Variable Papiamentu pluralization strategies in a written corpus

Margaret Glide & J. Clancy Clements
Indiana University

Variable noun phrase (NP) pluralization in Papiamentu has been described by researchers such as Dijkhoff (1980, 1983, 1993) and Kester & Schmitt (2007). In these studies, there is general agreement that definiteness correlates with overt plural marking (presence of –nan) and the presence of a quantifier in an NP with no overt plural marking (absence of –nan). Kester & Schmitt (2007) suggest that animacy has no effect on the presence / absence of –nan and that this morpheme might mark discourse-specific (as in discourse-old) information. In this study, we test these descriptive accounts with a variationist analysis of NP pluralization in Papiamentu using a corpus compiled from newspapers published in Aruba, Bonaire, and Curaçao. In our corpus, we test the effect quantifiers, definiteness, animacy, and discourse status have on overt pluralization. We corroborate the observations of previous researchers that definiteness favors the presence of –nan, while quantifiers disfavor it. Moreover, we find that animacy also favors the presence of –nan, but that discourse status shows no consistent patterning in our data.

Keywords: Papiamentu; information structure; noun phrase; pluralization; quantitative analysis

1. Introduction

Papiamentu (PA) uses the third person plural pronoun as a plural morpheme in NPs (Dijkhoff 1980, 1983, 1993; Goilo, 1998; Kester & Schmitt, 2007; Kouwenberg, 2007:323; Maurer, 1998), just as do other Atlantic creoles, such as Jamaican, Gullah, Saramaccan, Sranan, Krio, and Guyanese, although its position varies across languages (Alleyne, 1980: 100). In PA the third person

---

1 We would like to thank the three anonymous reviewers of this paper for their helpful input and literature recommendations. We owe a debt of gratitude to our colleague, Anna Wilson, for aiding us with statistical computations present in this paper. Of course, all errors remain our own.
plural pronoun functions as an enclitic morpheme \(-nan\), as in *instruktornan* “instructor-PL”.

Overt plural marking in these creoles is shown to be variable (see Baptista 2003), as is also the case in PA, where overt pluralization has been described as being sensitive to three factors: the presence of a quantifier, definiteness, and specificity (Dijkhoff, 1980, 1983, 1993; Kester & Schmitt, 2007).

Kester & Schmitt (2007: 139) define specificity as “that which the speaker has in mind”, which in our view corresponds to the notions of hearer-old information (information activated in the mind of the hearer) and discourse-old information (information previously mentioned in the discourse context). Kester & Schmitt (2007) found no correlation between animacy of an NP and overt pluralization, which is interesting given that in other creoles, such as Cape Verdean Creole, Guinea Bissau Creole, and Cassamance creole, such a correlation appears to exist (Baptista, 2003).

The quantitative study by Poplack & Tagliamonte (1994) compares pluralization in historical varieties of African American English (AAE) and normative Nova Scotian English. Of particular importance for the present paper, they found that AAE displayed sensitivity to both definiteness and quantifiers, as is the case for the studies on PA overt pluralization. They do not consider animacy or information status in their study.

Following Poplack and Tagliamonte, this study seeks to test quantitatively the impact of definiteness and quantifiers on PA overt pluralization, and adds the factors of animacy and information status. Following Haspelmath (2013), we expect the presence of animacy to correlate with overt pluralization (the presence of \(-nan\) on plural NPs). Following findings in Prince (1992), we also expect discourse-old NPs to correlate with PA overt pluralization.

2. PA pluralization

The enclitic plural morpheme \(-nan\) has its origins in a West African language; however, from which specific language remains unclear (Jacobs, 2012: 94). Parkvall (2000:104) presents a case for a Wolof origin given the analogous Wolof third person plural pronoun *na fn(u)*. See also Ndiaye (2004: 28). Unlike English, which uses obligatory pluralization in all nouns, PA displays variable plural marking (Haspelmath 2013), which according to previous research is
determined by a number of contextual and syntactic factors, to be reviewed in
the following paragraphs.

All researchers who have worked on PA pluralization have found that
the presence of a quantifier strongly disfavors the use of -nan with plural NPs
(Dijkhoff, 1980, 1983, 1993; Kouwenberg, 2007; Maurer, 1998). Indeed, one
pedagogical grammar of PA informs learners that –nan is not necessary when
a numeral is mentioned (Goilo, 1997). One would therefore anticipate that this
tendency would be near categorical in cases such as the following:

(1) dos kas
    two house
* dos kasnan
    two house,PL
    ‘two houses’ (Dijkhoff, 1983: 218)

Using the examples in (2), Dijkhoff (1983) also shows that quantified NPs that
have a definite referent allow for overt pluralization while the lack of –nan in
such cases is unacceptable.

(2) e dos kasnan
    DEF two house,PL
*e dos kas
    DEF two house
    ‘the two houses’ (Dijkhoff, 1983: 221)

Formal definiteness has been reported to be an important factor in PA plural
example, argues that there is a relationship between the presence of the PA
variable plural marker –nan and the referential status of the corresponding NP.
She provides three semantic categories of NP referentiality according to their
overt and non-overt pluralization tendencies, summarized in the table below:
Table 1: NP Referentiality and plural marking adapted from Dijkhoff (1983:217-227)

<table>
<thead>
<tr>
<th>Term</th>
<th>NP type</th>
<th>Plural marking</th>
<th>Example from Dijkhoff (1983)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Existentially presupposed</td>
<td>Definite NPs</td>
<td>Must carry –nan</td>
<td><em>E kas</em> ‘the house’</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><em>E kasnan</em> ‘the houses’</td>
</tr>
<tr>
<td>Existentially asserted</td>
<td>Indefinite NPs</td>
<td>Overtly marked only if plurality not indicated elsewhere</td>
<td><em>Kachónan a keda grita henter anochi.</em> ‘Dogs kept barking all night.’</td>
</tr>
<tr>
<td>Existentially hypothesized</td>
<td>Generic and non-specific NPs</td>
<td>Never marked with the plural marker</td>
<td><em>Mi a kumpra baki.</em> ‘I have bought a book/books.’</td>
</tr>
</tbody>
</table>

This classification suggests that NPs containing definite determiners strongly prefer overt pluralization. On the other hand, Dijkhoff (1983:217-227) adds that generic and non-specific entities tend not to be overtly marked for plural, and indefinite referents permit variation in plural marking according to contextual factors, such as previous discourse and the presence of a numeral.

With respect to definite marking, Kester & Schmitt (2007) associate overt plural marking of the NP with its specificity, in that plural NPs without any determiner, quantifier, or other modifier (also known as bare plurals) can be overtly plural when they refer to a pre-specified set of referents. Their concept corresponds to the notions of discourse-old and hearer-old, as these pertain to the information status of the NP. An example of this is shown in (3), taken from Kester & Schmitt (2007: 115):

(3) a. *Despues ku *hende/hendenan a keha, nan

   After that person/persons past complain they

   *a drecha e pelicula.

   past fix the film

   ‘After some people complained, they fixed the film’.

b. *Si hende/*hendenan ta keha, no wori

   If person/*persons present complain no worry

   ku nan.

   with them
‘If people complain, don’t worry about them’. (Kester & Schmitt 2007: 115)

The NP *hendenan* ‘people-PL’ in (3) refers to a set of people that have already been activated in discourse, either by explicit statement or by the logical entailment that there are people at a movie theater. Given that –*nan* can be understood as marking specificity in this case, -*nan* is present. In (3), the NP *hende* ‘person/people’ has a generic interpretation, is thus non-specific, and the overt plural marker –*nan* is not present. Dijkhoff (1980: 3) briefly mentions the role of NP specificity, but her statement seems to contradict what has been illustrated in (3). That is, she states that “when there is talk of a plural (indicating more than one object in reality), *nan* is not used as a matter of course”. In other words, if the plurality of a referent is clear from the discourse situation, overt plural marking with -*nan* may not be found. This study seeks to clarify this confusion by studying the effect of the specificity of an NP on the overt plural marking.

Lastly, Kester & Schmitt (2007) conclude that PA pluralization strategies are not sensitive to animacy, unlike the situation in Cape Verdean creole, Guinea Bissau creole, and Casamance creole discussed in Baptista (2003). While it is possible that the role of animacy of the NP in PA pluralization strategies differs from that of other creoles, it merits quantitative evaluation.

Table 2 summarizes the factors that have been proposed in previous studies as conditioning overt plural marking in PA.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantifiers</td>
<td>Dijkhoff, 1980; Dijkhoff, 1983; Dijkhoff, 1992; Maurer, 2002; Kouwenberg, 2007</td>
</tr>
<tr>
<td>Specificity</td>
<td>Kester &amp; Schmitt, 2007</td>
</tr>
<tr>
<td>Formal definiteness</td>
<td>Dijkhoff, 1980; Dijkhoff, 1983; Dijkhoff, 1992; Maurer, 2002; Kouwenberg, 2007</td>
</tr>
</tbody>
</table>

In addition to these variables, the current paper addresses the effect of information status or plural marking. In the following section we operationalize information structure, with definitions of related terms, and we describe why this factor is important to consider in PA pluralization strategies.
3. Information Structure

While the literature on variable plural marking in PA and in other creoles identified a set of factors that potentially condition plural marking, it has not investigated how a plural NP behaves on its first mention and subsequent mentions in a text. This section introduces Prince’s (1992) conceptualizations of hearer status, discourse status, and definiteness, which will be used to operationalize the analysis of such variables in this study.

Prince (1992) defines indefiniteness and definiteness as formal categories rather than semantic categories. That is, the presence of syntactic features, shown in Table 3 below, are indicators of the NP’s categorization as indefinite or definite.

Table 3: Formal definiteness features as described in Prince (1992: 299)

<table>
<thead>
<tr>
<th>Indefinite NP</th>
<th>Indefinite NP</th>
<th>Definite NP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indefinite article</td>
<td>Definite Article</td>
<td>Demonstrative article</td>
</tr>
<tr>
<td>Zero article</td>
<td>Possessive adjective</td>
<td>Personal pronouns</td>
</tr>
<tr>
<td>Other quantifiers (<em>some, any one, six, many few, numerals, etc.</em>)</td>
<td>Proper nouns</td>
<td>Certain quantifiers (<em>all, every</em>)</td>
</tr>
</tbody>
</table>

The indefinite vs definite NP distinction, with their respective features listed in Table 3, is related to the hearer and discourse status of an NP –either discourse-new information or discourse-old information. The latter tends to correspond to a definite subject. Given that definiteness is one of the variables claimed to condition pluralization strategies in PA (Dijkhoff 1980, 1983, 1993; Maurer 2002), it is important to describe and operationalize the concepts, hearer status and discourse status.

Hearer status relates to the hearer’s knowledge of the referent (Prince 1992). As Kruijff-Korbayová & Steedman (2003: 250) elaborate, it is the “utterance-internal structural and semantic properties reflecting the relation of an utterance to the discourse context”, such as “the actual and attributed attentional states of the discourse participants, and the participants’ prior and changed attitudes (knowledge, beliefs, intentions, expectations, etc.).” In this type of information status, items can be hearer-new or hearer-old. If the hearer has previous knowledge of the entity, it is considered hearer old and is typically definite. On the other hand, entities that are not previously known to
the hearer are considered hearer-new. Moreover, NPs that are able to appear in presentational *there*-clauses are typically hearer-new. In contrast, generic NPs are considered hearer-old because it is assumed that the hearer is aware of the generic entity as a class of entity types (Prince 1992). Returning to the idea that definite NPs tend to be overtly marked for pluralization in PA and that NPs without definite marking tend not to be, it follows that hearer-old referents in discourse would be more likely to be overtly marked than hearer-new referents. However, the relationship between hearer status and definiteness is not categorical (Prince, 1992).

This relationship becomes more complex when discourse status is also considered. Discourse-old and discourse-new referents are not necessarily hearer-old and hearer-new. Discourse status relates to the previous mention of an NP in the text or the contextual allusions toward it (Prince, 1992). This term refers to “all aspects of the internal or organizational structure of a discourse”, such as “segmentation, relations between segments (informational and intentional), anaphoric relations, modal subordination, discourse topic, thematic progression, etc.” (Kruijff-Korbayová & Steedman, 2003: 250). There are three components to discourse status: discourse-new, discourse-old, and inferred. Discourse-old referents represent NPs that have been previously mentioned in the discourse. Conversely, discourse-new referents have no prior mention or presuppositions in the text (Prince, 1992). That is, the first mention of an NP is considered discourse-new and the subsequent repetitions or pronominalizations of the NP are considered discourse-old. Lastly, inferable entities are considered the middle ground on the continuum between discourse old and new. These entities can be logically parsed from the context and depend on the hearer’s knowledge and understanding of the world, as illustrated in the following example: “He passed by the Bastille and the door was painted purple” (Prince 1992: 305)

In this example, the hearer must know or assume that the Bastille is a building, and also that buildings typically have doors. Therefore, the door would be considered inferred from the previous discourse about a building. Prince (1992: 305-306) considers inferable entities to be “like Hearer-old entities in that they rely on certain assumptions about what the hearer does know”; however, they are also similar to hearer-new entities as “the hearer is not expected to already have in his/her head the entity in question”. In this study, inferred entities are considered inferred for both the hearer and discourse status if these are recoverable through context. For example, in (4) below, *properties* was considered both hearer and discourse inferred because the article previously mentioned Belasting as one of the large companies
schedule to hold an auction. Therefore, it could be inferred that this company would have something to auction, particularly properties.

\[(4) \quad \text{Belasting a pone beslag riba tur e propiedatnan} \]

\[\text{Belasting PAST put fittings on all DEF properties} \]

\[\text{mobili i e indishi ta planeá pa dia 30 di desèmber.} \]

\[\text{mobile and DEF auction COP planned for day 30 of December.} \]

‘Belasting put fittings on all of the mobile properties and the auction is planned for the 30th of December.’\(^2\)

4. Methodology

The data for the present study was collected from four different newspaper sites online from all three of the ABC islands (see Table 4). Each plural NP, regardless of overt or non-overt marking with \textit{nan}, was considered for analysis. The newspapers in question report specifically local news as well as some international news; however, only local news stories were analyzed (e.g. stories about members of the community, obituaries, local politics, car accidents, etc.).

To constitute the data base, approximately 200 tokens were taken from each island for this preliminary analysis (197 tokens from Bonairean newspapers, 204 tokens from Curaçaoan newspapers, and 169 tokens from Aruban newspapers for a total of 570 tokens). The newspapers under analysis are listed in Table 4. The numbers in parentheses represent the number of articles analyzed from each newspaper.

Table 4: Newspapers analyzed by territory

<table>
<thead>
<tr>
<th>Aruba</th>
<th>Bonaire</th>
<th>Curaçao</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solo Di Pueblo (7)</td>
<td>Extra (16)</td>
<td>La Prensa i Ultima Noticia (14)</td>
</tr>
<tr>
<td>Bon Dia Aruba (6)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The following research questions are explored in this data:

1. Are the factors found to affect the overt pluralization of noun phrases in the qualitative studies also found to be significant in a quantitative analysis?

2. What is the effect of discourse status (barely addressed in previous studies) on overt pluralization on PA?

Our hypotheses regarding these questions are as follows: The alternation between overt and no marking is hypothesized to constitute a binary dependent variable with two variants: overt plural marking with \( -nan \), and no plural marking with \( -nan \), which may be conditioned by external factors. The various factors proposed in the research literature as conditioning overt plural marking, plus the additional factor of discourse status, are also adopted as hypotheses, and treated as independent variables potentially conditioning plural marking. Table 5 lists the dependent and independent variables.

### Table 5: Dependent and independent variables considered.

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Independent variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plural marker (presence vs. absence of ( -nan ))</td>
<td>Quantifier (presence, absence)</td>
</tr>
<tr>
<td></td>
<td>Animacy of noun nucleus of NP (animate, inanimate)</td>
</tr>
<tr>
<td></td>
<td>Formal definiteness of NP (presence, absence)</td>
</tr>
<tr>
<td></td>
<td>Hearer status of NP (new, old, or inferred)</td>
</tr>
<tr>
<td></td>
<td>Discourse status of NP (new, old, or inferred)</td>
</tr>
</tbody>
</table>

The data was coded as follows: first, plural nouns were coded for the dependent variable of overt or non-overt pluralization, in which the presence of the plural morpheme, \( -nan \), is considered as overt and its absence as non-overt. Then, each plural NP was coded for the independent variables listed in Table 5 above. Specifically, the quantifiers coded for were the equivalents of ‘all, some, any, every’. Animate or human NPs were coded as animate. Formal definiteness of the NP was coded as per the definition and operationalization provided above in Prince (1992). Hearer status of the NP was coded as previously unknown (new), known (old) to the hearer, or inferred from the context (inferred). Similarly, the discourse status of the NP was coded as the lack of previous mention of the NP within the article (new), the presence of a

---

3 In our review of the relevant literature, only Dijkhoff (1980) has addressed discourse status, but only briefly and only in terms of context.
previous mention (old), or the existence of the NP being somehow inferable from the context of the discourse (inferred). However, during the analysis, hearer and discourse status caused a major interaction in the data based on their similarity of coding. Therefore, these two variables were collapsed into discourse status.

The effects of these four independent variables were tested using logarithmic multivariate regression analysis by means of Goldvarb (Sankoff et al. 2005), in order to establish which set of factors significantly predicts overt plural expression. The identified set of independent variables constitutes a model of the variation.

The preliminary hypotheses are that the results will support the finding of the previous studies: that NPs with a definite determiner will be overtly marked and quantifier-marked NPs will be not overtly marked with the plural morpheme, -nan. Moreover, discourse-old and hearer-old NPs without a definite determiner are expected to be less likely to be overtly pluralized. This hypothesis is consistent with ideas in Dijkhoff (1980), in that the repetition of a discourse-old plural entity without overt marking will be sufficient, as the context or prior mention will represent a direct antecedent.

Finally, although one of the authors did the actual data coding, all difficult cases were coded together. Additionally, an interrater reliability test was conducted to ensure that this coding was consistent. After coding 59 plural NPs, representing approximately 10% of the entire data set of 570 tokens, both authors agreed 94.23% of the time; in 49 of 52 cases there was agreement. The cases requiring further research included partitive phrases such as como representante di Corona e ta un di e entidadnan mas importante di nos pais ‘As a representative of the Crown, it is one of our country’s most important entities’. In similar cases, both the status of the overtly marked plural entity as an independent NP and the whole NP’s interpretation being plural were unclear.

5. Results

Initial distributional analysis revealed that the majority of tokens in this data were overtly marked with the enclitic plural marker, -nan (71.9% overt with 410 tokens, 28.1% non-overt with 160 tokens). Given this distribution, the analysis of the dependent variable focused on the application value of overt pluralization. The following table summarizes the distribution overt and non-overt plural marking across each of the independent variables:
Table 6: Distribution of overt and non-overt plural marking among all factors

<table>
<thead>
<tr>
<th>Group</th>
<th>Overt N</th>
<th>Overt %</th>
<th>Non-overt N</th>
<th>Non-overt %</th>
<th>Total N</th>
<th>Total %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantifier Presence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>379</td>
<td>86.9</td>
<td>57</td>
<td>13.1</td>
<td>436</td>
<td>76.5</td>
</tr>
<tr>
<td>Yes</td>
<td>31</td>
<td>23.1</td>
<td>103</td>
<td>76.9</td>
<td>134</td>
<td>23.5</td>
</tr>
<tr>
<td>Animacy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>244</td>
<td>67.2</td>
<td>119</td>
<td>32.8</td>
<td>363</td>
<td>63.7</td>
</tr>
<tr>
<td>Yes</td>
<td>166</td>
<td>80.2</td>
<td>41</td>
<td>19.8</td>
<td>207</td>
<td>36.3</td>
</tr>
<tr>
<td>Definiteness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>213</td>
<td>62.1</td>
<td>130</td>
<td>37.9</td>
<td>343</td>
<td>60.2</td>
</tr>
<tr>
<td>Yes</td>
<td>197</td>
<td>86.8</td>
<td>30</td>
<td>13.2</td>
<td>227</td>
<td>39.8</td>
</tr>
<tr>
<td>Discourse Status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New</td>
<td>209</td>
<td>68.5</td>
<td>96</td>
<td>31.5</td>
<td>305</td>
<td>53.5</td>
</tr>
<tr>
<td>Old</td>
<td>175</td>
<td>75.1</td>
<td>58</td>
<td>24.9</td>
<td>233</td>
<td>40.9</td>
</tr>
<tr>
<td>Inferred</td>
<td>26</td>
<td>81.2</td>
<td>6</td>
<td>18.8</td>
<td>32</td>
<td>5.6</td>
</tr>
<tr>
<td>TOTAL</td>
<td>410</td>
<td>71.9</td>
<td>160</td>
<td>28.1</td>
<td>570</td>
<td>100</td>
</tr>
</tbody>
</table>

The Goldvarb program selected as statistically significant <quantifiers>, <formal definiteness>, and <animacy> as variables contributing to overt pluralization. Table 7 lists the results.

Table 7: PA overt pluralization. Multivariate analysis*

| Corrected mean | .794 |
| Log likelihood | -213.402 |
| P = 0.001 |

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>N</th>
<th>%</th>
<th>Factor Weight</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantifier presence</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>379</td>
<td>86.9</td>
<td>.69</td>
<td>.62</td>
</tr>
<tr>
<td>Yes</td>
<td>31</td>
<td>23.1</td>
<td>.07</td>
<td></td>
</tr>
<tr>
<td>Definiteness</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>197</td>
<td>86.8</td>
<td>.77</td>
<td>46</td>
</tr>
<tr>
<td>No</td>
<td>213</td>
<td>62.1</td>
<td>.31</td>
<td></td>
</tr>
<tr>
<td>Animacy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>166</td>
<td>86.8</td>
<td>.59</td>
<td>14</td>
</tr>
<tr>
<td>No</td>
<td>244</td>
<td>67.2</td>
<td>.45</td>
<td></td>
</tr>
<tr>
<td>Discourse Status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New</td>
<td>[209]</td>
<td>[68.5]</td>
<td>[.46]</td>
<td></td>
</tr>
<tr>
<td>Old</td>
<td>[175]</td>
<td>[75.1]</td>
<td>[.56]</td>
<td></td>
</tr>
<tr>
<td>Inferred</td>
<td>[26]</td>
<td>[81.2]</td>
<td>[.48]</td>
<td></td>
</tr>
</tbody>
</table>

*Factors not selected as significant in square brackets

In this model, within the variable <quantifier>, a value favoring overt pluralization was the absence of a quantifier. In the presence of a quantifier,
overt pluralization was strongly disfavored (factor weight of .07). The typical option of quantifier presence is illustrated in (5) below.

(5) Siman pasá a firma 6 konbenio⁴.

Week past PAST sign six agreement

‘Last week they signed six agreements.’

Consistent with previous studies, <formal definiteness> was selected by the model as a significantly favoring an overtly marked plural NP. While the effect of this variable was not as strong as <presence of a quantifier>, definite NPs registered a factor weight of .77 in favor of overt pluralization. Formally definite plural NPs are more typically marked with the overt plural morpheme, as illustrated below:

(6) Gobernador di Aruba lo inicia un investigacion riba e

Governor of Aruba FUT start INDEF investigation over DEF

realidad di e calculationnan den e presupuesto 2014.⁵

reality of DEF calculationPL in DEF budget 2014

‘The governor of Aruba will initiate an investigation on the reality of the calculations in the 2014 budget.’

Additionally, when the definite article is present, it overrides the tendency toward non-overt marking in the presence of a quantifier as described in Dijkhoff (1983). An illustrative example is given in (7) below.

---


Counter to Kester & Schmitt’s (2007) claims that animacy is not a contributing factor to PA pluralization strategies, the current data is sensitive to this variable. However, this tendency appears to be only slight, with animate entities only registering a factor weight of .59 in favor of overt pluralization. An example of an overtly marked animate entity is provided in (8) below:

(8) *Ku avisonan den medionan di komunikashon i spòtnan na radio, e kompania ta purba mishi ku e sentimentu di siudadanony.*

‘With advertisements in the media and radio spots, the company is trying to stay in touch with the feelings of the citizens.’

<Discourse status>, on the other hand, was not selected as a significant conditioning variable in this data. However, discourse-old referents registered a factor weight of .56, suggesting a marginal favoring of overt pluralization. Nonetheless, a Chi-square test, yielding a Pearson Chi-square value of .117 (see Table 6) confirmed that discourse status does not contribute to the overt pluralization model for the current set of data.
6. Conclusion

This study was motivated by the following research questions: (1) Are the factors found to affect the overt pluralization of noun phrases in qualitative studies found to be significant in a quantitative analysis?, and (2) What is the effect of discourse status, barely addressed in previous studies, on overt pluralization on PA? In the analysis presented, <quantifier>, <formal definiteness>, and <animacy> were the variables selected as statistically significant factors conditioning for overt plural marking. Unsurprisingly, the two variables that most favor plural marking are precisely those that are the most frequently mentioned in the majority of descriptive studies on PA (Dijkhoff, 1980, 1983, 1993; Maurer, 1998). The observed marking tendencies coincide with the observations of Poplack & Tagliamonte (1994), in their comparative study on standard Nova Scotian English and various, historical varieties of African American Vernacular English, who find that such tendencies are prototypical of English creoles.

Contrary to the observations of Kester & Schmitt (2007), the analysis of the data set in the present study identified animacy as exerting a statistically significant effect, with animate entities slightly favoring overt pluralization. In order to evaluate whether or not this finding is unique to the data set considered, a larger study using the same approach should be carried out. If the findings of a larger study were to hold, then further quantitative studies with different genres could be conducted in future research. Despite the significance of this variable, it would be difficult to regard PA as an animate-sensitive language in terms of Haspelmath’s (2013) taxonomy of plural expressions. Given the relatively low factor weight of this variable, it very well may vary due to other factors, for example, due to genre. It could be, for example, that in oral dialogic discourse, because of interlocutor-related issues, animacy might have a different impact on marking pluralization. For instance, Corbett’s (2000) animacy hierarchy shown in (9) below states that the leftmost animates are more salient in discourse. Likewise, they are more likely to receive overt pluralization. On the other hand, other more inanimate entities would be less likely candidates for overt pluralization.

(9) Corbett’s (2000) animacy hierarchy

speaker > addressee > 3rd person > kinship terms > other humans > “higher” animals > “lower” animals > discrete inanimates > nondiscrete inanimates.
Finally, discourse status does not contribute to the understanding of the distribution of overt plural marking in the analysis of the current data set. Rather than supporting the hypothesis that further mentions of a referent in discourse would be less lexically complex, this variable appears to pattern slightly toward overt pluralization, but this may be due to the relationship between certain discourse-old NPs formal definiteness rather than their discourse status. However, it might be insightful to carry out a qualitative analysis of subject continuity in addition to a quantitative approach like this one. This remains for future research.

References:


Variable Papiamentu pluralization strategies in a written corpus


