# EXTERNAL VOWEL SANDHI PROCESSES IN LUNG'IE CONSIDERING STRESS AND TONE 

PROCESSOS DE SÂNDI VOCÁLICO EXTERNO EM LUNG'IE CONSIDERANDO ACENTO E TOM

Ana Lívia Agostinho | Lattes | a.agostinho@ufsc.br<br>Universidade Federal de Santa Catarina


#### Abstract

This paper investigates the contexts in which stress ${ }^{1}$ allows or blocks external vowel sandhi processes in Lung'Ie, a Portuguese-based Creole language spoken in Príncipe Island, São Tomé and Príncipe. By analyzing the relationship between stress and tone in the sandhi processes, this research intends to demonstrate that the tonal pattern of Lung'Ie, such as described by Maurer (2009), does not block the phenomenon, which occurs in certain contexts of stress. We thus hope to further contribute to the understanding of the suprasegmental system of Lung'Ie and of the relationship between stress and tone in this language.


Keywords: Lung’Ie; Sandhi; Stress; Tone; Creole languages.

Resumo: Este trabalho investiga os contextos em que ocorrem processos de sândi vocálico externo em Lung’Ie, uma língua crioula de base portuguesa falada na Príncipe Island, São Tomé e Príncipe, e sua relação com o sistema suprassegmental da língua. Através da análise da relação entre acento e tom nos processos de sândi, esta pesquisa pretende demonstrar que o padrão tonal do Lung'Ie, tal como descrito por Maurer (2009), não bloqueia o fenômeno, que é bloqueado em determinados contextos de acento. A contribuição desta pesquisa envolve a compreensão do sistema suprassegmental do Lung'Ie e da relação das categorias de tom e acento na língua.

Palavras-chave: Lung'Ie; Sândi; Acento; Tom; Línguas crioulas.

[^0]
## 1 INTRODUCTION

The aim of this article is to present contexts of external vowel sandhi in Lung'Ie (ISO code 639-3: PRE) and the relation with its suprasegmental system in order to show that the tonal patterns ${ }^{2}$ do not contribute to the blocking of the sandhi process while stress does. Lung'Ie is a Creole language spoken by around 200 people (AGOSTINHO, 2015) in the Democratic Republic of São Tomé and Príncipe (STP) in the Gulf of Guinea region. There are four autochthonous and genetically related creole languages spoken in the Gulf of Guinea: Santome (or Forro), Angolar, Lung'Ie and Fa d'Ambô (FERRAZ, 1979; GÜNTHER, 1973; HAGEMEIJER, 2009; MAURER, 2009; SCHUCHARDT, 1889). Currently, the first three are spoken in São Tomé and Príncipe and the latter is spoken on the islands of Ano Bom and Bioko in Equatorial Guinea. The four languages, though related, are currently unintelligible to each other.


Figure 1: The Gulf of Guinea
Source: http://mapmaker.nationalgeographic.org/.

The islands of the Gulf of Guinea were uninhabited before the arrival of the Portuguese at the end of the 15 th century. According to Cardoso (2007), the islands of São Tomé and Príncipe were supposedly discovered by Portuguese navigators João de Santarém and Pedro Escobar, who arrived in São Tomé on December 21, 1470 and in

[^1]Príncipe on January 17, 1471. According to Hagemeijer (2009), the settlement of São Tomé encompassed a population of speakers of several African languages, mainly from regions such as the Niger Delta, where languages of the Edo group are spoken, as well as from Congo and Angola, a region home to the Bantu language family (SILVEIRA, 2013). In this multilingual scenario, an emergent language developed from the contact between settlers and enslaved people (ARAUJO et al., 2013, p. 29), which, in turn, gave rise to the Proto-Creole of the Gulf of Guinea (henceforth PGG) (BANDEIRA, 2017; see FERRAZ, 1979; HAGEMEIJER, 2011). Subsequently, a contingent of slaves from the Niger Delta region was transplanted directly to Príncipe Island while Príncipe received a small number of prisoners from Bantu regions, unlike São Tomé (HAGEMEIJER, 1999). After the PGG was formed, the geographical separation of its speakers began (see BANDEIRA, 2017), resulting in the four languages currently spoken: Santome, Angolar, Lung'Ie and Fa d'Ambô. Thus, these languages are the result of the linguistic contact between stress and tonal languages.

## 2 TONE AND STRESS IN LUNG'IE

There is no consensus in the literature regarding classification of the suprasegmental system of Lung'Ie (AGOSTINHO, 2015, 2016; GÜNTHER, 1973; MAURER, 2009; TRAILL; FERRAZ, 1981).

The very distinction between tonal and accentual languages is controversial. In addition, there are languages that are classified as tonal and accentual by the same authors (ODDEN, 1996). According to McCawley (1978), there are languages that are not wholly accentual, and, on the other hand, attested cases of languages that lose or acquire lexical tone (GUSSENHOVEN, 2004). Hymam (1977, p.38) mentions that tone and stress are not mutually exclusive and Hyman (2014a, p.19) mentions that there is evidence of mixed prosodic systems in the literature. However, these systems are scantly studied (MICHAEL, 2011) and continue to represent descriptive and theoretical challenges for the area (HYMAN, 2006).

The theme of the suprasegmental system of Lung'Ie was approached by Günther (1973), Ferraz and Traill (1981), Maurer (2009) and Agostinho (2015). The Lung'Ie stress was mentioned by Maurer (2009) and Günther (1973), but these authors did not delve into the subject. Agostinho (2015) demonstrates the importance of the stress on the phonology of Lung'Ie by examining phonological processes that are only possible if stress is considered, such as deletion, insertion, diphthongization, nasalization and external vowel sandhi.

Günther (1973) states that Lung'Ie is a tonal language with three tones: high (H), low (L, unmarked) and rising (R). According to Maurer (2009), Günther interprets stressed syllables as having a high tone, and unstressed syllables as having a low tone, so there would be no words with two identical adjacent tones, that is, HH or LL. Günther (1973) attributes the origin of the tones to diachronic processes: the high tones would originate from stressed syllables, and the low tones from unstressed syllables of Portuguese. The ascending tone would appear in words of Portuguese origin in which an intervocalic syllable was elided. The author does not make the correlation between increasing dynamic tone and long vowels, such as Ferraz and Traill (1981) and Maurer (2009). Thus, Günther (1973) states that Lung'Ie tones come from the accentual system of Portuguese, that is, the high tone would correspond to the lexical stress, while the low tone would correspond to the unstressed syllables. In Ferraz and Traill's (1981) and Günther's (1973) analysis, therefore, there could only be words with one high or rising tone, and it would not be possible to have words with only low tones.

Ferraz and Traill (1981), based on data from Günther and data collected in their fieldwork in 1969 and 1970, state that Lung'Ie cannot be considered a tonal language. They consider it as a free pitch-accent language. The difference between a tonal language and a pitch-accent language would be that at first it is possible to have, for example, two high tones in one word or one word with only low tones, and the second one requires only one high tone by word or that tones are predictable by the stress of the word. Thus, in pitch-accent languages there would be no HH or LL words in which the tonal pattern was not predictable by the stress (see GORDON, 2014). A pitch-accent language in which the stress is idiosyncratic, i. e., not fixed, is then called a free pitch-accent stress (FERRAZ; TRAILL, 1981). However, the pitch-accent classification has been criticized for not being a prototype, because it cannot be analyzed as an intermediate system between tonal and accentual, and because of its classification (HYMAN, 2009, 2014).

Ferraz and Traill (1981) present four pitches for Lung'Ie: high, low, rising and falling. The authors point out that dynamic tones (rising and falling) only appear in long vowels or vowel sequences. For them, in some cases, this "extra amount" is resolved by being performed as a rearticulated vowel.

According to Ferraz and Traill (1981), this analysis also differs from Günther's (1973) in that the stress in Portuguese does not always correspond to a high tone in their data. In addition, the authors suggest that "stress", that is, intensity, is present, and that it is always possible to determine a prominent syllable in the words, which can receive high,
rising or falling pitches. Only syllables without prominence can receive low pitch. In addition to intensity and high pitch, prominent syllables are longer. Thus, we can observe the distribution of pitch sequences in Ferraz and Traill's (1981: 209) analysis with high (H), low (L), increasing (R) and decreasing (F) tones for CV.CV words:
(1) a. HL, RL, FL, LH, LR, LF
b. *HH, ${ }^{*}$ LL, ${ }^{*}$ RR, ${ }^{*} \mathrm{FF}$
c. ${ }^{*} \mathrm{HR},{ }^{*} \mathrm{HF},{ }^{*} \mathrm{RF},{ }^{*} \mathrm{RH},{ }^{*} \mathrm{FH},{ }^{*} \mathrm{FR}$

In (1a), it is obvious that the possible sequences necessarily need a high or dynamic tone. In (1b), we cannot have two equal high or dynamic tones co-occurring in the same word and at least one loud or dynamic tone is needed in a word, since we cannot have LL. In (1c), we cannot have two loud or dynamic tones co-occurring in the same word. The same happens with words with three syllables, where there should be only one syllable with high or dynamic tone. Ferraz and Traill (1981) also discuss the relation of described tones with Portuguese and substrate languages.

For Maurer (2009), Lung'Ie is a tonal language. The main difference between his analysis and the previous ones is that Lung'Ie has only two tones: high and low, and that dissyllabic words can have all possible tone combinations: HH, HL, LH and LL. The fact that there are words HH and LL in his analysis would demonstrate that this language is tonal and not a pitch-accent language. However, the tonal pattern offered by the author is not totally unpredictable by the stress: if a final stressed word has a pitch H on the stressed syllable, its pattern will be LH; if there is an L on the stressed syllable, the pattern will be LL; in a pre-final stressed word, it would be possible to say that the standard HH and HL is realized phonologically in the stressed syllable and the second tone is realized phonetically in the unstressed syllable.

The same could be said for trisyllabic words: the final stressed words with H or L on the stressed syllable receive LL in the previous syllables, the antepenultimate stressed words always have an HHL pattern, prefinal stressed words receive LHH and LHL, always having an H tone in the stressed syllable. For Maurer (2009), long vowels in stressed syllable of final stressed words receive two tones phonologically attributed to the stressed syllable, which may be LH , LL or HH , and a predictable pitch H is assigned to the second syllable. The monosyllables with a long vowel would receive two tones phonologically, being each attributed to each mora.

There is much disparity in the notation of tone by the three authors, which makes it difficult to analyze its relationship with stress. In addition, Maurer (2009) explains that he recorded the sentences with the target word followed by the $T M A^{3}$ ka particle, which has a variable tone H or L that can influence the assignment of the tones. There is no information about the methodology of the recording in Ferraz and Traill (1981). Here are some examples of the differences in their analysis:

- [ko'sa] 'to scratch'
- Günther (1973): LH
- Ferraz e Traill (1981): HL to ['kosa]
- Maurer (2009): LL
- [pe'ru] 'turkey'
- Ferraz e Traill (1981): HL
- Maurer (2009): LL
- [fu'ta] 'to steal'
- Günther (1973): LH
- Ferraz e Traill (1981): FL
- Maurer (2009): LL
- ['fu:ta] 'bread-fruit'
- Günther (1973): LH
- Ferraz e Traill (1981): RL
- Maurer (2009): LHH

Maurer (2009) states that properties such as intensity can be observed in Lung'Ie, what would be explained by the fact that $90 \%$ of Lung'Ie's lexicon is derived from Portuguese, a language of lexical stress system. He states that the stressed syllables seem to correspond to the original Portuguese stress, although this is not necessarily true, since there are words that have been altered, such as the pre-final stressed word in Portuguese 'alma', which has a penultimate stress in Lung'Ie ['alima]. In addition, Maurer, like Günther (1973) and Ferraz and Traill (1981), assume that Lung'Ie descends directly from Portuguese, while Bandeira (2017) argues that it descends from PGG.

[^2]According to Agostinho (2015), Lung'Ie has, besides the tonal system - since there is indication of pitch in the lexical realization (see HYMAN, 2006) - a free lexical stress system. That is, each word has an idiosyncratic stress and it is not possible to establish a phonological rule for the attribution of this stress in the original stressed syllable. In this way, the Agostinho (2015) assumes the primary stress as part of the lexical information of each word, not being metrically attributed (see HAYES, 1989).

This fact is interesting, since it distances Lung'Ie from a prototypical stress pattern language, that is, a language in which the lexical stress is given by a metric rule. As the language is in direct contact with Portuguese and does not have monolingual speakers, the recent loans of Portuguese origin usually maintain the original stress. Thus, it is difficult to establish the default lexical stress of the language. In the case of verbs, the default stress for new words is final. According to Agostinho (2015), it is necessary to observe how the tone of new words and recent borrowings occurs and if the tonal variation proposed by Maurer (2009) occurs in these words. Finally, the discussion of tonal analysis in this language among the authors cited reveals that this system is not fully understood.

Agostinho's (2015) analysis is in line with Good's (2004) proposal that Atlantic creole languages may have lost their tonal characteristics through contact with stress languages, such as Portuguese, which remains in contact with Lung'Ie and is the main language of the media and schooling and the most prestigious language in the country (ARAUJO; AGOSTINHO, 2010, AGOSTINHO, 2016).

In this work, I demonstrate that stress needs to be considered when analyzing the phonology of Lung'Ie, as already argued by Agostinho (2015), by examining how tone and stress interact with the processes of external vowel sandhi. That way, it will be possible to show which suprasegmental category allows and blocks the process. Finally, I propose that the suprasegmental Lung'Ie system is mixed, and that both stress and tone should be considered. Hyman $(2006,2014)$ states that there are languages that do not fit into the prototypical categories of accentual pattern and tonal pattern and that there are languages with stress and tone, as I argue to be the case of Lung'Ie.

## 3 EXTERNAL VOWEL SANDHI IN LUNG'IE

The process of external vowel sandhi occurs when there is a word ended in a vowel with a word initiated by another vowel in an utterance, in which these two vowels are realized as being one, or in which diphthongization occurs (FREIRE; PAIS, 2006). Bisol (1996) observes that the ideal context for the occurrence of external vowel sandhi in

Brazilian Portuguese (BP) is when the vowels in the environment are both in unstressed syllables. The process domain is the syllable nucleus. In the case of external vowel sandhi, we also have a process of post-lexical resyllabification, whose domain is the phonological phrase (BISOL, 1999). We can observe some examples in Lung'Ie:
(2) a. ['livu]['owo] $\rightarrow$ [li'vowo] 'your (pl.) book'
b. ['livv]['inc] $\rightarrow$ [li'vinc] 'their book'
c. [ Ja 'ma]['in $\varepsilon] \rightarrow$ [ $\int \mathrm{a}$ 'min $\left.\varepsilon\right]$ 'call them'
d. [fe'ze]['ẽnv] $\rightarrow$ [fe'zẽnv] 'have a birthday'
e. ['minv]['ũa] $\rightarrow$ [mi'nũe] 'one child'
f. ['bwege][u'mẽ] $\rightarrow$ [bwegu'mẽ] 'hand palm'

There are three options for solving the hiatus between words in Lung'Ie: fusion, elision and diphthongization. The fusion (FS) occurs with two vowels of the same quality, as in (3a); the two vowels of the same quality merge into one vowel of the same quality; the elision (EL) occurs with vowels of different quality in which the second vowel is held, as in (3b, c) and the diphthongization (DT) occurs when one of the vowels is / u / or / i / as in (3d), where these vowels are performed as glides [j] or [w] ${ }^{4}$ :
(3) a. ['pi $\kappa \mathrm{k}][\mathrm{a}$ 'di] $\rightarrow$ [piКa'di] 'many palm fruits'
b. ['na][u'matu] $\rightarrow$ [nu'mato] 'in the bush'
c. ['bwege][u'mẽ] $\rightarrow$ [bwegu'mẽ] 'hand pal'
d. ['na][u'matu] $\rightarrow$ [naw'matv] 'in the bush'

In Lung'Ie, elision of the first or the second vowel can occur within the phonological word or phrase ${ }^{5}$ :
(4) a. ['piК̌]['ugbe] $\rightarrow$ [pi'Kugbe] 'many fences'
b. ['tava]['ugba] $\rightarrow$ [ta'vugbe] 'was/were in the fence'
c. [u'pe][u'ka] $\rightarrow$ [upu'ka] 'ocá tree'
d. [ta'ma][u'bwe] $\rightarrow$ [tamu'bwe] 'catch the ox'

In (4a, b), where the first vowel is unstressed and the second vowel is stressed, there is elision of the unstressed vowel and the stress remains in the stressed vowel of the sec-

[^3]ond word. In ( $4 \mathrm{c}, \mathrm{d}$ ), where the first vowel is stressed and the second is unstressed, the stressed vowel is elided and the stress remains in the stressed vowel of the second word. From these examples, we can observe that, regardless of its tonicity, elision always occurs with the first vowel. Casali (1997) notes that there is a universal tendency to preserve segments at the beginning of the word, given their acoustic prominence. This tendency is always valid in the processes of external vowel sandhi in Lung'Ie.

Tenani (2007) shows that BP and EP also obey this trend, although this can only be seen when sandhi is not blocked and elision occurs. This is a crucial difference, since we may argue that in these cases the stress plays a more important role in BP and EP, whereas acoustic prominence at the beginning of words is more important in Lung'Ie, since we may have the deletion of the stressed vowel in the final syllable of the first word.

According to Tenani (2007), the elision is blocked in Brazilian and European Portuguese when the first vowel is stressed. The author affirms that it is common that the stressed vowel is not erased in cases of sandhi. Differently from what would be expected, therefore, in Lung'Ie, elision will occur if the first or the second vowel is stressed within the domain $\phi$, as we can observe above.

The elision by external vowel sandhi occurs when a word finished with an open syllable is followed by a word initiated by a different vowel, in which the first is elided. In examples ( $5 \mathrm{a}, \mathrm{b}$ ), we can observe that the /maa/ e/bii/ nucleus is completely elided, being replaced by $[\mathrm{u}]$ of /umaN/ and /ubaaku/, which undergoes compensatory lengthening. Thus, this process demonstrates that the two segments are in the nucleus and are part of the same syllable, since they can be replaced.
(5) a. /'maa u'maN/ [mu: 'mã] 'to marry'
b. /'bii u'baaku/ [bu:'ba:kv] 'to cave a hole'

In cases of external vowel sandhi, we may have diphthongization occurring within the noun phrase. The examples in (6) show cases where we have a vowel /e, $\varepsilon, \mathrm{a}, \mathrm{o}, ~ \rho /$ before or after unstressed $/ \mathrm{i} /$ or $/ \mathrm{u} /$. In these cases, the $/ \mathrm{i} /$ and $/ \mathrm{u} /$ vowels are performed as glides [j] or [w] and the stress of the second word is maintained:
a. /i'nhemi 'ofo/
b. /i'nhemi u'saN 'loNgo/
[inhe'mjofo] 'tipo de inhame'
c. /i'Sima u'buka/
[inhemju'sã 'lõgo]
'tipo de inhame'
'buço'

Diphthongization as a result of external vowel sandhi can also occur within the prosodic phrase. We have, in (7), two prosodic phrases in which the first word ends with the vowel [a] and the second begins with the vowel [ $u$ ]. In these cases, the /i/ and /u/vowels are also performed as glides [j] or [w] and the stress of the second word is maintained:
(7) a. /'mwa u'maN/ [mwaw'mẽ] 'to pay'
b. /pa'sa u'buka/ [pasaw'buke] 'to eat'

In (8a), we have fusion of the two vowels of the same quality as a result of the sandhi process. In (8b), we have the sandhi occurring with vowels of different quality and the result is the deletion of the first vowel and the maintenance of the second. In (8c), we have two vowels of the same quality, but the first is preceded by a glide. The result of the process, in this case, is the merging of the two vowels and the maintenance of the glide. As the external vowel sandhi occurs in the nucleus, the onglide cannot be deleted because it is part of the onset ${ }^{6}$. In (8e), the onset is not deleted, but the $/ \mathrm{w} /$ plus $/ \mathrm{u} /$ sequence is performed as a long vowel [u:]. The same is also true for (8f), where the onset is also not deleted, but the sequence of $/ \mathrm{j} / \mathrm{plus} / \mathrm{i} /$ is performed as a long vowel [i:]. In ( $8 \mathrm{~g}-\mathrm{h}$ ), we can observe that the branched nucleus can undergo diphthongization and elision of the first vowel.
(8) a. /ke'ba 'arika/ $\rightarrow$ [ke'barike]
b. /pi'Ka 'ugba/ $\rightarrow$ [pi'Kugbe]
c. /fi'gwa 'arika/ $\rightarrow$ [fi'gwarıke], *[fi'ga:rıke]
'to break the arch'
'many fences'
d. /'kwa i'se/ $\rightarrow$ ['kwise], *['ki'se] (Günther 1973) 'this thing'
e. /fi'gwa 'ugba/ $\rightarrow \quad[$ fi'gu:gbe $]$, *[fi'gwugbe $] \quad$ 'figure of a fence'
f. /kaN'dja i'zeti/ $\rightarrow \quad\left[\right.$ kẽdi:'zet $\left.f_{\mathrm{I}}\right]$, *[kẽdji'zet $\left.\mathrm{f}_{\mathrm{I}}\right] \quad$ 'oil lamp'
g. /'maa u'maN/ $\rightarrow$ [maw'mã], [mu:'mã] 'to marry'
h. /'bii u'baaku/ $\rightarrow$ [biw'ba:kv], [bu:'ba:kv] 'to cave a hole'

The same is true for $(9)^{7}$ :

[^4](9) a. ['pwe u'kwato] $\rightarrow$ [pu: 'kwato] (Maurer 2009)
a. ['pwe u'mato] $\rightarrow$ [pu: 'mato] (Günther 1973)
b. ['pwe 'uga] $\rightarrow$ ['pu:ge]
c. [ka'dja i'ni] $\rightarrow$ [ka'di:nı] (Maurer 2009)
d. [u'fja iko' ko] $\rightarrow$ [ufi:ko'ko] (Günther 1973)
'to put in the room' 'to put in the bush' 'to out in the stress' 'elephant buttocks’ 'yam leaf'

In ( $9 \mathrm{a}, \mathrm{b}, \mathrm{c}$ ) we have / w / + / u / being realized as [u:] and in (9d, e), we have / j / + / i / being performed as [i:] on the word boundary. Although Maurer (2009) did not make this relation, Günther (1973) does: "the combination / wu / is phonetically realized as [u:] (...); the / ji / combination is phonetically realized as [i:] <. In this way, we can say that / w / can be realized as [ u ] in post-lexical phonology.

The process of vowel sandhi can also initiate a process of nasalization, as in the following examples:
(10)

| a. /ta'ma 'omi/ | ['tã: 'omi] $\sim$ ['tõ:mı] | 'tomar homem' |
| :--- | :--- | :--- |
| b. /'sunu ta'ma ' $\varepsilon /$ | ['sunv 'tã: ' $\varepsilon] \sim[$ 'sunv 't $\tilde{\varepsilon}$ :] | 'o sono tomou-lhe' |
| c. /'t $\varepsilon N$ u'buka/ | ['t $\tilde{\varepsilon}$ u'buke $] \sim[$ tũ'buke $]$ | 'até a boca' |

In (10a, b), the word/tama/ is performed as ['tã:] and the vowel [ã:] undergoes elision. The stressed vowel that initiates the next word is elongated by compensatory lengthening and the nasal feature of [ã:] passes to the vowel remaining after elision. In (9c), the vowel $[\varepsilon]$ is elided and the nasal trait is passed to the first vowel of the next word [u]. This process can also be seen as an argument in favor of VN or $\mathrm{V}+$ underlying nasal feature, since the nasal trait remains in the word. If the nasal vowel was phonemic in the language, it would be completely elided in cases like this and we would have *['to:mı], *['tz:] and *[tu'buke].

In Table 1, the first column and the first line contain Lung'Ie vowels, demonstrating the possible contact sequences. Each cell brings the possible hiatus resolution options for each sequence, where each vowel is in different prosodic words or phonological phrases. For vowels of same quality, the resolution will always be fusion; for different quality vowels in which there is no $/ \mathrm{u} /$ and $/ \mathrm{i} /$, the resolution will be the elision process; for different vowels in which one of them is unstressed $/ \mathrm{u} /$ or $/ \mathrm{i} /$, we may have elision or diphthongization. If $/ \mathrm{u} /$ or $/ \mathrm{i} /$ is stressed, the diphthongization process is not possible.

Table 1: Possible processes for each combination of two vowels.

|  | $/ \mathbf{i} /$ | $/ \mathbf{\varepsilon} /$ | $/ \mathbf{e} /$ | $/ \mathbf{a} /$ | $/ \mathbf{y} /$ | $/ \mathbf{0} /$ | $/ \mathbf{u} /$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $/ \mathbf{i} /$ | FS | $\mathrm{EL} / \mathrm{DT}$ | $\mathrm{EL} / \mathrm{DT}$ | $\mathrm{EL} / \mathrm{DT}$ | $\mathrm{EL} / \mathrm{DT}$ | $\mathrm{EL} / \mathrm{DT}$ | $\mathrm{EL} / \mathrm{DT}$ |
| $/ \boldsymbol{\varepsilon} /$ | $\mathrm{EL} / \mathrm{DT}$ | FS | EL | EL | EL | EL | $\mathrm{EL} / \mathrm{DT}$ |
| $/ \mathbf{e} /$ | $\mathrm{EL} / \mathrm{DT}$ | EL | FS | EL | EL | EL | $\mathrm{EL} / \mathrm{DT}$ |
| $/ \mathbf{a} /$ | $\mathrm{EL} / \mathrm{DT}$ | EL | EL | FS | EL | EL | $\mathrm{EL} / \mathrm{DT}$ |
| $/ \mathbf{y} /$ | $\mathrm{EL} / \mathrm{DT}$ | EL | EL | EL | FS | EL | $\mathrm{EL} / \mathrm{DT}$ |
| $/ \mathbf{0} /$ | $\mathrm{EL} / \mathrm{DT}$ | EL | EL | EL | EL | FS | $\mathrm{EL} / \mathrm{DT}$ |
| $/ \mathbf{u} /$ | $\mathrm{EL} / \mathrm{DT}$ | $\mathrm{EL} / \mathrm{DT}$ | $\mathrm{EL} / \mathrm{DT}$ | $\mathrm{EL} / \mathrm{DT}$ | $\mathrm{EL} / \mathrm{DT}$ | $\mathrm{EL} / \mathrm{DT}$ | FS |

## 4 STRESS AND SANDHI BLOCKING

The process of sandhi in Lung'Ie is neither blocked between prosodic words, nor between phonological phrases ( $\phi$ ), as described by Agostinho et al. (2012). According to the authors, who adopt the prosodic theory of Nespor e Vogel (1986) and the experiment elaborated by Tenani (2007), the sandhi is only obligatorily blocked when the vowels are in different intonational phrases (I) and one of them is stressed. The quality of the vowels (different or not) does not influence the blocking. The authors also demonstrate that Lung'Ie tends to preserve segments at the beginning of the word and that what triggers the sandhi process is a combination of prosodic context and stress. Agostinho et al. (2012) state that "tone did not prove to be a triggering or blocking element of vowel sandhi" (AGOSTINHO et al., 2012: 295).

The authors also analyzed the vowel quality of the following phonological context (V): same quality /a/ and different quality $/ \mathrm{a} /$ and $/ \mathrm{u} /$; stress contexts: two unstressed vowels, first stressed vowel and second unstressed vowel, first unstressed vowel and second stressed vowel, and two stressed vowels; and the prosodic context in which they appeared: within the same phonological phrase ( $\phi$ ), between two phonological phrases not branched, between two phonological phrases branched, and between two intonational phrases (I).

The following table summarizes the results obtained by Agostinho et al. (2012: 301) in each of these contexts.

Table 2: Processes of sandhi, in which shading marks blocking

|  | $\mathrm{V}=\mathrm{V}$ |  |  |  | $\mathrm{V} \neq \mathrm{V}$ |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathrm{V}+\mathrm{V}$ | 'V + V | $\mathrm{V}+{ }^{\text {'V }}$ | 'V + 'V | $\mathrm{V}+\mathrm{V}$ | 'V + V | $\mathrm{V}+\mathrm{V}$ | 'V + 'V |
| same $\phi$ | FS | FS | FS | FS | EL | EL | EL | EL |
| $\phi+\phi$ non-branched | FS | FS | FS | FS | EL | EL | EL | EL |
| $\phi+\phi$ branched | FS | FS | FS | FS | EL | EL | EL | EL |
| $\mathrm{I}+\mathrm{I}$ | FS | * FS | *FS | *FS | EL | ${ }^{*}$ EL/ ${ }^{*}$ DT | ${ }^{*}$ EL/ ${ }^{*}$ DT | ${ }^{*}$ EL/ ${ }^{*}$ DT |

Source: Agostinho et al. (2012:302).

With vowels of the same quality and in any context of stress, we can have fusion within the phonological phrase and between two phonological phrases, branched or not. The same happens with vowels of different quality, in any context of stress: we can have elision within the phonological phrase and between two phonological phrases, branched or not. However, between two intonational phrases, we can only have elision or fusion if the two vowels in question are unstressed. If either or both are stressed, the processes of fusion, elision and diphthongization are blocked. Blocking is therefore conditioned by the position of the vowels in the prosodic domain and the stress. Agostinho et al. (2012) state that the elision is more productive than the diphthongization.

We can observe the sandhi process occurring also with stressed monosyllables in (11), which is different from what occurs in Portuguese. In (11a, b) we have sandhi occurring with the interrogative particle ['a] and in (10c, d), with the third person singular pronoun [' $\varepsilon]$ :
(11) a. ['tudu]['pz]['a] $\rightarrow$ ['tudu]['pa] 'everything?'
b. ['te]['djo]['a] $\rightarrow$ ['te]['dja] 'do you have money?'
c. $\left[\int \mathrm{Ja}^{\prime} \mathrm{ma}\right][$ [ $\varepsilon] \rightarrow\left[\int \mathrm{a} \mathrm{m} \varepsilon\right] \quad$ 'call him'
d. $[\mathrm{fe}$ 'ze][ $[\varepsilon] \quad \rightarrow$ [fe'z $\varepsilon] \quad$ 'make it'

As seen above, the only context where there will be no application of the sandhi process is when the domain of the intonational phrase is combined with at least one stressed vowel. Tenani (2007) also states that BP and EP may have strategies for solving stress clusters, such as inserting a rhythm beat into or between the phonological phrase and retracting the stress within the phonological phrase, while Lung'Ie solves this recurring to the sandhi process or inserting pauses within the intonational phrase. In addition, we can state that the phonological domain $\phi$ is not active in processes of sandhi in Lung'Ie.

## 5 SANDHI BLOCKING AND TONE

Taking the data from the experiment by Agostinho et al. (2012) and new data, we will analyze the sandhi according to the underlying tones proposed by Maurer (2009), in order to corroborate the analysis that the tonal pattern does not contribute to the blocking of the sandhi process (see AGOSTINHO 2015). The underlying tonal patterns of each word were taken from Maurer (2009), except for the words [lu' 3 a] 'Luzia', ['t $\varepsilon: 3 \mathrm{z}$ ] 'Teresa' and ['arike] 'ark'. As put by Maurer (2009: 27), more research in this area is needed to establish the rules of external vowel sandhi, as well as how tone and stress systems interact.

In Table 3, each tone ( H : high, L : low) represents one of the adjacent vowels, so in ['piKe a'di] we have HH LH, represented in the table by H-L.

When there are two patterns in a cell, such as $\mathrm{H}-\mathrm{H} / \mathrm{L}-\mathrm{H}$, it means that there was more than one pattern for this combination of vowels and prosodic context.

Table 3: Tonal patterns, in which shading marks blocking

|  | $\mathrm{V}=\mathrm{V}$ |  |  |  | $\mathrm{V} \neq \mathrm{V}$ |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathrm{V}+\mathrm{V}$ | 'V+V | $\mathrm{V}+{ }^{\text {'V }}$ | 'V + 'V | $V+V$ | 'V+V | $\mathrm{V}+{ }^{\text {'V }}$ | 'V + 'V |
| same $\phi$ | H-L | H-L | $\begin{gathered} \mathrm{H}-\mathrm{H} / \\ \mathrm{L}-\mathrm{H} \end{gathered}$ | H-H | $\begin{gathered} \mathrm{H}-\mathrm{L} / \\ \mathrm{H}-\mathrm{H} \end{gathered}$ | H-L | H-1 | $\begin{gathered} \mathrm{H}-\mathrm{H} / \\ \mathrm{H}-\mathrm{L} \end{gathered}$ |
| $\phi+\phi$ nonbranched | L-L | L-L | L-H | L-H | L-L | $\begin{gathered} \text { L-H/ } \\ \text { L-L } \end{gathered}$ | $\begin{gathered} \text { L-H/ } \\ \text { L-L } \end{gathered}$ | L-L |
| $\begin{gathered} \phi+\phi \\ \text { branched } \end{gathered}$ | L-L | L-L | L-H | L-H | L-L | $\begin{gathered} \text { L-H/ } \\ \text { L-L } \end{gathered}$ | $\begin{gathered} \text { L-H/ } \\ \text { L-L } \end{gathered}$ | L-L |
| $\mathrm{I}+\mathrm{I}$ | L-L | $\begin{gathered} \text { H-L/ } \\ \text { L-L } \end{gathered}$ | L-H | H-H | L-L | $\begin{gathered} \text { H-L/ } \\ \text { L-L } \end{gathered}$ | L-H | H-H |

It is possible to observe that blocking can occur with all tone combinations between intonational phrases: H-H, H-L, L-H, L-L. Thus, we can state that the trigger for blocking is not the tonal pattern of the vowels in question, since all combinations of tones can be blocked.

We can note that the context for the blocking of the sandhi process is always between intonational phrases, both in the analysis made from stress patterns, and in the analysis made from tonal patterns. However, in the case of tones, any combination may be blocked.

It is possible to notice that all the combinations of tone in the I +I can be found in other positions, as in ${ }^{8}$ :

[^5]

Combining the information from the previous tables, we can analyze the data according to the contexts of vowel quality, tone sequence and prosodic context. It is possible to say that, with both vowels of the same quality and with vowels of different quality, the only prosodic context that can block sandhi is the intonational phrase. However, in taking out the stress context from the analysis, all tone contexts may or may not be blocked by the intonational phrase.

Table 4: Sandhi processes with tonal patterns, in which shading marks blocking.

|  | $\mathrm{V}=\mathrm{V}$ |  |  |  | $\mathrm{V} \neq \mathrm{V}$ |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | H-H | H-L | L-H | L-L | H-H | H-L | L-H | L-L |
| same $\phi$ | FS | FS | FS | FS | EL/DT | EL/DT | EL/DT | EL/DT |
| $\phi+\phi$ nonbranched | FS | FS | FS | FS | EL/DT | EL/DT | EL/DT | EL/DT |
| $\begin{gathered} \phi+\phi \\ \text { branched } \end{gathered}$ | FS | FS | FS | FS | EL/DT | EL/DT | EL/DT | EL/DT |
| $\mathrm{I}+\mathrm{I}$ | FS/ ${ }^{*} \mathrm{FS}$ | $\mathrm{FS} /{ }^{*} \mathrm{FS}$ | FS/* ${ }^{\text {FS }}$ | FS/ ${ }^{*} \mathrm{FS}$ | EL/ ${ }^{*} \mathrm{EL}$ | $\begin{aligned} & \mathrm{EL} / * \mathrm{EL} \\ & \mathrm{DT} /{ }^{*} \mathrm{DT} \end{aligned}$ | $\begin{aligned} & \mathrm{EL} /{ }^{*} \mathrm{EL} \\ & \mathrm{DT} /{ }^{*} \mathrm{DT} \end{aligned}$ | $\begin{aligned} & \mathrm{EL} / * \mathrm{EL} \\ & \mathrm{DT} /{ }^{*} \mathrm{DT} \end{aligned}$ |

We may also note that in every tone combination (H-H, H-L, L-H, L-L) sandhi or blocking can occur, which demonstrates that blocking is not tone dependent. In addition, the $\mathrm{V}+\mathrm{V}$ context allows the sandhi process in all possible tonal combinations for $/ \mathrm{a} /+/ \mathrm{a} /$ and $/ \mathrm{a} /+/ \mathrm{u} /$. This again suggests that it is the stress rather than the tone that determines the application or blocking of the sandhi process in Lung'Ie, since we have the same tone combination with different results.

Another way to look at the data is by examining the processes that may or may not occur with the possible combinations of underlying tone and stress. When looking at the
underlying tones and stress, the sequence of two unstressed vowels is the only context in which there is no blocking with any combination of tone. However, in contexts with at least one stressed vowel $(V V+V, V+V V$ and $V V+V V)$, there may be blocking with any combination of tone (H-H, H-L, L-H, L-L). Blank cells in Figure 2 occur due to the impossibility of finding a word with at least two syllables in which the first tone is a stressed $L^{9}$ :

Figure 2: Sandhi processes with stress and tonal patterns, in which shading marks possibility of blocking.

|  | $\mathrm{V}=\mathrm{V}$ |  |  |  | $\mathrm{V} \neq \mathrm{V}$ |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | H-H | H-L | L-H | L-L | H-H | H-L | L-H | L-L |
| $V+V$ | FS | FS | FS | FS | EL | EL | EL | EL |
| 'V + V | FS/ ${ }^{*} \mathrm{FS}$ | FS/*FS | FS/*FS | FS/ ${ }^{*} \mathrm{FS}$ | $\begin{aligned} & \mathrm{EL} / * \mathrm{EL} \\ & \mathrm{DT} /{ }^{*} \mathrm{DT} \end{aligned}$ | $\begin{aligned} & \mathrm{EL} /{ }^{*} \mathrm{EL} \\ & \mathrm{DT} /{ }^{*} \mathrm{DT} \end{aligned}$ | $\begin{aligned} & \mathrm{EL} / * \mathrm{EL} \\ & \mathrm{DT} /{ }^{*} \mathrm{DT} \end{aligned}$ | $\begin{aligned} & \mathrm{EL} / * \mathrm{EL} \\ & \mathrm{DT} /{ }^{*} \mathrm{DT} \end{aligned}$ |
| $\mathrm{V}+{ }^{\text {'V }}$ | FS/ ${ }^{\text {F }}$ S | - | FS/*FS | - | $\begin{gathered} \mathrm{EL} /{ }^{*} \mathrm{EL}, \\ { }^{*} \mathrm{DT} \end{gathered}$ | - | $\begin{gathered} \mathrm{EL} /{ }^{*} \mathrm{EL}, \\ { }^{*} \mathrm{DT} \end{gathered}$ | - |
| 'V + 'V | FS/*FS | - | FS/*FS | - | $\begin{gathered} \mathrm{EL} /{ }^{*} \mathrm{EL}, \\ { }^{*} \mathrm{DT} \end{gathered}$ | - | $\begin{gathered} \mathrm{EL} /{ }^{*} \mathrm{EL}, \\ { }^{*} \mathrm{DT} \end{gathered}$ | - |

I could not find any word whose first syllable was an unstressed /a/ with $L$ tone, namely, LL and ' $\sigma \sigma$. All HL words correspond to ' $\sigma \sigma$ in my data, and almost all LL words in Maurer's data correspond to $\sigma$ ' $\sigma$. The exceptions are baa 'to burn' and bii 'to open', but we can observe that both are CVV, which could be an argument for explaining their difference. Now we can make a correlation between LL and $\sigma^{\prime} \sigma$ : Maurer (2009) states that almost all the verbs are LL, except for daka 'to give', fedi 'to stink', mendu 'to be afraid' and vika 'to come' which are HH. It is not a coincidence that almost all verbs in Lung'Ie are final stressed (and LL), and that all HH exceptions quoted here have a pre-final stress.

We may notice that there is allowing/blocking for every tone combination, which means that blocking does not depend on tone. Now it is clear that $\mathrm{V}+\mathrm{V}$ allows sandhi in all possible tone combinations for $/ a+a /$ and $/ a+u /$, which means that it is the stress, and not the tone, what determines the allowing or blocking of sandhi process, since there can be the same tone combination with different results.

[^6]
## 6 FINAL REMARKS

In Lung'Ie, elision of the first vowel and fusion may occur if the first vowel or the second vowel is stressed within the domain of $\phi$. This means that there is a tendency of preserving segments at the beginning of a word in Lung'Ie and sandhi processes apply from left to right. The sandhi process of fusion, elision and diphthongization can only be blocked in Lung'Ie in I + I.

We can also observe that with fusion and elision processes, Lung'Ie is more permissive than BP and EP (TENANI, 2007), whereas these languages are more susceptible to diphthongization than Lung'Ie. I interpret the process of diphthongization as an intermediate process, and argue that Lung'Ie seems to have either elision/fusion, or blocking.

From the data, I argue that what triggers or blocks the sandhi process in Lung'Ie is a combination of prosodic context plus stress. We have seen that we only have blocking with $\mathrm{I}+\mathrm{I}$ and with ' $\mathrm{V}+\mathrm{V}, \mathrm{V}+{ }^{\prime} \mathrm{V}$ and ' $\mathrm{V}+{ }^{\prime} \mathrm{V}$. In verifying the underlying tones described by Maurer (2009), we can observe that $\mathrm{I}+\mathrm{I}$ is a blocking environment and that all combinations of tones can be blocked if they are in the I + I prosodic context. In other words, tone does not play a role in permitting or blocking processes of external vowel sandhi in Lung'Ie.

I believe that a more in-depth analysis of the tonal system and its interaction with the stress will be necessary for the suprasegmental Lung'Ie system to be better understood. It is noteworthy that this hypothesis is valid for the underlying tones described by Maurer (2009), but that can be modified by a new analysis of Lung'Ie tones, since all of the tone information in this paper comes from his description. The disparity of the tonal analysis in this language among the cited authors demonstrates that more research is needed in the area.

## References

AGOSTINHO, A. L., ARAUJO, G. A., FREITAS, S. Resolução de hiato externo em lung'Ie. Papia 22(2): 295-305. 2012.

AGOSTINHO, A. L.. Fonologia e método pedagógico do lung'Ie. Tese (Doutorado em Letras) - Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo. 2015.
$\qquad$ . Fonologia do lung'Ie. LINCOM Studies in Pidgin \& Creole Linguistics, 15. 2016.
ARAUJO, G.A. (Org.). O acento em Português: abordagensfonológicas. São Paulo: Parábola, p. 135-156. 2007.

ARAUJO, G. A.; AGOSTINHO, A. L. Padronização das línguas nacionais de São Tomé e Príncipe. Língua e Instrumentos Linguísticos 26, pp. 49-81, 2010.

BANDEIRA, M. Reconstrução fonológica e lexical do protocrioulo do Golfo da Guiné. Tese de Doutorado, Universidade de São Paulo, 2017.

BISOL, L. Sandhi Externo: O Processo e a Variação. In KATO, M. A. (Org.). Gramática do Português Falado: Convergências. São Paulo: UNICAMP/FAPESP. 1996, p. 55-96.
$\qquad$ . A sílaba e seus constituintes. In NEVES, M. H. de M. (org.). Gramática do Português falado VII, 701-742. Campinas: Editora da UNICAMP. 1999.

CASALI, Roderic. F. Vowel elision in hiatus contexts: which vowel goes? Language 73 (3), p. 493-533. 1997.

FERRAZ, L.; TRAILL, A. The interpretation of tone in Principense Creole, Studies in African Linguistics 12(2): 205-15. 1981.

FREIRE, C. C. S. S.; PAIS, C. T. Aspects of Co-Articulation Phenomenon in External Sandhi in the Brazilian Portuguese: Elision, Diphthong, Contamination. 2006.Disponível em: https://www.cefala.org/issp2006/camera-ready/freire.pdf

GÜNTHER, W. Das portugiesische Kreolisch der Príncipe Island. Marburg an der Lahn: Im Selbstverlag. 1973.

FERRAZ, L. The Creole of São Tomé. Johannesburg: Witwatersrand University Press, 1979.

GORDON, M. Disentangling stress and pitch-accent: A typology of prominence at different prosodic levels. In: HULST, H. Van Der (Ed.). Word Stress: Theoretical and Typological Issues. Cambridge: Cambridge University Press, 2014.

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

GUSSENHOVEN, C. The Phonology of Tone and Intonation. Cambridge: Cambridge University Press, 2004.

HAGEMEIJER, T. As Línguas de S. Tomé e Príncipe. Revista de Crioulos de Base Lexical Portuguesa e Espanhola, Lisboa, v. 1, n. 1, p. 1-27, 2009.

HAYES, B. The prosodic hierarchy in meter. In: KIPARSKY, P.; YOUMANS, G. (Ed.). Rhythm and Meter. Orlando: Academic Press, 1989.

HYMAN, L. M. Do all languages have word accent? In: HULST, H. Van Der (Ed.). Word Stress: Theoretical and Typological Issues. Cambridge: Cambridge University Press, 2014.
$\qquad$ . How (not) to do phonological typology: the case of pitch-accent. Language Sciences, Amsterdam, v. 31, n. 2-3, p. 213-238, 2009.
$\qquad$ .Word-prosodic typology. Phonology, Cambridge, v. 23, n. 2, p. 225-257, 2006.

MAURER, P. Principense (Lung'Ie). London: Battlebridge Publications, 2009.

SCHUCHARDT, H. Beiträge zur Kenntnis des kreolischen Romanisch. Zum Negerportugiesischen der Ilha do Principe. Zeitschrift für Romanische Philologie, Saarbrücken, v. 13, p. 461-475, 1889.

MAURER, P. Lung'Ie. Londres: Battlebridge Publications. 2009.
NESPOR, M.; Vogel, I. Prosodic Phonology. Dordrecht: Foris. 1986.

TENANI, L. E. Acento e processos de sandhi vocálico no Português. In ARAUJO, G. A. (Org.). 2007. O acento em Português: abordagens fonológicas. São Paulo: Parábola, 2007, p. 135-156.


[^0]:    ${ }^{1}$ I understand stress here as an abstract category of prominence at the lexical level (HULST, 2014, HULST, 2012), which has the potential to receive a pitch-accent (see BOLINGER, 1958, apud GUSSENHOVEN, 2004).

[^1]:    ${ }^{2}$ All tonal patterns are taken from Maurer (2009).

[^2]:    ${ }^{3}$ Tense-Mood-Aspect.

[^3]:    ${ }^{4}$ Examples from Agostinho et al. (2012).
    ${ }^{5}$ Examples from Agostinho et al. (2012).

[^4]:    ${ }^{6}$ See Agostinho (2015) for a discussion on the positions of glides within the syllable in Lung'Ie.
    ${ }^{7}$ My transcriptions.

[^5]:    ${ }^{8}$ Examples from Agostinho et al. (2012) and tone patterns from Maurer (2009).

[^6]:    ${ }^{9}$ Disyllabic words initiated by L are pre-final stressed and trisyllabic words initiated by L may be stressed on the final or pre-final syllable.

