



Null subject pronouns in Papiá Kristang (Malacca Creole Portuguese)

Larissa de Santana Silva, Centro de Linguística da Universidade de Lisboa, MO, santtannalarissa@hotmail.com

Alan Norman Baxter, Universidade de São José, Macau, alan.baxter@usj.edu.mo

This paper investigates variation in the subject pronoun expression (SPE) in Papiá Kristang (Malacca Creole Portuguese), a language without verb inflection. Previous accounts of Kristang (Baxter, 1988, 2012b; Hancock, 1969, 1973; Knowlton, 1964; Rêgo, 1942) present the language as predominantly displaying overt subject pronouns. However, the overall null subject pronoun rate in the oral corpus contemplated is substantial, registering a frequency of 45.7% (input *pr.* 0.32). As such, the paper focusses on the distribution of null subject pronouns and the constraints directing that distribution. The study considers the effects of priming, structural location, verb semantic class, and tense-mood-aspect (TMA), and addresses the effect of the individual speaker. It is found that while null subject pronoun use in Papiá Kristang resembles, with minor differences, many attributes of this mechanism in other languages, it displays one notable trait: TMA particles favor null subject pronoun use. Extralinguistic constraints on SPE within the Papiá Kristang speech community are seen to concern more the characteristics of individual speakers, and clusters of individuals, rather than predictors based on a pan-community age and gender classification.



1. Introduction

Papiá Kristang (PK)¹ is spoken by a small community in Malacca, in West Malaysia, and in Singapore (Baxter, 1988, 2012a). The language has its roots in the 16th and 17th century Portuguese colonial presence in Malacca, and it has traditionally co-existed with other local languages, including Malay, Baba Malay and Hokkien Chinese. Today its speakers are fluent in Malay and English.

Typologically, PK belongs to the Southeast Asian subgroup of Asian Creole Portuguese, referred to as Malayo-Portuguese (Holm, 1988, pp. 290–298), characterized by a highly analytic grammar. Unlike its European Portuguese lexifier, PK is devoid of verb inflection, yet it has referential and non-referential null subject pronouns, and permits null subject pronouns with all person-numbers.² The fact that previous accounts of PK (Baxter, 1988, 2012b; Hancock, 1969, 1973; Knowlton, 1964; Rêgo, 1942) describe the language as predominantly using overt subject pronouns makes the null subject pronoun an item of interest. In this sense, PK is also unlike Portuguese (or Spanish, for that matter), which are traditionally viewed as displaying predominant use of the null subject pronoun and where it is the distribution of the overt subject pronoun that is generally the topic of research interest.

The current paper studies the distribution of null subject pronouns in PK from a variationist sociolinguistics perspective, bearing in mind certain aspects of the typological and formal bases of the phenomenon. It builds on the explanatory model proposed in the study of PK pro-drop by Silva (2020), and refines aspects of the analysis of the linguistic and extralinguistic conditioners of the variation. Particular interest concerns the place of TMA marking and the relationship between individual speaker, and speaker age and gender in the community profile of null subject pronoun use.

2. Background

Research over the last three decades concerning the typological and formal underpinnings of languages displaying null subject pronouns (NSLs) has demonstrated that the characterization

¹ The language is commonly referred to by its speakers as Papiá Kristang, or simply Kristang, and in English as Malacca Portuguese. They refer to themselves as Kristang or Malacca Portuguese.

² In typological and formal syntax studies (e.g. Biberauer et al., 2010), languages that permit a null subject pronoun, are referred to as Null Subject Languages (NSL(s)), or pro-drop languages. Sociolinguistic studies that contemplate variation between overt subject pronouns and null subject pronouns generally refer to this phenomenon as variation in Subject Pronoun Expression (SPE). Nevertheless, the discussion of the null subject phenomenon via a combination of typological, formal and variationist approaches, a perspective first promoted by Tarallo and Kato (2007[1989]), often leads to a crossover of terminology (e.g. Lucchesi 2009, among others). In the current study, which is essentially variationist but which will necessarily refer to aspects of typological and formal accounts, we will refer to Null Subject Languages alternatively as pro-drop languages, and we will refer to the null subject pronoun (NSP) alternatively as a case of pro-drop.

of such languages within the parametric perspective of Rizzi (1986) is not well supported cross-linguistically (Gilligan, 1987, Newmeyer, 2005). D’Alessandro (2015) identifies four broad classes of NSL: (i) *canonic*, languages with rich verb inflection, permitting pro-drop of all subjects (e.g. Italian, European Portuguese); (ii) *radical*, lacking verb inflection but permitting pro-drop of all subjects (e.g. Chinese); (iii) *partial*, permitting a degree of subject pro-drop according to specific conditions (e.g. Brazilian Portuguese); and, (iv) *expletive*, permitting only pro-drop of expletives (e.g. Dutch).³ PK, having no verb agreement morphology and permitting null subjects with all person-number subjects, thus displays certain characteristics of the radical type of NSL. Indeed, it will be seen that its overall rate of pro-drop is within the 40% range, which brings it near the 47% rate observed for Mandarin by Jia and Bayley (2002, p. 108), yet also near the highest rate observed in a variety of Brazilian Portuguese: 43% (Novaes 2007, p. 156). In languages that lack verb inflection, the distribution of overt and null subject pronouns has mainly been researched within the perspective of generative grammar, invoking the role of discourse information. This type of approach views topic prominence as being interrelated with subject properties, whereby a topic may cease to be overtly expressed in topic chains, and functions as a discourse antecedent for pro-drop (Huang, 1984; Yang, 2010, among others).

Recent research on Mandarin Chinese by Frascarelli and Casentini (2019) proposes that licensing of pro-drop involves interpretation strategies based on the informational structure of the discourse. It is proposed that the interpretation of a referential pronoun, null or overt, depends on a corresponding relationship with a specific type of topic: an A-Topic (Aboutness-shift topic), introducing a new topic into discourse.⁴ The A-Topic embodies illocutionary force, and proposes what the sentence is about: “it is an *initiating* speech act providing the “entry” under which the subsequent proposition (an assertion, a question, a command, etc.) will be stored” (Frascarelli & Casentini, 2019, pp. 3–4). If an A-Topic is maintained for more sentences, it needn’t be performed openly (i.e. phonetically) in each subsequent C domain. This gives rise to topic chains in which the null A-Topics permit a relationship between the continuous A-Topic and each subsequent null subject.

Other generative approaches have concentrated on the role of formal features associated with the nominal category. Thus, Neeleman and Srendöi (2007), proposed that pro-drop will apply to those languages in which formal features are not all present to permit full pronoun insertion at the case phrase (KP) level, and which are distributed in lower structural levels. This

³ Extensive discussion of the different types of pro-drop languages is beyond the scope and size limitations of the present paper. More details on the classification of types of pro-drop languages are presented by Biberauer et alii (2010) and Borges-Gonçalves (2023).

⁴ This perspective depends on the concept of the split C domain in generative syntax (Rizzi 1986), according to which the COMP (Complementizer) involves at least two functional categories – Illocutionary Force (or speech-act modality), and Finiteness.

is exemplified by Japanese (in which pronouns carry a case morpheme) and Chinese (in which number is realized as an independent morpheme). Subsequent generative grammar research associates null subject pronouns with formal features in TENSE, although there are differences in the details of the proposals: for partial null subject languages, the null subject pronoun is orientated by a definiteness feature associated with phi features (Holmberg & Roberts, 2013), whereas proposals regarding radical null subject languages invoke the presence of different features, for example general nominal features and an unvalued reference feature (Phimsawat, 2011), and phi features and/or discourse features (Miyagawa, 2017).⁵

In variationist studies, a great deal of attention has been given to the identification of linguistic characteristics that direct the licensing of the subject as a null or overt pronoun, investigating morphosyntactic, semantic and discourse-related dimensions. These studies have mainly concerned Portuguese and Spanish. A large body of research addresses pro-drop in Brazilian Portuguese (Carvalho & Besset, 2015; Genuino, 2017; Novaes, 2007; Paredes Silva, 2003; Silveira, 2012; among many others), São Tomé Portuguese (Bouchard, 2018), and diverse varieties of Spanish (Erker & Guy, 2012; Orozco, 2016; Silva-Corvalán, 1997; among many others). A number of frequently observed linguistic constraints are reported (Carvalho, Orozco & Shin, 2015, xiv–xv):

- (i) Person and number: singular pronouns are more frequent in use than plural pronouns; 2SG is more favorable to the null pronoun (\emptyset) than 1SG
- (ii) Change in reference favors a subject pronoun
- (iii) Priming: a prior subject pronoun favors a subsequent subject pronoun; a prior \emptyset favors a subsequent \emptyset
- (iv) Ambiguous verb morphology promotes more pronoun use
- (v) Verb semantic class and aktionsart: 'dynamic' verbs favor \emptyset , 'cognitive' verbs favor *pro*
- (vi) Clause structure: \emptyset is less likely in main clauses, more likely in dependent and coordinated clauses

These variation studies involve languages where the subject is represented redundantly in the verb morphology. Yet, as they are the principal variation studies on SPE, they provide substantial methodological and descriptive bases for addressing SPE in non-inflecting languages such as PK. Furthermore, for PK, historically derived through contact involving Portuguese, it seems

⁵ Owing to word-count limitations, and the overall variationist approach of the current paper, we will not provide details of the workings of the feature-based mechanisms proposed in these references and sketched in Borges-Gonçalves (2023).

reasonable to bear in mind the possibility that regular patterns of null subject pronoun use in earlier varieties of European Portuguese could have exerted influences on PK.⁶

Little variation research has addressed languages of the radical null subject pronoun type. Notable exceptions are Jia and Bayley (2002), Li, Chen & Chen (2012), and Li & Bayley (2018), who study Mandarin Chinese data and confirm the relevance of the above constraints (i), (ii) and (iii).⁷ Furthermore, where Portuguese-lexified creole languages are concerned, null subject pronouns have been studied from a variationist perspective only with regard to Cape Verde Creole, where Rodríguez-Ricelli (2019, 2021) also confirms the effects of constraints (ii) and (iii). However, the creole Portuguese varieties of the Atlantic region that display null subject pronouns are typologically quite different from PK since they all have mechanisms redundantly representing the subject: subject clitics in Cape Verde Creole, and weak pronouns (behaving like subject clitics) in Papiamentu and Saramaccan (Hagemeyer, 2009; Kowenberg, 2007; Kowenberg & Scott, 2010; Veenstra, 1994, 2006).

In a first study of pro-drop in PK, Silva (2020) also confirmed the effects of reference continuity, priming and verb semantic class but registered certain differences in other factors when compared with other studies. For example, subject person-number was not selected as significant, and structural location of the variable revealed syndetic coordination as disfavoring null subjects, as also did dependent clauses. However, TMA marking had an effect on null subject presence potentially aligned with narrative structure, a point initially raised earlier by Bailey and Pease-Álvarez (1997) and Silva-Corvalán (1997). Furthermore, as in many studies, PK appeared to be only marginally conditioned by extralinguistic factors. The current study refines and extends the earlier analysis to re-address these topics.

3. Methodology

The data for analysis were extracted from sociolinguistic interviews conducted in the Malacca Portuguese Settlement⁸ community during 1980–1981 (Baxter, 1988). The significance of this material lies in the fact that, at that time, PK still had many creole-dominant speakers, and was much more widely spoken by the creole community than it is today.⁹ Our data comes from

⁶ Unfortunately, there are as yet no variationist studies of SPE in earlier varieties of European Portuguese.

⁷ During research for the current paper, no variationist studies of SPE in Malay or Hokkien Chinese were located.

⁸ The Portuguese Settlement (known as *Padri sà chang* ‘the priest’s land’) is essentially an exclusive Kristang ethnolinguistic domain. It was established in 1933, in Ujong Pasir, to the south of Malacca town, in order to re-house poorer members of the Kristang community. By law, only Kristang families may reside in the Settlement. In the early 1980s it had a population of approximately 1100 (Baxter, 1988, p. 10), whereas today it has approximately 2300 (p.c. Philomena Singho, 3/1/2024).

⁹ The Kristang community today shows a strong shift to English as its dominant language, with shrinkage of use of PK (Baxter, 2012a; Lee, 2004, among others). Nevertheless, considerable effort is being put into language revitalization (Pillai, Phillip, & Soh, 2016).

interviews with eighteen speakers, male and female, in three age groups.¹⁰ All speakers were born to Kristang parents and had spent the greater part of their lives living in Malacca. With one exception (F3L – born in Kelang), all were born in Malacca and, except for M3W (from the former Kristang cluster in Trankera¹¹) and F1K (born of parents from Trankera), all were from families traditionally residing in the Banda Hilir Kristang community.

The study follows Labovian variationist methodology (Guy & Zilles, 2007; Tagliamonte, 2012). The realization of the subject is treated as a binary variable with two variants – a null subject and a phonetic pronominal subject, potentially conditioned by linguistic and extralinguistic factors. The analysis focusses on the null variant.¹² The hypotheses regarding the conditioning of the null subject pronoun were instrumentalized as predictors.

Instances of realization and non-realization of the subject pronoun with human referents were extracted from the participant's contribution in each interview. These were reduced by elimination of non-variable instances, non-specific references, instances of V + V, modals, and existential *teng* 'be', and formulaic expressions. The current analysis also excluded instances such as (1) and (2) below, where the subject of the non-finite verb within the direct object or purposive clause is deleted under identity with the referent of the dative object.¹³

- (1) *Eli já beng mandá ku yo bai skola tona.*
 3SG PFV come send DAT 1SG go school again
 'He came and ordered me to go to school again' [F1A]
- (2) *Eli mandá Ø bai kaza*
 3SG order (1SG) go house
 'He ordered (me) to go home' [F1A]

¹⁰ The data were recorded by the second author, who is an L2 speaker of PK and Malay. Consent to use the recordings for research purposes was obtained at the outset, on the condition that the identities of the participants and other persons mentioned in the recordings be suppressed. The interviews were recorded after eight months of residence in the community and were conducted in the presence of additional PK speakers. After transcription, the interviews were attributed a code, and all identities rendered anonymous.

¹¹ At the time when the data in this study were collected, the majority of the Kristang population lived in the southern coastal area outside Malacca town, at Banda Praya, in Banda Hilir and in the Portuguese Settlement at Ujong Pasir, the latter location having the largest concentration. A further offshoot of the community was formerly located at Trankera, in the northern part of the old Malacca town area (Baxter, 1988). In earlier times, in the more ethnically diverse residential settings of Banda Praya and Trankera, PK speakers were in closer contact with local speakers of Malay, and in particular with the Baba Chinese, speakers of Baba Malay, a pro-drop language (Lee, 2014, pp. 257, 263).

¹² While many studies of Brazilian Portuguese and Spanish concerned with SPE focus on the use of the overt subject pronoun, the current study is not alone in its focus on the null subject pronoun, as may be seen from the research of Rodríguez-Ricelli (2019, 2021) on Cape Verdean Creole, Lucchesi (2009) on Afro-Brazilian Portuguese, and Bouchard (2018) on São Tomé Portuguese.

¹³ In the first study of subject pronoun expression in PK (Silva, 2020), *kù* in such structures was treated as a complementizer and the PRO + verb sequence was treated as a single clause. For the revised coded data-base, these data were excluded, as *kù* is analysed as a dative preposition, and the subsequent verb is in a separate clause.

Following CROSSTAB inspection, the data-base was further reduced through exclusion of all tokens containing empty cells within each predictor, resulting in a final total of 2490 tokens. Factor size within predictors was restricted to 30 tokens. A mixed-effects logistic regression was applied, using RBRUL (Johnson, 2009). Individual speaker and verb group were treated as random intercepts.

3.1. Coding

The data for the present analysis were coded for the following predictors:¹⁴

(i) Person-number: The effect of subject person-number on pro-drop has been reported in numerous studies. These include partial pro-drop languages such as Spanish and Brazilian Portuguese, and radical pro-drop languages such as Mandarin Chinese. The effects observed display differing patterns. In a number of studies of Spanish varieties, plural person-numbers favor the null form, whereas singular person-numbers tend to favor the overt pronoun (Carvalho & Child, 2011, p. 19; Orozco & Hurtado, 2021, p. 14; Otheguy & Zentella, 2012, p. 790).¹⁵ Yet, in Afro-Brazilian Portuguese, Lucchesi (2009, p. 177) found plural person-number and second-person singular to favor the overt pronoun, whereas the rest favored the null form. Duarte (1995, p. 48), studying Rio de Janeiro Portuguese, found a similar result, with the exception that third-person plural also favored the null form. Finally, in studies of the radical pro-drop language Mandarin Chinese, the null form is favored by second person plural and third person plural, whereas singular person forms and first person plural generally favor the overt pronoun (Jia & Bayley, 2002, p. 110; Li & Bayley, 2018, pp. 151–153; Li, Chen & Chen, 2012, pp. 103–105).

In the study of PK subject pronoun expression by Silva (2020, pp. 89–91), person-number was not selected as statistically significant, although its profile appeared suggestive of a conditioning effect. The present study re-considers person-number with the smaller and more orthogonal database in order to clarify the constraint status of person-number, and re-assess the hypothesis of Neeleman and Srendöi (2007) that agglutinating pronominal morphology and pro-drop are connected. In the PK pronoun system, the plural forms of the second and third person may be considered non-fusional. These are, respectively, *bolotu* (<*bo(s)* + *otru* ‘other’) and *elot(r)u* (>*eli* + *otru*), and are derived from the corresponding singulars, *bos* and *eli* by addition of the segment *otru*, that might be interpretable as a plural marker. In the present analysis, person-number values were coded as singular versus plural forms.

¹⁴ In addition to these predictors, <sentence function> was also coded, following Silva (2020) and Novaes (2007). However, as it was defectively distributed across predictors, it was excluded from the analysis.

¹⁵ Yet, among the singular pronouns, different profiles have also been reported, for example with third person pronouns as the favoring factor: Spanish (Torres-Cacoullós & Travis, 2019, p. 25), São Tomé Portuguese (Bouchard, 2018, p. 17), and Cape Verde Creole (Rodríguez-Ricelli, 2021, pp. 147–148).

(ii) Structural location of the variation: The effect of structural position of the subject has been noticed in a number of studies on different languages, including for example Hebrew (Melnik, 2007), Russian (Bizarri, 2015), and in variationist studies of Spanish (Guy & Orozco, 2008; Otheguy & Zentella, 2012, Torres-Cacoullous & Travis 2019, among others), Portuguese (Genuíno, 2017; Menon, 2000, among others) and Cape Verde Creole (Rodríguez-Ricelli, 2021). Differential effects are reported for subjects of main clauses, and coordinate and subordinate clauses, with the null variant frequently reported as prominent in the latter two structures.

Torres-Cacoullous and Travis (2019) have drawn attention to a graded effect of different types of clause-linking on null subjects in Spanish and English, noting differential roles for syntactic and prosodic linking in both languages. In the present study, re-assessment of this factor group investigates the possibility of a graded effect, and is further motivated by the exclusion of tokens coding instances such as those of examples (1) and (2), involving non-finite clause direct objects or purposive clauses of bitransitive verbs. The effect of linking is assessed by means of the following factors: (i) antecedent in an independent sentence, as in (3); (ii) conjunction by formal link, as in (4); (iii) first juxtaposed clause, conjunction by prosodic link alone, as in (5); (iv) complement subordination by prosodic link only, as in (6),¹⁶ and adverbial subordination, as in (7):

- (3) *Yo fai sibisu assistant store keeper. Ø ja bai sibisu konta di ungua anu.*
 1SG do work assistant store keeper SG1 PFV go work count of one year
 ‘I worked as an assistant shop keeper. I went to work for about a year’. [M3T]
- (4) *Nus papiá malayu kù olotu, mas nus prendé tudu na ropianu.*
 1PL speak Malay COM 3PL but 1PL learn all in English
 ‘We speak Malay with them, but we study everything in English’ [F2S]
- (5) *Nus fika juntado Ø ngka kazá.*
 1PL stay together 1PL NEG marry
 ‘We lived together (but) (we) didn’t marry’. [F2I]
- (6) *Yo lembrá eli ja parí na Malaca*
 1SG think 3SG PFV born in Malacca
 ‘I think he was born in Malacca’ [F1K]
- (7) *yo pun bai mar, yo mesu, kora yo já largá yo sa sibrisu na Jasin*
 1SG also go sea 1SG same when 1SG PFV leave 1SG GEN work in Jasin
 ‘I also went fishing, I did, when I left my job in Jasin’ [M3T]

¹⁶ In PK, it is rare for complement clauses to employ a subordinating conjunction.

(iv) Priming – Previous mentions of the subject and form of the reference: Since the early 1990s, many studies have noted that repetition is a significant factor for recall maintenance (Lucchesi, 2009; Paredes Silva, 2003; Scherre & Naro, 1991; Travis, 2007). Although in variation studies of subject pronoun expression in Portuguese and Spanish co-reference is rarely the first selected independent variable, variation studies of Mandarin-Chinese are unanimous in this respect (Jia & Bayley, 2002, p. 109; Li & Bayley, 2018, p. 150; Li, Chen & Chen, 2012, pp. 103–104). Yet, these studies do not consider the actual form of the antecedent. They merely report that antecedents favor the null subject, whereas a new reference favors a pronoun (Li & Bayley, 2018, p. 150; Li, Chen & Chen, 2012, p. 104).

Where the form of the prior reference is concerned, while variation studies of Spanish and Portuguese report subject pronouns or null subjects to be favored, respectively, by the prior same representation, Lucchesi (2004) and Novaes (2007) both concerning varieties of Brazilian Portuguese, found a prior NP or a prior null subject to favor a subsequent null subject pronoun. For Cape Verde Creole, Rodríguez-Ricelli (2021, p. 150) also confirmed this effect.

Similar findings to those of Lucchesi (2004) and Novaes (2007), were reported for PK in Silva (2020, pp. 73–76). Nevertheless, in that study, <form of antecedent> and <occurrence of an antecedent> were treated as separate predictors. In contrast, in order to improve orthogonality, the current study coded for a single predictor addressing both antecedence and form of antecedent according to whether the subject was previously mentioned, a pronoun, example (8), as an NP, example (9), or as a null subject, example (10), or whether not mentioned previously, example (11):

- (8) *Nu bai skola, Ø sabe papiá ingrés.*
 1PL go school, 1PL know speak English
 ‘We go to school, (we) know how to speak English’ [F2I]
- (9) *Yo sa abó ja muré, Ø ja largá aké chang kù sà familia*
 1SG GEN grandmother PFV die 3SG PFV leave DEM land DAT GEN children
 ‘My grandmother died, (and she) left the land to her children’ [F2S]
- (10) *Yo ngka mutu chadu papiá, Ø bai skola Ø ngka mutu chadu, Ø ngka tomá*
 1SG NEG very clever speak, 1SG go school 1SG NEG very clever, 1SG NEG take
buku prendé.
 book learn
 ‘I was not very clever at speaking, (I) went to school but (I) was not very smart,
 (I) didn’t take the book to learn’ [F3M]

- (11) *Eli membes teng mpoku doda, aké ora nus anumbés ngka fai nus sa sibrisu*
 3SG sometimes be little crazy DEM hour 1PL maybe NEG do 1PL GEN work
retu.
 correct
 ‘She was a bit crazy sometimes, those times maybe we didn’t do our work properly’
 [F1C]

(v) Verb semantic class: The effect of verb semantic class on subject pronoun expression has been widely observed in studies of Spanish (e.g. Bentivoglio, 1989; Erker & Guy, 2012; Orozco, 2015; Travis, 2007) and Brazilian Portuguese (e.g. Carvalho & Child, 2011; Silveira, 2012)¹⁷ Speech verbs, epistemics, perception and stative verbs favor the subject pronoun, whereas all other verbs favor the null variant (Silveira, 2012, pp. 95–100; Orozco & Hurtado, 2021, pp. 3–4).¹⁸ Silva (2020, pp. 83–85) confirmed these broad findings for PK with a predictor containing six verb classes, some of which registered similar weights. Following the revisions to the database, it was decided to re-run the analysis limiting the range of classes.

The present study codes verb classes following the approach of Erker and Guy (2012), reducing this factor group to three factors: a mental activity class (e.g. *lembrá* ‘to think, to imagine’, *intendé* ‘to understand’, *keré* ‘to want, to desire’, a stative class (e.g. *teng* ‘to have’, *fiká* ‘to stay, to live’, *gostá* ‘to like’, *sabé* ‘to know’) and an external activity class (e.g. *kuré* ‘to run’, *komprá* ‘to buy’, *skribé* ‘to write’, *prendé* ‘to teach’). This coding takes into account the broad effect of the stative/dynamic opposition observable in PK across the more complex factor group analyzed by Silva (2020), plus the notion that verbs of the mental activity type may promote individualization of the subject via the pronoun (Enríquez, 1984; Erker & Guy, 2012).

(vi) The presence of a TMA marker: In variation research of pro-drop in Spanish and Portuguese, a number of studies have found that different tense-aspect forms effect pronoun variation, with the following frequently observed ranking profile conditioning the null subject: preterite > present > imperfect > conditional/subjunctive (e.g. Bayley & Pease-Álvarez, 1997; Carvalho, Orozco & Shin, 2015; Silva-Corvalán, 1997; Silveira, 2012). Silva-Corvalán (1997) proposed that Spanish TMA forms play a conditioning role by virtue of their pragmatic function in discourse: both the preterite and present are factual and assertive, but the preterite is dynamic, event-focussed,

¹⁷ The effect of verb semantic classes has also been addressed in studies of pro-drop in other languages. These include Bizzari (2015), on Russian, and Frascarelli (2018), on Italian and Finnish, who examine the effect of matrix factive verbs in opposition to matrix ‘bridge’ verbs (i.e. opinion or assertive verbs) with regard to the subordinate clause subject. In Arabic, Altamimi (2015) finds action verbs to favor pro-drop more so than psychological verbs.

¹⁸ Nevertheless, inconsistencies have been noted. For example, Orozco (2020), reporting research on the Barranquilla variety of Colombian Spanish, draws attention to the contradictory behavior of verbs classed within a single semantic category. Such findings point to the need to consider the role of the verb by means of alternative predictors, a point we mention in section 5.

and foregrounded, whereas the present tense is not always dynamic or focal. In contrast, other TMA forms, such as the imperfect, conditional and subjunctive are non-assertive, non-dynamic, irrealis and backgrounded. It is hypothesized that it is in the foreground that more null subjects are expected, where the event is focused, yet in the background, where the event is not focused, more subject pronouns are expected (Silva-Corvalán & Enrique-Arías, 2017, pp. 180–181).

Whereas Spanish is a tense-prominent language, PK is aspect-prominent. The Kristang verb functions primarily in terms of aspect, represented by four pre-verbal TMA markers,¹⁹ which perform six basic functions (Baxter, 2012b, pp. 124–125). Two markers have unique aspect functions: *ja* ‘perfective’, as in (12), *ta* ‘progressive’, as in (13); and the other two have double functions: (i) *lo(gu)* ‘future-irrealis’ and ‘habitual aspect’, the latter is shown in (14); and, (ii) \emptyset ‘imperfective continuous’ and ‘perfective’, as in (15). The markers *ta*, ‘progressive’, *lo(gu)* ‘habitual’, and \emptyset ‘imperfective continuous’ are transparent for past and present. However, in negated sentences, the aspectual markers are replaced by the negator *ngka* occurring with verbs representing perfective or imperfective aspect. In example (10) above, the sentence \emptyset *ngka tomá buku prendé* ‘I didn’t take the book to learn’, presents one such instance, where the aspectual value of the verb, as perfective or imperfective, would be recuperated from the prior discourse.

(12) \emptyset *já* *nasé* *na* *Padri* *sà* *Chang*
 1SG PFV birth PREP priest GEN land
 ‘I was born in Padri sà Chang’ [F2S]

(13) *Yo* *ta* *fala* *ungua* *stori*.
 1SG PROG speak one story
 ‘I was telling a story’ [M2B]

(14) *Ungua* *bela* *lo* *bai*. \emptyset *lo* *bai* *fala* *ku* *nus* *sa* *pai* *mai*
 one old-FEM HAB go 3SG FUT go speak DAT 1PL GEN father mother
 ‘An old lady will go. (She) will go and speak to our parents’ [F3M]

(15) *Eli* \emptyset *da* *binti* *sen*
 3SG IPFV give twenty cent
 ‘He gave (lit. used to give) twenty cents’ [F2X]

Silva (2020, pp. 85–88) observed a conditioning effect in PK null subject pronoun use according to the four forms of the TMA marker, potentially related to TMA function and discourse structure. In order to clarify whether the variation is due to the presence of an overt TMA marker, or

¹⁹ The TMA markers (and the negator) occur immediately adjacent to the verb.

whether it is functionally based, the present study adopts a slightly different approach, and considers the individual TMA marker functions plus the effect of the negator *ngka*.

(vii) Individual speaker: This factor group assesses the stability of the predictors among individual speakers regardless of the potential predictions of <age-group> or <gender> (Gorman, 2010; Johnson, 2009). In the study of PK pro-drop by Silva (2020, pp. 92–93), it was found that the effect of speaker sex was not statistically significant, although, numerically, women deleted marginally more than men, whereas speaker age-group was found to exert a very weak conditioning effect suggestive of an apparent-time change towards less use of the null subject.

4. Analysis

The analysis focusing on the null subject pronoun revealed an overall proportion of 45.7% null subjects, corresponding to an input of *pr* 0.32. In a converging run, 6 predictors were identified as conditioners of the variation, shown in **Table 1**:

Predictors	Result
Prior reference and form	$p < 5.57e-67$
Structural location relative to antecedent	$p < 5.22e-07$
TMA marking	$p < 3.28e-11$
Person-number	$p < 0.011$
Individual speaker	random
Verb semantic class	random
Run details: Input prob.	0.332
Total tokens	2486
Loglikelihood	-1427.835
Degrees of freedom	12

Table 1: Overall results of analysis of predictors.

The results of these predictors are presented in the following sections. Of the non-selected predictors, only age-group and gender will be considered in section 4.6 in relation to individual speaker profile.

4.1 Co-reference and prior form

The effects of priming are presented in **Table 2**, where it is seen that two factors favor the null subject: a null antecedent is quite favorable (pr. 0.814), whereas an NP is moderately favorable (pr. 0.623).

Presence of prior form	Total tokens	% null subject	Centered Factor weight
∅ (null)	684	74.9%	0.814
NP	213	54.9%	0.623
PRO	1453	34.4%	0.452
No prior mention	136	9.6%	0.144

Table 2: Effect of co-reference and prior form.

However, a prior pronoun is slightly unfavorable (pr. 0.452), perhaps because of the possibility of ambiguity of reference. In contrast, the absence of a prior mention is quite unfavorable to the null subject pronoun (pr. 0.144), as it is expected that the subject of a first mention would be realized so that the interlocutor can properly identify the reference, unless this identification is facilitated by pragmatic considerations.

On the one hand, these results confirm the role of priming in conditioning the dependent variable, with null antecedents leading to subsequent null subjects and pronominal antecedents leading to subsequent pronominal subjects. This coincides with the findings of variation studies by other researchers. Unfortunately, existing variation studies of subject pronoun expression in Chinese have not addressed the specific prior subject forms in priming. However, in studies of Brazilian Portuguese, the null antecedent favors a subsequent null subject, whereas an antecedent subject pronoun registers a weak disfavoring of a subsequent pronominal subject or registers a neutral value (Lucchesi, 2009, p. 180; Novaes, 2007, p. 170). Furthermore, variation studies of varieties of Spanish reveal a similar result, with some researchers reporting parallelism for both factors: null antecedents prime subsequent null subjects while pronominal antecedents prime the subsequent pronominal subjects (Carvalho & Bessett, 2015, p. 158; Orozco, 2015, p. 29; Torres-Cacoullos & Travis, 2019, p. 17; Travis, 2007).

Basically, the subsequent subject is licensed on the basis of the identity strength of the previous subject, constituting a referential sequence. Having already activated a particular representation of the subject, the speaker, accessing the previous reference, has no need to activate a new full (i.e. phonetic) representation for the subsequent identical subject, unless clarification of the reference were required.

Inspection of the prior NPs in the corpus reveals that there are only three instances in which an NP is not a first reference. Hence, the overwhelming majority of prior NPs constitute switch references. While we noted above that the prior NP may favor a subsequent null subject, it is obvious that references subsequent to an NP may also be represented by pronouns. Thus, in PK, the initial stage of a chaining effect anchored by an NP can be immediate, licensing the null subject, or prolonged, licensing first a pronoun and subsequently a null subject.

4.2 Structural location relative to the antecedent

Table 3 presents the results of the analysis of the five factors of the predictor addressing location of the subject. Two factors are favorable to subject pronoun omission: asyndetic coordination, registering a strong relative weight of 0.788, and the presence of the antecedent in an independent sentence, registering a weak favoring effect (weight 0.546).

Structural location	Total tokens	% null subject	Centered Factor weight
1 ^a juxtaposed clause in asyndetic coordination (within same intonation period)	112	68.8%	0.788
Different macro sentence (different intonation period)	2126	47.6%	0.546
Complement clause (nominal subordination, asyndetic)	89	24.7%	0.436
1 ^a clause following a syndetic coordinator (e.g. <i>mas</i> 'but')	90	21.1%	0.409
Adverbial clause	69	17.4%	0.296

Table 3: Effect of structural position on null subject.

In the instances of asyndetic coordination, the antecedent is in the first clause immediately preceding the subject of a first juxtaposed clause, within the intonation contour of the complex clause. The tight linkage between the two clauses leads to a high degree of accessibility (*pr.* 0.788). However, when the preceding reference is in a different macro clause (i.e. clause separated by a pause and clause-final intonation), the effect is weaker (*pr.* 0.546). In this case, a previously referenced subject, if still fully accessible, could lead to a subsequent \emptyset , whereas a prior reference not entirely accessible (for example, moderated by factors such as the presence of a competing referent, or a change of action involving the same referent) could lead to a subsequent pronoun.

The remaining three factors are unfavorable to the null subject. Indeed, the results for subjects in coordinate and subordinate clauses suggest that omission may be restricted by structural considerations. In syndetic coordination, a null subject is disfavored with a relative weight of 0.409 when the subject is located in the first clause after a coordinator. Where complement and adverbial subordinate clauses (with the antecedent in the matrix sentence) are concerned, a weight moderately more unfavorable to null subject is recorded: 0.296. In this case, the syntactic embedding involving a subject with the same person-number as the subject of the matrix may be affected by the need to clarify reference.

There are certain similarities with the results of other studies. Research on Portuguese reports pronouns most likely to be expressed in nominal subordinate clauses, and either close to neutral or disfavored in coordinate clauses (Carvalho & Bessett, 2015, p. 158; Genuino, 2017, pp. 105–109; Novaes, 2007, p. 169; Silveira, 2012, p. 80). Some studies of Spanish varieties also report subordinate clauses as favoring pronoun expression, while coordinate clauses favor omission (Carvalho, 2015, p. 158; Orozco, 2015, p. 24). The preference for null subject in asyndetic coordination has also been noted in a variety of Brazilian Portuguese (Novaes, 2007, p. 169). Where independent clauses are concerned, a number of studies of Brazilian Portuguese and Spanish several studies also report independent clauses as registering a result close to neutral (Carvalho, 2019, p. 158; Carvalho & Bessett, 2015, p. 158; Genuino, 2017, pp. 105–109; Menon, 2000, pp. 170–171; Novaes, 2007, p. 169; Orozco, 2015, p. 24; Orozco & Guy, 2008, p. 77; Silveira, 2012, pp. 100–101).

4.3 Verb class

The results of the analysis of verb semantic class, presented in **Table 4**, reveal only a mild effect on the subject pronoun variable, with dynamic verbs slightly favoring the null subject, statives close to the neutral point²⁰, and mental activity verbs mildly unfavorable. The PK weighting profile is somewhat similar to that observed by Torres-Cacoulllos & Travis (2019:671) for Colombian (Cali) Spanish,²¹ and Bouchard (2018:17) for São Tomé Portuguese, with the null subject occurring most with dynamics (Cali: *pr.* 0.57 (67%); São Tomé *pr.* 0.55 (72.1%)) and statives and mental activities unfavorable, respectively, in that order (Cali: *pr.* 0.42 (53%), *pr.* 0.31 (35%); São Tomé *pr.* 0.53 (64.3%), *pr.* 0.43 (61.8%)).²²

²⁰ Possibly because of the inclusion of semi-statives in this factor.

²¹ Torres-Cacoulllos and Travis (2019) studied variation in 1st and 3rd singular subjects.

²² Erker & Guy (2012, p. 542) report a similar profile for New York Spanish: dynamics are favorable to the null subject, 69%; statives and mental activities unfavorable, respectively, 64%, 55%.

Verb type	Total tokens	% null subject	Centered factor weight
Dynamic	1908	48.2%	0.544
Stative	347	40.9%	0.488
Cognitive	231	34.6%	0.468

Table 4: The effect of verb class.

We noted earlier that researchers have recently begun to question the validity of a predictor based purely on verb semantic class. Torres-Cacoullós and Travis (2019, pp. 21–23) take the discussion a step further, pointing to the effect of verb semantics in discourse structure. In relation to temporal sequencing, they propose that the favorability of dynamic verbs to null subjects involves subjects referring to temporarily sequential or simultaneous central events or situations relating to *foreground* discourse, whereas statives and epistemics relate more to *background* discourse (Torres-Cacoullós & Travis (2019: 21–23).²³ This proposal gains support from research on narrative structure and lexical aspect in relation to tense-aspect morphology (Bardovi-Harlig (2000; among others), that reveals the natural predominance of activity, achievement and accomplishment aktionsart classes in narrative foreground, as opposed to statives.

4.4 TMA marking

Initially, five factors were included for analysis according to TMA marker functions, yielding the results shown in **Table 5**. The future-irrealis function of *lo(gu)* was excluded because of low token numbers in the revised data-base.

The profile in **Table 5** lacks clear indications of a discourse constraint on TMA relative to subject pronoun expression. For example, the result for the habitual aspect marker *logu* seems anomalous in terms of the proposal by Silva-Corvalán (1979) that imperfective values are more expected in narrative background where the overt subject pronoun would be preferred (Silva-Corvalán & Enrique-Arías, 2017, pp. 180–181).

²³ Discourse *foreground* consists of events that belong to the essential structure of the discourse and that advance the discourse in the temporal axis, in a sequential manner (Hopper, 1979; Dry, 1983) such that that the time reference point of each event in *foreground* follows the previous event. The *foreground* comprises the events central to the discourse and its protagonists, so there is a strong relationship between topics, topic chains (topic chain or topic continuity) and the foreground (Givón, 1983). In contrast, the role of the *background* is to support the *foreground*, furnishing material that elaborates or evaluates *foreground* events (Hopper, 1979).

TMA marker	Total tokens	% null subject	Centered factor weight
Habitual <i>lo(gu)</i>	145	65.5%	0.589
Progressive <i>ta</i>	138	48.6%	0.568
Negator <i>ngka</i>	312	47.8%	0.558
Perfective <i>ja</i>	570	51.1%	0.521
Continuous \emptyset	1129	41.6%	0.396
Perfective \emptyset	192	36.5%	0.370

Table 5: Effect of TMA marking (Run 1).

However, comparing the overt TMA marked instances with the \emptyset -marked instances, the close proximity of values is suggestive of a structural effect, whereby overt filling of TENSE favors a null subject and covert filling favors a pronominal subject.²⁴ As such, a second analysis was conducted, reducing the group to two factors only: ‘overt’ versus ‘covert’. The result is shown in **Table 6:**

TMA marker	Total tokens	% null subject	Centered factor weight
Overt marker	1165	51.7%	0.577
Covert marker	1321	40.9%	0.423

Table 6: Effect of overt and covert TMA marking (Run 2).

On comparison of the log-likelihood values of the two analyses, by chi² test, it was found that the difference between the two analyses is not significant.²⁵ In this case, the more economical analysis is preferred, whereby it appears that the recuperation of the prior referent of the null subject could be aided by the presence of the overt aspect marker. In the latter case, the listener may expect that the prior reference is continued.

²⁴ We thank Greg Guy for preliminary discussion of earlier results pointing in this direction. Naturally, responsibility for the interpretation here and in section 5 is entirely ours.

²⁵ Run 2: log.likelihood -1427.425.

4.5 Person-number

In the study of PK pro-drop by Silva (2020, pp. 89–91), person-number was not a significant conditioner of PK pro-drop. However, in the current analysis, person-number coded as singular versus plural was selected, yet it is the least influential of the linguistic variables considered. The results in **Table 7** reveal only a slight favoring of null subject by plural person forms:

Person-number	Total tokens	% null subject	Centered factor weight
PLURAL	866	48.0%	0.532
SINGULAR	1620	44.8%	0.468

Table 7: Effect of person-number.

The profile is reminiscent of some of the findings of the Brazilian and Chinese studies noted in section 3.1, where singular person-numbers disfavor the null subject, whereas plural person-numbers favor it. Nevertheless, this result for PK does not provide clear support for the Neeleman and Srendői (2007) hypothesis that a language only allows radical pro-drop if its personal pronouns manifest agglutination or some other nominal feature. Although PK presents agglutination in the contrast between both second and third person singular and plural pronouns, the second-person plural forms were few (15 tokens, yet all with pronominal subjects) and were excluded from the analysis. Thus, only non-agglutinating first-person plural *nus* and agglutinating third-person plural *olot(r)u* were considered.

4.6 The individual speaker

The results from the analysis of this predictor clarify the curiosities noticed by Silva (2020, pp. 88–89, 92–93) concerning the weak tendencies in the distribution of PK pro-drop according to speaker sex and age-group. **Figure 1** presents the profile of the individual speaker weights:

The individual rate of null subject pronoun expression ranges from 0.282 (F1C, female, age-group 1)²⁶ to 0.667 (FM3, female, age-group 3). There are 11 speakers with scores above 0.5 and 7 with scores below this point. Looking more closely at the sample, several interesting tendencies emerge, as do also certain notable individual differences.

Scores above 0.5: Among the 11 speakers above the 0.5 mark, there are four speakers of age-group 3: the three women of this group, and one man. The three women had lives restricted to the home domain. Speaker F3M, with the highest pro-drop rate, was born into a traditional fishing family in the Banda Praya community, and had no formal education. During interview,

²⁶ Speakers are identified by sex (F, M), age-group (1, 2, 3) and an individual identification code, thus F1C = female, age-group 1, identification code C.

she commented that she had scant knowledge of English. PK was clearly her dominant language, although she also knew local Malay.²⁷ The other two women in this group (F3L and F3E) had fathers who were in employment connected with the British colonial presence, and had lived for short periods in other locations in East Malaysia. They had minimal education primary level, as did the male MT3 (considered below). For this age-group, it is important to bear in mind that all speakers, men and women, were born in other locations in Malacca, and principally in Praia Lane and Trankera, traditional residential locations where Kristangs lived in contact with Baba Chinese, speakers of Baba Malay, a pro-drop language (Lee, 2014, pp. 257, 263).

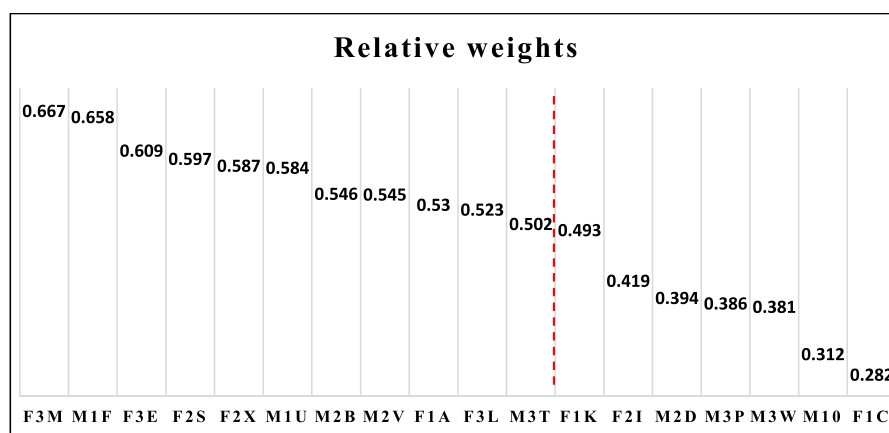


Figure 1: Individual speaker, effect on null subject pronoun.

Also, within the above 0.5 range, there are four age-group 2 speakers: F2S, M2V, F2X and M2B. F2S and M2V stand out as they are ‘professionals’ (primary school teacher, and nurse) with secondary and specialist education, yet their pro-drop scores are above the 0.6 point. Both F2S and M2V had always lived in the Ujong Pasir Kristang community. F2S is from a traditional fishing family. M2V’s father was employed by the British administration as a lighthouse attendant, and was often absent from the home setting. The other two speakers in age-group 2, F2X and M2B, are from fishing families. F2X had a traditional home-domain existence and M2B had worked as a fisherman and as a gardener in government employ in Independent Malaysia. Both had minimal primary education. The fact that there is minimal difference in the scores of these members of age-group 2 might be explained by the fact that they grew up in a period in which PK was more widely spoken in the community, when there was greater community network cohesion and when there were more Kristang-dominant speakers.²⁸

²⁷ In the early 1980s, lack of knowledge of English among elderly Kristang women was evident (Baxter, 1988, p. 13).

²⁸ It should also be noted that prior to the 1980s, there was more involvement in fishing, more poverty, and generally less adhesion to education (Chan, 1969).

Of the remaining two members of the above 0.5 group, M1F has the second highest pro-drop score in the data. This speaker was from a traditional fishing family and resided in a house with 16 members of extended family in a street that was known to most speak PK.²⁹ At the time of the interview, he worked in an electronics factory. The second age-group 1 speaker is M1U, from a traditional fishing family, with a father versed in Kristang oral traditions. He had primary education to level 3, was employed in a soft-drink company, yet also worked as a fisherman.

In sum, in the set of speakers above the 0.5 cut-off, there are some suggestions of an age-related tendency, as 8 of the ten speakers are in age-groups 2 and 3.

Scores below 0.5: Now, turning to the 8 speakers in this range, two facts stand out: (i) two of the three men of age-group 3 (M3P, M3W) score below 0.5, whereas MT3 scores just above the neutral point; and, (ii) four of the six speakers of age-group 1 are in this range, including all three females of this age-group. In age-group 3, the three men had primary education to level 3 or 4. However, they had once had employment with colonial institutions and businesses (M3T – rubber plantation clerk; M3P health inspection assistant; M3W – attendant, vehicle import company) where they interacted with a wider community and in use of English was important. M3T and M3W were also musicians, and occasionally worked away from the community. M3T and MP3 also participated in the British colonial military volunteer battalion during the Japanese occupation of former British Malaya and Singapore.

In age-group 2, speaker F2I (daughter of M3P) completed education to level 4 secondary,³⁰ subsequently married a Kristang government clerk and was thereafter occupied in the home domain. In contrast, M2D completed level 4 primary, spent 7 years in the army (under the British) and subsequently worked in the Malacca Public Works department before turning to full-time fishing.

Among the speakers of age-group 1, while F1K had only primary level 3 education, the educational level of the others is relatively higher, with F1A, M1O, F1C attaining secondary school level. F1A and F1K were occupied in the home domain, whereas M1O (government clerk) and F1C (teaching assistant, and sportswoman) were active beyond the community.

Overall, among the speakers with scores below 0.5, some facts are suggestive of a conditioning of pro-drop by age-difference and sex: the strong presence of age-group 1 speakers, especially females, whereas the presence of two men of age-group 3 (but absence of women of this age-group) might relate to a time when the role of Kristang women was very much restricted to the home-family and immediate community domains. On the other hand, exposure beyond

²⁹ A decade before the recording of the Kristang corpus considered here, Chan (1969, pp. 258–261) noted the importance of the extended family for use of Kristang.

³⁰ This speaker was one of several from age-group 2 who pointed out that it was common in her generation, especially in families that valued education and improvement, for parents to encourage the use of English in the home in the belief that this would give their children better chances of good employment.

the community, notably in occupations involving interaction and requiring language skills are suggestive of a negative influence on pro-drop.

5. Discussion

The distribution of null subject pronouns in PK resembles in certain aspects the profiles detected in variation studies of consistent null subject pronoun languages such as Brazilian Portuguese, São Tomé Portuguese, and Spanish, yet also, to a lesser extent the results of variation studies of languages devoid of verb inflection, such as Cape Verde Creole and Chinese.

As such, in PK, priming was found to be the foremost conditioner of the null subject pronoun. Second to priming, the structural location of the antecedent also exerts a significant constraint, with the commonly observed disfavoring of NSPs in subordinate clauses and slight disfavouring in syndetic coordination. However, the influence of a third predictor, the presence of a TMA marker, is somewhat of a novelty. So far, the few variation studies of NSP languages devoid of verb inflection have found this predictor to be insignificant, for example in Chinese (Jia & Bayley, 2002, p. 108) and also in Cape Verde Creole although in the latter case a switch in TMA reference was found to be significant, yet inhibiting NSP use (Rodríguez-Ricelli, 2019, pp. 355–356,405).

The thread common to these three predictors in PK is the licensing of a null subject pronoun based on the quality of representation, referential continuity and accessibility of the antecedent, factors that may obviate the activation of a fuller (i.e. phonetic) pronominal derivation of the subsequent subject (cf. Roberts, 2010; Frascarelli & Casentini, 2019). The structural location of the antecedent acts as a further constraint on the form of the subject pronoun, facilitating referential continuity, and again obviating the need for building a phonetic representation of the subject pronoun.

Where the TMA predictor is concerned, it is useful to consider the implications of the proposals for a TP wherein T hosts a set of formal features such as definiteness, reference and discourse, that facilitate the building of the co-referential pronominal (Roberts, 2010; Phimsawat, 2011; Holmberg & Roberts, 2013; Miyagawa, 2017). As such, we might hypothesize that the favoring of NSPs by the overt realization of TMA markers implicates a feature, say definiteness, enabling the identification of the referent and yielding the NSP. Nevertheless, this hypothesis will need testing in future research by investigation of the effect of TMA switching.

The effect of verb semantic class proved to be quite weak, yet its profile does resemble that found in some previous variationist research (Bouchard, 2018; Torres-Cacoullous & Travis, 2019). Although verb semantic class has been incorporated as a predictor in many studies, its results are quite diverse and generally mild. As hinted in section 4.3, it may be useful to assess the effect of the verb according to an aktionsart classification as used in recent studies of tense-aspect acquisition. Further alternatives that require exploration are a frequency-count classification

(e.g. Erker & Guy, 2012; Li & Bayley, 2012; Orozco & Hurtado, 2018), and a classification in terms of valency (e.g. Orozco & Hurtado, 2021).

In PK, person-number proved to be the least significant of the linguistic predictors, although its results appeared at first glance to offer some support to the hypothesis of Neeleman and Srendöi (2007) that NSPs in radical pro-drop languages are facilitated if pronouns carry agglutinating morphology (i.e. features not present for pronoun insertion at case-phrase level). The results showed plural pronouns to favor NSPs, as in certain other variation studies. However, support for the Neeleman and Srendöi hypotheses was undermined by the uneven distribution of agglutinated plural person-number data. Here too, further research is necessary, and will require a larger data base.

Finally, our analysis of individual speaker profiles revealed some indications of age-based and gender-based constraints on NSP use, oriented by differing life experiences of interaction in domains such as school, the home, and occupation. In this respect, the results obtained in Silva (2020), pointing to the effect of age in PK as well as a marginal relevance of gender, are not without reason. However, the motivations for those findings are better accounted for by individual speaker characteristics, which reveal a richer set of extralinguistic factors underlying the observed variation in NSP use.

6. Conclusion

The predictors considered here effect PK null subject pronouns in ways similar to those reported in studies of other null subject languages. Most noticeably, null subject pronouns in PK are governed strongly by internal factors rather than by extralinguistic factors. The null subject pronoun is directed by priming effects, structural location, verb semantics, and by overt marking of the pre-verbal tense-aspect component. The differences noted concerning TMA marking appear to be partly structural, with pronoun omission favored when the pre-verbal slot, TENSE, is filled with an overt TMA marker, and partly functional, connected with referential accessibility.

Where the community profile of the variation is concerned, null subject preference among older speakers (in this sample) may have its roots both in a past when Kristang women had a more socially restricted role, and to the community's origins in the localities of Praya Lane and Trankera, settings where there was a greater immediate contact with local varieties of Malay, in particular Baba Malay. In contrast, the slightly increased use of pronominal subjects among the younger age groups may come from the adoption of English as the community's dominant language.

Competing Interests

The authors have no competing interests to declare.

References

- Altamimi, M. I. (2015). Arabic pro-drop. (Unpublished master's thesis). Eastern Michigan University.
- Bardovi-Harlig, K. (2000). *Tense and Aspect in Second Language Acquisition: Form, Meaning and Use*. Blackwell Publishers.
- Baxter, A. N. (1988). *A Grammar of Kristang: Malacca Creole Portuguese*. Pacific Linguistics.
- Baxter, A. N. (2012a). The Creole Portuguese language of Malacca: a delicate ecology. In L. Jarnagin (Ed.), *Culture and Identity in the Luso-Asian World: Tenacities and Plasticities* (pp. 115–142). Institute of Southeast Asian Studies. DOI: <https://doi.org/10.1355/9789814345514-010>
- Baxter, A. N. (2012b). Kristang – Malacca Creole Portuguese. In S. M. Michaelis, P. Maurer, M. Haspelmath & M. Huber (Eds.), *Atlas of Pidgin and Creole Language Structures* (pp. 122–130). Oxford University Press.
- Bayley, R., & Pease-Álvarez, L. (1997). Null pronoun variation in Mexican-descent children's narrative discourse. *Language Variation and Change*, 9(3), 349–371. DOI: <https://doi.org/10.1017/S0954394500001964>
- Bentivoglio, P. A. (1989). Función y significado de la posposición del sujeto nominal en el español hablado. [Function and meaning of the postposition of the nominal subject in spoken Spanish.] *Estudios sobre Español de América y Lingüística Afroamericana*, 83, 40–58.
- Biberauer, T., Holmberg, A., Roberts, I., & Sheehan, M. (2010). *Parametric Syntax: Null Subjects in Minimalist Theory*. Cambridge University Press.
- Bizarri, C. (2015). Russian as a Partial Pro-Drop Language – Data and Analysis from a New Study. *Annali di Ca' Foscari. Serie occidentale*, 49, 335–362.
- Borges-Gonçalves, H. (2023). *Sujeitos nulos: uma revisão do estado da arte*. [Null subjects: a review of the state of the art.] *Fórum Linguístico*, 20(3), 9254–9279. DOI: <https://doi.org/10.5007/1984-8412.2023.e85486>
- Bouchard, M., (2018). Subject Pronoun Expression in Santomean Portuguese. *Journal of Portuguese Linguistics*, 17(1), 1–29. DOI: <https://doi.org/10.5334/jpl.191>
- Carvalho, A. M., & Bessett, R. M. (2015). Subject Pronoun Expression in Spanish in Contact with Portuguese. In A. M. Carvalho, R. Orozco & N. L. Shin (Eds.), *Subject pronoun expression in Spanish: a cross-dialectal perspective* (pp. 145–167). Georgetown University Press.
- Carvalho, A. M., & Child, M. (2011). Subject Pronoun Expression in a Variety of Spanish in Contact with Portuguese. In J. Michnowicz & R. Dodsworth (Eds.), *Selected Proceedings of the 5th Workshop on Spanish Sociolinguistics* (pp. 14–25). Cascadilla Proceedings Project.
- Carvalho, A. M., Orozco, R., & Shin, N. L. (2015). Introduction. In A. M. Carvalho, R. Orozco & N. L. Shin (Eds.), *Subject Pronoun Expression in Spanish: A Cross-Dialectal Perspective* (pp. xiv–xv). Georgetown University Press.

- Chan, K. E. (1969). A study of the social geography of the Malacca Portuguese Eurasians. (Unpublished master's thesis.) University of Malaya.
- D'Alessandro, R. (2015). Null subjects. In A. Fabregas, J. Mateu & M. Putnam (Eds.), *Contemporary Linguistic Parameters* (pp. 201–226). Bloomsbury.
- Dry, H. (1983). The movement of narrative time. *Journal of Literary Semantics*, 12, 19–53. DOI: <https://doi.org/10.1515/jlse.1983.12.2.19>
- Duarte, M. E. L. (1995). A perda do princípio “evite pronome” no português brasileiro. [The loss of the “avoid pronoun” principle in Brazilian Portuguese.] (Unpublished doctoral dissertation.) Universidade de Campinas.
- Enríquez, E. V. (1984). *El pronombre personal sujeto en la lengua española hablada en Madrid*. [The personal pronoun subject in the Spanish language spoken in Madrid.] Consejo Superior de Investigaciones Científicas.
- Erker, D., & Guy, G. R. (2012). The role of lexical frequency in syntactic variability: Variable subject personal pronoun expression in Spanish. *Language*, 88(3), 526–557. DOI: <https://doi.org/10.1353/lan.2012.0050>
- Frascarelli, M., & Casentini, M. (2019). The Interpretation of Null Subjects in a Radical Pro-drop Language: Topic Chains and Discourse-semantic Requirements in Chinese. *Studies in Chinese Linguistics*, 40(1), 1–45. DOI: <https://doi.org/10.2478/scl-2019-0001>
- Genuíno, W. R. A. (2017). A expressão do sujeito pronominal no português falado em Vitória/ES. [The expression of the pronominal subject in the Portuguese spoken in Vitória/ES.] (Unpublished master's thesis.) Universidade Federal do Espírito Santo.
- Gilligan, G. (1987). A cross-linguistic approach to the pro-drop parameter. (Unpublished doctoral dissertation.) University of Southern California.
- Givón, T. (1983). Topic continuity in discourse: An introduction. In T. Givón (Ed.), *Topic Continuity in Discourse: A Quantitative Cross-Language Study* (pp. 5–41). John Benjamins. DOI: <https://doi.org/10.1075/tsl.3>
- Gorman, K. (2010). The Consequences of Multicollinearity among Socioeconomic Predictors of Negative Concord in Philadelphia. *University of Pennsylvania Working Papers in Linguistics*, 16(2), 66–75.
- Guy, G. R., & Orozco, R. (2008). El uso variable de los pronombres sujetos: ¿qué pasa en la costa Caribe colombiana? [The variable use of subject pronouns: what is happening on the Colombian Caribbean coast?] In M. Westmoreland & J. A. Thomas (Eds.), *Selected Proceedings of the 4th Workshop on Spanish Sociolinguistics* (pp. 70–80). Cascadilla Proceedings Project.
- Guy G. R., & Zilles, A. (2007). *Sociolingüística quantitativa – Instrumental de análise*. [Quantitative sociolinguistics – Analytic tools.] Parábola.
- Hagemeyer, T. (2009). Creole languages and pro-drop. Unpublished paper, Centro de Linguística da Universidade de Lisboa (CLUL).
- Hancock, I. F. (1969). The Malacca Creoles and their language. *Afrasian*, 3, 38–45.
- Hancock, I. F. (1973). Malacca Creole Portuguese: a brief transformational account. *Te Reo*, 16, 23–44.

- Holm, J. (1988). *Pidgins and Creoles. Vol II Survey*. Cambridge University Press.
- Holmberg, A., & Roberts, I. (2013). The syntax–morphology relation. *Lingua*, 130, 111–131. DOI: <https://doi.org/10.1016/j.lingua.2012.10.006>
- Hopper, P. (1979). Aspect and foregrounding in discourse. In T. Givón (Ed.), *Syntax and Semantics*. Vol. 12. *Discourse and Syntax* (pp. 213–241). Academic Press. DOI: https://doi.org/10.1163/9789004368897_010
- Huang, C. T. (1984). On the distribution and reference of the empty categories. *Linguistic Inquiry*, 15, 531–574.
- Jia, L., & Bayley, R. (2002). Null pronoun variation in Mandarin Chinese. *University of Pennsylvania Working Papers in Linguistics*, 8(3), 103–116.
- Johnson, D. E. (2009). Getting off the GoldVarb Standard: Introducing Rbrul for Mixed-Effects Variable Rule Analysis. *Language and Linguistics Compass*, 3(1), 359–383. DOI: <https://doi.org/10.1111/j.1749-818X.2008.00108.x>
- Knowlton, E. G. (1964). Malaysian Portuguese. *The Linguist*, 26, 211–213, 239–241.
- Kowenberg, S., & Scott, J. (2010). Null subjects in Papiamentu: a reassessment. In N. Faraclas, R. Severing, C. Weijer & E. Echteid (Eds.), *Crossing shifting boundaries – Language and changing political status in Aruba, Bonaire and Curaçao* (pp. 75–84). Proceedings of the ECICC-conference, Dominica 2009, Volume 1. Fundashon pa Planifikashon di Idioma & University of the Netherlands Antilles.
- Lee, E. (2004). Language shift and revitalization in the Kristang community, Portuguese Settlement, Malacca. (Unpublished doctoral dissertation.) University of Sheffield.
- Lee, N. H. (2014). A grammar of Baba Malay with sociophonetic considerations. (Unpublished doctoral dissertation.) University of Hawaii.
- Li, X., & Bayley, R. (2018). Lexical frequency and syntactic variation – Subject pronoun use in Mandarin Chinese. *Asia-Pacific Language Variation*, 4(2), 135–160. DOI: <https://doi.org/10.1075/aplv.17005.li>
- Li, X., Chen, X., & Chen, W-H. (2012). Variation of subject pronominal expression in Mandarin Chinese. *Sociolinguistic Studies*, 6(1), 91–119. DOI: <https://doi.org/10.1558/sols.v6i1.91>
- Lucchesi, D. (2004). Contato entre línguas e variação paramétrica: o sujeito nulo no português afro-brasileiro. [Language contact and parametric variation: the null subject in Afro-Brazilian Portuguese.] *Língua(gem)*, 1(2), 63–91.
- Lucchesi, D. (2009). A realização do sujeito pronominal. [Realization of the pronominal subject.] In D. Lucchesi, A. N. Baxter & I. Ribeiro (Eds.), *O português afro-brasileiro* (pp.167–183). EDUFBA. DOI: <https://doi.org/10.7476/9788523208752.0008>
- Melnik, N. (2007). Extending Partial Pro-drop in Modern Hebrew: A Comprehensive Analysis. In S. Müller (Ed.), *Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar, Stanford Department of Linguistics and CSLI's LinGO Lab* (pp. 173–193). CSLI Publications. DOI: <https://doi.org/10.21248/hpsg.2007.11>
- Menon, O. P. da S. (2000). Uso do pronome sujeito de primeira pessoa no português do Brasil. [Use of the first-person subject pronoun in Brazilian Portuguese.] *Organon*, 14(28/29), 157–177. DOI: <https://doi.org/10.22456/2238-8915.30202>

- Miyagawa, S. (2017). *Agreement beyond Phi*. The MIT Press. DOI: <https://doi.org/10.7551/mitpress/10958.001.0001>
- Neeleman, A., & Szendrői, K. (2007). Radical Pro-drop and the Morphology of Pronouns. *Linguistic Inquiry*, 38(4), 671–714. DOI: <https://doi.org/10.1162/ling.2007.38.4.671>
- Newmeyer, F. (2005). Against a Parameter-setting approach to language variation. *Linguistic Variation Yearbook*, 4, 181–234. DOI: <https://doi.org/10.1075/livy.4.06new>
- Novaes, J. C. A. (2007). O Parâmetro do Sujeito Nulo no português popular do interior do estado da Bahia. [The Null Subject Parameter in vernacular Portuguese in the interior of the state of Bahia.] (Unpublished master's thesis.) Universidade Federal da Bahia.
- Orozco, R. (2015). Pronominal Variation in Colombian Costeño Spanish. In A. M. Carvalho, R. Orozco & N. L. Shin (Eds.), *Subject pronoun expression in Spanish: A cross-dialectal perspective* (pp. 19–39). Georgetown University Press.
- Orozco, R. (2016). Subject pronoun expression in Mexican Spanish: ¿Qué pasa en Xalapa? *Proceedings of the Linguistic Society of America* (pp. 1–15), Volume 1, Article 7. DOI: <https://doi.org/10.3765/plsa.v1i0.3703>
- Orozco, R., & Guy, G. R. (2008). El uso variable de los pronombres sujetos: ¿qué pasa en la costa Caribe colombiana? [The variable use of subject pronouns: what happens on the Colombian Caribbean coast?] In M. Westmoreland & J. A. Thomas (Eds.), *Selected Proceedings of the 4th Workshop on Spanish Sociolinguistics* (pp. 70–80). Cascadilla Proceedings Project.
- Orozco, R., & Hurtado, L. M. (2021). A Variationist Study of Subject Pronoun Expression in Medellín, Colombia. *Languages*, 6(5), 1–29. DOI: <https://doi.org/10.3390/languages6010005>
- Otheguy, R., & Zentella, A. C. (2012). *Spanish in New York: Language Contact, Dialectal Leveling, and Structural Continuity*. Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780199737406.003.0001>
- Paredes Silva, V. L. (2003). Motivações funcionais no uso do sujeito pronominal: Uma análise em tempo real. [Functional motivations in the use of the pronominal subject: A real-time analysis.] In M. C. Paiva & M. E. Duarte (Eds.), *Mudança linguística em tempo real* (pp. 97–114). FAPERJ/Contracapa.
- Phimsawat, O.-U. (2011). The syntax of pro-drop in Thai. (Unpublished doctoral thesis). University of Newcastle.
- Pillai, S., Phillip, A., & Soh, W. Y. (2016). Revitalizing Malacca Portuguese Creole. In P. P. Trifonas & T. Aravossitas (Eds.), *Handbook of Research and Practice in Heritage Language Education* (pp. 1–17). Springer. DOI: https://doi.org/10.1007/978-3-319-38893-9_27-1
- Rêgo, A. da S. (1942). *Dialecto Português de Malaca*. [The Portuguese Dialect of Malacca.] Agência Geral das Colónias.
- Rizzi, L. (1986). Null objects in Italian and the theory of pro. *Linguistic Inquiry*, 17, 501–557.
- Roberts, I. (2010). A deletion analysis of null subjects. In T. Biberauer, A. Holmberg, I. Roberts & M. Sheehan (Eds.), *Parametric Syntax: Null Subjects in Minimalist Theory* (pp. 58–87). Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511770784.002>

- Rodríguez-Riccelli, A. (2019). The Subject Domain in Cabo-Verdean Creole: Combining variationist sociolinguistics and formal approaches. (Unpublished doctoral thesis). The University of Texas at Austin.
- Rodríguez-Riccelli, A. (2021). Variable subject pronoun expression in Cabo-Verdean Creole – Some language-internal constraints. *Journal of Pidgin and Creole Languages*, 36(1), 109–174. DOI: <https://doi.org/10.1075/jpcl.00071.rod>
- Scherre, M. M. P., & Naro, A. J. (1991). Marking in discourse: “Birds of a feather”. *Language Variation and Change*, 3(1), 23–32. DOI: <https://doi.org/10.1017/S0954394500000430>
- Silva, L. de S. (2020). O apagamento do sujeito no Crioulo Português de Malaca (Malásia). [Subject deletion in the Portuguese Creole of Malacca (Malaysia).] (Unpublished master’s thesis). Universidade Federal da Bahia.
- Silva-Corvalán, C. (1997). Variación sintáctica en el discurso oral: Problemas metodológicos. [Syntactic variation in oral discourse: Methodological problems.] In F. M. Fernández (Ed.), *Trabajos de sociolingüística hispánica* (pp. 115–135). Universidad de Alcalá.
- Silva-Corvalán, C., & Enrique-Arias, A. (2017). *Sociolingüística y pragmática del español*. [Sociolinguistics and pragmatics of Spanish.] Segunda edición. Georgetown University Press.
- Silveira, A. (2012). Subject expression in Brazilian Portuguese: Construction and frequency effects. (Unpublished PhD dissertation.) The University of New Mexico.
- Tagliamonte, S. (2012). *Variationist sociolinguistics – Change, observation, interpretation*. Wiley-Blackwell.
- Tarallo, F., & Kato, M. (2007[1989]). Harmonia trans-sistêmica: variação intra- e inter-lingüística. [Trans-systemic harmony: intra- and inter-linguistic variation.] *Diadorim*, 2, 13–42. DOI: <https://doi.org/10.35520/diadorim.2007.v2n0a3849>
- Torres Cacoullós, R., & Travis, C. E. (2019). Variationist typology: Shared probabilistic constraints across (non-)null subject languages. *Linguistics*, 57(3), 653–692. DOI: <https://doi.org/10.1515/ling-2019-0011>
- Travis, C. E. (2007). Genre effects on subject expression in Spanish: Priming in narrative and conversation. *Language Variation and Change*, 19(2), 101–135. DOI: <https://doi.org/10.1017/S0954394507070081>
- Veenstra, T. (1994). The acquisition of functional categories: The creole way. In D. Adone & I. Plag (Eds.), *Creolization and Language Change* (pp. 99–117). Niemeyer. DOI: <https://doi.org/10.1515/9783111339801.99>
- Veenstra, T. (2006). Syntaxis pur: expletiva in Papiamentu. In G. Mensching & E.-M. Remberger (Eds.), *Deutsche Romanistik-minimalistisch* (pp. 61–82). Günther Narr.
- Yang, C. T. B. (2010). Null Subject Revisited. In L. E. Clemens & C-M. L. Liu (Eds.), *Proceedings of the 22nd North American Conference on Chinese Linguistics (NACCL-22) & the 18th International Conference on Chinese Linguistics (IACL-18)*, Vol 2 (pp. 408–416). Harvard University.

