http://dx.doi.org/10.31513/linguistica.2023.v19n2a61298

Irrealis mood in Lung'Ie: *ka*

Modo irrealis em Lung'Ie: ka

Núbia Ferreira Rech[1]

Ana Lívia Agostinho[2]

ABSTRACT

This paper describes and analyzes the *ka* particle in constructions with dynamic predicates in Lung'Ie, an endangered Portuguese-lexifier creole language spoken in Príncipe Island. In contrast with the literature on Lung'Ie (GÜNTHER, 1973; MAURER, 2009), we argue that *ka* plays one single role in Lung'Ie: marking the irrealis mood (cf. AGOSTINHO; RECH, 2023). Our data show that *ka* is found in hypothetical, conditional, and counterfactual constructions, as well as in future and habitual aspect constructions. We also verified that *ka* is not used in past constructions nor in those that describe events at the time of the speech – prototypical realis contexts.

KEYWORDS: Creole languages. TMA. São Tomé and Príncipe. Gulf of Guinea.

RESUMO

Este artigo descreve e analisa a partícula *ka* em construções com predicados dinâmicos em Lung'Ie, uma língua crioula de base lexical portuguesa em ameaça de extinção, falada na ilha de Príncipe. Em contraste com a literatura sobre o Lung'Ie (GÜNTHER, 1973; MAURER, 2009), nós argumentamos que *ka* desempenha uma única função em Lung'Ie: marcador de modo irrealis (cf. AGOSTINHO; RECH, 2023). Nossos dados mostram que *ka* é encontrado em construções hipotéticas, condicionais e contrafactuais, assim como em construções no futuro e no aspecto habitual. Nós também verificamos que *ka* não é usado em construções no passado nem naquelas que descrevem eventos no momento da fala – contextos *realis* prototípicos.

PALAVRAS-CHAVE: Línguas crioulas. TMA. São Tomé e Príncipe. Golfo da Guiné.

Introduction

This study[3] focuses on describing and analyzing the *ka* particle in grammar structures with dynamic predicates in Lung'Ie.[4] We show that the *ka* particle is solely employed to mark a context

[1] Universidade Federal de Santa Catarina (UFSC), nubiarech1971@gmail.com, https://orcid.org/0000-0002-9278-2702.

[2] Universidade Federal de Santa Catarina (UFSC), a.agostinho@ufsc.br, https://orcid.org/0000-0002-2395-4961.

[3] According to the CRediT taxonomy (https://www.casrai.org/credit.html), the roles of each author were: Author 1 – Conceptualization, Investigation, Data curation, Methodology, Writing – original draft, Writing – review & editing; Author 2 – Conceptualization, Investigation, Data curation, Methodology, Writing – original draft, Writing – review & editing.

[4] We focused on dynamic verbs because stative verbs show a different behavior in Lung'Ie, as in other creole languages (cf. MAURER, 2009, p. 71; BICKERTON, 1981, p. 58, 1984). Winford (2017, p. 13) observes that "[i]t has now been well established that the distinction between stative and nonstative or dynamic predicates is central to all creole TMA systems, although there are differences across creoles concerning which verbs are assigned to one or the other category".

Recebido em 04/10/2023- Aceito em 07/03/2024.

as Irrealis (cf. AGOSTINHO; RECH, 2023), i.e., it does not play multiple functions, as stated by Günther (1973) and Maurer (2009). This study also contributes to the description of mood systems in creole languages since, according to Winford (2018, p. 203), "the areas of mood and modality remain the most neglected aspects of the study of creole TMA systems."

Our investigation is based on the available literature on Lung'Ie (GÜNTHER, 1973; MAURER, 1997, 2009; AGOSTINHO, 2015, 2016b; AGOSTINHO; ARAUJO, 2021; AGOSTINHO; RECH, 2021) and original fieldwork data (AGOSTINHO, 2009, 2016a, 2019; AGOSTINHO; RECH, 2016).[5] This paper builds upon the analysis presented in Agostinho and Rech (2023), incorporating a different and expanded set of data as well as new contexts to provide a more comprehensive examination.

The corpus of research about the Lung'Ie linguistic system is still quite small, within which we can highlight the works of Günther (1973), Maurer (1997, 2009), Agostinho (2015, 2016b), and Agostinho and Rech (2023). According to Günther (1973), *ka* marks the habitual aspect, future tense, generic reference, and iterative contexts. Maurer (2009) associates it with several other functions, such as a marker for counterfactuality, modality, future tense, habitual aspect, imperfective aspect, and generic reference, among others. This paper argues that such contexts are considered irrealis in Lung'Ie, and for this reason, the *ka* particle is used. On the other hand, in realis contexts, the particle appears as an unmarked form – similar to what is observed in other languages that oppose realis and irrealis moods (cf. MITHUN, 1995, 1999; ELLIOTT, 2000). Even though Günther's fieldwork was conducted nearly half a century ago, we did not find any changes regarding *ka* when comparing his data to Maurer's and Agostinho's. This is probably because the Lung'Ie speakers portrayed in the more recent work are generally elders.[6]

This paper consists of four sections: Section 1 introduces historical information regarding the languages spoken in the Gulf of Guinea based on the works of Schuchardt (1889), Günther (1973), Ferraz (1979), Cardoso (2007), Hagemeijer (2011), Araujo et al. (2013), Agostinho and Hyman (2021), among others. In Section 2, we analyze the *ka* particle in Lung'Ie, investigating its employment in contexts classified as prototypically irrealis (subsection 2.1); prototypically realis (subsection 2.2); and also, in contexts that can be marked as realis and/or irrealis depending on cultural (or even diachronic) aspects (subsection 2.3); in the end of the section, we make considerations regarding the irrealis mood in Lung'Ie (subsection 2.4). Lastly, Section 3 summarizes the main contributions of this research and suggests other topics for future investigation.

[5] The transcriptions are left as provided by the original authors. The data from Agostinho (2009, 2016a, 2019) and Agostinho and Rech (2016) was transcribed by the second author using the *Alfabeto Unificado para as Línguas Nativas de São Tomé e Príncipe* (ALUSTP) (PONTÍFICE et al., 2010).

[6] One of Günther's informants was also the second author informant.

1. Historical Context

Lung'Ie [lungíɛ], also known as Principense, is a Portuguese-lexifier creole language spoken in Príncipe Island by less than 200 people when considering speakers with different levels of proficiency (AGOSTINHO, 2015; AGOSTINHO; HYMAN, 2021), part of the Democratic Republic of São Tomé and Príncipe (hereinafter STP), a multilingual country located in the Gulf of Guinea. Its origin is linked to the Portuguese exploration and the kidnapping and confinement of African populations from the 15th to the 19th century (AGOSTINHO; HYMAN, 2021).

According to Agostinho et al. (2016), based on the typology by Crystal (2000), Lung'Ie can be categorized as an endangered language due to several factors: (1) lack of children learning it as a first language; (2) the fact that the community in social circumstances does not primarily use it; and (3) the impact of other languages, especially Portuguese (the official language, most commonly used among all social groups) and Kabuverdianu, spoken by many descendants of workers that arrived in the Príncipe Island by the end of the 19th century and the beginning of the 20th century. Portuguese has been the country's official language since 1975 and is the mother tongue of most of its inhabitants. According to the 2012 Census (INE, 2012), 98.4% of the STP population speaks Portuguese; 36.2% speak Santome; 1% speak Lung'Ie; 6.6% speak Angolar, and 8.5% speak Kabuverdianu. The Census does not provide data regarding bilingualism and multilingualism, nor does it define whether such languages are spoken as L1 or L2 (ARAUJO; AGOSTINHO, 2010). At the same time, linguists have considered the number of Lung'Ie speakers portrayed in the Census to be super estimated (cf. MAURER, 2009, AGOSTINHO, 2015; AGOSTINHO, 2016b, among others).

Four autochthonous, genetically related creole languages are spoken in the Gulf of Guinea: Santome (also known as Forro or Lungwa Santome), Angolar, Lung'Ie, and Fa d'Ambô (SCHUCHARDT, 1889; GÜNTHER, 1973; FERRAZ, 1979; MAURER, 2009; HAGEMEIJER, 2009). The first three are currently spoken in STP and have the status of national languages; the latter is spoken in the island of Annobón and Bioko in Equatorial Guinea. The four languages, although related, are currently mutually unintelligible.

The islands in the Gulf of Guinea were uninhabited until the arrival of the Portuguese at the end of the 15th century. According to Cardoso (2007), the Islands of São Tomé and Príncipe would have been discovered by Portuguese explorers João de Santarém and Pedro Escobar, who arrived in São Tomé on December 21, 1470, and in Príncipe on January 17, 1471.

Figure1: The Gulf of Guinea

Source: http://mapmaker.nationalgeographic.org/

São Tomé was populated by the Portuguese, other European people, and enslaved people in obedience to an order from the Portuguese Royalty. It became a permanent settlement in 1493 (CARDOSO, 2007). The Island of São Tomé was the first to be populated between 1480 and 1493. The settlement on Príncipe Island began in 1500 (MAURER, 2009), probably by people from São Tomé. Portuguese settlers and enslaved people from São Tomé were the first to settle in the islands of Príncipe and Annobón in 1500 and 1503, respectively.

According to Hagemeijer (2009), São Tomé was populated by African people who spoke multiple languages, mainly from the regions of the Niger Delta (where Edoid languages are spoken) and Congo and Angola (where Bantu languages are spoken). A group of enslaved people from the Niger Delta – who spoke Edoid languages – was subsequently transported to Príncipe Island. At the same time, the island received a small number of prisoners from a Bantu region, as opposed to São Tomé (HAGEMEIJER, 1999). In this multilingual scene, the contact between settlers and enslaved people gave rise to a pidgin language (ARAUJO et al., 2013) that, once expanded, originated the Proto-Creole of the Gulf of Guinea (PGG) (cf. FERRAZ, 1979; HAGEMEIJER, 2011; BANDEIRA, 2017). After the creation of PGG, the speakers were geographically isolated (cf. BANDEIRA, 2017). In this sense, it is believed that the enslaved people taken from São Tomé to Príncipe already spoke PGG.

2. The use of the *Ka* Particle in Lung'Ie

Semantic contexts associated with the hypothetical context are marked as irrealis in languages distinguishing between realis and irrealis. These are potential events whose occurrence depends upon certain circumstances (conditional), including counterfactuals, as well as modal and imperative constructions (ELLIOTT, 2000, pp. 69-70). Overall, the literature on realis/irrealis showcases counterfactual and conditional constructions as typically irrealis, while past tense constructions, as well as those that describe an event that happens at the time of enunciation, are typically realis. Modal, imperative, interrogative, aspectual, and future tense constructions, in their turn, can be marked as irrealis or realis depending on grammatical factors – such as being a finite sentence or not –, diachronic factors, or even cultural factors that may interfere with the speaker's expectations concerning the realization of the event (MITHUN, 1995, 1999; ELLIOT, 2000).

The analysis we will present throughout this section is based on the available literature on Lung'Ie (GÜNTHER, 1973; MAURER, 1997, 2009; AGOSTINHO, 2015, 2016b, AGOSTINHO; ARAUJO, 2021; AGOSTINHO; RECH, 2023) and on original fieldwork data collected in Príncipe Island (AGOSTINHO, 2009, 2016a, 2019; AGOSTINHO; RECH, 2016). All glosses are from the original authors unless otherwise noted.

We verified the occurrence of 37 counterfactual and conditional contexts in our corpus[7]. The *ka* particle was used in all of them. Such data corroborate the hypothesis investigated in this research, namely, that the *ka* particle functions as a marker for irrealis (AGOSTINHO; RECH, 2023) since there is consensus in the literature that counterfactual and conditional contexts are marked as irrealis across languages (MITHUN, 1995, 1999; ELLIOTT, 2000).

2.1. Prototypical Irrealis Contexts

Considering that counterfactual and conditional constructions are more closely related to thoughts than reality, it is expected such contexts to be prototypically irrealis in languages in general. For this reason, we shall begin our analysis of irrealis marking in Lung'Ie in these contexts. The following examples were transcribed from Maurer (2009) and illustrate, respectively, counterfactual and conditional contexts:

(1)	**Xi**	**ôzê**	**n**	**ka**	**tava**	**tê**	**dyô,**	**n**	**ka**	**tava**	**kopa**
	if	today	1SG	COUNT	MOD	have	money	1SG	MOD	PST	buy
	kaxi	**ũa.**									
	house	one									

'If today I had had money, I would have bought a house.' (MAURER, 2009, p. 89)

[7] Initially, we searched for *xi* 'if', which is present in these constructions; in a second moment, we considered information from the context to verify whether, in fact, they were hypothetical, conditional, and counterfactual constructions.

(2)	**Xi**	**ê**	**sa**	**ladran,**	**n**	**ka**	**fala**	**kôli**	**fa.**
	if	3SG	COP	thief	1SG	FUT	speak	with.3SG	NEG

'If he is a thief, I won't speak to him.' (MAURER, 2009, p. 99)

In (1), the first instance of *ka* is associated with counterfactuality marking; the second is associated with modality marking. In (2), *ka* is identified as a future tense marker. Example (1) corresponds to a counterfactual context since it describes an event that did not happen (the house purchase), but that could have happened should a condition have been met (having money today). According to Maurer (2009), in (1) the first instance of the *ka* corresponds to a counterfactuality marker, while the second instance corresponds to a modal item. In our view, both *ka* mark this counterfactual context as irrealis in Lung'Ie, the same way counterfactual contexts are marked as irrealis in languages that use markers to demonstrate opposition between realis/irrealis.

Example (2) corresponds to a conditional context equally marked as irrealis in languages that carry such marking. In Maurer's gloss, *ka* is identified as a future tense marker. Although the sentence with *ka* describes a future event concerning the situation described in the previous phrase, we suppose that the conditional context justifies the employment of *ka*, indicating there is no certainty about the 'I will not speak to him' event since it is conditioned to the situation described in the previous phrase ('if he is a thief'), which could be true or not. The realization of the event described in the second phrase depends on the circumstance described in the first phrase. It is a probable – but not certain – event, hence the use of *ka* to indicate the irrealis mood.

Example (3) was transcribed from Maurer (2009) and illustrates a hypothetical situation, thus also being characterized as a prototypical irrealis context:

(3)	**Ningê**	**ki**	**ka**	**panha**	**ufya**	**sê**	**ka**	**daru**	**na**
	Person	REL	GENER	take	leaf	DEM	GENER	rub	LOC
	uman	**zunta**	**awa fyô**	**pasa**	**na**	**uwê,**	**ê**	**ka**	**vê**	**maxi**
	hand	add	water cold	pass	LOC	eye	3SG	FUT	see	more
	dêkê	**ki**	**pekadô**	**tava**	**vê.**					
	than	REL	person	was	see					

'The person who takes this leaf, rubs it in its hands, adds cold water to it and puts it on his eyes will see better than how he used to see.' (MAURER, 2009, p. 55)

According to Maurer, the first two instances of *ka* in (3) are associated with a generic reference, while the third indicates the future tense. The context described is hypothetical since it is not a narrative of a real situation. The speaker hypothesizes that an unspecified person (which may have led to associating *ka* with a generic reference) may perform a series of actions (take the leaf, rub it in

its hands, add cold water, put it in the eyes) to improve vision. There are three instances of *ka* in this construction, two of which are associated with hypothetical events – irrealis context – that shall be realized by people who wish to improve their vision. The third relates to a future situation, conditioned to realizing the sequence of events listed in the context. Should *ka* function as an irrealis mood marker, it would be expected that the event were associated with *having better vision than before* situation (*ê ka vê maxi dêkê ki pekadô tava vê*), as it indeed happens. The state resulting from the hypothetical events will also be a hypothetical situation. Henceforth, we assume that the three *ka* instances in this context play one single role: marking contexts as Irrealis (cf. AGOSTINHO; RECH, 2023).

2.2. Prototypical Realis Contexts

According to Mithun (1995), prototypical realis contexts describe situations that occurred in the past or while the enunciation takes place. In (4) and (5), we shall transcribe sentences in Lung'Ie that illustrate both these contexts, respectively:

(4) **Êli** **ki** **migu** **me** **baya** **ontxi.**
3SG.DES CONJ friend 1SG.POSS dance yesterday.
'He and my friend danced yesterday.' (AGOSTINHO, 2015, p. 178)

(5) **A: Sabiina,** **kwa** **txi** **sa** **fêzê** **wo** **sê** **a?**
Sabrina thing 2SG COP do moment this INT
'Sabrina, what are you doing right now?'
B: Ami **a?** **N** **sa** **kuxi.**
1SG INT 1SG COP cook
'Me? I am cooking.'
A: Kwa **txi** **sa** **kuxi** **a?**
thing 2SG COP cook INT
'What are you cooking?'
B: N **sa** **kuxi** **kumê** **Baji**.
1SG COP cook food Brazil.
'I am cooking Brazilian food.' (AGOSTINHO, 2015, p. 203, our gloss)

The event described in (4) takes place before the moment of the speech, as indicated by the adverb *ontxi* (yesterday). Lung'Ie does not have a morphological marker to indicate the past tense with dynamic verbs, so bare dynamic verbs always indicate the past (cf. MAURER, 2009). The *ka* marker is absent from this example and other past tense cases in our data. The absence of *ka* in these

constructions is expected, as past events occurred and hence are marked as realis (AGOSTINHO; RECH, 2023). In turn, the sentences in example (5) refer to the time of the enunciation. Speaker A asks Speaker B what she is doing at the time of the enunciation, which is indexed with the adverb *wo sê* (right now). It is worth noting that the phrases in (4)–(5) are built upon constructions using copula *sa* followed by a dynamic verb, and *ka* cannot be used, which corroborates our hypothesis that this particle indicates the irrealis mood and thus is not present in constructions that describe events that are past or happening occurring concomitantly to the enunciation.

2.3. Contexts Marked as Either Realis or Irrealis

From this point on, we shall consider some contexts that can be marked as realis or irrealis (future tense, negation, and habitual aspect) across languages or situations where the same context may be marked as realis or irrealis.

2.3.1. Future Tense

According to our data, Lung'Ie marks future tense contexts in affirmative sentences as Irrealis (cf. AGOSTINHO; RECH, 2022). The following examples illustrate this use:

(6) **No ka vê amanhan ô!**
1PL IRR see tomorrow ENF
'We will see each other tomorrow.' (AGOSTINHO, 2015, p. 288, our gloss)

(7) **Amanhan n sa ke posan.**
Tomorrow 1SG PROG IPFV.go town
'Tomorrow I will go to town.' (MAURER, 2009, p. 82, ex. 477)

In (6) and (7), *ka* is employed in future constructions, as the adverb *amanhan* 'tomorrow' demonstrates. This context has been associated with the irrealis mode in many languages, which explains the *ka* use in Lung'Ie. In (7), *ke* results from the agglutination of *ka we* (*ka* + *we* 'go') (cf. MAURER, 2009, p. 69). Our data suggest that the combination *sa ke* marks the near future.[8] De Haan (2006, p. 41) holds that it "can be argued that future is a prototypical irrealis category because it refers to events that have not yet happened and are therefore unreal. [...] However, in others [languages] it is treated as a realis category". Thus, Lung'Ie would be another example of a language, such as Amele and Muyuw (cf. DE HAAN, 2006, p. 41), in which the future is marked as irrealis. In the second author's data (AGOSTINHO, 2009, 2016a, 2019, among others), future tense affirmative constructions (see section 2.3.2) always have *ka* or another form of it, such as (*sa*) *ke* or *keka* (agglutination of *sa* + *ka* + *vika* 'come' (cf. MAURER, 2009, p. 81)).

[8] However, this issue requires further investigation.

2.3.2. Negation

According to the data we analyzed, negative constructions, marked by the realization of the morpheme *fa* sentence-finally, correspond to a realis context in Lung'Ie. The examples below display two examples of negation:

(8) **N** **mêsê** **urumu** **fô.**[9]
1SG.SUJ want safu NEG.ENF
'I don't want *safu*, no!' (AGOSTINHO, 2015, p. 152)

(9) **N** **maxi** **sêbê** **baya** **uvungu** **afinaku** **fa.**
1SG.SUJ still know dance dance African NEG
'I still don't know how to dance African dances.' (AGOSTINHO, 2015, p. 170)

In the examples, the speakers are very clear about their refusal in (8) and their ineptitude for African dances in (9). These constructions seem to indicate a degree of certainty or the speaker's commitment to the spoken content. Hence, such constructions are expected to be marked as realis, which seems to correspond to the unmarked form in Lung'Ie.

Maurer's analysis uses *sa* "instead of *ka* in negated sentences" (MAURER, 2009, p. 83). To argue his viewpoint, the author uses examples (10) and (11), transcribed below:

(10) ***Amanhan** **n** **ka** **kume** **fa.**
tomorrow 1SG FUT eat NEG
'Tomorrow, I won't eat.' (MAURER, 2009, p. 83, ex. 484)

(11) **Amanhan** **n** **sa** **kume** **fa.**
tomorrow 1SG FUT eat NEG
'idem' (MAURER, 2009, p. 83, ex. 485)

According to Maurer, the ungrammatical construction observed in (10) results from employing the *ka* particle in negative contexts. The author does not directly state that *sa* is the negative variant of *ka*, supposedly constituting allomorphism. Since *sa* is associated with marking future tense (FUT) or habitual aspects (HAB) in his glosses of negative sentences, which are the same functions to which he assigns to *ka* in affirmative sentences, the realization of *ka* as *sa* in these cases causes two analytical problems: (i) there is no phonetic-phonological motivation for such variation, and (ii) it

[9] The formula *fô* is an agglutination of *fa* plus interjection *ô* (cf. MAURER, 2009; AGOSTINHO, 2015).

would not be expected that an allomorphic variation would result in a phonetic form identical to that of another language's functional morpheme. According to the analysis, the morpheme *sa* corresponds to the non-past copula, hence being found in negative constructions (see footnote 12). The future negative construction combines the copula *sa* with the lexical verb *kume* (eat), meaning something like 'I won't be eating tomorrow.' Languages such as Brazilian Portuguese (BP) and English can also express habitual aspect and/or future tense by employing copulas: 'Eu estou bebendo' ('I'm drinking') (habitual aspect) in BP or 'I'm drinking tomorrow' (future) in English.

It is worth noting, however, that *ka* appears in negative constructions when they correspond to conditional (counterfactual) or hypothetical contexts, characterized as prototypical irrealis contexts inter-linguistically. In these cases, there seems to be a hierarchy of contexts when marking a construction as irrealis, in which, prototypical irrealis contexts prevail over the negation context. Let us proceed to analyze the following examples:

(12)	**Xi**	**n**	**ka**	**sa**	**kume,**	**n**	**ka**	**rêgê**	**wo**	**sê**	**fa.**
	if	1SG	IRR	COP	eat	1SG	IRR	get.up	moment	DEM	NEG

'If I was eating, I wouldn't get up now.' (AGOSTINHO 2009, our gloss)

(13)	**Txi**	**ka**	**lega**	**n**	**fa,**	**n**	**ka**	**gita**	**ô!**
	2SG	MOD	set.free	1SG	NEG	1SG	FUT	shout	VAL

'If you don't let go of me, I'll shout!' (MAURER 2009, p. 84)

In (12) and (13), *ka* is used in negative constructions, as indicated by the presence of the *fa* morpheme. While analyzing our data, we observed that the use of *ka* in negative contexts is restricted to prototypical irrealis contexts, such as hypothetical constructions, like (12), or conditional (counterfactual) constructions, like in (13).

2.3.3. Habitual Aspect

De Haan (2012, p. 121) observes that "[h]abitual aspect is perhaps a strange category to discuss in a section about realis and irrealis. As this category denotes that an action is or was done habitually, there would seem to be little doubt that such actions are real and any marker of habitual aspect would fall into the realis camp. Nevertheless, there are languages in which habitual aspect is marked either identical to other irrealis categories or has a separate irrealis morpheme attached". Along the same lines, the habitual aspect seems to correspond more to irrealis contexts when compared to the perfective aspect, "... since habitual aspect describes an event type that is instantiated from time to time by actual events" (PAYNE, 2007, p. 245). Thus, the occurrence of *ka* in habitual constructions is not necessarily a counter-argument for the analysis we are proposing for Lung'Ie since habitual

constructions could correspond to irrealis contexts in this language (cf. AGOSTINHO; RECH, 2023). Examples (14) to (16) correspond to aspectual constructions in which *ka* is employed:

(14) **Ôtô ka fala lunge Ie, ôtô ka fala putugêzê.**
Other HAB speak language island other HAB speak Portuguese
'Some speak Principense, some speak Portuguese.' (MAURER, p. 40, ex. 159)

(15) **Ora ki a ka fala bê na salasa, a ka**
time REL PRO.INDF IRR speak hello in friends PRO.INDF IRR
kudi 'malimentê', ora ki a ka fala bê pa ningê
answer so-so time REL PRO.INDF IRR speak hello to person
tamwin a ka kudi 'bensa di dêsu'.[10]
adult PRO.INDF IRR answer bless of god
'When people greet between friends, they answer "so-so", when they greet grown-ups, they answer "God bless".' (AGOSTINHO, 2015, p. 146, our gloss)

(16) **Dêxa bayu sê ê axi: minu mye ka bixi seya ki kimoni**
dêxa dance this 3SG like.so child woman IRR wear skirt with *kimoni*
ô buluza, ulensu baanku ô kô ôtô i minu
or blouse handkerchief white or color other and child
omi ka bixi kalisan ope, palito ki kazaku.[11]
man IRR wear pants foot jacket with coat
'*Dêxa*, this dance, it's like this: young women wear a skirt and a *kimoni* or top, a white or another color handkerchief and young men wear long pants, a suit and a coat.' (AGOSTINHO, 2015, p. 173, our gloss)

In examples (14)-(16), we can see that no specific event is described; instead, we see the description of habitual events, in which one cannot identify a specific time or specific participants. In (14) and (15), the statements refer to the local reality, used as a general assumption based on the observation of real events. In (16), the situation is similar: the speaker describes the clothes worn by young men and women while dancing *dêxa*. For examples (14)-(16) and similar ones in our data, we assume that the *ka* particle is used to mark this context (habitual aspect) as irrealis in Lung'Ie.

[10] Example without a gloss in the original.

[11] Example without a gloss in the original.

We built our hypothesis upon Elliot's (2000, p. 79) statement regarding marking habitual/customary events as irrealis: "One possible explanation for this is that the non-specific nature of the events being discussed in this particular usage leads to the use of irrealis, since no particular realized occasion is being discussed."

It is important to mention that, in Lung'Ie, aspectual constructions can be found with or without the use of *ka*, as seen in the following examples:

(17) **Na wêtu ora n tava kume za.**
LOC eight hour 1SG PST.COMPL eat already
'At eight o'clock, I was already eating.' (MAURER, 2009, p. 86)

(18) **N goxta di palapala montxi fa, maji n ka kume li.**
1SG.SUJ like PREP palapala much NEG but 1SG.SUJ HAB eat 2SG.OBJ
'I don't like *palapala* very much, but I usually eat it.' (AGOSTINHO, 2015, p. 218)

The example in (17) describes a durative aspect (*n tava kume za*) with the use of a past copula (*tava*), followed by the dynamic verb (*kume*). *Ka* is not used in this sentence. According to our analysis, *ka* is absent in this construction because the example describes a particular event that occurred in the past (*Na wêtu ora*), hence marked realis. Nonetheless, the *ka* particle is found in (18). It is noteworthy that while the copula *tava* followed by the verb *kume* describes a specific event as durative in (17), example (18) describes an event that is instantiated occasionally, characterizing it as a habitual aspect. Thus, according to our hypothesis, the use of *ka* in (18) was already expected since no specific event is described in the example – which, according to Elliot (2000), marks it as irrealis.

Examples (19) and (20) below, transcribed from Maurer (2009), show *ka* in sentences associated with the habitual aspect, but unlike (14) to (16) and (18), the events described in (19) and (20) are no longer ongoing in the time of speech.

(19) **Dya tudu pe n tava ka rêgê na xink'ora di pemyan.**
day all IDEO 1SG PST HAB get.up at five.o'clock of morning
'I used to get up at five o'clock in the morning every day.' (MAURER, 2009, p. 87)

(20) **Ine na tava ka kume kani pôkô dyêxi tudu pe.**
3PL VAL PST HAB eat meat pig day. DEM IDEO
'They really used to eat pork every day.' (MAURER, 2009, p. 67)

The use of the past copula (*tava)* locates the habitual events described in these examples in a time before the moment of the utterance. These examples are important because they reveal that *ka* is not incompatible with the past tense. If the event is described as habitual, *ka* appears in past constructions, as seen in (19) and (20). On the other hand, sentences in which the event is marked with a progressive aspect in the past, as in (21), below, do not appear with *ka:*

(21)	**Ine**	**tava**	**sa**	**kirya**	**minu**	**sê**	**pobêmentê.**
	3SG	PST	COP	raise	child	3PL.POSS	poor=ly

'They were raising their children poorly.' (AGOSTINHO, 2016a, our gloss)

In (21), using the copula *sa* followed by a dynamic verb marks the past event as progressive: *tava sa kirya,* similar to the use of 'were -ing' in English. If the *ka* particle is, in fact, an irrealis marker, as we are proposing, its absence in this context is expected since the sentence describes just one progressive event in the past.

In example (22) below, there is also the occurrence of the functional items *tava* and *ka,* but, in this case, there is no description of a habitual event in the past. It is a prototypical irrealis context, as it describes a counterfactual situation whose event, which could have been ongoing at the time of the utterance, has not been initiated.

(22)	**Xi**	**non**	**ka**	**tava**	**sa**	**xivi**	**wosê,**	**non**	**ka**	**tava**
	if	1PL	COUNT	PST	PROG	work	now	1PL	COUNT	MOD
	tê	**dyô.**								
	have	money								

'If we were working now, we would have money.' (MAURER, 2009, p. 90)

Example (22) generates a progressive interpretation, but not a habitual one. The presence of the adverb *wosê* 'now' determines the time (simultaneous with the moment of the utterance) in which the event *sa xivi* 'to be working' should be in progress for the state *tê dyô* 'to have money' to be true. In (22), *tava* does not correspond to a past tense mark, unlike examples (19) and (20). Evidence thereof is that this marking would be incompatible with the one indicated by the adverb *wosê.*

In the following example, also transcribed from Maurer (2009), the sequence *ka>tava* appears in a counterfactual context with the occurrence of an adverb indicating future tense:

(23)	**Xi**	**amanhan**	**non**	**ka**	**tava**	**xivi,**	**non**	**ka**	**tava**
	if	tomorrow	1PL	COUNT	MOD	work	1PL	COUNT	MOD
	tê	**dyô,**	**maji**	**xivisu**	**tê**	**fa.**			
	have	money	but	work	there.be	NEG			

'If tomorrow we worked, we would get money, but there is no work.' (MAURER, 2009, p. 89)

In (23), similarly to (22), *tava* occurs after *ka* in a sentence whose event is not in the past, as evidenced by the adverb *amanhan* 'tomorrow'. The counterfactual reading is generated because the event *xivi* 'worked' is not likely to happen, not because it is located in the past, as is common in counterfactual constructions, but because there is no work, information given in the sentence *maji xivisu tê fa* 'but there is no work', which integrates the construction.

Concerning example (23), Maurer notes that "(t)he fact that counterfactuality can be expressed by *ka* alone and that *tava* co-occurs with time adverbs like *ôzê* 'today' and *amanhan* 'tomorrow' shows that in this context, the past marker *tava* does not exert a temporal function; it metaphorically reinforces the counterfactuality already expressed by *ka*" (2009, p. 90).

The following example illustrates a case in which the sequence *ka tava* occurs in a counterfactual context in the past, as evidenced by the presence of the adverb *ontxi* 'yesterday':

(24)	**Ontxi,**	**ora**	**txi**	**xiga,**	**xi**	**no**	**ka**	**tava**	**sa**	**xivi,**
	yesterday	hour	2SG	arrive	if	1PL	COUNT	MOD	PROG	work
	no	**ka**	**tava**	**vê**	**txi**	**fa.**				
	1PL	MOD	PST	see	2SG	NEG				

'Yesterday, when you came, if we had been working, we wouldn't have seen you.' (MAURER, 2009, p. 90)

In the sentence ... *xi no ka tava sa xivi* 'if we had been working', *ka* precedes the functional item *tava.* Then, the form of the progressive (*sa xivi*) occurs*,* resulting in the following ordering: *ka* > *tava* > *sa*> dynamic verb. Example (24) also corresponds to a prototypical irrealis context: the counterfactual. As we have argued, *ka* plays a single role in Lung'Ie: marking a context as irrealis. *Tava*, in turn, does not seem to correspond to a past tense marker when it appears after *ka*, as shown in examples (1) and (22) to (24)[12] and already noted by Maurer (2009, p. 90, quoted above). We are

[12] It is worth noticing that not all past events are marked with *tava* in Lung'Ie. For instance, sentence (i) describes the event in the past, which does not exhibit any morphological mark. If the non-past copula (*sa*) is used, as in (ii), the event will be described in the progressive aspect, constructed in Lung'Ie by the sequence *sa*>dynamic verb. The insertion of *tava* before *sa*, as illustrated in (iii), describes a progressive event in the past; therefore, in this construction, *tava* corresponds to a past tense marker, differentiating (iii) from (ii). Finally, in (iv), the occurrence of *ka* marks the sentence as irrealis, which can be interpreted as future tense or habitual aspect, depending on the context.

assuming that the interpretation associated with *tava* – as past tense or not – is directly related to this ordering and is a consequence of the position of the functional head to which *tava* corresponds in the sentence structure and the way this head relates to the other functional categories that integrate the Lung'Ie TMA system. In (25), we present part of the hierarchy of functional heads proposed by Cinque (1999) for the categories of mood, modality, tense, and aspect in natural languages:

(25) Functional Projections Hierarchy:

Moodspeech act > Moodevaluative > Moodevidential >Modepistemic>**T(Past)** >T(Future)> **Mood(irrealis)** > Modnecessity > Modpossibility > **Asphabitual** > Asprepetitive(I) > Aspfrequentative(I) > Modvolitional > … > Aspperfect > Aspretrospective > Aspproximative> AspPdurative >**AspPprogressive** > AspPprospective > … > >ModPobligation> ModPability > AspPfrustrative/success >ModPpermission>AspPconative > AspPcompletive (I) > VoiceP...

(CINQUE, 1999, p. 106, our emphasis)

On the other hand, the sequence *tava>ka,* in (19) and (20) above, generates an interpretation of the habitual past, with the item *tava* being interpreted in the category T(Past); and *ka,* in the Mood(irrealis) category. Note that the interpretation associated with the sequence *tava>ka* is in accordance with the ordering of the functional heads to which they correspond in the hierarchy proposed by Cinque: T(Past) >... > Mood(irrealis). If *ka,* as we have been arguing, corresponds to irrealis marking, then it was already expected that *tava* in the sequence *ka > tava* would not generate a past tense interpretation. Thus, Lung'Ie seems to provide further evidence for the ordering proposed by Cinque.

Finally, in (22) and (24), the functional items *ka* and *sa* seem to follow the ordering proposed by Cinque's hierarchy: Mood(irrealis) > … >... >... > AspPprogressive. Until this moment of the research, we do not know the precise interpretation position of the functional item *tava* when it appears after *ka.* Still, examples (22) and (23) show clear evidence that it does not correspond to the T(past) head. Considering the sequence *ka>tava>sa,* it is expected that *tava* corresponds to a head located in the hierarchy in (25), between Mood(irrealis) and AspProgressive.

(i) **N** **bêbê.**
1SG drink
'I drank'.

(ii) **N** **sa** **bêbê.**
1SG COP drink
'I am drinking'.

(iii) **N** **tava** **sa** **bêbê.**
1SG PST COP eat
'I was drinking'.

(iv) **N** **ka** **bêbê.**
1SG IRR eat
'I will drink' or 'I drink'.

2.4. Considerations Regarding the Irrealis Mood in Lung'Ie

Throughout this section, we showed the marking of irrealis mood in Lung'Ie based on the analysis of prototypically irrealis contexts: conditional, counterfactual, and hypothetical (section 2.1); prototypical realis contexts: events described in the past or simultaneously to the time of speech (section 2.2); and certain contexts marked as realis and/or irrealis across languages: future tense, negation, and habitual aspect (section 2.3).

Considering the economy principle that rules natural languages, it is unlikely that one sole form is associated with such distinct functions in the same language, as argued by Maurer (2009), who states that, in Lung'Ie, the *ka* particle would be used as a marker for counterfactuality, future tense, habitual past, and generic reference, among others. We observed that all contexts where the particle is found describe possible – i. e. not real – situations that exist "only as a conceptual idea, a thought, or a hypothetical notion" (ELLIOTT, 2000, p. 56). Hence, we argue that *ka* seems to function as an irrealis mood marker in Lung'Ie, as proposed by Agostinho and Rech 2023.

In this sense, Lung'Ie would not be the only creole language to equate notions of future and habitual with irrealis. According to Winford (2018, p. 203), "[c]reolists have generally interpreted irrealis (…) as a cover term for future, conditional, and subjunctive meanings. Creoles generally employ future markers in conditional clauses to express hypothetical or counterfactual notions associated with the conditional or subjunctive mood in other languages". Taylor (1971, 1977) cites, for example, creole languages in which the habitual aspect is indicated by the same future or irrealis marker, such as Kabuverdianu, and Negerhollands, in addition to Forro (genetically related to Lung'Ie).

3. Conclusion

Our analysis proposes that *ka* in Lung'Ie functions as an irrealis mood marker, corroborating Agostinho and Rech (2023). This contrasts with the analyses of Günther (1973) and Maurer (2009), who associate *ka* with various functions in the language, such as a marker of counterfactuality, modality, future tense, habitual aspect, imperfective aspect, generic reference, among others. The primary motivation for this paper is that *ka* always occurs in contexts described as prototypical irrealis mood, namely hypothetical and conditional/counterfactual constructions (cf. MITHUN, 1995, 1999; ELLIOTT, 2000).

We also looked into the use of *ka* in contexts that the linguistic literature describes as prototypical realis (constructions describing past events and those concomitant to the enunciation); and contexts that can be marked realis and/or irrealis in different languages (future, negation, and habitual aspect). Our data indicate that the *ka* particle is used in all prototypical irrealis contexts and future and habitual aspect contexts. Nevertheless, it is not used in past constructions nor in those that describe events concomitantly to the enunciation, which correspond to prototypical realis contexts. We also verified that *ka* is not used in negative contexts, which suggests it is also a realis context in Lung'Ie. Ultimately, we found that there is a hierarchy in mood marking in Lung'Ie, in which prototypical irrealis contexts prevail over others (see examples (12), (13), (22), and (24).

The similarities in time, mood, and aspects of creole languages across different lexifiers have been highlighted by several studies (SINGLER, 1990). For example, Muysken (1981) observes that creole languages tend to have three markers: past tense, irrealis mood, and durative aspect, whereas Bickerton argues that this is the prototypical creole TMA system (1981, p. 58; 1984). This structure is aligned with the analysis presented in section 2: the markers in Lung'Ie would be Ø/*tava* for past, *ka* for irrealis mood, and the non-past copula *sa* for durative aspect.

List of abbreviations

1	first person	IPFV	imperfective aspect
2	second person	IRR	irrealis
3	third person	LOC	locative
COMPL	complementizer	MOD	modal
CONJ	conjunction	NEG (1,2)	negative
COP	copula	OBJ	object
COUNT	counterfactual	PST	past
DEM	demonstrative	PL	plural
DES	desiderative	POSS	possessive
ENF	emphatic	PREP	preposition
FUT	future	PRO	pronoun
GENER	generic reference	PROG	progressive
HAB	habitual	REL	relative
IDEO	ideophone	SG	singular
INDF	indefinite	SUJ	subject
INT	interrogative	VAL	validator

Acknowledgements

We are very grateful to Gabriel Antunes de Araujo for his insightful comments in a preliminary version of this article. We would also like to thank Tainá Monique Schaffer for helping us with the data organization. Research by the second author was supported by the Brazilian National Council for Scientific and Technological Development (CNPq), grant 200519/2019-0.

References

AGOSTINHO, Ana Lívia. *Fieldnotes and data collection – São Tomé and Príncipe*. Manuscript, 2009.

AGOSTINHO, Ana Lívia. *Fonologia e método pedagógico do lung'Ie*. (Ph.D. Dissertation). São Paulo: University of São Paulo, 2015.

AGOSTINHO, Ana Lívia. *Fieldnotes and data collection: São Tomé and Príncipe*. Manuscript, 2016a.

AGOSTINHO, Ana Lívia. Fonologia do lung'Ie. München: Lincom, 2016b.

AGOSTINHO, Ana Lívia. *Fieldnotes and data collection: São Tomé and Príncipe*. Manuscript, 2019.

AGOSTINHO, Ana Lívia; ARAUJO, Gabriel Antunes de. *Lung'Ie, lunge no: Método para aprender lung'Ie*. São Paulo: FFLCH/USP, 2021.

AGOSTINHO, Ana Lívia; BANDEIRA, Manuele; ARAUJO, Gabriel Antunes de. O lung'Ie na educação escolar de São Tomé e Príncipe. *Trabalhos em Linguística Aplicada,* v. 55, n. 3, pp. 591-618, 2016.

AGOSTINHO, Ana Lívia; HYMAN, Larry M. Word Prosody in Lung'Ie: One System or Two? *Probus*, v. 33, n. 1, pp. 57-93, 2021.

AGOSTINHO, Ana Lívia; RECH, Núbia. *Fieldnotes and data collection – São Tomé and Príncipe*. Manuscript, 2016.

AGOSTINHO, Ana Lívia; RECH, Núbia Ferreira. Lessons from the field: Irrealis mood in Lung´Ie. In*:* KLOK, J. Vander; RECH, N. F.; GUESSER, S. (eds.), *Modality in Underdescribed Languages: Methods and Insights.* Berlin, Boston: De Gruyter Mouton, 2023, pp. 295-336.

ARAUJO, Gabriel Antunes de et al. Fa d'Ambô: língua crioula de Ano Bom. *Cadernos de Estudos Linguísticos*, v. 55, n. 2, pp. 25-44, 2013.

ARAUJO, Gabriel Antunes de; AGOSTINHO, Ana Lívia. Padronização das línguas nacionais de São Tomé e Príncipe. *Língua e Instrumentos Linguísticos, v.* 26, pp. 49-81, 2010.

BANDEIRA, Manuele. *Reconstrução fonológica e lexical do protocrioulo do Golfo da Guiné*. (Ph.D. Dissertation). Universidade de São Paulo, São Paulo, 2017.

BICKERTON, Dereck. *Roots of Language*. Ann Arbor: Karoma Press, 1981.

CARDOSO, Manuela. *Cabo Verde e S. Tomé e Príncipe.* Porto: IPAD, 2007.

CINQUE, Guglielmo. *Adverbs and Functional Heads*: a cross-linguistic perspective. New York: Editora Oxford University Press, 1999.

CRYSTAL, David. *Language Death.* Cambridge: Cambridge University Press, 2000.

DE HAAN, Ferdinand. Typological Approaches to Modality. *In*: *The Expression of Modality*, FRAWLEY, William (ed.). Berlin: Mouton de Gruyter, 2006, pp. 27-69.

DE HAAN, Ferdinand. 2012. Irrealis: Fact or Fiction? *Language Sciences,* v. 34, n. 2, pp. 107-30.

ELLIOTT, Jennifer. Realis and irrealis: forms of grammaticalisation of reality. *Linguistic Typology*, v. 5, pp. 55-90, 2000.

FERRAZ, Luiz Ivens. *The creole of São Tomé*. Johannesburg: Witwatersrand University Press, 1979.

GÜNTHER, Wilfried. *Das portugiesische Kreolisch der Ilha do Príncipe.* Marburg an der Lahn: Im Selbstverlag, 1973.

HAGEMEIJER, T. As ilhas de Babel: a crioulização no Golfo da Guiné. *Revista Camões*, v. 6, pp. 65-71, 1999.

HAGEMEIJER, T. Initial vowel agglutination in the Gulf of Guinea creoles. In*:* ABOH, E. O.; SMITH, N. (eds.), *Complex processes in new languages*. Amsterdam: John Benjamins, 2009.

HAGEMEIJER, Tjerk. The Gulf of Guinea Creoles: genetic and typological relations. *Journal of Pidgin and Creole Languages*, v. 26, n. 1, pp. 111-54, 2011.

HAGEMEIJER, Tjerk. The Gulf of Guinea creoles: genetic and typological relations. In: BHATT, P.; VEENSTRA, T. (eds.), *Creole languages and linguistic typology.* Amsterdam, Philadelphia: John Benjamins. 2013b, pp. 163-206.

HAGEMEIJER, Tjerk; MAURER-CECCHINI, Philippe; ZAMORA SEGORBE, Armando. *A Grammar of Fa d'Ambô*. Berlin: Mouton de Gruyter, 2020.

MAURER, Philippe. Tense-Aspect-Mood in Lung'Ie. In: SPEARS, A. K.; WINFORD, D. (eds.) *The structure and status of pidgins and creoles*. Amsterdam, Philadelphia: John Benjamins, 1997, pp. 415-35.

MAURER, Philippe. *Lung'Ie*. London: Battlebridge Publications, 2009.

MITHUN, Marianne. On the relativity of irrealis*.* In: J. Bybee; S. Fleischman (eds). *Modality in Grammar and Discourse* (pp. 367-88). Amsterdam, Philadelphia: John Benjamins, 1995.

MITHUN, Marianne. *The languages of Native North America*. New York: Cambridge University Press, 1999.

MUYSKEN, Pieter C. Creole Tense/Mood/Aspect systems: The unmarked case? In: MUYSKEN, P.C. (eds), *Generative Studies on Creole Languages*. Dordrecht: Foris, 1981, pp. 181-99.

PAYNE, Thomas Edward. *Describing Morphosyntax*: *A Guide for Field Linguists*. Cambridge: Cambridge University Press, 2007.

PONTÍFICE, J. et al. *Alfabeto Unificado para a Escrita das Línguas Nativas de S. Tomé e Príncipe* (ALUSTP). 2010. Unpublished manuscript.

SCHUCHARDT, Hugo. Beiträge zur Kenntnis des Kreolischen Romanisch IV. *Zum Negerportugiesischen der Ilha do Principe. Zeitschrift für Romanische Philologie,* v. 13, pp. 461-75, 1889.

SINGLER, Jonh Victor. Pidgin and Creole Tense/Mood/Aspect Systems. *Creole Language Library,* v. 6. Amsterdam, Philadelphia: John Benjamins, 1990. Available in: https://doi.org/10.1075/cll.6.

TAYLOR, Douglas R. Grammatical and lexical affinities of creoles. In: Dell Hymes (ed). *Pidiginization and creolization of languages*. Cambridge: Cambridge University Press, 1971, pp. 293-96.

TAYLOR, Douglas R. *Languages of the West Indies*. Baltimore: Johns Hopkins University Press, 1977.

WINFORD, Donald. Creole Tense-Mood-Aspect Systems*. Annual Review of Linguistics*, v. 4, pp. 193-212, 2018.