma*. Half of the time, speakers use *ma* for marking plurality (51%, $n = 223/437$). Prenominal *ma* may mark plural reference, but it is not an obligatory marker, as speakers can also use a bare form to express a plural meaning and do so approximately half the time. Nevertheless, when *ma* does occur, it is overwhelmingly plural (87%, $n = 223/257$). We turn

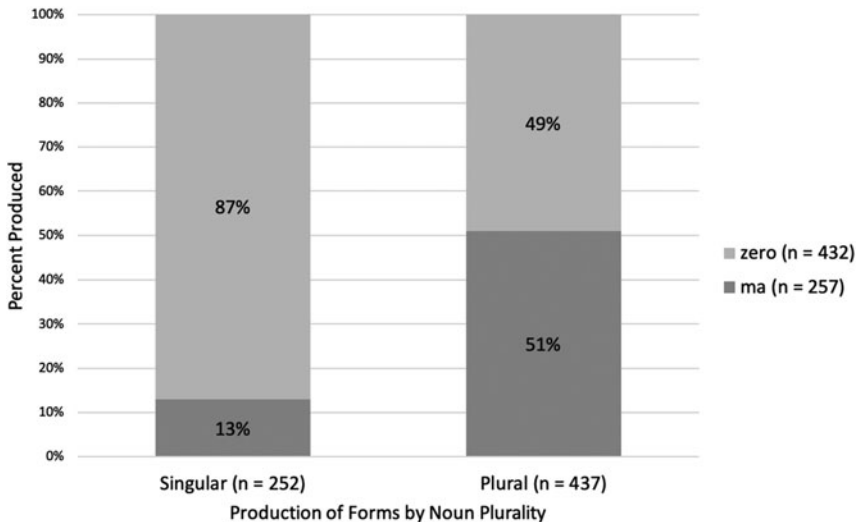


FIGURE 3. Production of forms by noun plurality ($n = 689$).¹²

now to the multivariate analyses of the factors conditioning *ma* versus zero, using Goldvarb Yosemite (Sankoff, Tagliamonte, & Smith, 2015), in three contexts: all contexts (singular and plural), plural contexts only, and in singular contexts.

Multivariate analysis 1: ma versus bare nouns in singular and plural contexts

Table 3 shows the factors contributing to the presence of *ma* (compared to bare nouns) in all contexts (i.e., both singular and plural). The total *n* of tokens analyzed was 689. As indicated by the corrected mean (.33), the overall likelihood of *ma* to occur before any factor groups are considered is about one-third of the time. The range between the largest and the smallest factor weights gives an indication of relative magnitude of effect; number shows the highest range, a finding that gives probabilistic support to the claim that *ma* expresses plural meaning (Moñino, 2007; Schwegler, 2007). Number is followed by specificity, syntactic role, and the presence of a locative, while animacy was not selected as significant.

Specificity is the second-most impactful factor group for the occurrence of *ma*. Its range of 24 lags far behind that of number (47), but clusters with the other two significant groups, syntactic role (21) and locative (18). Looking at

TABLE 3. Variable rule analysis of factors contributing to the presence of prenominal *ma* (versus zero)

Corrected Mean: .33			
Total <i>n</i> : 257/689			
	Factor Weight:	% <i>ma</i>	<i>n</i>
Number			
Plural	.68	51%	437
Singular	.21	13%	252
	<i>range: 47</i>		
Specificity			
Specific	.60	38%	350
Generic	.52	59%	92
Nonspecific	.36	28%	247
	<i>range: 24</i>		
Syntactic Role			
Subject	.64	53%	232
Object/Oblique	.43	29%	457
	<i>range: 21</i>		
Locative			
Absent	.52	40%	619
Present	.33	14%	70
	<i>range: 18</i>		
Total		37%	689

Note: Animacy was not selected as significant.

Goldvarb Yosemite: A multivariate application for Macintosh (Sankoff, Tagliamonte, & Smith, 2015)

direction of effect, factor weights show that *ma* is favored over zero with specific (.60) and, less so, by generic nouns (.52), but is disfavored in nonspecific contexts (.36). Additionally, *ma* expresses specific meaning 52% of the time ($n = 134/257$ tokens of *ma*). Regarding generic contexts, we see that 59% of these nouns employ prenominal *ma* ($n = 54/92$). We also note that the behavior of PL generic nouns is similar to Spanish (Torres Cacoullós & Aaron, 2003:310) in that they both disfavor zero, though the effect is stronger in Spanish. However, while the rate of *ma* in generic contexts (59%) is higher than in specific contexts, the factor weight is lower (.52). We look into this mismatch in the relative ordering of factor weights and percentages within the specificity factor group (Walker, 2010:42) by examining plural and singular contexts separately (Tables 4 and 8, below); it turns out, there are virtually no generic singulars ($n = 1$).

When we look at syntactic role, we observe that *ma* is favored in subject role (.64) and disfavored in all other argument roles (.43) (we combined objects, obliques, and “other” arguments). Put another way, objects/obliques favor zero-coded, or bare nouns. This latter finding is similar to what has been found for both Spanish and English, where bare nouns were favored in object role (Torres Cacoullós & Aaron, 2003:309).

Finally, *ma* is slightly favored in the absence of a co-occurring prenominal locative (.52), which is the overwhelming number of cases (90%, $n = 619/689$). Of the seventy prenominal locatives, only 19% ($n = 13$) (which does not include the eleven tokens of locative + *ma*) appear with an additional element modifying the noun, all of which were postnominal possessive pronouns (for example, *Majanasito a taba jundo otra be akí kasa mi* ‘The children were together again **here at my house**’). A post hoc analysis of the 173 prenominal determiners/particles (shown in Figure 2, and which were not part of the multivariate analysis) revealed that *ma* occurred with another element only 7% ($n = 12/173$) of the time, all cases being *ese ma* ($n = 10$) and *e ma* ($n = 2$). It could be that *ma* is disfavored with locatives because the use of more than one prenominal element is avoided.

As seen above, *ma* is not an exclusive or consistent plural marker. We present now the factors conditioning the variation in plural and singular contexts separately.

Multivariate analysis 2: ma versus bare nouns in plural contexts

Table 4 shows the multivariate analysis of factors contributing to the presence of *ma* (over zero) in *plural contexts only*. The significant factor groups are specificity and syntactic role. The difference in the range between these two factor groups is negligible. Thus, in plural contexts, we see that *ma* is favored in specific (.60) and generic contexts (.54), and in subject role (.63). Locative was not among the factor groups selected as significant. As with the first analysis, the factor group animacy was not significant.

TABLE 4. Variable rule analysis of factors contributing to the presence of prenominal *ma* (versus zero): plural contexts

Corrected Mean: .51			
Total n: 223/437			
	Factor Weight:	% <i>ma</i>	<i>n</i>
Specificity			
Specific	.60	60%	173
Generic	.54	59%	91
Nonspecific	.38	38%	173
	<i>range: 22</i>		
Syntactic Role			
Subject	.63	64%	164
Object/Oblique	.42	43%	273
	<i>range: 21</i>		
Total		51%	437

Note: Animacy and co-occurring locative were not selected as significant.

Goldvarb Yosemite: A multivariate application for Macintosh (Sankoff, Tagliamonte, & Smith, 2015)

Table 5 shows a cross-tabulation of the factor groups *specificity* and *syntactic role*. Cross-linguistically, valency roles and information flow parameters correlate (Dumont, 2016:76; Thompson, 1997:72–3). Here we observe that, while plural objects/obliques appear as specific ($n = 117$) or nonspecific ($n = 122$) nouns with nearly equal frequency (first row), those that are specific tend to have plurality expressed with *ma* (56%, $n = 66/117$), while nonspecific nouns in the same syntactic role are expressed more with a bare form (68%, $n = 83/122$). Objects/obliques tend not to be used as generics (which do not refer to specific entities), but when they are, plurality is also expressed more by zero (62%, $n = 21/34$). A chi-square test showed that there was a significant relationship between grammatical marking and specificity for nouns in object position ($p < .01$).

TABLE 5. Rate of plural *ma* versus bare nouns (zero) by specificity and syntactic role

	Specific	Generic	Nonspecific	Totals
Object/Oblique	56% (66/117)	38% (13/34)	32% (39/122)	43% (118/273)
Subject	66% (37/56)	72% (41/57)	53% (27/51)	64% (105/164)
Totals	60% (103/173)	59% (54/91)	38% (66/173)	51% (223/437)

Looking at plural nouns in subject role (second row), we find a (near) equal distribution of specificity types. Prenominal *ma* seems to be most strongly favored with generic subjects, with which it occurred 72% of the time ($n = 41/57$), and also with specific subjects, occurring there 66% of the time

TABLE 6. *Rate of plural ma versus bare nouns (zero) by animacy and syntactic role*

	Animate	Inanimate	Totals
Object/Oblique	44% (63/142)	42% (55/131)	43% (118/273)
Subject	61% (84/138)	81% (21/26)	64% (105/164)
Totals	53% (147/280)	48% (76/157)	51% (223/437)

($n = 37/56$), followed by nonspecifics (53%, $n = 27/51$). However, the differences were not statistically significant ($p = .1126$). In sum, plural subjects favor *ma*-marking, including when used as nonspecific nouns, though less so, whereas objects/obliques favor *ma* only when used as specific nouns.

A cross-tabulation of *animacy* and *syntactic role* is displayed in Table 6. Not surprisingly, subjects overwhelmingly tend to have animate referents, which make up 84% of that group ($n = 138/164$). In contrast, animate nouns comprise only 52% ($n = 142/273$) of nouns in a syntactic role other than that of subject. A closer look reveals that both animate and inanimate objects/obliques prefer zero coding, since the rate of *ma* is 44% and 42%, respectively (top row) (the difference is not statistically significant, $p = .6915$). Thus, we see that zero is favored with nouns that are objects/obliques, regardless of animacy (there is no animacy effect).

In a parallel fashion, both animate and inanimate subjects prefer overt marking (second row). Inanimate subjects overwhelmingly take *ma* 81% compared to animate subjects, which occur with *ma* 61% of the time. A chi-square test revealed no significant difference ($p = .3823$). We see now why animacy was not selected as significant in the multivariate analysis—but syntactic role and specificity were: *ma* is favored with subjects, and zero is favored with objects/obliques, regardless of whether they are animate or inanimate. In aggregate, however, animates tend to show a higher rate of *ma* than inanimates (53% versus 48%, bottom row), since subjects are disproportionately animate.

Table 7 is a cross-tabulation of *animacy* and *specificity*. We observe that bare forms are favored with nonspecific nouns (first row), regardless of whether the noun is animate (56%, $n = 54/97$) or inanimate (70%, $n = 53/76$). However, there is a tendency for more *ma* with nonspecifics that are animate (44%, $n = 43/97$) versus those that are inanimate (30%, $n = 23/76$), which would be consistent with typological predictions. A chi-square test revealed that the differences are not quite significant ($p = .0587$). On the other hand, *ma* is favored with specific nouns

TABLE 7. *Rate of plural ma versus bare nouns (zero) by animacy and specificity*

	Animate	Inanimate	Totals
Nonspecific	44% (43/97)	30% (23/76)	38% (66/173)
Specific	56% (57/101)	64% (46/72)	60% (103/173)
Generic	57% (47/82)	78% (7/9)	59% (54/91)
Totals	53% (147/280)	48% (76/157)	51% (223/437)

(second row), regardless of animacy ($p = .3249$). The overwhelming majority of generics were animate ($n = 82/91$), which favored *ma*.¹³ Again, we see that animacy is not the major driving force behind morphological coding.

Multivariate analysis 3: ma versus bare nouns in singular contexts

Table 8 now displays the conditioning factors for *ma* and zero in *singular contexts only*. Our results indicate that speakers *can* use *ma* with singular reference. In all, there were a total of 13 speakers (8 males, 5 females) of the 27 in the present study who used *ma* in singular contexts; the two youngest were 27 and 40 years of age. The rest of the speakers were in their 50s, 60s, 70s, and 80s. The factor groups with the largest impact on the occurrence of *ma* are presence of a locative (range = 45) (which is not significant for plurals), specificity (35), and syntactic role (28). Animacy was not selected as significant (as for plurals).

We observe that *ma* is strongly *disfavored* when there is a co-occurring locative present (.16). The locatives present in singular contexts (*Total n* = 46) were (in order of frequency) *andi* ‘in, at, there’ ($n = 17$), *al*, *ayá* ‘there’ ($n = 13$), and all others (such as *lentro* ‘inside’, *en* ‘in’, *a* ‘at’, *po* ‘by’) ($n = 12$), *akí* ($n = 2$), and *akí andi* ‘here at’ ($n = 1$); co-occurring with *ma* was one token of *po* ‘by.’ As mentioned earlier, it may be that *ma* is disfavored with locatives because the presence of more than one pronominal element is avoided.

Singular *ma* is favored when in specific contexts and in subject role, which raises the question of whether *ma* is functioning as a singular determiner. A

TABLE 8. Variable rule analysis of factors contributing to the presence of pronominal *ma* (versus zero): singular contexts

Corrected Mean: .09			
Total n: 34/252			
Factor	Weight:	% <i>ma</i>	<i>n</i>
Locative			
Absence	.59	16%	206
Presence	.16	2%	46
range: 43			
Specificity			
Specific	.61	18%	177
Nonspecific	.26	4%	74
range: 35			
Syntactic Role			
Subject	.70	28%	67
Object/Oblique	.42	8%	184
range: 28			
Total		13%	252

Note: Animacy was not selected as significant.

Goldvarb Yosemite: A multivariate application for Macintosh (Sankoff, Tagliamonte, & Smith, 2015)

TABLE 9. Rate of singular *ma* versus bare nouns (zero) by syntactic role and specificity

	Specific	Nonspecific	Totals
Subject	29% (16/56)	27% (3/11)	28% (19/67)
Object/Oblique	12% (15/121)	0% (0/0)	8% (15/169)
Totals	18% (31/176)	27% (3/11)	14% (34/236)

cross-tabulation of the factor groups *syntactic role* and *specificity* (Table 9) revealed that, though zero is overwhelmingly more frequent than *ma* in singular contexts, when *ma* does appear, it is favored according to the following hierarchy: specific subjects (29%) > nonspecific subjects (27%) > specific objects (12%), and does not occur at all with nonspecific objects. This would be consistent with definiteness, in that subjects tend to have given rather than new referents more than objects/obliques do (cf., Thompson, 1997:72) and, in turn, given referents are identifiable (Thompson, 1997:68). However, a larger dataset of singular *ma* is needed to further investigate an association with definiteness.

CONCLUSION

As we have seen, the favoring of *ma* is not dependent on the plurality of the noun, all by itself, but on whether the NP is specific and in subject role. Multivariate analysis of plural contexts showed us that when all factor groups are considered simultaneously, specificity and syntactic role are the clear winners. We see different results for objects and subjects in the interplay with specificity. The cross-tabulations revealed that, for nouns in object role, specificity of reference matters with regard to the structural coding of the noun, such that *ma* is favored only with plural specific objects, whereas subjects favor *ma* across the specificity categories, although somewhat less with nonspecific than specific and generic uses. Clearly, the stronghold of zero is nonspecific objects.

In our data, animacy was consistently not chosen as significant. This seems surprising, given the role of an animacy hierarchy in plural marking cross-linguistically (e.g., Corbett, 2000; Croft, 2003:128–32). Closer analyses showed that objects tended to be expressed by bare forms, regardless of noun animacy, whereas subjects were expressed by *ma* again across animacy categories. In like manner, bare forms were favored with nonspecifics, regardless of noun animacy; in turn, *ma* was favored with specifics. The appearance that animates favor plural marking may be due to the tendency for subjects to be animate (Table 6) and for specific uses to occur with animate nouns proportionally more than nonspecific uses do (Table 7).

That *ma* is favored with plural count nouns that are both specific and in subject position make sense from a discourse-typological perspective, given that there is an intersectionality between valency, discourse/information flow parameters, and morphological coding (Dubois, 1980; Givón, 1981; Hopper & Thompson, 1980, 1984; Thompson, 1997). For example, Hopper and Thompson (1980:288) found that the crucial element in the way an object is marked is not its definiteness but

having the property of a specific referent. Nevertheless, this finding is not inconsistent with an association between definiteness and *ma*, but, rather, supports it, since subjects (in contrast with objects and obliques) tend to have given, identifiable referents (cf., Thompson, 1997:68, 72).

The key insight of Hopper and Thompson's (1984:703) cross-linguistic survey is that "the closer a form is to signaling [the] prime function [of a noun, namely, 'discourse manipulable participant'], the more the language tends to recognize its function through morphemes typical of the category." Discourse referential NPs are those that are used "for the purpose of either introducing [a referent] for further tracking or continuing tracking an earlier mention" (Thompson, 1997:69). Thus, subjects, which tend to be animate and specific and, therefore, tend to be used referentially in discourse (Travis & Torres Cacoullos, 2020:134, and references therein), should predictably show more marking of nounhood, that is, of nominal categories, including number. It is therefore not surprising that correlations between plural marking and specificity of reference together with syntactic role exist in Palenquero.

In this study, we found that not only do bare forms depend on the context for their interpretation (Moñino, 2007; Schwegler, 2007), so do forms coded with *ma*. However, the distributions for both *ma* and zero are largely predictable and cohere with broad typological predictions. Apparent anomalies, such as was found for animacy, which was not selected as significant, also trend in the right direction and can be explained in terms of plural marking developing first with specific nouns in subject position, which are mostly animate.

In employing systematic quantitative analysis of linguistic data from members of a well-defined speech community, the major contributions of this study are the use of an exhaustively transcribed corpus of Palenquero speech, thus bringing variation theory to bear on issues in Palenquero's grammar (Morton, 2005; Smith, 2013, 2018). Using this approach, we demonstrated that, despite its classification as a creole, and its putative African-origin plural marker, this creole language neatly conforms to broad typological patterns that have been observed in other world languages.

That said, an important goal of this project was to have community members participate in the scientific research of their own community at every stage, from data collecting, transcription, research, and writing; thus, we put science at the service of the community. More generally, this study contributes to the documentation of Afro-Hispanic language varieties, which have been socially stigmatized and are endangered, and this study aims to contribute to the larger debate about the nature of grammatical phenomena in creole languages, which, for a long time, were compartmentalized as exotic.

ACKNOWLEDGMENTS

We would like to especially thank John Lipski and Alex McAllister for their feedback and comments on earlier versions of this paper. We also extend thanks

for the insightful comments from other colleagues and two anonymous reviewers which greatly improved the paper at various stages of its development.

NOTES

1. Within parentheses following examples are the speaker's sex and age, followed by the recording number and time stamp (minute:second) of the example.
2. The corpus was supported by NSF (National Science Foundation) grant BCS-1226655.
3. Lipski (2012:33) has noted, in particular, that the use of *ma* and *e ma* among some young speakers appears to behave like (or be moving in the direction of) the Spanish singular articles *el* and *la* (e.g., *ma puetta ta celao* 'the door is closed').
4. Each of us independently extracted the first 100 tokens from the same materials, with 75% agreement. We subsequently cross-checked samples of each other's coding of the tokens for all factors.
5. The rates of the other pronominal elements are as follows: demonstratives *ese/ete/eta* (5%, $n = 63$), indefinite articles *un/un ma* (5%, $n = 60$), numerals and quantifiers (3%, $n = 33$), and definite articles (1%, $n = 17$).
6. Separate analysis of *jende* 'people' ($n = 213$) indicates the same tendencies as in the general dataset for the favoring of *ma* in subject role, for example, *Ma jende a ta kumo, asina, un poko loko* 'People are like, a little crazy like that', however, the factor group specificity was not pertinent.
7. Palenquero does not have gender agreement.
8. While the literature claims that Palenquero has no definite articles, it is clear that forms like *e* in example (5), here glossed as DET, and which possibly derives from Spanish *el* 'the' or through demonstrative grammaticalization, may function as such.
9. Spanish glosses are presented in brackets.
10. For example, Lyons's (1999:2–12) view of definiteness correlates to *familiarity*, *identifiability*, and *uniqueness*. His concept of uniqueness, or the singularity of the noun phrase, that is, "that there is just one entity satisfying the description used" (Lyons, 1999:8), is interchangeable with the definition of *specificity* adopted here. Another information flow category for investigating definiteness may be *identifiability* (Thompson, 1997:66).
11. *Ma mamá suto* 'our mothers' is coded as plural because the speaker, when describing her relationship with her first cousin (who was also present during the interview), was pointing out that both of their mothers (who were sisters) are dead.
12. Mass nouns ($n = 78$) and ambiguous nouns ($n = 27$) are excluded ($n = 105$). There were too few inanimate generics for a reliable *ma* rate.
13. The difference with animate generics is not significant ($p = .2355$).

REFERENCES

- Ashby, William J., and Bentivoglio, Paola. (1993). Preferred Argument Structure in Spoken French and Spanish. *Language Variation and Change* 5:61–76.
- Corbett, Greville. G. (2000). *Number*. Cambridge: Cambridge University Press.
- _____, and Mithun, Marianne. (1996). Associative Forms in a Typology of Number Systems: Evidence from Yup'ik. *Linguistics* 32:1–17.
- Croft, William. (2003). *Typology and Universals*, 2nd ed. Cambridge: Cambridge University Press.
- Dubois, John W. (1980). Beyond definiteness: The trace of identity in discourse. In W. Chafe (Ed.), *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production*. Norwood: Ablex. 203–74.
- Dumont, Jenny. (2016). *Third person references: Forms and functions in two spoken genres of Spanish*. Amsterdam/Philadelphia: John Benjamins.
- Faingold, Eduardo D. (1994). The genesis of the article system in creolization and historical change. *Papere zur Linguistik* 50:52–63.
- Friedemann, Nina de, and Patiño Roselli, Carlos. (1983). *Lengua y sociedad en el Palenque de San Basilio*. Bogotá: Instituto Caro y Cuervo.
- Givón, Talmy. (1981). *Topic continuity in discourse: A quantitative cross-language study*. Amsterdam/Philadelphia: John Benjamins.
- Granda, Germán de. (1971). Sobre la procedencia africana del habla criolla de San Basilio de Palenque. *Thesaurus* 26:84–94.
- Guthrie, Malcolm. (1970). *Comparative Bantu: An Introduction to the Comparative Linguistics and Prehistory of the Bantu Languages, vol. 4*. Farnborough: Gregg Press.

- Hopper, Paul J., and Thompson, Sandra A. (1980). Transitivity in grammar and discourse. *Language* 56:251–99.
- Hopper, Paul J., and Thompson, Sandra A. (1984). The discourse basis for lexical categories in universal grammar. *Language* 60:703–52.
- Jacobs, Bart, and Parkvall, Mikael. (2020). How ‘Portuguese’ are Palenquero and Chabacano really? Published online: 27 March, 2020, <https://doi.org/10.1075/rro.19001.jac>. Amsterdam: John Benjamins.
- Labov, William. (1969). Contraction, deletion, and inherent variability of the English. *Language* 45:715–62.
- . (1972). *Sociolinguistic patterns*. Philadelphia: University of Philadelphia Press.
- . (1984). Field methods of the project on linguistic change and variation. In J. Baugh and J. Sherzer (Eds.), *Language in use: Readings in sociolinguistics*. Englewood Cliffs: Prentice Hall. 28–53.
- . (2008). Quantitative analysis of linguistic variation. In U. Ammon, N. Dittmar, K. Mattheier, and P. Trudgill (Eds.), *HSK Sociolinguistics/Sociolinguistik, vol. 1*. Berlin: de Gruyter. 6–21.
- Lausberg, H. and H. Sloetjes. (2009). Coding gestural behavior with the NEUROGES-ELAN system. *Behavior Research Methods, Instruments, & Computers* 41:841–9.
- Lipski, John. (2005). *A History of Afro-Hispanic Language: Five Centuries, Five Continents*. Cambridge: Cambridge University Press.
- . (2012). The “new” Palenquero: revitalization and re-creolization. In R. J. File-Muriel and R. Orozco (Eds.), *Colombian Varieties of Spanish*. Vervuert: Iberoamericana. 30–8.
- . (2020). Language revitalization as L2 shadow boxing: the case of Palenquero plural-marking. *Studies in Second Language Acquisition*. Published online: 5 August, 2020. <https://doi.org/10.1017/S0272263120000339>. 1–16.
- Lyons, Christopher. (1999). *Definiteness*. Cambridge: Cambridge University Press.
- Megenney, William W. (1986). *El palenquero: un lenguaje post-criollo de Colombia*. Bogotá: Instituto Caro y Cuervo.
- Meyerhoff, Miriam. (2009). Replication, transfer, and calquing: Using variation as a tool in the study of language contact. *Language Variation and Change* 21:297–317.
- Milroy, Lesley. (1980). *Language and social networks*. Baltimore: University Park Press.
- Moñino, Yves. (2007). Convergencias lingüísticas iberocongolesas en palenquero: ¿integrarse a la sociedad mayoritaria o distinguirse de ella? In W. Mihatsch and M. Sokol (Eds.), *Language Contact and Language Change in the Caribbean and Beyond (Lenguas en contacto y cambio lingüístico en el Caribe y más allá)*. Bern: Peter Lang. 37–59.
- . (2013). Nuevas luces sobre el uso del plural en Palenquero y sus orígenes kikongo. *Vistas al Patio* 7:39–58.
- . (2016). Past, present, and future of Palenquero Creole. In A. Schwegler, B. Kirschen, and G. Maglia (Eds.), *Palenque (Colombia): Orality, Identity, and Resistance. An interdisciplinary approach*, Creole Language Library. Amsterdam/New York: John Benjamins. 15–50.
- Morton, Thomas. (2005). *Sociolinguistic variation and change in El Palenque de San Basilio (Colombia)*. Doctoral dissertation, University of Pennsylvania.
- Navarrete, María Cristina. (2007). Las memorias de San Basilio de Palenque. Bogotá: Instituto Colombiano de Antropología e Historia—ICAH, Área de Historia Colonial.
- . (2008). *San Basilio de Palenque: memoria y tradición*. Cali: Programa Editorial, Universidad del Valle.
- Pérez Tejedor, Juana Pabla. (2006). *El criollo de San Basilio de Palenque: una visión estructural de su lengua*. Bogotá: Edición Uniandes-CCELA, Descripciónes 14.
- Poplack, Shana. (1993). Variation theory and language contact. In D. Preston (Ed.), *Variation theory and language contact: American dialect research*. Amsterdam/Philadelphia: John Benjamins. 251–86.
- Sankoff, Gillian. (1990). The Grammaticalization of Tense and Aspect in Tok Pisin. *Language Variation and Change* 2:295–312.
- Sankoff, David, Tagliamonte, Sali A., and Smith, Eric. (2015). Goldvarb Yosemite: A variable rule application for Macintosh. Department of Linguistics, University of Toronto. <http://individual.utoronto.ca/tagliamonte/goldvarb.html>
- Schwegler, Armin. (2007). Bare nouns in Palenquero. In M. Baptista and J. Guéron (Eds.), *Noun phrases in creole languages*. Amsterdam/Philadelphia: John Benjamins Publishing. 205–22.
- . (2011). Palenque (Colombia): Multilingualism in an extraordinary social and historical context. In M. Diaz-Campos (Ed.), *The Handbook of Hispanic Sociolinguistics*. Oxford: Oxford University Press. 446–71.

- _____. (2013). Palenquero structure dataset. In S. M. Michaelis, P. Maurer, M. Haspelmath, and M. Huber (Eds.), *Atlas of Pidgin and Creole Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://apics-online.info/contributions/48>. Accessed on 9/01/2020.)
- _____. (2016). Combining population genetics (DNA) with historical linguistics. In S. Sessarego and F. Tejedo (Eds.), *Spanish language and sociolinguistic analysis*. Amsterdam/Philadelphia: John Benjamins.
- _____. (2017). On the African origin(s) of Palenquero. In A. Schwegler, B. Kirschen, and G. Maglia (Eds.), *Palenque (Colombia): Orality, Identity, and Resistance. An interdisciplinary approach*. Creole Language Library. Amsterdam/New York: John Benjamins. 51–119.
- Smith, Hiram L. (2011–2014). *The Bilingual Corpus of Palenquero Creole: San Basilio de Palenque, Cartagena, and Barranquilla*.
- Smith, Hiram L. (2013). Habitual Aspect Marking in Palenquero: Variation in Present Temporal Reference. In A. M. Carvahlo and Sara Beaudrie (Eds.), *Selected Proceedings of the 6th Workshop on Spanish Sociolinguistics*. Somerville, MA: Cascadilla Proceedings Project. 97–108.
- _____. (2018). Addressing questions of grammaticalization in creoles: It's all about the methodology. In H. Narrog and B. Heine (Eds.), *Grammaticalization from a typological perspective*. Oxford: Oxford University Press. 372–93.
- Tagliamonte, Sali A. (2006). *Analysing Sociolinguistic Variation: Key topics in Sociolinguistics*. Cambridge: Cambridge University Press.
- Thompson, Sandra A. (1997). Discourse motivations for the core-oblique distinction as a language universal. In A. Kamio (Ed.), *Directions in functional linguistics*. Amsterdam: Benjamins. 59–82.
- Torres Cacoullós, Rena, and Aaron, Jessi. (2003). Bare English-origin nouns in Spanish: Rates, constraints, and discourse functions. *Language Variation and Change* 15:289–328.
- Torres Cacoullós, Rena, and Travis, Catherine E. (2018). *Bilingualism in the Community: Code-switching and Grammars in Contact*. Cambridge: Cambridge University Press.
- Torres Cacoullós, Rena, and Vigil, Neddy A. (2003). Sustantivos de origen inglés en discurso español: ¿cuál es su gramática? *Río Bravo. A Journal of the Borderlands* (Special editor: Glenn Martínez) 2:69–84.
- Travis, Catherine E., and Torres Cacoullós, Rena. (2020). The role of pragmatics in shaping linguistic structures. In J. Cesar Félix-Brasdefer and D. Koike (Eds.), *The Routledge Handbook of Spanish Pragmatics*. New York: Routledge. 129–47.
- Walker, James A. (2010). *Variation in Linguistic Systems*. New York: Routledge.
- Wheat, David. (2011). The First Great Waves: African provenance zones for the Trans-Atlantic Slave Trade to Cartagena de Indias, 1570–1640. *Journal of African History*. 52:1–22.