Exploring accessibility in direct object constructions in Santomean Portuguese: A spoken corpus-based study

Rita Gonçalves
Universidade de Lisboa
ORCID ID: 0000-0002-0101-6261
rita.gamito@edu.ulisboa.pt

Ana Madeira
Universidade NOVA de Lisboa
ORCID ID: 0000-0001-9870-007X
ana.madeira@fcsh.unl.pt

Abstract

Null and pronominal direct objects have been extensively studied in both European Portuguese (EP) and Brazilian Portuguese (BP). We aim to characterize the properties and distribution of null/pronominal objects in STP, by comparing them with those of EP/BP. We assume that accessibility can be determined by the syntactic role, animacy and explicitness of the antecedent, as well as by the presence of other potential antecedents. The study, based on the PALMA-STP spoken corpus, shows that there are fewer pronominal objects than null objects in islands and non-islands. Regarding the syntactic function of the antecedent, structural parallelism is more relevant for null than for pronominal objects. In this respect, STP seems to be closer to BP than to EP. As for overtness of the antecedent and intervention effects, STP is less restrictive than EP/BP, allowing referential chains with several intervening null objects. Moreover, there appears to be a trend towards an extension of the null object along the Referential Hierarchy, as it may be less constrained by animacy than EP/BP. Hence, null objects

in STP exhibit high productivity and are subject to fewer syntactic and semantic restrictions than in other varieties of Portuguese.

1. Introduction

São Tomé and Príncipe, located in Western Africa, is a former Portuguese colony and, historically, a creole-speaking country. Although Portuguese is the official language and the only language of education, three autochthonous creole languages, with origins dating back to the 15th and 16th centuries, are also spoken on the islands: Forro (also known as (Lungwa) Santome), Lung' Ie (or Principense) and Ngola (or Angolar). For historical reasons, there is also a community of Cape Verdean Creole speakers. However, the relatively stable situation of active bilingualism (and multilingualism) that characterized the islands until the 19th century changed significantly with the arrival of laborers from other Portuguese colonies who came to work on cacao and coffee plantations and adopted Portuguese as a second language. Furthermore, after independence in 1975, Portuguese gradually became the L1, a trend that has continued over the years (e.g., Bouchard 2019; Hagemeijer 2018, 2024). Portuguese is spoken by 98.4% of the population, both as a L1 and a L2, according to the latest available census data, from 2012.

Studies conducted mainly in the last 15 years have highlighted some grammatical features of this Portuguese variety, reflecting both language contact effects and the historical past of L2 acquisition. In the syntactic domain, for instance, the restructuring of the verbal argument structure has been the focus of research (see the work on Goal arguments verbs of inherently directed motion (Hagemeijer et al. 2022) and ditransitive verbs (Gonçalves 2010, 2016; Gonçalves et al. 2022). It has also been shown that Santomean Portuguese (STP) shows a wide spectrum of inter- and intralinguistic variation, since grammatical features that differ from the standard European variety can be related to sociolinguistic factors, such as the level of education. This is the case, for instance, of double object structures, which have been considered as one of the core syntactic features of Santomean Portuguese and have been associated with speakers with a lower level of education (Gonçalves 2016; Gonçalves et al. 2022).

In addition to the tendency to lose (functional) prepositions, which leads to a tendency towards (di)transitivity, STP also tends to drop the 3rd person accusative clitic *o(s), a(s),* which can be replaced by the 3rd person strong pronouns *ele(s), ela(s)* or by the dative clitic *lhe(s)* (Gonçalves et al. 2024). Given the low production of pronominal objects observed in this study, the use of null anaphoric objects may also be a strategy followed in STP, although this has not been investigated so far. Note that the paradigm of reflexive and inherent clitics is also affected, with inherent clitics showing a high rate of omission (Madureira 2023). Finally, it has also been observed that within the nominal domain, STP has a consistent tendency to drop definite articles (e.g., Afonso 2008, Lima Afonso 2009). As pointed out by Hagemeijer (2024), this has not been systematically studied, despite being a striking feature of the available language data.

Overall, STP is a relatively under-explored variety in linguistic research and offers a unique perspective on the dynamics of language change. Research will benefit greatly from comparing STP with other well-documented varieties of Portuguese. Such comparisons will provide valuable insights into the specific ways in which STP diverges from or converges with other varieties, thus offering a broader understanding of the mechanisms driving linguistic variation and change in the Portuguese-speaking world.
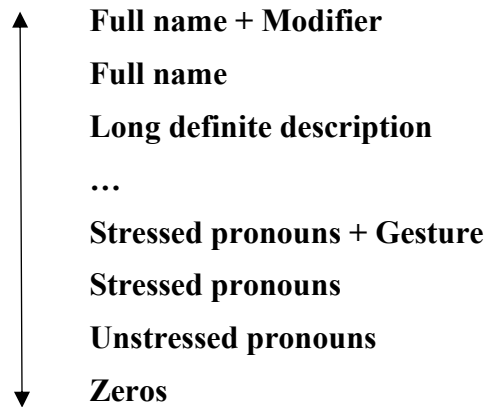
Considering the small number of pronominal objects attested in previous studies on STP, and the availability of null objects in other varieties of Portuguese – namely, European Portuguese (EP) and Brazilian Portuguese (BP) –, in the present study we aim to determine whether null objects are available in STP and, if so, to characterize their distribution and properties, comparing them to those of EP and BP. Given the role played by accessibility factors (Ariel 1990) in the licensing and identification of null and pronominal objects (e.g., Cyrino & Matos 2016; Kato et al. 2023), our analysis will focus mainly on properties which are known to determine the degree of accessibility of the antecedent, namely, grammatical function, animacy, explicitness and intervention of other referential expressions.

Section 2 provides an overview of the main properties of null and pronominal objects in both EP and BP, as well as our research questions and hypotheses. The methodology is presented in Section 3. In section 4 we describe the results and compare them with EP/BP. Finally, section 5 discusses the data, addressing our research questions and hypotheses, and concludes the paper.

2. Null objects and accessibility

2.1. Accessibility factors

It is widely recognized that how a referring expression is realized depends on the degree of accessibility of its antecedent. Following Ariel's (1990) Accessibility Theory, the more accessible or prominent the antecedent, the less explicit the expression tends to be. The accessibility of the antecedent (which is related to its availability in memory) may be determined by factors such as its grammatical function/structural position, its explicitness, its distance from the referring expression (recency of mention and intervention of potential null or overt antecedents), as well as the semantic feature of animacy (e.g., Ariel 1990; Fukumura & van Gompel 2011). The Accessibility scale (Ariel 1990, 1991) (1) arranges referring expressions according to the degree of accessibility of the antecedent. Hence, null expressions and clitics are both High Accessibility Markers, but null forms should require a more accessible antecedent than clitics, which, in turn, are predicted to require a more accessible antecedent than stressed pronouns.

(1) **Accessibility scale (**Ariel 1991: 449, abbreviated)

LOW ACCESSIBILITY

    Full name + Modifier

    Full name

    Long definite description

    …

    Stressed pronouns + Gesture

    Stressed pronouns

    Unstressed pronouns

    Zeros

HIGH ACCESSIBILITY

According to Ariel (1991), this scale represents *degrees of relative accessibility*, which are predicted to be universal. However, languages may differ in the *absolute degrees of accessibility* of their referring expressions, which may vary not only cross-linguistically, but also intra-linguistically, both synchronically and diachronically. This is precisely what has been claimed regarding the distribution of null and pronominal forms in BP by Cyrino et al. (2000). They propose that, in a language like BP, the referential status of the referent (its specification for person and features such as animacy and specificity) influences the realization of the object, determining whether it may be realized as a pronoun or a null form. This follows from the Referential Hierarchy (2a), together with the Implicational Mapping Hypothesis (2b), which allow us to account for the fact that pronouns in BP may bear 1st / 2nd person features, but not null objects, which are restricted to 3rd person referents; and, while pronouns may refer to animate ([+ anim]) or inanimate ([-anim]) entities, null objects have a strong tendency to refer to [-anim] entities. According to Cyrino et al. (2000), the diachronic changes observed in the distribution and properties of null objects in BP since the 19th century (Cyrino 1997) are a consequence of the extension of the null form along the Referential Hierarchy, whereby it became gradually more referential.

(2)   a. **Referential Hierarchy** (Cyrino et al. 2000):

non-argument	proposition	[-human]		[+human]	
			3rd p.	2nd p.	1st p
-specif					+specif
[-ref] <--> [+ref]					

    b. **The Implicational Mapping Hypothesis**:

        (i) The more referential, the greater the possibility of a non-null pronoun.
        (ii) A null variant at a specific point on the scale implies null variants to its left in the Referential Hierarchy.

In the next section, we will characterize the distribution and properties of null and pronominal objects both in BP and in EP, seeking to identify the differences and similarities between these two varieties of Portuguese.

2.2. Null objects in European and Brazilian Portuguese

It is well-known that both EP and BP allow 3rd person null direct objects (DOs). The reference of the null object may either be recovered anaphorically from a previous discourse antecedent (3a) or deictically from the extralinguistic context (3b). Null objects may alternate with a pronominal form, which is realized as a clitic in EP and (predominantly) as a strong pronoun in BP (4).

(3)   Portuguese, Cyrino & Matos (2016: 295)
        a. Ele experimentou **o   casaco**$_i$, mas   não comprou [-]$_i$.
           he try.PST.3SG   the coat   but   not buy.PST.3SG
           'He tried on the coat but he did not buy (it).'
        b. [Situation: Someone sees a famous star in a restaurant and makes the
           following comment:]
           Eu vi [-]   na   TV ontem.
           I   see.PST.1SG on.the TV yesterday
           'I saw (him) on TV yesterday.'

(4)   a. Eu vi-   o   na   TV   ontem. [EP]
           I   see.PST.1SG CL.ACC.M.3SG on.the TV   yesterday
           'I saw him on TV yesterday.'
        b. Eu vi   ele   na   TV ontem. [BP]
           I   see.PST.1SG PRON.ACC.M.3SG on.the TV yesterday

The null object construction should be distinguished from other constructions involving constituent deletion such as VP ellipsis, illustrated in (5), where all the verbal complements and optionally VP adjuncts are deleted:

(5)   *Compraste   o   casaco nesta   loja? – Comprei* [VP *o casaco nesta loja*].
        buy.PST.2SG the coat   in.this shop   buy.PST.1SG
        'Did you buy the coat in this shop? – I did.'

Unlike other Romance languages such as Spanish (Campos 1986), where the equivalent to the Portuguese sentences in (3) and (4) would be ungrammatical, as illustrated in (6), Portuguese allows definite null objects (Raposo 1986), even though these are more likely to be indefinite (Sainzmaza-Lecanda & Schwenter 2017; Schwenter 2014).

(6)   a. Él se probó   la   chaqueta, pero no *(la)   compró.
           he SE try.PST.3SG the coat   but   not CL.F.3SG buy.PST.3SG
           'He tried on the coat but he did not buy *(it).'

     b. *(Lo)   vi   ayer   en la   televisión.
        CL.M.3SG see. PST.1SG yesterday on the television
        'I saw *(him) on TV yesterday.'

Although null objects display very similar characteristics in EP and BP, there appear to be some differences between the two varieties at least regarding the productivity and distribution of null objects (Cyrino & Matos 2016). In the next subsections, we will provide a brief characterization of the properties of null objects in these two varieties of Portuguese.

2.2.1. Productivity
It has been claimed that, in EP, null objects may be predominantly associated with informal registers in both oral and in written speech (Duarte & Costa 2013) or occur mostly in the spoken language (Rinke 2025). Although a number of corpus-based studies have shown that null objects are more frequent in BP than in EP, in both written (Cyrino 2001; Kato & Raposo 2001) and oral data (Figueiredo 2014; Sainzmaza-Lecanda & Schwenter 2017; Schwenter 2014), there are indications that they may also be more frequent in oral than in written speech in BP (Kato & Raposo 2001). Table 1 below shows a synthesis of the findings reported in Schwenter (2014)[1] (see also Sainzmaza-Lecanda & Schwenter 2017) and Figueiredo (2014)[2] regarding the proportion of null and pronominal objects in oral corpus data in EP and BP. The results of the monolingual EP participants of Flores et al. (2017)[3] (see also Rinke et al. 2018), as well as those of Rinke (2025)[4], which are likewise based on oral spontaneous data, are also shown in Table 1. Note that, although these results show higher rates of null objects in BP than in EP, there are notable differences among the four studies in the proportion of null objects and clitic pronouns in EP: null objects are considerably more frequent than clitics in Schwenter (2014), whereas the percentage of clitics is substantially higher than that of null objects in Flores et al. (2017) and Rinke (2025); in Figueiredo (2014), however, speakers do not appear to favor either option, despite there being a slight preference for clitics. The differences found among these studies

---

[1]   The EP data from Schwenter (2014) (and Sainzmaza-Lecanda & Schwenter 2017) are taken from the oral component of the CRPC - *Reference Corpus of Contemporary Portuguese* (Center of Linguistics of the University of Lisbon). The BP data are extracted from the Programme of Studies on the Use of Language of the Federal University of Rio de Janeiro (PEUL) corpus, a collection of sociolinguistic interviews conducted with speakers from lower socioeconomic groups in the 1980s. The percentages indicated in Table 1 were calculated considering only the occurrences of null and pronominal objects.

[2]   In Figueiredo's (2014) study, both sets of data come from varieties of Portuguese spoken in rural areas. The EP data come from CORDIAL-SIN - *Syntax-oriented Corpus of Portuguese Dialects* (Center of Linguistics of the University of Lisbon) and the EP data from the corpus of the project *Strands of the Popular Portuguese from the State of Bahia* (*Vertentes do português popular do Estado da Bahia*) (Federal University of Bahia).

[3]   The corpus data from Flores et al.'s (2017) study presented in Table 1 come from oral interviews with two groups of monolingual speakers of EP (the study also included two bilingual groups): the first group (G1_mon) consists of speakers with a mean age of 64.25 and with a low level of education; and the speakers in the second group (G2_mon) have a mean age of 25.25 and a higher level of education.

[4]   Rinke's (2025) corpus data are drawn from the *Sociolinguistic Profile of the Braga Speech* corpus (University of Minho).

may be due to differences in the sociolinguistic profile of the participants (mainly related to dialectal characteristics and level of education). A comparison between the two monolingual groups in Flores et al.'s (2017) study also reveals that the older monolingual speakers produce more clitic pronouns than the younger monolinguals, who in turn display higher rates of object omission than the older ones. Although these differences are not statistically significant, they suggest that null objects may be becoming more frequent in present-day EP.

Table 1. Distribution of null and pronominal objects in EP and BP.

	Schwenter (2014)		Figueiredo (2014)		Flores et al. (2017)	Rinke (2025)
	EP	**BP**	**EP**	**BP**	**EP**	**EP**
null objects	64.3%	85.4%	47.8%	90.1%	24.3% (G1_Mon) 29.6% (G2_Mon)	30.2%
clitic pronouns	35.7%	0.4%	52.2%	-	64.4 % (G1_Mon) 55.8% (G2_Mon)	69.8%
strong pronouns	-	14.2%	-	9.9%	-	-

It is interesting to note that several experimental studies based on elicited production data (Flores et al. 2020; Teixeira et al. 2024; Zhao 2020) have found that EP speakers show a consistent preference for clitic pronouns (or lexical DPs) over null objects (which are produced in small number), even in contexts where all the conditions for object omission are met. Considering both the corpus-based and the experimental data, we may conclude that there is a preference for clitic pronouns over null objects in EP, whereas BP exhibits a clear preference for null over pronominal objects (at least in the spoken language).

2.2.2. Locality restrictions
The distribution of null objects also appears to be more restricted in EP than in BP. BP allows null objects in strong islands (Castro et al. 2017; Cyrino & Matos 2016; Kato et al. 2023) as long as they are [-anim] (see the contrast in (7) involving a complex DP containing a relative clause).

(7)   BP, Cyrino & Matos (2016: 301)
        a. *O   José conheceu   a   mulher que beijou [-]. [BP]
           the José meet.PST.3SG the woman that kiss.PST.3SG
           'José met the woman who kissed (him).'
        b. O   José conheceu   a   mulher que comprou [-]. [BP]
           the José meet.PST.3SG the woman that buy.PST.3SG
           'José met the woman who bought (it).'

Cyrino (1997) proposes a DP-ellipsis analysis for null objects in BP, whereby their content is recovered through identity with an antecedent. This analysis allows us to explain not only why null objects are found in island contexts, but also, for example,

the availability of strict and sloppy readings (see (8)), which is a property that characterizes elliptical structures.

(8)   BP, Cyrino & Matos (2016: 301)
         De noite, João abriu   a   janela, mas Pedro preferiu   fechar [-].
         at night   João open.PST.3SG the window but Pedro prefer.PST.3SG close.INF
         'At night, João opened the window, but Pedro preferred to close it.'
         __ = João's window (*strict reading*)
         __ = Pedro's window (*sloppy reading*)

Unlike BP, EP has been argued to disallow null objects in island contexts (Raposo 1986). Hence, sentences with null objects occurring in island contexts, such as complex DPs (9) and adverbial clauses (10), are judged to be ungrammatical in EP.

(9)   EP, Raposo (1986: 382)
         [Context: someone talking about the cake that the boy has left on the table]
         *O rapaz que trouxe   mesmo agora da   pastelaria   era
         the boy   that bring.PST.3SG right   now   from.the pastry.shop be.IPFV.3SG
         o   teu   afilhado. [EP]
         the your godson
         'The boy who brought (it) from the pastry shop right now was your godson.'

(10)   EP, Raposo (1986: 382)
         [Context: someone is talking about a map showing the location of the treasure]
         *O pirata partiu   para as Caraíbas   depois de ter   guardado
         the pirate leave.PST.3SG for   the Caribbean after   of have.INF put.PTCP
         cuidadosamente no   cofre. [EP]
         carefully   in.the safe
         'The pirate left for the Caribbean after putting (it) carefully in the safe.'

Following Huang's (1984) analysis for Chinese, Raposo (1986) proposes that the null object in EP should be analysed as a variable bound by a null operator and identified by a null discourse topic (see also Duarte 1987).
     However, some authors have claimed that null objects are at least marginally possible in island contexts in EP. Duarte & Costa (2013), for example, propose that they may be allowed in islands in contexts in which they recover [-anim] antecedents, as shown in (11).

(11)   EP, Duarte & Costa (2013: 2345)
         a. A – E   então, o   carro novo?
                and so   the car   new
                'So, what about the new car?'
           B – A   minha mulher está furiosa porque comprei   sem   ela
                the my   wife   is   furious because buy.PST.1SG without   she
                saber.
                know.INF
                'My wife is furious because I bought (it) without her knowing.'

b. A – E   então, a   Maria?
    and so   the Maria
    'So, what about Maria?'
  B – *A minha mulher ficou   furiosa porque eu beijei
    the my   wife   become.PST.3SG furious because I   kiss.PST.1SG
    na festa.
    at.the party
    'My wife became furious because I kissed (her) at the party.'

However, in an experimental study based on an acceptability judgement task administered to EP and BP speakers, Castro et al. (2017) found that their EP-speaking participants rejected null objects in islands, regardless of animacy; by contrast, the BP-speaking participants judged the equivalent BP sentences to be acceptable, but only when the null object was [-anim]. These judgments thus confirm the asymmetry between EP and BP described in the literature.

Nevertheless, a small number of occurrences of [-anim] null objects in islands have been found in corpus data in EP (Figueiredo 2014; Rinke et al. 2018). See the example from Figueiredo (2014) in (12), with an adjunct clause:

(12)   EP, Figueiredo (2014: 110)
E   eu arranjei-lhe,   então, umas febras$_i$,   e   disse-
and I   prepare.PST.1SG-CL.DAT.3SG then   some pork.steaks and say.PST.1SG-
-lhe: (…)   enquanto eu tiver [-]$_i$,   vem.
CL.DAT.3SG while   I   have.SBJV.1SG come.IMP
'And so, I prepared her some pork steaks and told her: (…) while I have (them), come.'

Overall, the results from corpus-based and experimental studies appear to support the conclusion that the acceptability of null objects in island contexts in EP may be subject to variation. This is acknowledged by Raposo (2004), who reassesses the judgements of sentences such as (9) and (10) above, judging them to be slightly degraded but not fully ungrammatical. To account for these revised judgements, Raposo (2004), while maintaining his characterization of the null object as a variable derived by A'-movement, proposes that the object DP is headed by a null determiner lacking number and gender features, which selects a *pro* complement. He further suggests that, in order for its content to be identified, *pro* must adjoin to F, a functional category in the left periphery of the sentence. The resulting structure is shown in (13):

(13)   EP, Raposo (2004: 55)
(*esse livro*) [$_{FP}$ pro$_i$+F [$_{TP}$ *eu só   encontrei* [$_{DP}$ D t$_i$] *na   FNAC*]]
that book   I   only find.PST.1SG   in.the FNAC
'(That book,) I only found (it) in FNAC.'

2.2.3. The position of the antecedent
As discussed in 2.1., one of the factors which plays a role in determining the accessibility of the antecedent is its syntactic position. Defective forms (such as null and clitic forms) tend to require antecedents that are more prominent syntactically (Ariel 1990; Carminati 2002). In the case of null subjects, for instance, Carminati

(2002) proposes the Position of Antecedent Hypothesis, according to which null subjects show a bias towards subject antecedents (i.e., in SpecIP), whereas pronominal subjects tend to recover antecedents in other (less prominent) positions. Extending this hypothesis to null objects, these would be predicted to prefer antecedents in subject position, with pronominal objects being more flexible in their choice of an antecedent.

However, in intrasentential contexts, only BP allows null objects with an antecedent in subject position; in EP, this is possible only if the subject antecedent is not in the same sentence as the null object (see (14)). This contrast follows from the differences between EP and BP regarding the status of the null object: if, in EP, but not in BP, null object constructions involve A'-movement to the left periphery of the sentence, the ungrammaticality of (12a) in EP may be explained as a crossover effect, since A'-movement is blocked by the intervening coreferential subject (Kato et al. 2023).

(14)   Portuguese, Kato et al. (2023: 353)
   a. [Esse brinquedo]$_i$ permite   que as   crianças   montem [-]$_i$
      this   toy   allow.PRS.3SG that the children assemble.SBJV.3PL
      sem   ajuda. [ok BP; * EP]
      without help
      'This toy allows children to assemble (it) without help.'
   b. Esse brinquedo$_i$ está partido. – Sim, as   crianças partiram [-]$_i$ sem
      this toy   is   broken   yes the   children break.PST.3PL without
      querer. [ok BP and EP]
      want.INF
      'This toy is broken. – Yes, the children broke (it) accidentally.'

Previous corpus-based studies have indicated that null objects in both EP and BP may display a structural parallelism effect, which results in a preference for antecedents in object position (Figueiredo 2014; Rinke & Kerezova 2024). Figueiredo (2014), for example, observed that antecedents in DO position in BP are mostly recovered by a null object, while antecedents in subject position are more easily recovered by (strong) pronouns; as for EP, both null objects and clitic pronouns tend to recover a DO antecedent. These latter findings were confirmed by the results of Rinke & Kerezova's (2024) study (which is also based on the CORDIAL_SIN corpus), who also found no significant difference between null objects and clitics in their preference for an antecedent in object position.

2.2.4. Animacy of the antecedent

Another factor which plays a role in determining the accessibility of the antecedent is the semantic property of animacy. It has long been noted that there is a (non-categorical) division of labor between null and pronominal objects with regard to animacy in both EP and BP (e.g., Cyrino 1997; Cyrino & Matos 2016; Cyrino et al. 2000, Kato et al. 2023), as [-anim] objects are significantly more likely to be null and [+anim] objects tend to be realized as pronominal forms. See (15) and (16):

(15)　BP, Cyrino (1997: 146)
　　　a. João descascou　　a　banana*i*,　mas Pedro não　comeu [-]*i*. [BP]
　　　　　João peel.PST.3SG the banana　　but Pedro not　eat.PST.3SG
　　　　　'João peeled the banana but Pedro didn't eat (it).'
　　　b. ?João trouxe　　　　a Maria*i*,　mas Pedro　não beijou*i*.
　　　　　João　bring.PST.3SG the Maria　　but Pedro　not kiss.PST.3SG
　　　　　'João brought Maria but Pedro didn't kiss (her).'
　　　c. João trouxe　　　　a　Maria*i*, mas　Pedro não beijou　　　ela/
　　　　　João bring.PST.3SG the Maria　　but　Pedro not kiss.PST.3SG PRON.F.3SG/
　　　　　não a*i*　　　　　beijou.
　　　　　not CL.ACC.F.3SG　kiss.PST.3SG
　　　　　'João brought Maria but Pedro didn't kiss her.'

(16)　EP Duarte & Costa (2013: 2345)
　　　a. *Quando encontro　　o　Pedro*i*, beijo [-]*i*　　com ternura.
　　　　　when　　meet.PRS.1SG the Pedro　kiss.PRS.1SG with tenderness
　　　　　'When I meet Pedro, I kiss (him) tenderly.'
　　　b. Quando encontro　　uma gralha*i*, corrijo [-]*i*　　　imediatamente.
　　　　　when　　find.PRS.1SG a　　typo　　correct.PRS.1SG immediately
　　　　　'When I find a typo, I correct (it) immediately.'

Nevertheless, [+anim] null objects are possible in both varieties in certain contexts: according to Duarte & Costa (2013) and Kato et al. (2023), [+anim] null objects are possible when the antecedent is in the discourse (17) or can be recovered from the situational context (18).

(17)　EP, Duarte & Costa (2013: 2345)
　　　a. [Context: People are discussing who is taking whom to the party.]
　　　　　A: E　o　João*i*?
　　　　　　　and the João
　　　　　　　'What about João?'
　　　　　B: A　Maria leva [-]*i*.
　　　　　　　the Maria take.PRS.3SG
　　　　　　　'Maria will take (him).'
　　　　　　　　　　(Kato et al. 2003: 337–8)
　　　b. A: E　a　Ana*i*?
　　　　　　　and the Ana
　　　　　　　'What about Ana?'
　　　　　B: Encontrei　　ontem [-]*i* na　　exposição da　　Paula Rego.
　　　　　　　meet.PST.1SG yesterday in.the exhibition of.the Paula Rego
　　　　　　　'I met (her) yesterday at Paula Rego's exhibition.'

(18)　EP, Duarte & Costa (2013: 2345)
　　　A　[olhando para a fotografia de um rapaz na secretária de B]:
　　　　　[looking at the photo of a young man on B's desk]:
　　　　　Conheceste [-]　em Itália?
　　　　　meet.PST.2SG　　in Italy
　　　　　　'Did you meet (him) in Italy?'

The claim that there is a division of labor with respect to animacy between null and pronominal objects is supported by a number of corpus-based and experimental studies.

Let us consider first the corpus-based studies. In her analysis of BP and EP corpora, Figueiredo (2014) observed identical patterns in the distribution of null and pronominal objects in the two varieties, i.e., preference for a null form with [-anim] antecedents and for a pronominal form with [+anim] antecedents. However, she noted that null objects were used to recover both [-anim] and [+anim] referents in both BP and EP. Hence, in the BP corpus, 79% of all [+anim] referents were realized as null (the percentage is 94% for the [-anim] referents), while, in the EP corpus, 35% of all [+anim] objects (and 50% of [-anim] objects) were realized as null. There were also occurrences of [+anim] null objects in Rinke et al.'s (2018) monolingual EP data, although at much lower rates than those found by Figueiredo (2014) (under 10%) and Rinke (2025) (12%).

As for Schwenter (2014) (see also Sainzmaza-Lecanda & Schwenter 2017 and Schwenter & Silva 2003), his findings also provide evidence for a clear division of labor between (clitic/strong) pronouns and null objects in EP and BP regarding animacy, with a clear preference for [-anim] referents to be realized as null and [+anim] referents to be realized as overt. Nevertheless, 45.6% and 22.5% of all [+anim] objects are null in the BP corpus and in the EP corpus, respectively. An interaction between animacy and definiteness/specificity was observed in both varieties: pronouns (whether strong or clitic) appear to be prototypically [+anim, +definite, +specific] (see also Dickinson et al. 2024), whereas null objects tend to be [-anim, -definite].

Turning now to the experimental data, Castro et al. (2017) found similar animacy effects in BP and EP in question/answer pairs, in which the antecedent of the null object in the answer was in the question (*contra* Duarte & Costa 2013 and Kato et al. 2023). In their study, speakers of both varieties judged sentences where the null object recovered a [-anim] antecedent to be significantly more acceptable than those in which the antecedent was [+anim]. No such effect was found in either BP or EP with pronominal objects, which were consistently assigned high acceptability ratings than the sentences with null objects.

Similar results were reported for two other studies on EP that also used acceptability judgement tasks: Rinke (2025), who, similarly to Castro et al. (2017), employed a design where the antecedent was in a different sentence from the null/clitic object, and Teixeira et al. (2024 who used experimental sentences where the antecedent was in the same sentence as the null/clitic object. Both studies found a significant animacy effect with null objects, which were rated as significantly more acceptable when the antecedent was [-anim], but not with clitic pronouns; moreover, clitics consistently received higher acceptability ratings than null objects, regardless of animacy. Finally, despite the low rates of object omission observed in Flores et al.'s (2020) study (based on elicited production data), they also found an effect of animacy in EP, as fewer null objects and more clitics were produced when the antecedent was [+anim].

The division of labor which characterizes null and pronominal objects in EP and BP is consistent with Cyrino et al.'s (2000) Referential Hierarchy, as [-anim] referents (which are less referential) are mostly realized as null objects and [+anim]

referents (which are more referential) tend to be expressed as overt pronouns. Note, however, that the possibility of the null object referring to [+anim] (and [+human]) antecedents suggests that this form may be expanding its uses in both varieties.

2.2.5. Explicitness of the antecedent and intervention effects

Besides the grammatical function and the animacy of the antecedent, two other factors which have also been found to be relevant for accessibility are the explicitness of the antecedent and the intervention of other referring expressions between the form and its antecedent.

Several studies have concluded that null objects in both EP and BP tend to retrieve antecedents that are explicitly mentioned in the immediately preceding discourse or in the situational context, i.e., highly accessible antecedents (Duarte & Costa 2013, Kato et al. 2023). Hence, in (19a) the null object refers back to the closest antecedent (in the immediately preceding clause), whereas in (19b) the intervention of several other referents makes it hard to recover the antecedent *os óculos* ('the glasses'), rendering the sentence ungrammatical (or, at least, anomalous).

(19)   EP, Duarte & Costa (2013: 2346)
      a. O   João tirou   **os óculos**$_i$ e   guardou **[-]**$_i$ na   gaveta.
         the João take.PST.3SG the glasses and keep.PST.3SG in.the drawer
         'João took his glasses off and put (them) away in the drawer.'
      b. O   João tirou   **os óculos**$_i$. Foi   buscar   o   jornal e
         the João take.PST.3SG the glasses go.PST.3SG fetch.INF the paper and
         fazer   um chá. *Ligou   a   televisão e,   finalmente,
         make.INF the tea   switch.PST.3SG the television and finally
         guardou **[-]**$_i$ na   gaveta.
         keep.PST.3SG in.the drawer
         'João took his glasses off. He went to get the paper and make some tea. He turned on the television and put finally (them) away in the drawer.'

Hence, factors such as the intervention of other potential antecedents condition the accessibility of the antecedent and influence the acceptability of the null object.

These conclusions are corroborated by several corpus-based and experimental studies which have considered the effect of accessibility on the realization of the object in EP. For example, Rinke & Kerezova (2024) shows that both null objects and clitic pronouns tend to recover antecedents that are explicitly mentioned either in the same sentence or in the immediately preceding sentence. However, they stress that this is a "strong tendency", rather than a "categorical criterion" (p. 192), as there are instances in the corpus of null objects and clitics with a more distant antecedent, which is separated from the null/clitic object by other potential antecedents (20). Nevertheless, there appears to be a tendency for both null objects and clitics not to be included in long referential chains; hence, when a null/clitic form is employed, either the referent tends not to be maintained or further occurrences of the same referent tend to be expressed as lexical DPs (but see the example in (21)).

(20)   EP, Rinke & Kerezova (2024: 192)
As morcelas_i,   como é que se comem?   Com **pãozinho**. Fritas ou
the blood.sausages how is that SE eat.PRS.3PL with bread   fried   or
cozidas, quem gosta.   E   é com **pão**. Também há   pessoas que
cooked who like.PRS.3SG and is with bread also   there.is people that
comem   **[-]**_i com batata. Mas eu não gosto. (CRV03-45)
eat.PRS.3PL   with potato but   I   not like.PRS.1SG
'How do you eat blood sausages? With bread. Fried or boiled, whoever likes
them. And it's with bread. There are also people who eat (them) with potatoes.
But I don't like (them).'

(21)   EP, Rinke & Kerezova (2024: 191)
A   canastra_i está cheia, vão   levar **[-]**_i para uma tina que está em
the basket   is full   go.PRS.3PL take.INF   to   a   vat that is   in
cima dum carro. E   depois o   carro vai   transportar **[-]**_i para a
top of.a car   and then   the car   go.PRS.3SG transport.INF   to   the
adega. (ALC16-62)
cellar
'The basket is full; they are going to take (it) to a vat that is on top of a car.
And then the car will take (it) to the cellar.'

The strong preference of null objects for referring back to highly accessible
antecedents has been confirmed by several experimental studies on EP. For example,
using an elicited production task to investigate the effect of distance of the discourse
antecedent, Flores et al. (2020) concluded that higher accessibility of the discourse
antecedent (mentioned in the immediately previous utterance) increases the likelihood
of either a clitic or a null object (although there is a consistent preference for clitics
over null objects) and lower accessibility of the antecedent (mentioned in the previous
discourse but not in the immediately previous utterance) makes it more likely for the
object to be realized as a lexical DP. Similarly, the results of an acceptability
judgement task employed by Teixeira et al. (2024) showed that the acceptability of the
null object increases significantly when the antecedent is immediately accessible
(when there are no intervening referents and the antecedent occurs in the immediately
preceding clause), although clitic pronouns remain the preferred option for the
realization of the object.

2.3. Research questions

The characterization provided in 2.2. clearly demonstrates that the occurrence of null
objects in Portuguese is constrained by a number of syntactic, semantic and pragmatic
conditions. The two varieties of Portuguese described above, EP and BP, share some
properties but they also display some important differences. Hence, in both varieties,
null objects tend to recover highly accessible antecedents, i.e., antecedents that are
explicitly mentioned in the immediately preceding discourse; and (ii) [-anim] referents
are significantly more likely to be realized as null than [+anim] referents. On the other
hand, null objects appear to be a more productive option in BP than in EP, as, in EP,
there is evidence that object (clitic) pronouns are significantly more frequent than null
objects, which appear to be associated mostly with colloquial registers; and the

syntactic distribution of null objects appears to be more restricted in EP than in BP, as only BP clearly allows null objects in island contexts, as well as in non-islands with an intrasentential antecedent in subject position. The different locality restrictions on null objects in EP and BP indicate that this form may have a different status in the two varieties: in EP, it has been analysed as a variable bound by a null operator and identified by a null discourse topic (Raposo 1986) or as involving a null definite D that selects a *pro* complement, which must be identified by movement to a functional projection (Raposo 2004); in BP, on the other hand, it may result from DP ellipsis, being recovered through identity with an antecedent (Cyrino 1997).

Unlike pronominal objects, null objects have not been systematically studied in the African varieties of Portuguese. In the case of STP, it is known that 3rd person clitic/strong pronouns are not frequent (Gonçalves et al. 2024), which makes it likely that speakers object omission is a more productive option for object realization than in EP, for example, where clitics have been found to be highly productive. Hence, the main goal of the present study is to characterize the distribution and properties of null objects in STP, comparing them to those of EP and BP.

Given what is known about null and pronominal objects in EP and BP, and also considering the findings of research concerning (clitic and strong) pronominal objects in STP, the present study will address the following research questions and investigate the following hypotheses:

Q1. What is the preferred option for object realization in non-island and island contexts in STP: null or pronominal objects?

> **H1.** Following the findings of previous studies (Gonçalves et al. 2024), it is predicted that null objects will be more productive than pronominal objects in both non-island and island contexts in STP.

Q2. Does the grammatical function of the antecedent condition the realization of the DO in STP?

> **H2a.** If null objects recover highly accessible antecedents (Ariel 1990, 1991), they will prefer antecedents in subject position, whereas clitic/strong pronouns will be more flexible in their choice of antecedent.
>
> OR
>
> **H2b.** If null objects and clitic/strong pronouns show an effect of structural parallelism (Rinke & Kerezova 2024), they will prefer antecedents in object position.

Q3. Does animacy of the antecedent condition the realization of the DO in STP?

> **H3.** According to the predictions of the Referential Hierarchy (Cyrino, Duarte & Kato 2000), null objects are expected to prefer [-anim] antecedents, while [+anim] objects tend to be realized as clitics or full pronouns.

Q4. Do null and pronominal objects differ in their preference for antecedents that are explicitly mentioned in the immediately preceding discourse (i.e., without the intervention of other referring expressions)?

> **H4a.** According to the Accessibility scale (Ariel 1990, 1991), a difference is expected: null objects are predicted to prefer explicit antecedents with no

intervening potential antecedents, whereas pronominal objects will not show any preference.

OR

H4b. Considering the findings of Rinke & Kerezova (2024), both null objects and pronominal objects are predicted to prefer explicit antecedents with no intervening potential antecedents.

3. Methodology

This study is based on a spoken urban corpus of STP (*PALMA Corpus São Tomé e Príncipe*) (Gonçalves et al. 2021), collected between 2008 and 2012 in the city of São Tomé, through semi-structured interviews. The corpus was prepared and published as part of the project *Possession and Location: microvariation in African varieties of Portuguese* (PALMA) (PTDC/LLT-LIN/29552/2017), funded by the Portuguese Foundation for Science and Technology (FCT).

The corpus consists of interviews with 77 informants, 32 hours of recordings and 322,999 tokens. It is equally balanced regarding the main sociolinguistic variables: gender, age and schooling (see Table 2 below). The speakers are either monolingual in Portuguese or bilingual (with active or passive knowledge of Santome, the dominant Portuguese-based Creole), but for the vast majority Portuguese is the most widely spoken language. The interviews were conducted by researchers from the Centre for Linguistics at the University of Lisbon, who are speakers of European Portuguese, and/or a speaker of the Santomean variety. The corpus was annotated with part-of-speech (POS) and lemma information, and it has been made searchable on the CQPweb platform (cf. Hagemeijer et al. 2022 for further details).

Table 2. Distribution of informants according to variety, gender, age, and schooling (adapted from Hagemeijer et al. 2022).

		m	**f**	**TOTAL**
Age	17-25	9	10	19
	26-35	10	11	21
	36-45	10	6	16
	46 >	11	10	21
	average age	**37**		**77**
Schooling	0-4 years	5	8	13
	5-9 years	11	12	23
	10-12 years	19	10	29
	higher education	5	7	12
				77

For the purposes of this study, a specific set of data was organized in an Excel file. Using lemma and POS tags at the query node, we extracted from the corpus all the occurrences of null and pronominal 3rd person DOs, with a subset of transitive verbs. The following contexts were excluded: (i) VP-ellipsis constructions; (ii) propositional objects; (iii) 'unspecified object sentences' (Raposo 1986: 376); (iv) causative structures; and (v) unclear contexts, due to hesitations and reformulations,

as well as cases that were ambiguous between null anaphoric objects and VP-ellipsis. Furthermore, null objects without a linguistic antecedent were excluded from the analysis. Since we have identified several examples with referential chains formed by a lexical head and various null objects distributed over several sentences, all the identified verbs that occur in a null object structure and were not considered in the first phase of extraction have thus undergone a similar process of extraction. A total of 134 transitive verbs occurring in the corpus with 3rd person DOs were considered for analysis. The following variables were used to classify the DO structures:

(i)   type of DO: null object/clitic/full pronoun;
(ii)   syntactic context: island/non-island;
(iii)   syntactic role of the antecedent: subject/object/other (topics not included);
(iv)   animacy of the antecedent: [+human]/[+anim]/[-anim]
(v)   type of antecedent: explicit/null;
(vi)   DO (explicit/null) intervening between null object and antecedent: yes/no.[5]

As sociolinguistic variables were not the focus of the research, factors such as gender, age and level of education were not considered.
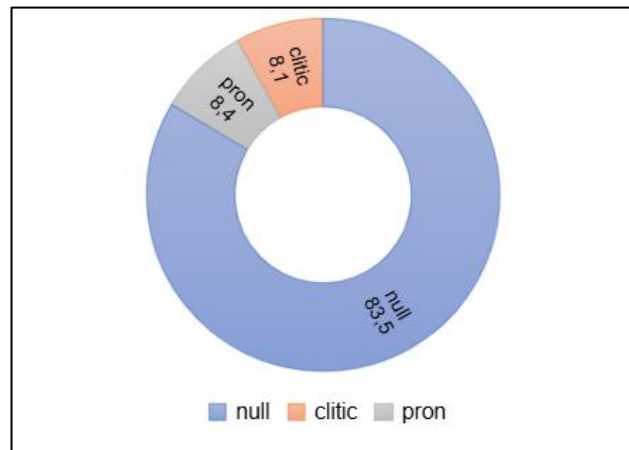
Data were organized and accounted for using an Excel spreadsheet. Further statistical testing was conducted using the online software Jamovi (The Jamovi Project), applying chi-square and logistic regression analyses to determine the significance of the observed patterns and reinforce the reliability of the descriptive trends. For statistical purposes, we compared the results for null and pronominal objects, combining clitics and full pronouns into a single category.

4. Results

4.1. Type of DO realization

A total of 632 DO structures were coded and analyzed. Figure 1 shows the distribution of null and pronominal objects (both clitic and full pronouns).
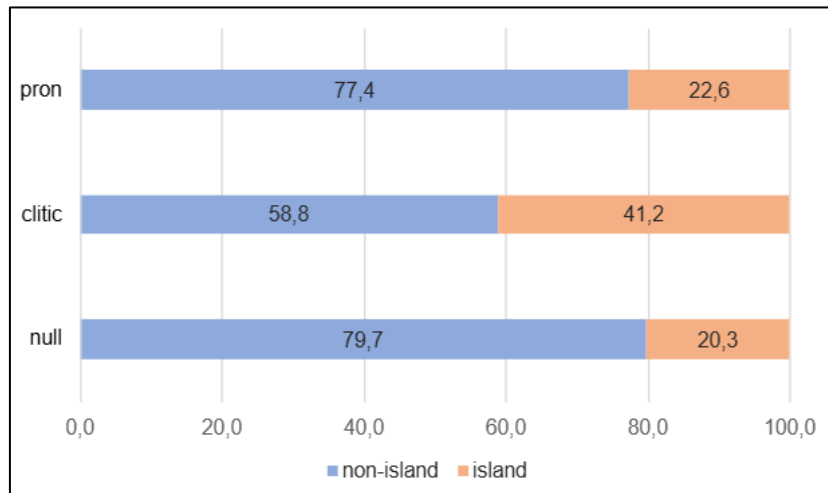
---

[5]   Given that both null and clitic DO in EP favor antecedents in object position (Rinke & Kerezova 2024), in this study only nominal expressions with the grammatical function of DO were considered as interveners.

Figure 1. Distribution of null and pronominal objects.

As the percentages reveal, STP has a significantly higher percentage of null objects (83.5%, i.e., 528/632) than clitics (8.1%, i.e., 51/632) and full pronouns (8.4%, i.e., 53/632). The low rates of null objects are consistent with the findings of Gonçalves et al. (2024), which also found low rates of pronominal objects, with a preference for full pronouns over clitics. Examples (22–23) below illustrate the intraspeaker variation between null, clitic and full pronominal forms.

(22)   Há   cliente que prefere   mais **peças** assim   desmontada (…).
        there.is client that prefer.3SG more pieces like.this disassembled
        Quando a patroa   vende   **[-]**,   nós também
        when   the mistress sell.3SG   (it)   we also
        carregamos-**a**,   pomos **[-]** em casa.
        carry.1PL-CL.ACC.F.SG put.1PL (it) in house
        'There are customers who prefer parts to be disassembled. When the mistress sells (it), we also take it and put (it) in the house.

(23)   Se eu **lhe**   vejo,   se eu vejo   **eles**   na   estrada,
        If I   CL.DAT see.1SG if I   see.1SG PRON.M.3PL   on-the   street
        eu chamo **[-]**   porque   normalmente é   família.
        I call   (them) because usually   be.3SG   family
        'If I see them on the street, if I see them on the street, I call (them) because they're usually family.'

Figure 2 shows that null objects occur predominantly in non-island contexts (79.7%, i.e., 421/528), but they also occur in island contexts (20.3%, i.e., 107/528). The same pattern is observed for pronominal objects: while 58.8% of clitics (i.e., 30/51) occur in non-island contexts and 41.2% (i.e., 21/51) in island contexts, 77.4% (i.e., 41/53) and 22.6% (i.e., 12/53) of full pronouns occur in these contexts, respectively. The asymmetry between the overall number of non-islands (492/632 occurrences) and islands (140/632 occurrences) may explain the higher number of both null and pronominal objects in non-islands.

Figure 2. Distribution of null and pronominal objects by syntactic context.

Examples (24-26) illustrate null objects, clitic and full pronouns in both non-island and island contexts (see also 22).

(24)   a. Outras   seitas   consegue   pegar   **pessoa**   na   rua,
          other   sects   manage.3SG   pick.up.INF person   in.the   street
          pôr   **[-]**   na   igreja.
          put.INF   (them) in.the   church
          'Other sects manage to pick people up off the street and put (them) in church.'
       b. depois trazem   **caderno**. Os   pais,   quando compram **[-]**,
          then   bring.3PL   notebook   the   parents when   buy.3PL   (it)
          trazem   aquele caderno   que têm   poucas folhas.
          bring.3PL   that   notebook   that have.3PL   few   pages.
          'Then they bring notebooks. When parents buy (them), they bring notebooks with few pages.'

(25)   a. uma   rádio precisa   gentes   que vai   colher   **informação**,
          a   radio need.3SG   people   that go.3SG   gather.INF information
          tratar   **informação**, depois pô-**la**   no   ar.
          process.INF   information, then   put.ACC.F.3SG in.the   air
          'A radio station needs people to do the work of gathering information, processing it and then broadcasting it.'
       b. O **político**   tem   esse lado que **o**   caracteriza.
          the politician   have.3SG   this side that ACC.M.3SG caracherize.3SG
          'Politicians have this side that distinguishes them.'

(26)   a. O **príncipe**   quando vai   à   corte   para falar   com o
          the prince   when   go.3SG to.the court   to   talk.INF with the
          pai,   vai   visitar o   pai   e   acusam   **ele**.
          father go.3SG visit.INF the   father   and   accuse.3PL PRON.M.3SG
          'When the prince goes to court to talk to his father, he goes to visit his father, and they accuse him'.

b. A **cobra** veio, parou mesmo de frente a mim.
the snake come.PST.3SG stop.PST.3SG right of front to me
Parou mesmo também a ver-me mesmo também.
stop.PST.3SG right also to see.INF-CL.ACC.3SG right also
Depois ele disse: se você não atacar **ela**,
then he say.PST.3SG if you.SG not attack.INF PRON.F.3SG,
ela não ataca você.
she not attack.3SG you.SG
'The snake came and stopped right in front of me. She stopped watching me too. Then he said: If you don't attack it, it won't attack you.'

c. Mulher, quando foi acompanhar **ele**, encontrou
woman when go.PST.3SG accompany PRON.M.3SG find.PST.3SG
ele já a sair.
PRON.M.3SG already to leave.INF
'The woman, when she went to accompany him, found him already leaving.'

These data show that STP clearly allows null objects in strong islands, which is consistent with the properties of BP, in which null objects are said to be allowed (Kato & Raposo 2001, Schwenter 2014). However, note that null objects have been argued to be also (marginally) possible in these contexts in EP (Raposo 2004, Duarte & Costa 2013) and have been attested in corpus data (Figueiredo 2014, Rinke et al. 2018).

Zooming in on the STP data, we can see that null objects are allowed in different types of syntactic islands. Examples (22) and (24b) have temporal adverbial subordinate clauses, but null objects are not exclusive to this syntactic context. We also found null objects in relative clauses, as in (27-28).

(27)   Quando **mais velha** era mais pequenina, (…) eu estava em
when more old be.IPFV.3SG more small, I be.IPFV.1SG in
casa, não trabalhava, só trabalhava é tomar
home not work.IPFV.3SG only work.IPFV.3SG be.3SG take.INF
conta dela. Uma altura assim que dava para pôr **[-]**
care of.her. A time like.this that give.IPFV.3SG to put.INF (her)
no jardim, eu pus **[-]**.
in.the garten I put (her).
'When the older one was smaller, I was at home, I didn't work, I just looked after her. When the time came to send her to nursery school, I did.'

(28)   – INT[6]: qual é a tua opinião sobre essa língua, o
what be.3SG the your opinion of this language the
dialeto, não é, que vocês chamam, **o lungwa santome**?
dialect not be.3SG that you.PL call.3PL the language Santome

---

[6]   Interviewer and informant are represented by INT and INF, respectively.

– INF: não   se   liga   muito. Há   casa   que   pai
          not   SE   connect much   There.is   house   that   father
          proíbe   filho   de   aprender**[-]**.
          forbid.3SG   son   from   learn.INF   (it)
          '- INT: What do you think of this language, the dialect, isn't it, that you call Lungwa Santome?
          - INF: People don't care much. There are houses where the father forbids his son to learn.'

In conclusion, the null object is the preferred option for DO realization in STP, with both clitic and pronominal forms being less used. This preference for null objects extends to strong islands, a syntactic context that was considered to not allow (or only marginally allow) null objects in EP and BP.

To examine whether syntactic context affects the occurrence of a null object, a chi-square test was performed, showing a significant association between syntactic context and null object realization (χ^2 (1, N = 632) = 6.62, p = 0.010). Null objects were more frequent in non-island contexts than in island contexts.
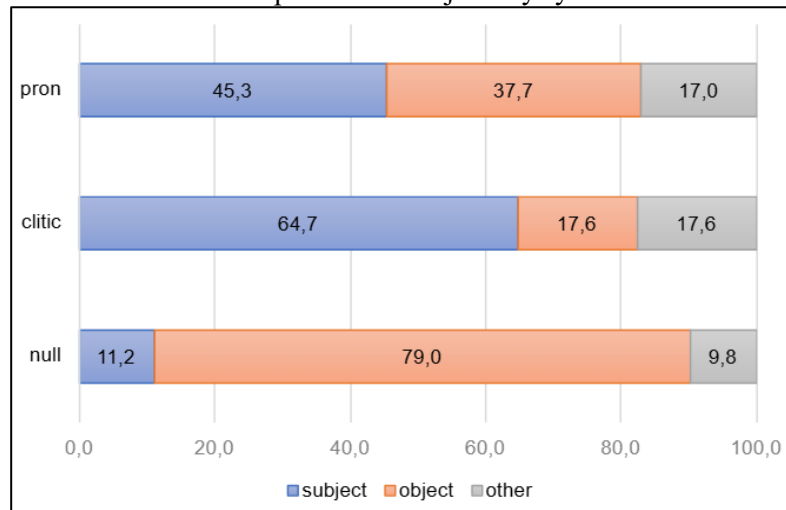
To investigate this further, we conducted a binomial logistic regression analysis, taking the realization of the object as the dependent variable and syntactic context as the predictor variable. The model confirmed that null objects were significantly more prevalent in non-island contexts than in island contexts (β = 0.604, p = 0.011). However, the model only explained a small proportion of the variance (R^2 = 0.0109), potentially due to the imbalance in the data (most occurrences were in non-island contexts).

Let us consider next the syntactic role of the antecedent.

4.2. Syntactic role of the antecedent

Regarding the syntactic role of the antecedent, Figure 3 shows that null objects in STP prefer antecedents in object position (79.0%, i.e., 417/528) (see (22-24)), whereas clitics tend to have antecedents in subject position: 64.7% (i.e., 33/51) (see (25b)). Full pronominal forms, on the other hand, seem to allow antecedents in both subject (45.3%, i.e., 24/53) (see (26a-b)) and object position (37.7%, i.e., 20/53) (see (26c)).

Figure 3. Distribution of null and pronominal objects by syntactic role of the antecedent.

The occurrence of an antecedent in a position other than subject or object (topics were not considered) is less expressive in the corpus. Nevertheless, some examples are given below in (29).

(29) a. se   fosse   um país   que   fazia   uso d**essa língua**,
         if   be.SBJV.3SG a country that be.IPFV.3SG use of.that language
         talvez eles já   podiam   aprender **[-]** como também
         maybe they already can.IPFV.3PL learn.INF (it) as   also
         aprenderam   português.
         learn.PST.3PL Portuguese
         'If it had been a country that used this language, they might have been able to learn it in the same way that they learned Portuguese.'

      b. Há   instituições aqui fica aberto até altas horas, com
         there.is   institutions here stay open until early hours   with
         dois três funcionários só   lá   dentro, não tem   ninguém   lá
         two three employees   only there inside not there.is no.one   there
         para ir   assaltá-**los**.
         to   go.INF   rob.INF-CL.ACC.M.3PL
         'There are institutions here that stay open until the early hours of the morning, with only two or three people inside, and there's no one there to rob them.'

      c. convivi   um caso de **um colega**   que
         live.together.PST.1SG a case of a   colleague   that
         trabalhávamos há   muito tempo, mas depois chegou
         work.IPFV.1PL there.is   much time   but after   arrives.PST.3SG
         um determinado tempo, quando já   não queriam   **ele**,
         a certain   time   when   already not want.IPFV.3PL PRON.M.3SG
         tiraram   **ele**   como fosse   uma brincadeira.
         take.PST.3SG PRON.M.3SG   how   be. SBJV.3SG a   joke
         'I've seen the case of a colleague who we had worked with for a long time, but after a while, when they didn't want him anymore, they took him away as a joke.'

These data show that structural parallelism is more relevant for null objects than for clitics/full pronouns.

On the one hand, STP differs from EP in that in EP clitics and null objects clearly pattern alike; both prefer antecedents in object position, showing a bias for structural parallelism (Rinke & Kerezova 2024). On the other hand, STP is consistent with (rural Afro-descendant varieties of) BP. Figueiredo (2014), considering data from rural Portuguese in the state of Bahia, mentions that the use of a pronominal form is the preferred strategy available to recover antecedents in the subject position. In STP, both clitic and full pronouns are preferred to null objects in this syntactic context, despite the observed variation.

A chi-squared test was performed to examine whether the syntactic role of the antecedent affects the occurrence of a null object. The results revealed a significant association between syntactic role and object realization: $\chi^2(2, N = 632) = 127$, $p < .001$. Null objects appeared most frequently with antecedents that had the syntactic

role of 'object' (79.0%), whereas pronominal objects appeared more frequently with antecedents that were not DOs.
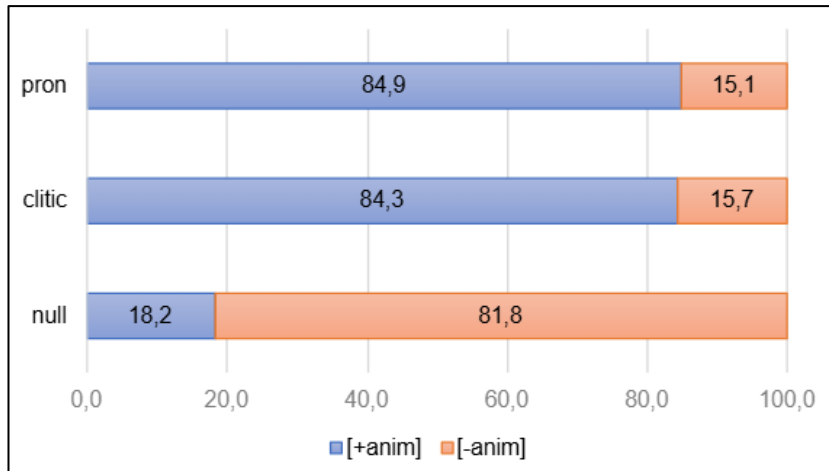
A binomial logistic regression confirmed this pattern. Compared to the 'object' category, the probability of using a null object was significantly lower in both the 'other' ($\beta = -1.60$, $p < .001$) and 'subject' ($\beta = -2.63$, $p < .001$) categories. This indicates that null objects mainly occur with antecedents that are syntactically DOs.

The model explained approximately 19.5% of the variance ($R^2 = 0.195$), demonstrating the strong effect of syntactic function on the realization of DOs as null or pronominal.

4.3. Animacy of the antecedent

With respect to animacy, Figure 4 shows that there is an animacy effect in the STP data, since null objects are mostly [-anim] (81.8%, i.e., 432/528); on the other hand, clitics (84.3%, i.e., 43/51) and full pronouns (84.9%, i.e., 45/53) are mostly [+anim].

Figure 4. Distribution of null and pronominal objects by animacy of the antecedent.

Note that, according to the extensive available literature on null objects, [+anim] referents favor the use of a full pronominal in BP, whereas [-anim] referents favor the occurrence of an anaphoric null object. Moreover, data from the EP spontaneous corpus analyzed by Rinke (2025) also show that [+anim] referents are predominantly realized by clitics and [-anim] referents by null objects. Hence, STP patterns with EP and BP in this respect.

Additionally, STP allows [+anim] null objects in both island and non-island syntactic contexts (see (30)), suggesting that STP is less restrictive in this respect than EP/BP.
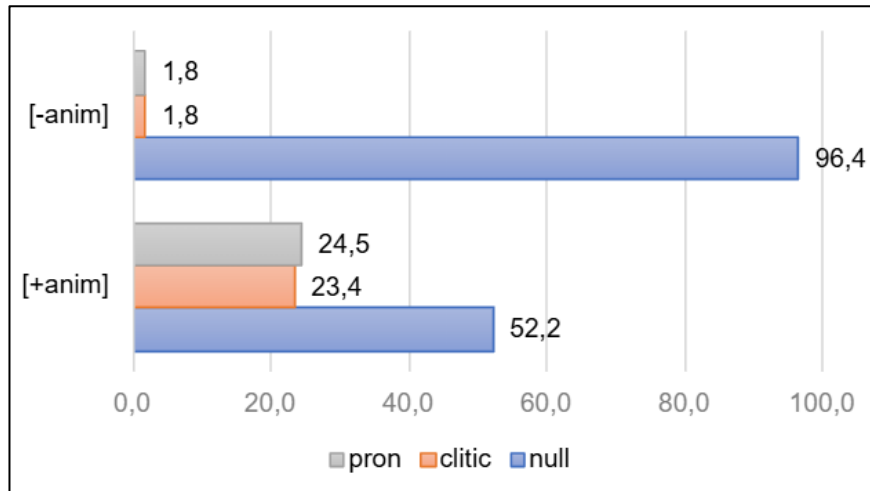
(30) a. um outro colega   ou amigo ou vizinho (…) faz cria d**os porcos**,
          a   other colleague or friend or neighbor   do   breed of.the pigs
          não está   a   ver? Quando abate   **[-]**,   se você
          not be.3SG to see? When   slaughter.3SG (them), if you.SG
          tem   possibilidade, você   pode comprar   um   quilo.
          have.3SG possibility,   you.SG may buy   one   kilo

'A colleague or friend or neighbor... he breeds pigs, do you see it? When he slaughters them, if you can, you can buy a kilo.'

b. banco aqui **as** **pessoas** saem   com plástico de dinheiro, saca
   bank here the   people   go.out.3PL with plastic of money   bag
   dinheiro na   mão na   rua   a andar. (…) Não tem
   money   in.the   hand in.the street   to walk.   not have.3SG
   nenhum assalto, pois   como vai   chegar   um   dia   desses
   no   robbery because how go.3SG come.INF   one day   of.those
   assim   que eles vão   vir   assaltar   **[-]**.
   this.way that they go.3PL come.INF rob.INF   (them)

'People here leave the bank with the money in a plastic bag, they carry this plastic bag in their hands and walk down the street. There are no robberies, but there will be a day when they'll come and rob them.'

However, it should be noted that there are more [+anim] null objects in STP than reported in most BP and EP studies (but see, e.g., Schwenter 2014, and Figueiredo 2014). In fact, if we consider only the [+anim] objects of the corpus, more than half of the [+anim] objects are null (52.2%, i.e., 96/184); [-anim] objects are almost always null (96.4%, i.e., 432/448) (see Figure 5).

Figure 5. Distribution of null and pronominal objects according to the animacy of the antecedent.

A chi-squared test was performed to examine whether animacy affects the occurrence of null objects. The results revealed a significant association between animacy and object type: $\chi^2(1, N = 632) = 186$, $p < .001$. Null objects appeared most frequently when the referent was [-animate] (96.4%), whereas only 52.2% of [+animate] referents were realized as null objects. In contrast, pronominal objects were more prevalent among [+animate] referents (47.8%).
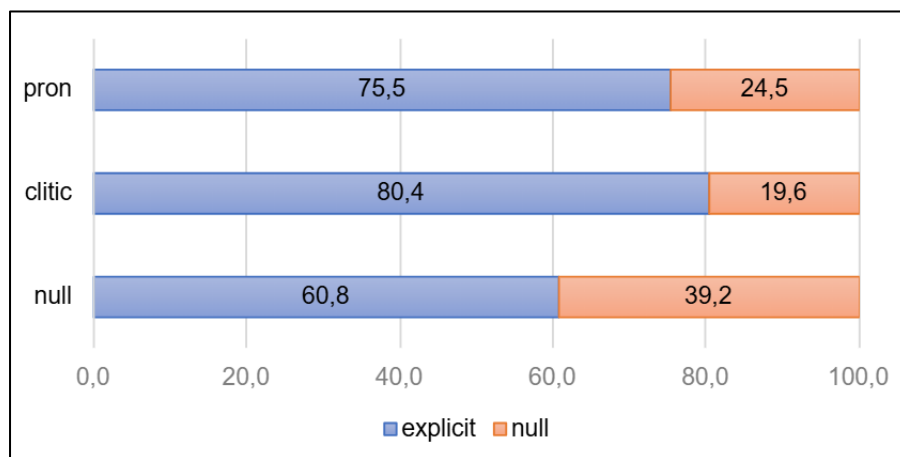
A binomial logistic regression confirmed this pattern. Compared to [+animate] referents, the probability of using a null object was significantly higher when the referent was [-animate] ($\beta = -3.21$, $p < .001$). This indicates that animacy plays a key role in object realization, with null objects being strongly favored when the referent is inanimate.

The model explained approximately 30.5% of the variance (R² = 0.305), demonstrating that animacy is a strong predictor of whether a DO is realized as null or pronominal.

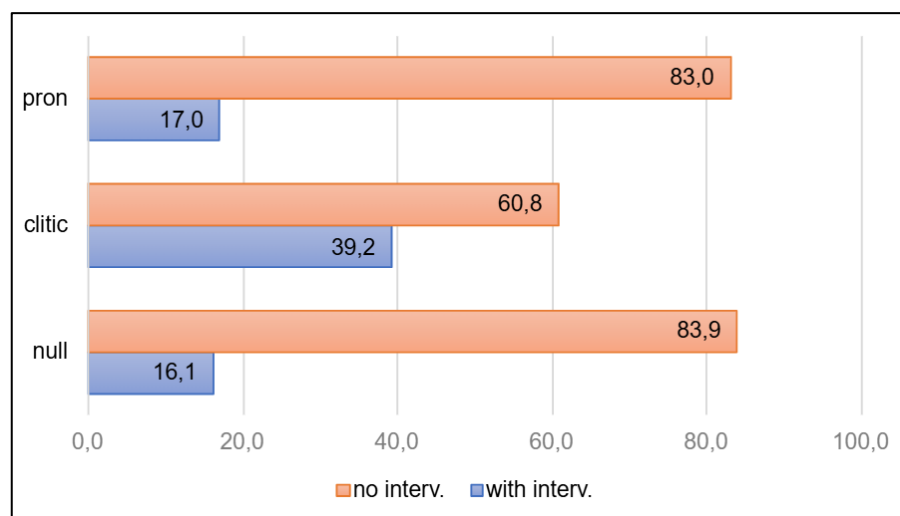## 4.4. Explicitness and distance of the antecedent

Null objects occur with both explicit (60.8%, i.e., 321/528) and null antecedents (39.2%, i.e., 207/528), as do clitics (80.4%, i.e., 41/51 *versus* 19.6%, i.e., 10/51) and full pronouns (75.5%, i.e., 40/53 *versus* 24.5%, i.e., 13/53). However, they all prefer explicit antecedents, as shown in Figure 6.

Figure 6. Distribution of null and pronominal objects according to the explicitness of the antecedent.

Moreover, both null objects and clitics/full pronouns tend to appear with no other intervening potential DO (see Figure 7 below).

Figure 7. Distribution of null and pronominal objects considering the intervention of a DO.

Null objects are considered first. 83.9% (i.e., 443/528) occur without an intervening DO (see (22) and (24)), and only 16.1% (i.e., 85/528) occur with an intervening DO (see (31) below).

(31) a. – INT: mas há   algumas organizações não-governamentais aqui, até
                 but there.is some   organizations not-governmental   here even
                 de voluntários que vêm   apoiar   **o hospital**$_i$,
                 of voluntaries that come.3PL support.INF the hospital
                 não sei   se conhece **algumas dessas situações**$_j$.
                 not know.1SG if know. 3SG some   of.these situations
           – INF: eu não conheço **[-]**$_j$ mas há   **[-]**$_j$. Há   assim   que
                 I not know.1SG (it) but there.is (it) there.is like.this that
                 apoia   **[-]**$_i$ sempre.
                 support.3SG (it) always
                 '– INT: But there are some non-governmental organizations here,
                 even voluntary ones, that come to support the hospital, I don't know
                 if you know about some of these situations.
                 – INF: I don't know, but there are. They always support.'
       b. nesse   capim é   que **eles**$_i$ vêm,   **esse cardume todo**$_i$,
          in.that grass be.3SG that they come.3PL that shoal   all
          para   vir   desovar **os   ovos**$_j$. Daqui é   que temos
          to   come.INF lay.INF the   eggs from.here be.3SG that have.3PL
          uma rede$_z$, tipo dum arco, então com essa rede é   que
          a   net   like of.a bow so   with that net be.3SG that
          vamos apanhar **[-]**$_i$,   despejar **[-]**$_i$, pôr   **[-]**$_i$
          go.1PL catch.INF (them) dump.INF (them) put.INF (them)
          na   canoa.
          in.the canoe
          'This is where they come to lay their eggs. From here we have a net, like a
          bow, so we use it to catch them, dump them and put them in the canoe.'
       c. **Criança**$_i$ cai,   parte   **braço**$_j$, quando gente vê…   **ou**
          child   fall.3SG break.3SG arm   when we   see.3SG or
          o   dedo$_z$ com   pedra, a gente tem de   levar   **[-]**$_i$
          the finger with   stone   we   have.3SG to take.INF (him/her)
          para ir   curar   **[-]**$_{i/j/z}$.
          to   go.INF heal.INF (him)
          'If a child falls and breaks his arm or finger with a stone, we must take him
           and heal him.'

In (31a), the speaker not only recovers the immediately explicit antecedent - *algumas dessas situações* ('some of these situations') – with two anaphoric null objects – but also uses an anaphoric null object to recover the explicit DO mentioned earlier. The interpretation is guaranteed by the repeated verb. However, in the case of (31b) and (31c), we have to rely on our knowledge of the world to identify the most plausible antecedent(s); hence, the interpretation of the object of *curar* ('heal') remains ambiguous, as both the child, his/her arm or his/her finger can be healed.

Null objects and full pronominals are similar in this respect. 83% (i.e., 44/53) of the full pronouns have no potential DOs intervening between them and their

antecedent (see (26a)). Only 17% (i.e., 9/53) have explicit intervening DOs (see (26b) and (32) presented below). There are no full pronominals with intervening null DOs in the data.

(32) a. se **o meu colega**$_i$ está fraco, **ele**$_i$ não consegue,
         if the my colleague be.3SG weak he not get
         portanto, dominar **aluno**$_j$, tenho que chamar **ele**$_i$ para
         therefore control student have.1SG that call PRON.ACC.M.SG to
         dizer: epá, por esse caminho que você está a ir,
         say.INF hey for that way that you be.3SG to go.INF
         não vai…
         not go.3SG
         'If my colleague is weak and can't deal with the students, I have to call him and say: - Hey, the way you're doing it, you're not going to go...'
     b. Em São Tomé chama moço carta, que é   o rapaz da
         in São Tomé call.3SG boy letter that be.3SG the boy of.the
         carta. Então, **ele**$_i$ leva **carta**$_j$ e pegaram **ele**$_i$.
         letter so he take.3SG letter and catch.PST.3SG PRON.ACC.M.3SG
         'In São Tomé they call him the letter boy. So, he carried the letter, and they caught him.'

In contrast to pronominals and null objects, structures with clitics are more flexible in allowing potential DO intervention. Although 60.8% (i.e., 31/51) of the clitics have no intervening DOs (see (22) and (25a)), 39.2% (i.e., 20/51) occur with one or more DOs intervening between the antecedent and the clitic; these are mostly explicit DOs (see (33a-b)), with only one occurrence of a null intervening DO (see (33c)).

(33) a. depois quando se criou   **o instituto**$_i$, que terminou
         then when SE create.PAST.3SG the Institute   that finish.PST.3SG
         em noventa e quatro por aí, então houve
         in ninety and four by there so there.is.PST
         necessidade$_j$ de pô-**lo**$_i$ em funcionamento.
         need   of put.CL.ACC.M.SG in functioning
         'Then when the Institute was set up, which ended in 1994 or so, there was a need to get it up and running.'
     b. **ele**$_i$ esqueceu-se   e deixou **a chave**$_j$ na ignição
         he forget.PST.3SG-SE and leave.PST.3SG the key in.the ignition
         do carro. Foi   para lá dentro para ir chamar **o**
         of.the car go.PST.3SG to there inside to go call the
         segurança$_z$ lá dentro para vir cá para ajudá-**lo**$_i$.
         security   there inside to come.INF here to help.3SG-CL.ACC.M.3SG
         'He forgot his keys and left them in the car. He went inside to call the security guard to come and help him.'

c. Quando **ele**_i reconheceu   **[-]**_j (…), quando o   presidente
when   he   recognize.3SG (him)   when   the president
passou,   mandou   outro agente prendê-**lo**_i
pass.PST.3SG   send.PST.3SG other officer arrest.INF-CL.ACC.M.SG
'When he recognized him and the president passed by, he sent another
officer to arrest him.'

In summary, both null and pronominals objects (clitic and full pronouns) prefer antecedents that are explicitly mentioned in the immediately preceding discourse, as previous studies have already shown for EP and BP. We must highlight, however, that, in STP, null objects may easily recover null antecedents, in contrast with what has been found for EP (Rinke & Kerezova 2024). As shown below, there are a number of examples of referential chains formed by a lexical head and various null objects distributed across several sentences (see (34-35)).

(34)   – INT: Quando coloca assim essa proteção, é   seguro? (...)
when   put.3SG this.way this protection be.3SG safe?
O   ladrão nunca tenta   cortar **[-]**?
the   thief   never   try.3SG cut   (it)
    – INF: Não corta   **[-]**, não. (...) A casa   só   de madeira assim
no   cut.3SG it   no   the house   only of wood   like.this
que eles pode   tenta   arrancar **[-]**,   porque à(s) vez(es)
that they can.3SG try.3SG rip.INF (it)   because sometimes
a gente prega   **[-]** com prego, eles arrancam   **[-]**.
we   nail.3SG (it)   with nail   they rip.3PL   (it)
'– INT: If you put this protection on, is it safe? (...) The thief will never try to cut it?
– INF: No, they don't cut it. If the house is only made of wood, so they try to rip it off because sometimes we nail it down, they rip it off.'

(35)   Há   cliente que prefere   mais **peças** assim   desmontada.
there.is client   that prefer.3SG more pieces   like.this disassembled
Como vem   encaixotada, há   cliente   que prefere
how come.3SG in.a.box   there.is customers that prefer.3SG
montar   **[-]** em casa   e   há   outro também compra
assemble.INF   (it) in   house   and there.is other also   buy.3SG
[-] aqui. Como já   está   montada   assim, levamos **[-]**
(it) here how already be.3SG assemble.PTCP   this.way take.3PL (it)
só. Pomos **[-]** em casa   do   cliente só, vimos   embora.
only put.1PL (it) in house of.the   client   only come.1PL   away
'There are customers who prefer parts to be disassembled. As it comes in a box, there are customers who prefer to assemble it at home and others who buy it here. Because it's already assembled, we just took it away. We just put it in the customer's house and left.'

A chi-squared test was performed to examine whether the presence of an intervening antecedent affects the realization of the DO as null or pronominal. The

results revealed a significant association between antecedent configuration and object type (χ^2(1, N = 632) = 8.16, p = .004).

It is noteworthy that 83.9% of null objects occurred in contexts without an antecedent, while only 16.1% occurred with an antecedent present. As for pronominal objects, 72.1% appeared without an intervening antecedent and 27.9% appeared with one intervening antecedent.

A binomial logistic regression confirmed this pattern. The likelihood of a null object being realized was significantly lower when an intervening antecedent was present (β = −0.70, p = .005), suggesting that distance plays a role in licensing null objects.

The model explained approximately 1.3% of the variance (R^2 = 0.013), indicating a relatively modest effect compared to other variables, such as animacy. Nonetheless, it points to the relevance of discourse configuration in shaping object realization.

In the next section we will return to our research questions and hypotheses and discuss the data described.

5. Discussion and conclusions

We start by discussing our first research question, which focuses on the productivity and distribution of null and pronominal objects in STP.

Q1. *What is the preferred option for object realization in non-island and island contexts in STP: null or pronominal objects?*

Our results confirm the findings of Gonçalves et al. (2024) regarding pronominal objects, which were shown to be considerably less productive than null objects. These are clearly the preferred option in STP. Null objects are allowed and widely used in both island and non-island contexts, although it was observed that they are significantly more prevalent in non-islands than in islands. However, this may be due to an imbalance in the data, as the non-island contexts considerably outnumber the island contexts in the corpus. Therefore, taking into account productivity and locality, the STP data tend to converge with BP and diverge from EP, where, according to the available studies, it is not controversial that clitics are the preferred option, with null objects being less productive in spoken corpora and less accepted in elicitation tasks; in BP, on the other hand, null objects are more productive than full pronouns and they are allowed in both islands and non-islands with antecedents in the same sentence in subject or object position. The data support our H1, which predicted that null objects would be more productive than pronominal objects in both non-island and island contexts in STP.

At this point we need to explain why we find higher rates of null objects in STP than in EP (and BP), irrespective of the syntactic context. On the other hand, we also need to explain why the STP grammar does not show pronominal forms in variation with null objects, as observed in BP, but instead shows consistently low percentages of pronominal forms (both clitic and full pronouns).

The preference for null objects in STP may be explained by one or more of the following factors. First, it has been proposed that null objects in non-native grammars

do not show as much sensitivity to syntactic and semantic constraints as native grammars (e.g., Teixeira et al. 2024). This could explain the widespread use of this structure in STP, since this variety historically emerged in the course of successive processes of L2 acquisition.

Furthermore, one could also hypothesize that the widespread use of null objects in STP is due to a general tendency to overgeneralize null objects. This is a trend that has been observed diachronically in BP, where null objects have become increasingly more frequent since the 19th century. On the other hand, the expanded use of non-propositional null objects may also point to internal language change, which finds support in the generational differences observed in present-day EP. As noted by Flores et al. (2017) and Rinke et al. (2018), younger generations show a higher number of null objects in their speech than older generations.

Typically, studies of European language varieties attempt to explain features that differ from the target language in terms of linguistic contact. For example, considering 3rd object clitics in STP, Gonçalves et al. (2024) hypothesize that the higher number of full pronouns instead of clitics could result from a reinterpretation of the pronominal system in STP, according to which full pronouns could be reanalyzed as weak pronouns, as in Santome, the dominant creole. However, our data, covering a larger number of structures in the PALMA-STP corpus, show that the STP grammar has a clear preference for null objects over both clitics and full pronouns. If the (historical) L1 – Santome, which is still the dominant creole in the archipelago – has played a direct role and still shapes the internal structure of DOs, one would expect STP speakers to use pronominal forms. However, language contact may not have directly affected DO constructions but may instead influence the use of syntactic structures that typologically distinguish Santome from EP, specifically serial verb constructions (Hagemeijer 2000). We refer to the examples given in (24a) e (31b), which seem to reflect those constructions with a locative/directional interpretation. We acknowledge that the overgeneralization of null objects in L2 acquisition may occur independently of L1 properties (Teixeira et al. 2024). The widespread use observed in the STP data provides further evidence for a complex acquisition process of these structures. Some additional and more in-depth research must be developed on this issue.

All in all, the historical process of nativization that STP has undergone in the last decades, the situation of language contact, as well as internal language changes, could converge to explain the data presented and described.

Our remaining research questions addressed the role played by factors which contribute to the accessibility of the antecedent in object realization. We consider that accessibility may reflect the interplay between syntactic, semantic and discourse constraints, and may be determined by factors such as the syntactic role of the antecedent, the semantic feature of animacy, the explicitness of the antecedent and the distance between the object and its antecedent (recency of mention and intervention of potential null or explicit antecedents). Our second research question focuses on the syntactic role of the antecedent.

Q2. *Does the grammatical function of the antecedent condition the realization of the DO in STP?*

We have seen that null objects in STP prefer antecedents in object position, whereas clitics tend to have antecedents in subject position and full pronominal forms allow both. Therefore, these data only partially confirm H2a, which, following the Accessibility scale (Ariel 1990, 1991), predicted that null objects would prefer antecedents in subject position, whereas clitic/strong pronouns would be more flexible in their choice of antecedent – these predictions are confirmed for strong pronouns, but not for clitics (which prefer subject antecedents) or null objects (which prefer object antecedents). As for H2b, according to which both null objects and clitic/strong pronouns show an effect of structural parallelism (Rinke & Kerezova 2024), it is again only partially confirmed, as this effect is only observed with null objects.

As a result, we can conclude that structural parallelism is more relevant than accessibility for null objects in STP. In this respect, they differ from clitics and full pronouns, showing that this African variety of Portuguese has similar features to rural varieties of Portuguese in Brazil (Figueiredo 2014).

Our third research question addresses the role played by animacy in object realization.

Q3. *Does animacy of the antecedent condition the realization of the DO in STP?*

Salience or prominence is also commonly perceived in terms of animacy. Following the Referential Hierarchy (Cyrino et al. 2000), our H3 predicted that null objects would tend to be [-anim], whereas [+anim] objects would be realized as clitics or full pronouns. This is confirmed by our data, which show that there is an animacy effect in DO realization in STP: the object is significantly more likely to be realized as a null form when the referent is inanimate, and [+anim] referents are predominantly realized by pronominal forms. In this respect, STP converges with EP and BP. However, we have also shown that there are more [+anim] null objects in STP than reported in most EP studies, suggesting that, as in BP, we may be observing an extension of null objects along the Referential Hierarchy in STP.

Considering animacy restrictions on null objects in EP, Rinke (2025) argues in favor of a colloquial referential pronominal system, in which animacy is the relevant factor. According to the author, "in the Standard language, the etymological system is prevalent, and clitics are used to refer to animate and inanimate object referents. In the colloquial language, the referential system is prevalent, and null objects are used to refer to inanimate referents whereas clitics refer to animate one". Evidence in favor of a generalization of the referential system in spontaneous data is provided by the STP data we have analyzed. However, we still need to explain why, in this variety of Portuguese, null objects are also – and widely – used with [+anim] referents.

In an attempt to explain (residual) cases of [+anim] null objects in EP, since they are not really ungrammatical in EP, Rinke (2025) hypothesizes that there is an internal hierarchy to [+anim] nouns, following Erteschik-Shir et al. (2013). According to this hierarchy, (i) humans are ranked higher than other animates; and (ii) among humans, children, for example, are ranked lower than adults. Indeed, the author found that most of the [+anim] null objects found in the EP that she analyzed have animals and children as antecedents. This proposal does not extend to our STP data. If we consider the total number of [+anim] null objects analyzed in the corpus (corresponding to 18.2%, i.e., 96/528 of the null objects) (see Figure 4), we find that 76 of these null objects involve [+hum] referents. Therefore, the vast majority of

[+anim] null objects are human, ranked higher in the internal hierarchy to [+anim] nouns previously mentioned. Moreover, these [+anim] antecedents can be singular (see (27)) and definite (see (36) below), triggering a specific interpretation.

(36) a. Eu não continuei   mais no   estudo porque estive   no
         I not continue.PST.3SG more in.the   study because be.3SG in.the
         hospital   com **aquele moço de cinco anos**. Ele quando tem
         hospital   with that   boy of five years he   when   have.3SG
         a febre   dentro, vai   querer dar   convulsão e   tem
         a fever   inside   go.3SG   want.INF give. INF convulsion and have.3SG
         que levar   **[-]**   no   hospital.
         that take.INF (it)   in.the   hospital
         'I didn't study anymore because I was in hospital with this five-year-old boy. If he has a fever inside him, he could have a convulsion and must be taken to hospital.'
      b. meu problema é   que   meus filho foi   batizado em   católico
         my problem be.3SG that   my   son   be.PST.3SG baptized in   Catholic
         e   a   **mulher** que eu vivo   com ela também é   católica
         and the woman   that I live.1SG with her also   be.3SG Catholic
         Eu já   tenho   tentado para   ver   se   consigo levar   **[-]**
         I   already   have.1SG   try.PTCP to   see.INF if   get.1SG take.INF (her)
         para lá,   mas está   a ser   difícil. Ela   disse
         to   there but be.3SG to be.INF difficult she   say.PST.3SG
         que ela não vai   sair.
         that she not go.3SG   leave.INF
         'My problem is that my son was baptized Catholic and the woman I live with is also Catholic. I've been trying to see if I can take her there (=to another church), but it's proving difficult. She says she doesn't want to go.'

Although there are still not many studies investigating the acquisition of null objects in L2 Portuguese, especially studies that address the role of animacy and accessibility, as pointed out by Teixeira et al. (2024), the studies that have been conducted so far show that L2 learners of Portuguese do not show the animacy constraints exhibited in the target grammar, even at advanced stages of acquisition (Teixeira et al. 2024; Zhao 2020). Therefore, the expanded use of animate null objects in STP may reflect both the colloquial referential pronominal system (Rinke 2025) and the internal language change resulting from its historical past as an L2.

Our fourth research question was concerned with the role of the explicitness of the antecedent and of the distance between the object and its antecedent (assessed in terms of the intervention of other potential antecedents) on object realization.

Q4. *Do null and pronominal objects differ in their preference for antecedents that are explicitly mentioned in the immediately preceding discourse (i.e., without the intervention of other referring expressions)?*

We observed that, in STP, null and pronominal objects mostly recover antecedents that are explicitly mentioned in the previous discourse: However, according to the Accessibility scale (Ariel 1990, 1991), pronominal objects would be

expected to allow null antecedents (which are less accessible that explicit ones) more easily than null objects. In fact, 39.2% of the null objects in the corpus recover null antecedents but only 24.5% of the pronouns and 19.6% of the clitics do so. As for intervention, again as expected, for most of the null and pronominal objects their antecedent is the closest nominal DO. This tendency is weaker with clitics, which appear to allow intervening DOs between them and their antecedent more easily. These results do not support H4a, which predicted different behaviors for null objects and clitic/strong pronouns, with the former showing a preference for more accessible antecedents (i.e., overtly realized antecedents and no intervention) and the latter showing no such preference. Instead, we found that null and pronominal forms display identical patterns of behavior. However, null objects appear to be more flexible than their pronominal counterparts regarding the realization of their antecedent (which is unexpected), while clitics appear to be slightly more 'tolerant' of intervening potential antecedents. All in all, our results confirm H4b, which, based on the findings of Rinke & Kerezova (2024), predicted that both null objects and pronominal DOs would prefer explicit antecedents with no intervening potential antecedents.

It is important to bear in mind, however, that it is hard to make comparisons between null and pronominal objects given the differences in sample size (there were 528 null objects to 51 clitics and 53 pronouns). Hence, it is difficult to reach any firm conclusions on whether these two types of forms in fact behave differently with respect to accessibility. In order to do so, we would need to enlarge the size of the corpus so as to include more occurrences of pronouns and clitics. However, given the small number of naturally occurring clitic and pronominal objects observed both in our study and in previous work (Gonçalves et al 2024), it might be more appropriate to employ experimental methodology.

We saw in 2.2.5 above that previous research on EP and BP has shown that anaphoric null objects in these two varieties tend to favor highly accessible antecedents, i.e., antecedents that are explicitly mentioned in the immediately preceding discourse (Duarte & Costa 2013; Kato et al. 2023). If we compare our findings on STP with the data from EP reported in Rinke & Kerezova (2024), it is clear that the "strong tendency" displayed by null objects to recover an explicit antecedent, which they observed in their corpus (their percentage was 82.4%), is not as strong in our STP corpus (60.8%). Moreover, they also observed that null objects tended not to participate in long referential chains. In our corpus, however, we noted a few cases where null objects occur in long referential chains consisting of a lexical head and several null objects distributed across several sentences (see the examples in (34) and (35) above). In this study, we did not conduct a systematic analysis of these cases, which we will leave for future work.

Our findings with respect to the effect of the explicitness of the antecedent and the intervention of other potential antecedents on null objects in STP are in line with what we found for the other properties investigated in our study, in particular, regarding syntactic distribution and animacy, thus supporting the conclusion that null objects are less constrained in STP than in EP or even BP. As suggested above, the distribution and properties of null objects in STP appear to suggest that they may be extending along the referential hierarchy (in a process similar to the one that, according to Cyrino 1997, occurred in the diachrony of BP) and thus becoming more referential than null objects in EP (and possibly also in BP).

We acknowledge that inter and intra-speaker variation may be relevant to the distribution of null and pronominal objects in STP. This is a question that needs to be addressed, considering in particular the role of sociolinguistic variables such as age and schooling level in the realization of the DO. Another aspect which was not considered in our analysis, but which is clearly highly relevant, as evidenced by the above literature review, is the role played by definiteness and specificity in the licensing of null and pronominal objects in STP, as well as the interaction between these features and other features such as animacy. We will come back to these questions in future work.

Acknowledgments

References

Afonso, Beatriz. 2008. *A problemática do bilinguismo e ensino da língua portuguesa em São Tomé e Príncipe* [The issue of bilingualism and Portuguese language teaching in São Tomé and Príncipe]. MA dissertation, Universidade Nova de Lisboa.

Ariel, Mira. 1990. *Accessing noun-phrase antecedents*. Routledge.

Ariel, Mira. 1991. The function of accessibility in a theory of grammar. *Journal of Pragmatics* 16: 443–463. DOI: http://doi.org/10.1016/0378-2166(91)90136-L

Bouchard, Marie-Eve. 2019. Becoming Monolingual: The Impact of Language Ideologies on the Loss of Multilingualism on São Tomé Island. *Languages* 4(3): 50. https://doi.org/10.3390/languages4030050

Campos, Héctor. 1986. Indefinite object drop. *Linguistic Inquiry* 17: 354–359.

Carminati, Maria N. 2002. *The processing of Italian subject pronouns.* PhD thesis, University of Massachusetts at Amherst.

Castro, Tammer, Rothman, Jason, & Marit Westergaard. 2017. On the directionality of cross-linguistic effects in bidialectal bilingualism. *Frontiers in Psychology* 8: 1382. https://doi.org/10.3389/fpsyg.2017.01382

Cyrino, Sonia. 1997. *O objeto nulo no Português do Brasil: um estudo sintático-diacrônico* [Null object in Brazilian Portuguese: a syntactic-diachronic study]. Editora da UEL.

Cyrino, Sonia. 2001. O objeto nulo no português do Brasil e no português de Portugal [Null object in Brazilian Portuguese and European Portuguese]. *Boletim da ABRALIN* 25: 173–181.

Cyrino, Sonia, & Gabriela Matos. 2016. Null objects and VP Ellipsis in European and Brazilian Portuguese. In W. L. Wetzels, J. Costa, & S. Menuzzi (eds) *The Handbook of Portuguese Linguistics*, 294–316. Wiley Blackwell.
 https://doi.org/10.1002/9781118791844.ch16

Cyrino, Sonia, Duarte, Maria Eugênia, & Mary Kato. 2000. Visible subjects and invisible clitics in Brazilian Portuguese. In M. Kato, & E. Negrão (eds) *Brazilian Portuguese and the Null Subject Parameter*, 55-73. Iberoamericana.

Dickinson, Kendra V., Lamberti, Luana, & Scott A. Schwenter. 2024. Evaluating the form of third-person anaphoric direct objects in Portuguese: A cross-dialectal study. *Journal of Portuguese Linguistics* 6: 1–29. https://doi.org/10.16995/jpl.16290

Duarte, Inês. 1987. *A Construção de Topicalização na Gramática do Português. Regência e Ligação e Condições sobre Movimento* [The Construction of Topicalization in Portuguese Grammar. Government and Binding and Conditions on Movement]. PhD thesis, Universidade de Lisboa.

Duarte, Inês, & João Costa. 2013. Objecto Nulo [Null object]. In E. P. Raposo *et al.* (eds) *Gramática do Português, vol. II*, 2339–2348. Gulbenkian.

Erteschik-Shir, Nomi, Ibnbari, Lena, & Sharon Taube. 2013. Missing objects as topic drop. *Lingua* 136: 145–169. http://doi.org/10.1016/j.lingua.2013.07.009

Figueiredo, Cristina. 2014. Objeto nulo versus estratégias pronominais no português rural do estado da Bahia e no de Portugal [Null object versus pronominal strategies in rural Portuguese in the state of Bahia and in Portugal]. *Confluência: Revista do Instituto de Língua Portuguesa* 46: 79–114.

Flores, Cristina, Rinke, Esther, & Cecília Azevedo. 2017. Object realization across generations. A closer look on the spontaneous speech of Portuguese first and second generation migrants. In E. Domenico (ed.) *Complexity in Acquisition*, 179–206. Cambridge Scholars.

Flores, Cristina, Rinke, Esther, & Aldona Sopata. 2020. Acquiring the distribution of null and overt direct objects in European Portuguese. *Journal of Portuguese Linguistics* 19(5): 1–20. https://doi.org/10.5334/jpl.239

Fukumura, Kumiko, & Roger van Gompel. 2011. The effect of animacy on the choice of referring expression. *Language and Cognitive Processes* 26(10): 1472–1504. https://doi.org/10.1080/01690965.2010.506444

Gonçalves, Rita. 2010. *Propriedades de subcategorização verbal no português de São Tomé* [Verbal subcategorization properties in the Portuguese of São Tomé]. MA dissertation, Universidade de Lisboa.

Gonçalves, Rita. 2016. *Construções ditransitivas no português de São Tomé* [Ditransitive constructions in the Portuguese of São Tomé]. Tese de Doutoramento, Universidade de Lisboa.

Gonçalves, Rita, Duarte, Inês, & Tjerk Hagemeijer. 2022. Dative Microvariation in African Varieties of Portuguese. *Journal of Portuguese Linguistics* 21(6): 1–39. https://doi.org/10.16995/jpl.8488

Gonçalves, Rita, Duarte, Inês, & Tjerk Hagemeijer. 2024. Objetos diretos em variedades africanas do Português: um estudo de caso de microvariação [Direct objects in African varieties of Portuguese: a case study of microvariation]. In S. F. Brandão, & S. R. Vieira (eds) *Para o estudo comparativo de variedades do Português: Questões teórico-metodológicas e análises de dados*, 53–84. De Gruyter. https://doi.org/10.1515/9783110670257-005

Gonçalves, Rita, Hagemeijer, Tjerk, Alcântara, C., Cornejo, Catarina, Madureira, Raquel, Généreux, Michel, & Amália Mendes. 2021. *PALMA Corpus São Tomé e Príncipe*. Centro de Linguística da Universidade de Lisboa.

Hagemeijer, Tjerk. 2000. *Serial verb constructions in São-Tomense*. MA dissertation, Universidade de Lisboa.

Hagemeijer, Tjerk. 2018. From creoles to Portuguese: Language shift in São Tomé and Príncipe. In L. Álvarez López, P. Gonçalves, & J. Avelar (eds) *The Portuguese Language Continuum in Africa and Brazil*, 169–184. John Benjamins. https://doi.org/10.1075/ihll.20.08hag

Hagemeijer, Tjerk. 2024. São Tomé and Príncipe. In U. Reutner (ed.), *Manual of Romance languages in Africa*, 609–624. De Gruyter. https://doi.org/10.1515/9783110628869-027

Hagemeijer, T. *et al.* 2022. The PALMA corpora of African varieties of Portuguese. In N. Calzolari *et al.* (eds) *Proceedings of the 13th Conference on Language Resources and Evaluation (LREC 2022)*, 5047-5053. European Language Resources Association.

Huang, James C.-T. 1984. On the distribution and reference of empty pronouns. *Linguistic Inquiry* 15: 531–575.

The Jamovi project. 2025. *Jamovi* (Version 2.6) [Computer Software]. Retrieved from https://www.jamovi.org

Kato, Mary, & Eduardo, P. Raposo. 2001. O objeto nulo definido no português europeu e no português brasileiro: convergências e divergências [The definite null

object in European Portuguese and Brazilian Portuguese: convergences and divergences]. *Atas do XVI ENAPL*, 673–685. Associação Portuguesa de Linguística.

Kato, Mary, Martins, Ana Maria, & Jairo Nunes. 2023. *The Syntax of Portuguese*. Cambridge University Press.

Lima Afonso, Helena. 2009. *Interferências linguísticas: um contributo para o ensino da língua portuguesa em São Tomé e Príncipe* [Linguistic Interference: A Contribution to the Teaching of Portuguese in São Tomé and Príncipe]. MA dissertation, Universidade de Lisboa.

Madureira, Raquel. 2023. Clíticos reflexos e inerentes na variedade de português de São Tomé e Príncipe [Reflexive and inherent clitics in the Portuguese variety of São Tomé and Príncipe]. *Revista da Associação Portuguesa de Linguística* 10: 182-199. https://doi.org/10.26334/2183-9077/rapln10ano2023a10

Raposo, Eduardo P. 1986. On the null object in European Portuguese. In O. Jaeggli, & C. Silva Corvalán (eds) *Studies in Romance Linguistics*, 373–390. Foris. https://doi.org/10.1515/9783110878516-024

Raposo, Eduardo P. 2004. Objetos nulos e CLLD. Uma teoria unificada [Null objects and CLLD. A unified theory]. *Revista da ABRALIN* 3(1/2): 41–73.

Rinke, Esther. 2025. Animacy restrictions on null objects in European Portuguese. In P. Barbosa, & C. Flores (eds) *Null Objects from a Cross-Linguistic and Developmental Perspective*, 114–140. John Benjamins. https://doi.org/10.1075/lfab.19.05rin

Rinke, Esther, & Nelli Kerezova. 2024. Que fatores condicionam a resolução anafórica de objetos nulos e clíticos em português europeu? Um estudo de corpus [Which factors condition the anaphoric resolution of null and clitic objects in European Portuguese? A corpus study]. *Revista da Associação Portuguesa de Linguística* 11: 180–195. https://doi.org/10.26334/2183-9077/rapln11ano2024a9

Rinke, Esther, Flores, Cristina, & Pilar Barbosa. 2018. Null objects in the spontaneous speech of monolingual and bilingual speakers of European Portuguese. *Probus* 30(1): 93–120. https://doi.org/10.1515/probus-2017-0004

Rinke, Esther, Flores, Cristina, & Aldona Sopata. 2019. Heritage Portuguese and Heritage Polish in contact with German. More evidence on the production of objects. *Languages* 4(3): 53. https://doi.org/10.3390/languages4030053

Sainzmaza-Lecanda, Lorena, & Scott A. Schwenter. 2017. Null objects with and without bilingualism in the Portuguese-and Spanish-speaking world. In K. Bellamy, M.W. Child, P. González, A. Muntendam, & M.C. Parafita Couto (eds) *Multidisciplinary Approaches to Bilingualism in the Hispanic and Lusophone World*, 95–119. John Benjamins. https://doi.org/10.1075/ihll.13.05sai

Schwenter, Scott A. 2014. Two kinds of Differential Object Marking in Portuguese and Spanish. In P. Amaral, & A. M. Carvalho (eds) *Portuguese/Spanish Interfaces*, 237–260. John Benjamins. https://doi.org/10.1075/ihll.1.12sch

Schwenter, Scott A. & Gláucia V. Silva. 2003. Anaphoric direct objects in spoken Brazilian Portuguese: semantics and pragmatics. *Revista Internacional de Lingüística Iberoamericana* 1(2): 99–123.

Teixeira, Joana, Fiéis, Alexandra, & Ana Madeira. 2024. Efeitos do input na aquisição de PE L2: o caso dos objetos nulos [The effects of input on L2 EP acquisition: the case of null objects]. *Revista da Associação Portuguesa de Linguística* 11: 196–214. https://doi.org/10.26334/2183-9077/rapln11ano2024a10

Zhao, LiangLiang. 2020. *Aquisição do objeto nulo e dos pronomes clíticos por falantes chineses de português língua segunda* [The acquisition of null objects and clitic pronouns by Chinese-speaking learners of Portuguese as a second language]. MA dissertation, Universidade do Minho.