DOI: 10.5433/2237-4876.2020v23n3p27

Vowel harmonization in Brazilian and in São Tomé Portuguese and the raising of the pre-stress /e/

Fabiane de Mello Vianna da Rocha Teixeira Rodrigues do **NASCIMENTO***

* PhD (2018) and Master's degree (2013) in Vernacular Letters at Universidade Federal Rural do Rio de Janeiro (UFRRJ). Professor at Universidade Federal Rural do Rio de Janeiro (UFRRJ). Contact: fabyufrrj@gmail.com.

Abstract:
This article draws on the tenets of Variationist Sociolinguistics (LABOV, 1972, 1994, 2001) to examine the influence of the process of vocal harmonisation on raising of the pre-stress / e / in the Brazilian and São Tomé varieties of Portuguese. Few studies have been published on the African varieties of Portuguese and Portuguese-based creoles. As a result, the characteristics of their phonological systems remain to be investigated. Rocha's findings (2018) on the urban variety of São Tomé Portuguese confirm the hypothesis of Marquilhas (2003) that it approaches "Brazilian Portuguese in vowel harmonisation and in insubordination to the general rule of reduction" recurrent in European Portuguese. Furthermore, they suggest similarities with what is found in Portuguese-based creoles and, especially in Forro, the creole most spoken in São Tomé. It was thus decided to compare these findings with those of Yacovenco (1993) and Rocha (2013), drawn from municipalities in Rio de Janeiro State, and with the descriptions by Ferraz (1979, 1987) and Hagemeijer (2009) of vocalism in Gulf of Guinea creoles. The aim is to demonstrate that a trend already confirmed in a number of languages – the influence of the following high vowel on raising of the pre-stress front vowel – exists in the non-European varieties of Portuguese.

Keywords:
Vowel harmonisation. Brazilian Portuguese. São Tomé Portuguese.

Signum: Estudos da Linguagem, Londrina, v. 23, i. 3, p. 27-42, Dec. 2020

Received on: 05/04/2020
Accepted on: 10/27/2020

Vowel harmonization in Brazilian and in São Tomé Portuguese and the raising of the pre-stress /e/

Fabiane de Mello Vianna da Rocha Teixeira Rodrigues do Nascimento

Introduction

In his descriptions of the vowel system of Brazilian Portuguese (BP), Camara Jr (1970) proposed that the number of vowel phonemes varied by the segment's position in relation to the stress, reaching an upper limit of distinctness in the stressed syllable. He argued that, in that position, there are seven distinct units, which are reduced by neutralisation to five, four and three phonemes in, respectively, the pre-stress, non-final post-stress and final positions.

The neutralisation or annulment of one or more distinct features occurs, in pre-stress syllables, among open and closed mid segments of both the front and back series. That pattern is found in words such as "*pesar*" and "*morar*", the open or closed pronunciation of whose vowels – p[ɛ]sar or p[e]sar and m[ɔ]rar or m[o]rar, respectively – does not entail any change in meaning, as occurs in the stressed context; they are merely variants of the same phonological element, the so-called archiphonemes /e/ and /o/. Note, however, that these four segments – [e], [ɛ], [o] and [ɔ] – can occur in speech, constituting dialect markers and, depending on the linguistic environment, acting together with high vowels. This latter possibility – classified as raising and readily found in all dialects of BP – culminates in a sporadic neutralisation among mid and high vowels, known as overlapping (CAMARA JR, 1970, p. 45).

As this issue involves different motivations and/or constraints, which depend on the individual, the communication situation, the speech community and the lexical item, it has been addressed, from either a phonetic or a phonological perspective, by Camara Jr. (1970, 1977), Bisol (1981, 2003), Callou and Leite (1986), Callou, Leite and Coutinho (1991), Silva (1991), Yacovenco (1993), Cardoso (1999), Brandão and Cruz (2005), Marques (2006), Oliveira (2008), Rocha (2013) and others. Most of these approaches confirm that processes of vowel harmonisation and reduction are operating. In the former case, the presence of a stressed high vowel following the pre-stress vowel would result in raising. In reduction, on the other hand, "vowels come to be closer in articulation to the adjacent consonant segments", because they form part of a phonological environment where "the word has no high vowel that can trigger the process of raising" (CARMO, 2009, p. 24).

As can be seen, the variable phenomenon in question is "one of the most important points, but also one of the most obscure, in the history of Portuguese" (TEYSSIER, 1980, p. 68), particularly when one compares the different varieties of Portuguese spoken around the world. Marquilhas (2003) argued that, as the official language in a multilingual setting, São Tomé Portuguese (STP) is close to the Portuguese-based African creoles, particularly Forro, and stands "alongside Brazilian Portuguese in displaying the phenomenon of vowel harmonisation and insubordination to the general rule of reduction" (p. 7) prevalent in European Portuguese (EP).

This study observed the influence of the vowel harmonisation process on raising of the pre-stress vowel /e/ in the Portuguese of Brazil and of São Tomé. The African varieties constitute systems whose phonetics and phonology "have practically yet to be explored" (HAGEMEIJER, 2016, p. 48). That is why it is intended here to compare the findings of Yacovenco (1993) and Rocha (2013), in municipalities in the State of Rio de Janeiro, and of Rocha (2018a), regarding the urban variety of São Tomé Portuguese, with the

analyses of Ferraz (1979, 1987) and Hagemeijer (2009), of the autochthonous languages spoken on the islands of São Tomé and Príncipe. These authors have commented on the vocalism of Gulf of Guinea creoles, while not exploring the pre-stress position in depth as much as the other studies selected. Variationist studies in Rio de Janeiro and São Tomé are notable for their findings with regard to the vowel type of the syllable following a pre-stress front vowel. The semi(categorical) contexts of raising and the peculiarities of each study aside, it is hoped to demonstrate that the tendency found in a number of languages, for the following high vowel to cause raising of the pre-stress /e/, is also found in the varieties of Portuguese studied here.

To that end, this Introduction is followed by sections on: i) tenets of Variationist Sociolinguistics (cf. LABOV, 1972, 1994, 2001), the theoretical approach underpinning most of the studies on the subject considered here (cf. Section 1); ii) brief remarks on the vowel harmonisation process; iii) descriptions of this process operating in the State of State of Rio de Janeiro (cf. YACOVENCO, 1993 and ROCHA, 2013), in Gulf of Guinea creoles (cf. FERRAZ, 1979, 1987 and HAGEMEIJER, 2009) and in the urban variety of STP (cf. ROCHA, 2018a); iv) some final remarks; and v) the bibliographical references that informed the study.

Variationist Sociolinguistics or Theory of Variation and Change

Variationist Sociolinguistics or the Theory of Variation and Change was proposed by Weinreich, Labov and Herzog (1968) and established by the studies of Labov in the 1970s.[1] This approach builds on the assumption that variation and change result from the action of linguistic and extralinguistic factors, thus directly relating social heterogeneity and linguistic heterogeneity. The analyses build on the notion of speech community to propose that, although variants present in dialectical and individual contexts may appear disordered, they are provided for by the system and, accordingly, are intrinsic to the language. As a methodology, Sociolinguistics seeks reliable data by prioritising, among other things, a combination of informants who represent significant realities and an interview context that encourages the most natural expressions of language possible.

Following Labov, informant selection is random, but stratified. What this means is that data are organised into cells that illustrate social variables. This enables each linguistic phenomenon studied to be interpreted as a set of real manifestations accessible to the speaker and/or speech community. It also demonstrates that choices among the many existing variants result from the direct effect of structural factors (the linguistic context in which the segment analysed is situated) and social factors (extralinguistic factors, such as age range, sex, schooling, degree of formality, area of residence and others).

In this way, Variationist Sociolinguistics – respecting axioms such as the ordered heterogeneity of language, the social realities of the individual speaker and the contextual realities of the segment studied, the degree of discursive formality and the existence of consistent, reliable data – has currently become established as one of the soundest methods for linguistic studies, particularly as regards unstressed vocalism in Portuguese.

The Importance of the Vowel Harmonisation Process to Applicability of the Raising Rule

As noted in the introduction to this article, most approaches to raising of pre-stress mid vowels in Portuguese emphasise the action of processes of vowel harmonisation and reduction. In line with the proposal

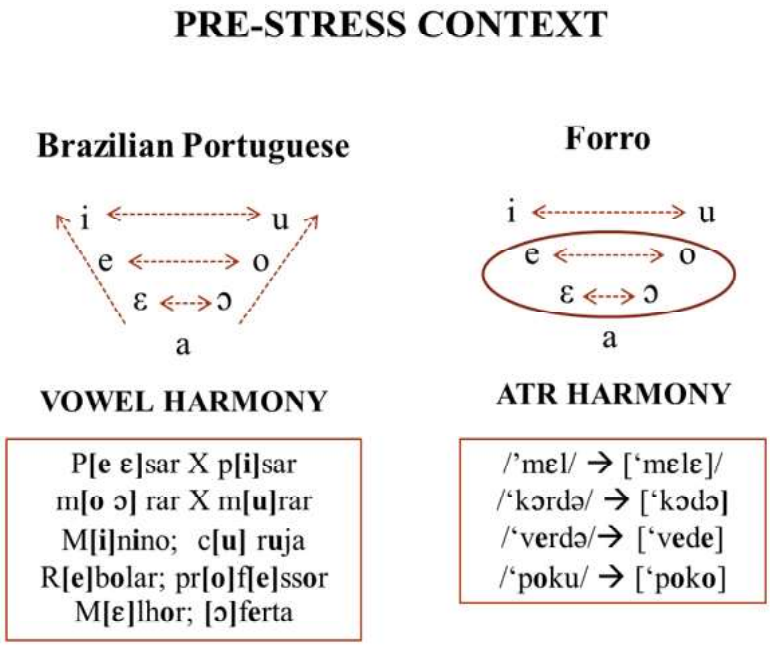
[1] Weinreich, Labov and Herzog (1968) set out the empirical foundations for a possible theory of language change. They clarified the concept of ordered heterogeneity, which was fundamental to development of Labov's Variation Theory or Quantitative Sociolinguistics, published in 1972.

of this study, this section highlights information about vowel harmonisation, given that raising/maintaining/opening the timbre of the pre-stress vowel is influenced by the nature of the vowel of the following syllable, a recurrent phenomenon in a number of languages that is also found in Portuguese and in creoles originating from it.

In his pioneering studies of BP, Camara Jr (1970) highlighted the occurrence of assimilation of the [+high] feature between the stressed vowel and pre-stressed mid vowel in the same word and interpreted the stress as a necessary condition for the relation between these segments. The mid-vowel raising seems to result from assimilation, because "it changes the quality of one or more vowels of the word so as to harmonise with another (...) present in the same word" (LEMLE, 1996, as in SILVEIRA, 2008, p. 29). The vowel neutralisation and harmony was thus seen to correspond to changes in the "aperture of the oral cavity, characterised by the binary [open] feature, which can be separated into gradations" (WETZELS, 1992, as in COLLISCHONN; SILVA, 2013, p. 6).

Many Variationist studies have pointed to vowel harmonisation between the pre-stress vowel and the following vowel. They have proven the tendency described by Lemle (1996, as in GRAEBIN, 2008, p. 116), that "high vowels favour raising – *pirigo*, *currida*; [...] mid vowels lead to maintenance of the mean pronunciation – *rebolá*, *professor*; and [...] low vowels tend to lower the height of the pre-stress vowel – *melhor*; *oferta*". Bisol (1981, as in BRANDÃO; ROCHA; SANTOS, 2012) attributed this process to the presence of an adjacent "trigger": the presence of a high (stressed or unstressed) vowel following the pre-stress vowel would drive raising. Words such as "*mexerica*" – which is pronounced as "*mex[i]rica*" or "*m[i]x[i]rica*", but never as "**m[i]xerica*" – would also indicate that "implementation of the rule does not jump about" and that "among the high vowels, [i] has more assimilatory force than [u], because of the very shape of the oral cavity, whose anterior portion, on the vertical axis, offers more articulatory space than the posterior portion" (p. 14).

As can be seen, these processes, in which features are transferred, culminate in filling, when motivated by a mid or low vowel, or in change, if resulting from a high vowel (cf. BISOL, 2003, p. 10). Actually, some regard the raising as involving a height dimension, "represented by the [+high] feature or by a scalar aperture feature" (CLEMENTS; HUME, 1995 as in COLLISCHONN; SILVA, 2013, p. 6). On the other hand, pre-stress vowel lowering is regarded as corresponding to a case of ATR harmony, involving the movement of the tongue root (cf. LEE; OLIVEIRA, 2003 as in COLLISCHONN; SILVA, 2013, p. 6), as occurs, for instance, in Forro, the creole most spoken in São Tomé (cf. Figure 1):



Source: Rocha (2018a).

Figure 1 – Vowel harmony and ATR harmony in Brazilian Portuguese and Forro

Lastly, with a view to elucidating the motivations guiding this extremely variable phenomenon in Portuguese, Collischonn and Silva (2013, p. 6), agreeing with Calabrese (1985), propose a combination of "height features" and "a feature relating to tongue root – [ATR] (Advanced Tongue Root) – position [...] used to characterise harmony in Bantu languages".

The Pre-Stress Position in Different Varieties of Portuguese and in Gulf of Guinea Creoles

By and large, BP and EP had different unstressed vowel systems: in the former, variation predominates, while in the latter, there is greater regularity in unstressed positions, particularly as a result of the process of vowel reduction. The findings of Rocha (2018a) regarding STP suggest a low incidence of open vowels and greater proximity to what is found in Brazil's South and Southeast regions. Accordingly, some of the information published in: i) Yacovenco (1993) and Rocha (2013) as to vowel harmonisation processes in the State of Rio de Janeiro; ii) Ferraz (1979, 1987) and Hagemeijer (2009) on vowel systems in Gulf of Guinea creoles; and iii) Rocha (2018a) on the urban variety of São Tomé Portuguese will be highlighted below.

A Vowel Harmonisation in the State of Rio de Janeiro

Variation in the pre-stress context is a very longstanding phenomenon in Portuguese, as described by philologists (NASCENTES, 1922; TEYSSIER, 1980; SILVA NETO, 1988) and dialectologists (AMARAL, 1920; MARROQUIM, 1934; TEIXEIRA, 1938). Attention to the issue has been intensifying since the 1970s, when Variationist studies were conducted in Brazil. By allying grammatical and social aspects, it was possible to outline an overall view of the characteristics governing implementation of the variants. Since then, a large number of studies have addressed the complexity of phonetic fluctuation inherent to the pre-stress syllable (cf. CAMARA JR, 1970, 1977; BISOL, 1981, 2003; SILVA, 1991; YACOVENCO, 1993; CALLOU; LEITE, 1986; CALLOU; LEITE; COUTINHO, 1991; CARDOSO, 1999; VIEGAS, 2001; SCHWINDT, 2002; BRANDÃO; CRUZ, 2005; HORA; SANTIAGO, 2006; MARQUES, 2006; OLIVEIRA, 2008; ROCHA, 2013, and several others included in the References section, below).

Of particular note as regards vowel harmonisation in raising the pre-stress /e/ in the State of Rio de Janeiro are Yacovenco (1993) and Rocha (2013). The former considered educated speech by investigating part of the NURC/RJ[2] *corpus*, while the latter studied both educated and popular speech in the metropolitan region from interviews of residents of the municipality of Nova Iguaçu.

Vowel Harmonisation in Educated Speech in Rio de Janeiro City

Yacovenco (1993) observed 18 informants of the NURC/RJ *corpus* from the 1970s, distributed equally by sex, three age brackets and three distinct zones of residence, to examine occurrences of maintenance, raising and lowering in the sample. While offering a more comprehensive view of pre-stress mid vowels, she aimed to establish the factors conditioning maintenance of the dialect's closed timbre, on the assumption that the maintenance rule was a "standard norm among the educated community of Rio de Janeiro" and represented

[2] Callou, Leite and Coutinho, as can be seen from the references, conducted the first studies of pre-stress mid vowels in the speech of Rio de Janeiro. However, it was decided to comment only on the Yacovenco study not only for its greater comprehensiveness, but also because it is based on some of the inquiries considered by those researchers.

"an effort to restore the linguistic system, [...] connected with an endeavour by educated speakers to approximate their pronunciation to the official spelling" (p. 172).

This claim is confirmed by comparing interviews of men and women residing in the North, South and suburban zones, divided equally among young, middle-age and elderly interviewees, since "the young age group, the female informants and residents in the North zone are those who most favour the maintenance rule, as against raising. On the other hand, the male informants and residents in the South zone are those that most suppress this rule" (YACOVENCO, 1993, p. 173).

In structural terms, she evaluated the distance between the vowel in question and another high vowel in the word, the timbre of the stressed and subsequent unstressed vowels, the nature of the atonicity of the syllable whose nucleus is pre-stress and the place or articulation of the adjacent consonants.

She first presented seven variants of the dependent variable, reduced at a later stage to five in view of the phonetic similarities between the various types of pre-established diphthongs and the scarcity of sample data. Table 1 shows the overall distribution of the tokens and the rationale for combining the cases of diphthongisation into the same group.

Table 1 – Distribution of pre-stress vowels in educated speech of Rio de Janeiro, by different variants of a single dependent variable

VOWEL TIPE	Tokens.	%
Front oral	2070	49.4
Back oral	1188	28.4
Front nasal	535	12.8
Back nasal	264	6.3
Full diphthong	**91**	**2.2**
Reduced diphthong	**37**	**0.9**
Reduced, raised diphthong	**4**	**0.1**

Source: Yacovenco (1993, as in ROCHA, 2018a, p. 63).

The oral vowels accounted for 3258 (78%) of the 4189 total tokens, with front vowels more recurrent than back vowels. By and large, this sample displays: i) higher rates of maintenance of mean timbre (63.8%), to the detriment of raised [i u] pronunciations (32.9%) and, at much lower levels (3.3%), open timbre vowels [ɛ, ɔ]; and ii) the cases of both lowering and raising seem to stem mainly from the action of vowel harmony. This section prioritises only raising of the oral vowel /e/, for its representativeness and in line with the aims of this article.

Observation of high pronunciations entails eliminating data where application of the rule nears or surpasses 90%. This involves word-initial front vowels closed by /S/ (89.2%), in the form of des-, whether prefixal or not, (87.5%) and also front vowels preceded by affricate consonants (100%). If such tokens are disregarded, the vowel type in the word's stressed syllable becomes the group of factors "most important to updating the variable rules, given that it is related to vowel harmonisation, that is, to the influence of the stressed on the pre-stressed vowel" (YACOVENCO, 1993, p. 96).

High segments act to favour raising of /e/, whether they are homorganic (RW 0.74)[3] or non-homorganic (RW 0.72). On the other hand, the [e] pronunciation seems to stem from the presence of segments of the same timbre (RW 0.59) and low timbre (RW 0.54) in the word (cf. Table 2).

[3] The relative weights obtained in the analyses described are given in brackets.

Table 2 – Rates of maintenance and raising of the pre-stress vowel /e/, by contiguous stressed vowel type, in educated speech of Rio de Janeiro

Contiguous stressed vowel type	Examples	Maintenance			Raising		
		Tokens	%	RW	Tokens	%	RW
Non-homorganic high	nenh**u**m	23	59	0.28	1	**38.5**	**0.72**
Homorganic high	conhec**i**mento	197	58	0.26	1	**42**	**0.74**
Low	hier**a**rquia	353	**78**	**0.54**	22	17	0.46
Mid	acontec**e**r, esp**e**ro, dep**o**is, senh**o**ra	488	**82**	**0.59**	18	15	0.41

Source: Yacovenco (1993 as in ROCHA, 2018a, p. 68, with alterations).

Vowel Harmonisation in the Municipality of Nova Iguaçu/RJ

Rocha (2013) examined pre-stress mid vowels in the speech of Nova Iguaçu, a municipality in the Rio de Janeiro metropolitan region. With a view to describing the structural, social, lexical and idiolectic motivations in both series, the research was divided into two stages. The first considered the action of linguistic and extralinguistic factors on the variable phenomenon. The second involved listing the items in which the tokens appeared, so as to evaluate the importance of lexicon and individual behaviour on the spread of change. The Variationist analysis involved 19079 tokens (11378 for /e/ and 7280 for /o/), which are separated by variant in Table 3.

Table 3 – Distribution of tokens of raising, maintenance, lowering and diphthongisation of /e/ and /o/ in the speech of Nova Iguaçu/RJ

Variants	Examples		Pre-stress /e/	Pre-stress /o/
[i u]	Cons**[i]**guir, c**[u]**meço	**Tokens**	3861	1885
		%	34	245
[e o]	p**[e]**ssoas, j**[o]**rnal	**Tokens**	7158	5314
		%	63	69
[ɛ ɔ]	B**[ɛ]**líssimas, B**[ɔ]**tafogo	**Tokens**	96	81
		%	1	1
[ej ow]	r**[ej]**spondeu, c**[ow]**ntrole	**Tokens**	263	421
		%	2	55
		Total tokens	11.378	7701

Source: Rocha (2013, as in ROCHA, 2018, p. 73, with alterations).

Once the (semi-)categorical cases and data on lowering were eliminated, the poor productivity of the [i] and [u] variants ratified the view of Yacovenco (1993) that maintenance is the norm in the dialect studied. Of the fourteen groups postulated as representing the concomitant structural and social influences on the phenomenon of pre-stress raising, eight were selected for /e/, viz.: the quality of the vowel of the following syllable, the manner of articulation of the preceding consonant, the manner of articulation of the following consonant, the grammatical class of the word and isolated items, nasality, the place of articulation of the preceding consonant, the place of articulation of the following consonant and age range.

Peculiarities apart, the [e] variant proved most likely when non-high segments followed a high vowel. On the other hand, demonstrating the vowel harmony process, contiguity with the, mainly stressed, [i] vowel stood out as a setting favouring high pronunciation (RW 0.95) (cf. Table 4).

Table 4 – Rates of raising of the pre-stress /e/ vowel, by contiguous vowel type, in the speech of Nova Iguaçu/RJ

Contiguous vowel type		Examples	[i]		
			Tokens	%	RW
Stressed vowel (oral/nasal/ nasalised)	Homorganic high	aprend**i**	460/882	**52**	**0.95**
	Non-homorganic high	seg**u**ndo	87/172	**51**	**0.73**
	Non-high	dife**r**ença, ent**r**egue, futeb**o**l, direç**ã**o	196/3.543	5.5	0.29
Unstressed vowel (oral/nasal/ nasalised)	Homorganic high	desperd**i**çar	104/550	**19**	**0.81**
	Non-homorganic high	ed**u**cação	29184	**16**	**0.74**
	Non-high	el**e**mentar, aer**o**porto, lev**a**ntamento	70/882	8	0.31

Source: Rocha (2013, p. 111, with alterations).

Vocalism in Gulf of Guinea creoles

Of the few published phonetic and phonological studies of the languages spoken on the islands of São Tomé and Príncipe, priority is given here to the descriptions by Ferraz (1979, 1987) and Hagemeijer (2009). The studies by Ferraz formulated generalisations about Forro or Santome (FERRAZ, 1979) and other Portuguese-based creoles found in Asia and West Africa (FERRAZ, 1987). These pioneering, overall analyses pointed up phonological, morphological, syntactical and lexical features of these systems. Exploring some of these observations in greater depth, Hagemeijer (2009) explained the process of initial vowel agglutination in Gulf of Guinea creoles and identified vowel harmony as one of the possible motivations for this phenomenon.

Ferraz (1979, p. 29) reported that Forro displayed a series of morphological and phonological mechanisms which enabled it to incorporate words from the superstrate language, European Portuguese. As with the other Gulf of Guinea creoles, its vowel system comprises 7 oral segments and is subject to the harmony process, a recurrent phenomenon in languages in general, and shared by Portuguese and other languages of the Bantu and Kwa groups (p. 49). More precisely, "In ST there are also examples of a type of vowel harmony which is preference rather than complete assimilation. Usually the preference is not completely systematic" (p. 19). This is an essential phonological feature, based on phonemic and morphophonemic factors: a not always systematic tendency "for the same vowel to occur in two consecutive syllables within a morpheme" (p. 25). The accent in the word incorporated into the creole corresponds to that of the original form in Portuguese. In this regard, respecting certain habitual rules of phoneme conversion, vowels in Forro, which determine harmony, are seen to be absorbed from the superstrate language, with their articulation tending to approximate to it.

Table 5 – Forro words displaying the vowel harmony process, selected by Ferraz (1979)

PORTUGUESE ORIGIN	FORRO	TRANSLATION
/ˈk ɔ rd ə /	[ˈk ɔ d **ɔ]**	"Corda"
/ˈd**o**r/	[ˈd**o**l**o**]	"Dor"
/iz ʁ'b ɛ l/	[ze'be]	"Isabel"
/'largu/	[ˈlalugu]	"Largo"
/'m ɛ l/	[ˈm ɛ l ɛ]	"Mel"
/' ɔ dyu/	[ˈɔ j ɔ]	"Ódio"
/p ə'd**i**r/	[p**i**'j**i**]	"Pedir"
/ˈp**o**ku/	[ˈp**o**k**o**]	"Pouco"
/pr ə's**i**zu/	[p**l**i**š**i**z**u]	"Preciso"
/s ʁ'b**e**r/	[s**e**'b**e**]	"Saber"
/ˈv**e**rd ə /	[ˈved**e**]	"Verde"
/v ə r'dad ə /	[v ɛ ˈd ɛ]	"Verdade"

Source: Rocha (2018a, p. 77).

Examples taken from Lung'Ie and Angolar, two creoles still spoken on the islands of São Tomé and Príncipe, indicate that, in these systems, albeit in differing proportions, the unstressed vowel also "copies the stressed vowel" harmoniously (FERRAZ, 1987, p. 343), thus:

Portuguese	→	Lung'Ie	Portuguese	→	Angolar
"mel"	→	/'m ɛ l ɛ /	"saber"	→	/se'be/

Complementing the descriptions by Ferraz (1979, 1987), Hagemeijer (2009) examined the recurrent agglutination of initial vowels in a more specific group of words that form part of the inventory of Gulf of Guinea creoles: items that etymologically begin with consonants in the superstrate language. Hagemeijer regarded Gulf of Guinea creoles as sharing not only a lexifier base, but also as deriving from the spread over time and place of a protolanguage, the Gulf of Guinea protocreole (GGPC). He considered agglutination of initial vowels in the protocreole to be determined by gender in Portuguese and by a reduced ATR system, apparently Edoid in origin. The phenomenon was thus motivated by morphological (gender) and phonological (vowel harmony) factors (p. 38).

Reiterating Ferraz (1979), Hagemeijer (2009) wrote that the Gulf of Guinea creoles used the same vowel inventory as standard Portuguese: seven oral vowels [a e ɛ i o ɔ u]. That tendency for the same vowel to occur in two consecutive syllables within a morpheme (FERRAZ, 1979, p. 25) should, nonetheless, be interpreted as a case of assimilation, because vowel harmony in these creoles is a phenomenon restricted more to the domain of mid vowels (HAGEMEIJER, 2009, p. 37). There are not many studies of the phonological constraints on agglutination of initial vowels, but such cases recur in Gulf of Guinea creoles, such as those taken from Santome and listed in Table 6.

Table 6 – Two-syllable words in Santome

V1 \ V2	i	u	**E**	**ɛ**	**o**	**ɔ**	a
i	Ligi	migu	izê (esteira,cama)	Mile	libo	Jinklo	Mina
u	Buli	mulu	Ubwê (boi)	kume	-	-	Uswa
e	Sêji	dêsu	Vêndê	-	-	**-**	Zema
ɛ	Peli	petu	-	vede	-	Tebo	Bega
o	Sôtxi	wôdu	Ômê	-	pôvô	-	Lopa
ɔ	Doxi	mosu	-	love	-	Kodo	Bola
a	mali (mar)	matu (mato)	Padê	manse	kasô	Avo	Faka

Source: Rocha (2018a, p. 79, adapted from HAGEMEIJER, 2009, p. 37).

These two-syllable words indicate that, in Santome, i) high and low vowels do not determine the quality of adjacent sounds; and ii) co-occurrence constraints are operating, motivated by the mid vowels. In this respect, at least in two-syllable items, what is found is a consistent harmony rule that restricts the co-occurrence of open and closed mid segments (cf. HAGEMEIJER, 2009: 37).

In the endeavour to clarify the origin of the process, he observed diachronic data and claimed that, at an early stage in the formation of the Gulf of Guinea protocreole, "agglutination is strongly associated to vowel height" (HAGEMEIJER, 2009 c: 38). This relationship, although not absolute, was found between vowels in the words of origin, in European Portuguese, and the segments agglutinated in the protolanguage, as illustrated in Table 7.

Table 7 – Vowel agglutination in two-syllable words of the Gulf of Guinea protocreole

Stressed vowels in the words of origin (EP)	**Agglutinated vowels in GGPC**	**Examples**	
		EP	**GGPC**
Low and low-mid [a ɛ ɔ]	Low and low-mid back [a ɔ]	"Pá" - ['pa] "fé"- ['f ɛ] "mar" - ['mar] "pé" - ['p ɛ] "nó" - ['n ɔ]	[a]pa [a]fe [ɔ]mali [ɔ]pe [ɔ]no
[ɛ i]	[o]	"céu" - ['s ɛ w] "rio" - ['r̆iw]	[o]sê; [o]lhô
Alta [u]	Alta [u]	"nu" - ['nu]	[u]nu

Source: Rocha (2018a, p. 79).

In addition to the phonological constraint, which gives rise to the correspondence between the height of the stressed vowel of the base-word and the quality of the word-initial vowel agglutinated in the protolanguage, the female and male definite articles (respectively, "a" and "o") appear to restrict such agglutination to the low vowel [a] and the low mid, high mid and high back vowels [ɔ o u]. Comparison among the systems resulting from this language's expansion over time indicates singular behaviour, however. In the endeavour to explain them, Hagemeijer (2009) warns:

> early proto-GGC is essentially the result of Edoid slaves acquiring Portuguese. Since this period corresponds by and large to the *société d'habitation*, there was arguably better access to the Target Language (TL) and therefore phonetic calquing upon Portuguese gender was additionally able to satisfy the Edoid constraint that nouns are vowel-initial (p. 43).

Today, the creoles spoken on the Island of São Tomé (Forro and Angolar) contain a considerable number of words beginning with consonants, indicating greater influence by the superstrate language and consequently poor productivity in the agglutination of word-initial vowels. In Forro, the phenomenon is limited to items agglutinated at that diachronic stage of the protolanguage. In Angolar, meanwhile, "the absence of this feature [...] suggests that this creole still remained in contact with Santome after the spread of the Proto-GGC to Principe and Annobon" (HAGEMEIJER, 2009, p. 44). Expansion is thought to have occurred first in Príncipe, giving rise to Lung'Ie, a creole that conserved a series of Edoid features, including agglutination, as a result of being "isolated at an early stage of creolization" (p. 44). As Fad'ambô subsequently moved away, the time gap between the separations of these two languages is considered to warrant a smaller number of agglutinated items in the protocreole that reached Annobón. In other words, the time interval corresponds to the expansion of the plantation society and the mass arrivals of slaves speaking Bantu group languages. Hagemeijer argued that the interruption or decrease in agglutination in the protocreole should not be attributed to those slaves. To a large extent, nouns of Bantu origin begin with a consonant, thus admitting the possibility that, in learning the protocreole, speakers of those languages related "agglutination to the [...] definite article *o* and therefore reanalyzed the morpheme boundary. It is, however, more likely that Portuguese was responsible for lexical restructuring" (p. 44). This proposal is confirmed by the observations of Zamora (as in HAGEMEIJER, 2009, p. 44), which attest to the variation common in items of contemporary Fad'ambô – "(o)po", "(ô)bôyô", "(o)man" e "(am)pan"[4] – in which agglutinated word-initial vowels may be present or absent.

In the language of Príncipe, the cases of [i] agglutination clearly involve harmony with the stressed front vowel [i and ɛ]. There are, however, items where a back high vowel is agglutinated, even when the stressed segment is frontal. To explain these, he suggested they involved a more recent diachronic process, apparently following the pattern: "If there is a front vowel in the stem and no round material (vowels or glides) elsewhere in the stem, [i]-agglutination is triggered; in all the other cases, [u]-agglutination occurs by default" (p. 39). This would ratify occurrences of variation in Lung'Ie, such as "idêntu~udêntu 'dentro', ifi~ufu~ifu 'fio, arame', usolu~isolu 'sol'" (p. 39), indicating the generalisation of other patterns and consequent replacement by [u].

Preference for [u] in Lung'Ie and for [ɔ] and [o] in Fad'ambô show that, in these systems, the Edoid constraint was conserved: "reanalysis phonetically based on the Portuguese gender distinction was lost and that the [...] masculine definite article was retained as input for agglutination" (p. 43-44). However, while in Fad'ambô, agglutination seems to relate to the same harmony process found in the Gulf of Guinea protocreole (p. 38), in Lung'Ie, the agglutination of [u] appears to reproduce other processes pre-existing in the Edo substrate.

Summarising, vowel assimilation and harmony (by height and/or ATR) are phenomena common to all the Gulf of Guinea creoles (p. 37) that coexist with Portuguese on the islands of São Tomé and Príncipe.

Vowel Harmonisation in São Tomé Portuguese

Rocha (2018a) set out to analyse the behaviour of pre-stress mid vowels in São Tomé Portuguese, by

[4] Examples cited in Hagemeijer (2009, p. 44) on the basis of the *Gramática descriptiva del fa d'ambô* drawn up by Zamora.

evaluating social, structural, lexical and multiple language contact influences on processes of timbre raising/ maintenance in the stress position in question, drawing on other approaches to the same issue in the Brazilian and European varieties of the Portuguese language. She also intended to determine whether the vowel system of STP lay within an Afro-Brazilian or Afro-European *continuum*, or constituted a particular linguistic system. The *corpus* that served as the basis for her study included data drawn from interviews that constituted the samples of the *VAPOR Project,* at the Centro de Linguística of the Universidade de Lisboa. This involved 17 Informant-Documenter Dialogue (IDD)-type surveys conducted by Professor Tjerk Hagemeijer, of L1 Portuguese-speaking residents of the Island of São Tomé, distributed by sex, three age ranges and three levels of schooling.

The overall sample comprised **11179** tokens, in which 6643 data, i.e., 59.4%, referred to the vowel /e/. By concentrating the investigation on the [i] and [e] pronunciations, the initial number was reduced to **5915** cases distributed by variant as in Table 8.

Table 8 – Distribution of tokens of raising and of maintenance of the vowel /e/ in the urban variety of STP

Variants	Examples	Tokens	%
[i]	**[i]**xemplo, p**[i]**ssoa, **[i]**mprego, **[i]**scola	3.371	57
[e]	dif**[e]**rente, **[e]**conomia, r**[e]**alidade	2.544	43

Source: Rocha (2018a, p. 125).

It was hypothesised that incidence of high pronunciation would be greater when the vowel is word-initial: i) in headless syllables or in coda; ii) in those whose head position is available and the coda is filled by /S/ or /N/; and iii) in the sequence /deS/, whether prefixal or not. In addition to these environments, raising was thought more likely in vowels in hiatus contexts; and in the items "depois" and "pessoa(s), pessoal". All the cases listed were disregarded in the final analysis, either for high rates of raising or maintenance, or for low productivity. The final sample thus evaluated **3334** tokens of mid non-back vowels, redistributed among the [i] and [e] variants in Table 9.

Table 9 – Redistribution of tokens of raising and of maintenance of the vowel /e/ in the urban variety of STP

Pre-stress front	Examples	Token	%
[i]	apar**[i]**cer, m**[i]**lhor	1.216	36.5
[e]	qu**[e]**rer, d**[e]**v**[e]**ria	2.118	63.5

Source: Rocha (2018a, p. 128).

The chosen run (.42 input and 0.010 significance) established that the variables most influencing the raising of /e/ were, respectively: the vowel of the following syllable, the manner of articulation of the preceding consonant, the manner of articulation of the consonant of the following head, the place of articulation of the following head consonant, the proximity of a high vowel in the word, the word's grammatical class, the informant's schooling, the nature of the atonicity, the informant's age range, the informant's sex/gender and the frequency with which they used Forro.

As can be seen, after the (semi)categorical contexts were disregarded, the most decisive factor in the raising of /e/ in the urban speech of São Tomé Portuguese was the quality of the vowel of the syllable contiguous to the pre-stress position, whose factors and respective indices are distributed in Table 10.

Table 10 – Rates of raising of the pre-stress vowel /e/, by contiguous vowel type, in the urban variety of STP

Contiguous Vowel Type		Examples	[i]		
			Token	%	RW
Stressed/ unstressed vowel (oral/ nasal/ nasalised)	**High (homorganic and non-homorganic)**	seg**ui**nte, Rep**ú**blica	550/1088	**51**	**0.55**
	High-mid	dife**re**nça, dep**o**is	457/1290	**35**	**0.54**
	Low (low-mid and low)	lev**a**ntamento, esp**e**ro, melh**o**ra	209/956	22	0.39

Source: Rocha (2018a, p. 129).

At first, it was thought that the different heights of the contiguous vowel, and also the homorganic and non-homorganic nature of the high vowels, would influence appearance of raising of the pre-stress /e/. However, the finding of similar behaviour patterns made it possible to redistribute these into three degrees of openness: high (homorganic and non-homorganic vowels), high-mid and low (low-mid and low). Although the relative weights were very close to neutral (RW 0.50), the [i] pronunciation was more likely when the contiguous syllable contained a high vowel (whether or not homorganic) (RW 0.55) or high-mid vowels (RW 0.54). On the other hand, contiguity with low segments favoured maintenance of timbre in /e/ (RW 0.38).

Final Remarks

Overall, the Variationist studies of pre-stress vocalism in the urban varieties of Brazilian and São Tomé Portuguese considered here indicated that timbre is predominantly maintained, to the detriment of raising. In Yacovenco (1993), Rocha (2013) and Rocha (2018a), once the contexts in which the [i] pronunciation is (semi)categorical have been discarded, the variable that most influences in favour of application of the raising rule is the contiguous vowel type. The three studies observed the timbre of the vowel of the syllable following the pre-stress /e/. However, while Yacovenco (1993) and Rocha (2013) distributed the high vowels into homorganic and non-homorganic, Rocha (2018a) grouped these segments into a single factor. Also, while Yacovenco (1993) limited herself to observing the influence of stressed vowels in a subsequent context and highlighted the high, low and mid timbres, Rocha (2013) distributed the variants by stress (stressed and unstressed) and restricted the timbres to high and non-high. Lastly, Rocha (2018a) grouped stressed and unstressed segments and offered three possible timbres: high, high-mid and low (low-mid and low).

Disregarding the peculiarities of each study as regards the variable examined, all the studies emphasise the contiguous high vowel as the factor that most influences raising of the front mid vowel, confirming the importance of harmonisation with the high feature of the contiguous vowel in application of the rule. In BP, homorganicity is another reference factor, because the vowel [i] stood out more than [u] in both Yacovenco (1993) and Rocha (2013). That distinction appears not to be that important in the urban speech of São Tomé, because these data were grouped into a single variant and raising became significant when, in the syllable following the pre-stress vowel, the vowels [i] or [u] (whether or not stressed) followed /e/ (0.55). Accordingly, the vowel harmonisation process found in the studies of BP and STP examined here, and the observations with regard to the Gulf of Guinea creoles, appear to confirm the hypothesis of proximity between Portuguese-based creoles and the speech of Rio de Janeiro and São Tomé, as regards their insubordination to the reduction rule that predominates in EP (MARQUILHAS, 2003, p. 7).

References

AMARAL, A. *O dialeto caipira.* 2. ed. São Paulo: HUICITEC; Secretaria da Cultura, Ciência e Tecnologia, 1976 [1920].

BISOL, L. A. *Harmonia vocálica:* uma regra variável. 1981. Thesis (Doctorate in Linguistics) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1981.

BISOL, L. A. A neutralização das átonas. *DELTA*, São Paulo, v. 19, n. 2, p. 267-276, 2003.

BRANDÃO, S. F.; CRUZ, M. L. C. Um estudo contrastivo sobre as vogais médias pretônicas em falares do Amazonas e do Pará com base nos dados do ALAM e do ALISPA. *In*: AGUILERA, V. de A. (ed.). *A geolinguística no Brasil:* trilhas seguidas, caminhos a percorrer. Londrina: EDUEL, 2005. p. 299-318.

BRANDÃO, S. F.; ROCHA, F. M. V. ; SANTOS, E. R. Vogais médias pretônicas em início de vocábulo na fala do Rio de Janeiro. *Letras & Letras:* Estudos em Fonologia, Uberlândia, v. 28, p. 244-257, 2012.

CALABRESE, A. Metaphony in Salentino. *Rivista di Grammatica Generativa,* Venice, v. 10, p. 9-10, 1985.

CALLOU, D.; LEITE, Y. As vogais pretônicas no falar carioca. *Estudos Linguísticos e Literários,* Salvador, v. 5, p. 151-162, 1986.

CALLOU, D.; LEITE, Y.; COUTINHO, L. Elevação e abaixamento das vogais pretônicas no Rio de Janeiro. *Organon,* Porto Alegre, v. 18, p. 71-78, 1991.

CAMARA JR, J. M. *Estrutura da língua portuguesa.* Rio de Janeiro: Vozes, 1970.

CAMARA JR, J. M. *Para o estudo da fonêmica portuguesa.* Rio de Janeiro: Padrão, 1977.

CARDOSO, S. As vogais pretônicas no Brasil: uma visão diatópica. *In*: AGUILERA, V. de A. (ed.). *Português no Brasil:* estudos fonéticos e fonológicos. Londrina: Editora da UEL, 1999. p. 93-124.

CARMO, M. N. *As vogais médias pretônicas dos verbos na fala culta do interior paulista*. 2009. Dissertation (Masters in Linguistics) – Universidade Estadual Paulista, São Paulo, 2009.

COLLISCHONN, G.; SILVA, M. Elevação das médias pretônicas por harmonia: questões teóricas e empíricas. *Working Papers em Linguística,* Florianópolis, v. 14, n. 2, p. 1-14, 2013. Available in: https://bit.ly/3qcUJqq. Access in: 20 nov. 2017.

FERRAZ, L. I. *The creole of São Tomé.* Johanesburgo: Witwatersrand University Press, 1979.

FERRAZ, L. I. Portuguese creoles of West Africa and Asia. *In*: GILBERT, G. G. (ed.). *Pidgin and creole languages:* essays in memory of John E. Reinecke. Honolulu: University of Hawai Press, 1987. p. 337-360.

GRAEBIN, G. de S. *A fala de Formosa/GO:* a pronúncia das vogais médias pretônicas. 2008. Dissertation (Masters in Linguistics) – Universidade de Brasília, Brasília, 2008.

HAGEMEIJER, T. Initial vowel agglutination in the Gulf of Guinea creoles. *In*: ABOL, E.; SMITH, N. (ed.). *Complex processes in new languages.* Amsterdam; Philadelphia: John Benjamins, 2009. p. 29-50.

HAGEMEIJER, T. O Português em contacto em África. *In:* MARTINS, A. M.; CARRILHO, E. (ed.). *Manual de Linguística Portuguesa.* Frankfurt; Leipzig: Deutsche Nationalbibliothek, 2016. p. 43-67.

HORA, D. da; SANTIAGO, S. Vogais pretônicas no Norte do Brasil: o falar de Macapá. *In*: RAMOS, J. (ed.). *Estudos sociolingüísticos:* os quatro vértices do GT da ANPOLL. Belo Horizonte: UFRMG, Faculdade de Letras, 2006. p. 21-36.

LABOV, W. *Sociolinguistics patterns.* Oxford: Blackwell, 1972.

LABOV, W. *Principles of linguistic change.* v.1: Internal factors. Cambridge: Blackwell, 1994.

LABOV, W. *Principles of linguistic change.* v. 2: Social factors. Cambridge: Blackwell, 2001.

MARQUES, S. M. O. *As vogais médias pretônicas em situação de contato dialetal.* 2006. Thesis (Doctorate in Portuguese Language) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2006.

MARQUILHAS, R. Mudança analógica e elevação das vogais pretónicas. *In*: CASTRO, I.; DUARTE, I. (ed.). *Razões e emoção.* Miscelânea de estudos em homenagem a Maria Helena Mira Mateus. Lisboa: Imprensa Nacional; Casa da Moeda, 2003. p. 7-18.

MARROQUIM, M. *A língua do Nordeste.* 3. ed. Curitiba: HD Livros, 1996 [1934].

NASCENTES, A. *O linguajar carioca.* 2. ed. Rio de Janeiro: Organização Simões, 1953 [1922].

OLIVEIRA, M. A. Variação fonológica: o indivíduo e a comunidade de fala. *In*: SEMINÁRIO DE PESQUISA DO PPGLLP – UNESP, 8., 2008, Araraquara. (Oral presentation).

ROCHA, F. de M. V. da. *O comportamento das vogais médias pretônicas na fala de Nova Iguaçu*. 2013. Dissertation (Masters in Vernacular Letters/Portuguese Language) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2013.

ROCHA, F. de M. V. da. *O sistema vocálico do Português de São Tomé e o comportamento das vogais médias em contexto pretônico.* 2018. Thesis (Doctorate in Vernacular Letters/Portuguese Language) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2018a.

ROCHA, F. de M. V. da. A pretônica /e/ na variedade urbana do Português de São Tomé. *In:* Grupo de Trabalho de Sociolinguística do XXXIII ENANPOLL, UFMT, Cuiabá, 2018b. (Oral presentation).

SCHWINDT, L. C. A regra variável de harmonização vocálica no RS. *In*: BISOL, L.; BRESCANCINI, C. (ed.). *Fonologia e variação:* recortes do Português Brasileiro. Porto Alegre: EDIPUCRS, 2002. p. 161-182.

SILVA, M. B. da. Um traço regional na fala culta de Salvador. *Organon,* Porto Alegre, v. 18, p. 79-89, 1991.

SILVA NETO, S. da. *História da língua portuguesa do Brasil.* 5. ed. Rio de Janeiro: Livros de Portugal, 1988.

SILVEIRA, A. A. M. *As vogais pretônicas na fala culta do noroeste paulista.* 2008. Dissertation (Masters in Linguistics Studies) – Universidade Estadual Paulista, São José do Rio Preto, 2008.

TEIXEIRA, J. A. O falar mineiro. *Separata da Revista do Arquivo municipal de São Paulo*, São Paulo, 1938.

TEYSSIER, P. *A história da língua portuguesa.* Lisboa: Sá da Costa, 1980.

VIEGAS, M do C. *O alçamento de vogais médias pretônicas e os itens lexicais*. 2001. Thesis (Doctorate in Letters) – Universidade Federal de Minas Gerais, Belo Horizonte, 2001.

WEINREICH, U.; LABOV, W.; HERZOG, M. Empirical foundations for a theory of language change. *In*: LEHMANN, W.; MALKIEL, Y. (ed.). *Directions for historical linguistics*. Austin: University of Texas Press, 1968. p. 97-195.

YACOVENCO, L. C. *As vogais médias pretônicas no falar culto carioca*. 1993. Dissertation (Masters in Portuguese Language) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1993.