Isogloss
Open Journal of Romance Linguistics

Received: 29-02-2024
Accepted: 04-11-2024
Published: 22-03-2025

Stressed Vowels in São Tomé and Príncipe Portuguese (STPP): Acoustic Space (F1 and F2) and [ə] Production

Amanda Macedo Balduino
Universidade Estadual de Campinas
amandamb@unicamp.br

Shirley Freitas
Universidade da Integração Internacional da Lusofonia Afro-Brasileira
shirleyfreitas@unilab.edu.br

How to cite: Balduino, Amanda Macedo & Shirley Freitas. 2025. Stressed Vowels in São Tomé and Príncipe Portuguese (STPP): Acoustic Space (F1 and F2) and [ə] Production. In *Romance minority languages: A challenge for linguistic theory*, eds. Michela Russo & Shanti Ulfsbjorninn. Special issue of *Isogloss. Open Journal of Romance Linguistics* 11(3)/10, 1–27.
DOI: https://doi.org/10.5565/rev/isogloss.417

Abstract

This paper describes 2 acoustic correlates of oral stressed vowels in Portuguese of São Tomé and Príncipe (STPP): First formant (F1) and second formant (F2). Based on data from Escudero et al. (2009), we compared STPP with urban varieties of Brazilian (BP) and European Portuguese (EP), such as those spoken in São Paulo and Lisbon, respectively. Based on the analysis of 1,776 occurrences of 74 lexical items produced by 6 male and 6 female participants, we attested the presence of 8 stressed vowels in STPP [i, e, ɛ, a, ə, ɔ, o, u] – a different result from Escudero et al. (2009), who found 7 oral stressed vowels in BP and EP [i, e, ɛ, a, ɔ, o, u]. [ə] as an allophone of /a/, in STPP, is also accompanied by a reduction in the acoustic space of this variety, which, in general, presents a more compressed acoustic space than BP and EP. This finding differentiates STPP – a minoritized Portuguese variety – from the urban varieties of BP and EP in question, as it produces a unique outcome among documented Portuguese varieties that lead to the observation of some phonological processes in stress position.

Keywords: São Tomé and Príncipe, formants, stressed vowels, vowel space, linguistic contact.

1. Introduction

São Tomé and Príncipe (STP) is a West African country characterized by a rich linguistic landscape that encompasses different languages. Notably, Portuguese serves as the mother tongue for 98.4% of the population, as reported in the last Census (Ine 2012). Consequently, STP stands out as the African nation where the Portuguese language exhibits the highest level of vitality. Studies such as Balduino's (2022) underscore that, while the current teaching model is based on European Portuguese – especially the norm of the Lisbon-Coimbra axis, which is the prestige norm taught in Santomean schools –, the archipelago has cultivated its distinct Portuguese variety known as São Tomé and Príncipe Portuguese (STPP). Balduino (2022) argues that, from a phonological point of view, STPP displays a certain degree of unity, with the same phonemes being shared between São Tomé Portuguese (PST)[1] and Príncipe Portuguese (PP), both of which are Portuguese microvarieties spoken within STP. However, when examined from a phonetic perspective, distinct behaviors emerge within each variety. In light of this context, this paper focuses on the stressed oral vowels within STPP with the objective of delineating specific acoustic characteristics, including the first and second formants.

This research is justified by the limited body of work on STPP – dedicated studies on this variety are relatively recent. Moreover, research focused on the acoustic aspects of African varieties of Portuguese and, more broadly, of creole languages that emerge in contact situations, remains scarce. This scarcity hinders our ability to deepen our understanding of these varieties and to make meaningful comparisons with others. Therefore, this paper serves as a crucial step towards expanding our knowledge of STPP phonetics and phonology. This knowledge, in turn, paves the way for future comparisons with well-documented varieties, such as Brazilian Portuguese (BP) and European Portuguese (EP).

The paper is organized as follows: Section 2 introduces STPP and its current sociolinguistic situation and briefly discusses the stress vowels in PST and PP. Section 3 describes the methodology used and, in section 4, we propose an acoustic characterization of stressed vowels of STPP. We considered the values of the first and second formants in Hertz, and the values found were compared with those of other urban varieties of Portuguese, such as European and Brazilian. It is worth noting that our comparison will occur primarily with the varieties spoken in São Paulo and Lisbon, which were included in the study by Escudero et al. (2009).[2] Therefore, although we

[1] We shall use the acronym PST in reference to São Tomé Portuguese, as the use of STP may cause confusion with that used in reference to the country (São Tomé and Príncipe - STP).

[2] It is important to emphasize, however, that the comparison presented in this paper is primarily based on the availability of literature. This is due to substantial differences between the methods employed in this study and those used by Escudero et al. (2009). Nonetheless, given that this is a pioneering work aimed at providing an initial phonetic description of stressed vowels in São Tomé and Príncipe Portuguese (STPP), and at discussing our findings in relation to the available literature on EP and BP, while examining similarities and

generally refer to these varieties as Brazilian Portuguese (BP) and European Portuguese (EP), we recognize the constitutive diversity of these varieties, which are composed of different and diverse dialects. Finally, section 5 brings the study's final considerations.

2. The Portuguese Spoken in São Tomé and Príncipe

The Democratic Republic of São Tomé and Príncipe is one of the 9 countries where Portuguese is spoken as the official language. According to the National Institute of Statistics (INE), Portuguese is spoken by 98.4% of the population, while Santome (ISO 639-3: cri) and Lung'Ie (ISO 639-3: pre) – local creole languages – are increasingly less used and transmitted in urban centers (Agostinho 2015; David 2018; Bouchard 2019).

After the independence of STP in 1975, earlier generations acquired creole languages at home and, subsequently, learned Portuguese through formal education (Bouchard 2019; David 2018; Araujo 2020). However, aware of the social role played by Portuguese, these Portuguese speakers (as a second language) began to pass it on to the new generations, who, in turn, acquired it as their mother tongue, but not without adding their own specificities along the way (Gonçalves 2010; Pereira, Hagemeijer; Freitas 2018; Bouchard 2019; Araujo 2020). As a consequence, not only is Portuguese present in STP, but it has also been "Santomized" (Balduino, Bandeira e Freitas 2022).[3]

São Tomé and Príncipe Portuguese (STPP) presents unique features, often resulting from contact with other languages spoken in the archipelago and the type of linguistic transmission that occurred there, therefore constituting a vast field of linguistic analysis yet to be explored (Baxter 2018; Figueiredo 2010; Bouchard 2017; 2018; David 2018; Brandão et al. 2017; Balduino 2018; 2022; Braga 2018; Pereira, Hagemeijer; Freitas 2018; Araujo 2020; Agostinho; Mendes 2021; among others).

Even in face of the emergence of a distinct local variety of Portuguese in STP – STPP –, this variety is often equated with the linguistic norm of the Lisbon-Coimbra axis of European Portuguese, commonly assessed negatively and, thus, delegitimized in its differences. This leads to a contrasting reality in STP: Even though Portuguese is the most widespread language, often to the detriment of other languages in the archipelago, there is still no standardized national variety recognized by official bodies.

The discrimination to which STPP and other non-European varieties of Portuguese are subjected reinforces their confinement to a marginal social space, as these varieties are denied official participation in schools, media, and other institutional spaces. They are politically and ideologically minoritized. The urgency of establishing a linguistic norm specific to STP is thus evident, and linguistic research plays a crucial role in changing the concept of social marginalization assigned to

differences among urban varieties of Portuguese spoken in diverse sociolinguistic contexts, we believe that this initial comparison can be productive.

[3] For a discussion on the features that distinguish STPP from other varieties of Portuguese, as well as from other languages resulting from linguistic contact, see Agostinho (2016); Bandeira (2017); Pereira, Hagemeijer, and Freitas (2018); Balduino (2018; 2022), Agostinho; Mendes (2021), among others.

STPP. In this regard, linguists play an important role, as evidenced by studies dedicated to African varieties of the Portuguese language in STP (Baxter 2018; Figueiredo 2010; Bouchard 2017; 2018; Brandão et al. 2017; Balduino 2018; 2022; Braga 2018; Pereira, Hagemeijer; Freitas 2018; Araujo 2020; Agostinho; Mendes 2021; among others). Among the cited studies, all dedicated to the Portuguese of São Tomé and Príncipe, the emergence of local varieties of Portuguese is an aspect that has been systematically suggested by different research from various theoretical and analytical linguistic perspectives. Therefore, this paper does not only bring forth kindred debate but also adds to it, by providing documentary support for the linguistic and social autonomy of different Portuguese language varieties through the description and analysis of stressed vowel acoustics.

Beyond the macrovariety we refer to as STPP and considering that Portuguese is widely spoken and transmitted in the country, STP is characterized by the local diversity of Portuguese language varieties (cf. Figueiredo 2011; Baxter 2018). In relation to urban varieties, this paper focuses on São Tomé Portuguese (PST) and Príncipe Portuguese (PP), varieties spoken in the cities of São Tomé and Santo António do Príncipe, respectively.

The distinction between these varieties is relevant because, in addition to the existing geographical distance between the islands of São Tomé and Príncipe (approximately 140 kilometers by sea), PST and PP display unique linguistic ecologies and are different in terms of diachronic and synchronic contact. PST, for instance, is primarily in contact with Santome in urban areas. This linguistic contact occurred diachronically, considering the formation and emergence of PST as a national variety, and it can also be currently observed, given that PST, once established, continues to be in contact with Santome. It is worth noting that PST is also in contact with: (i) Angolar (ISO 639-3: aoa), a creole language with great vitality that originated and is spoken in the interior of the island, such as in the Angolares region (Bandeira 2017) and (ii) Kabuverdianu (code ISO 639-3: kea), a creole language from the island of Cape Verde that arrived in STP through Cape Verdean contract workers during the second period of colonization (18th and 20th centuries). This contact arises from the mobility of speakers, involving daily movements between the interior of the island and the urban center that is the city of São Tomé.[4]

The formation of PP, in turn, refers to a diachronic contact with Lung'Ie. Nowadays, the contact mainly occurs with Kabuverdianu, since Príncipe Island received the largest contingent of Cape Verdean migrants. Besides, even in a scenario of widespread disuse of Lung'Ie, one must consider that, beyond the songs composed and performed in Lung'Ie that circulated on the island, this language was taught on the island of Príncipe as an optional subject without assessments between 2009 and 2015, covering from pre-school to the 11th grade. Since 2016, however, the discipline became mandatory from the 5th grade on (Agostinho; Bandeira; Araujo 2016), suggesting that, even at the level of a second language, PP is currently in contact with Lung'Ie.

Considering both scenarios, this article will first analyze PST and PP separately, paying attention to the possibility of structural change promoted by

[4] The native creole languages (Angolar, Santome, and Lung'Ie) and circulating creole (Kabuberdianu) exhibit phonological differences among themselves and in comparison to Portuguese.

linguistic contact or even by the process of acquiring Portuguese as a mother tongue by past generations, based on input that was originally a second language. Having done that, the data will be evaluated jointly. Prior to that, however, we shall present a brief discussion on the presence of stressed vowels in PST and PP according to the literature.

2.1. Stressed vowels in STPP

STPP has a 7-vowel system organized in terms of frontness, backness, and height: /i, e, ɛ, a, ɔ, o, u/ (Christofoletti; Araujo 2018; Balduino 2022; Santiago 2019; Balduino; Freitas 2022). Christofoletti and Araujo (2018) and Santiago (2019) examined PST and PP, respectively, and neither attested to the presence of [ə] in the phonetic inventory of the stressed vowels of these varieties. A different result came from Balduino (2022), who identified the realization of [ə] by PST speakers but not by PP speakers.

While Christofoletti and Araujo (2018) and Santiago (2019) carried out an auditory analysis, Balduino (2022) examined the vowel formants of stressed vowels to observe phonological processes. According to Balduino, /a/ in PST is often raised, particularly in unstressed syllables and, less commonly, in stressed syllables. In order to expand the studies dedicated to vowels in PST and PP, we propose an acoustic description of F1 and F2 of the stressed vowels in these varieties to later evaluate them jointly. In addition to producing a new acoustic description of these varieties, we seek to examine if /a/ exhibits allophones, as pointed out by Balduino (2022), and whether this production has influence over the acoustic space of stressed vowels in PST and PP – and even in STPP as the macrovariety of STP. For that, we will compare our results with those of Escudero et al. (2009).

3. Method

3.1. Fieldwork and materials

Data were collected during fieldwork in the cities of São Tomé, on the island of São Tomé, and Santo António, on the island of Príncipe, in 2019. Since the nineteenth century, fieldwork has been an integral part of linguistic research, characterized by its documentary and descriptive nature. This approach demands multidisciplinary scientific and interpersonal skills from the researcher (Brickell 2018). Additionally, fieldwork enables in situ data collection, allowing for the observation of the sociolinguistic contexts in which the target language is used.

Fieldwork presents several challenges. First, it demands careful preparation and the respectful immersion of the researcher into the speech community. This process is time-consuming and often requires multiple visits to the field. In our study, although the data was collected in 2019, the researcher, Amanda Macedo Balduino, had been engaged with the community since 2016. To ensure productive data collection during this fieldwork period and to gather a substantial amount of data within a limited time frame, it was crucial to take into account the cultural and historical context in which the target language varieties are spoken (Vaux & Cooper, 1999). Thus, during the fieldwork, it was necessary to invest time in engaging with the speech community, seeking insertion into various social, bureaucratic, and/or

everyday environments that facilitated interaction between the researcher and speakers (Vaux; Cooper 1999; Bowern 2008; Brickell 2018). After this stage, additional challenges emerged, particularly regarding the data collection environment, which lacked laboratories, specialized equipment, and sometimes even basic infrastructure that could enhance the quality of the recordings. However, being embedded in the Santomean and Príncipe communities enabled us to recognize local needs and plan the recording sessions accordingly, aiming to collect high-quality data despite the structural limitations of São Tomé and Príncipe. In the absence of acoustic laboratories, recordings were conducted in private rooms within public spaces provided by local institutions, such as the Príncipe radio station and the Brazilian Embassy in São Tomé. These locations were generally quiet and allowed the sessions to be conducted with only the informant present, reducing external noise interference. Additionally, considering potential issues with local electricity, all test materials were pre-printed and presented to the informants in hard copy to avoid disruptions.

Considering these issues related to fieldwork, for the analysis of the stressed vowel system, we examined a total of 74 distinct lexical items comprising one of the 7 vowels /i, e, ɛ, a, ɔ, o, u/ in a stressed syllable. This consisted of 14 words where the stress was placed on the vowel /a/ in the stressed syllable, and an additional 10 words in which the stress corresponded to the vowels /i, e, ɛ, ɔ, o, u/, as outlined in the Table of the Appendix.

Data collection involved an elicitation process. During the recording sessions, conducted using an Olympus DM-680 microphone, participants were presented with an image. Subsequently, they were instructed to name the action or object depicted in such image and then repeat the target word within the carrier phrase “Eu falo X baixinho,” “I speak X softly.” A total of 74 words were elicited through the repetition of this carrier phrase, with ‘X’ being replaced by the specific target item. Each participant was provided 3 instances of each word, and the initial round of recordings was excluded from the analysis.

We conducted an analysis of data collected from 12 participants, evenly split between 6 men and 6 women. Our dataset consisted of 74 words, each containing the target vowel in a stressed syllable. For each participant, we considered 148 occurrences (74 lexical items, each repeated twice). In total, we recorded 1,776 occurrences during the experiment (148 occurrences per participant, multiplied by 6 male participants and 6 female participants). However, due to fieldwork conditions, approximately 31% of the initial data either could not be collected or had to be excluded from the analysis. This was mainly due to issues such as external noise impacting audio quality or, in rare instances, participants providing responses with synonyms or entirely different words. Our final dataset consisted of 1,230 tokens – 609 for PST and 608 for PP. The fraction of discarded data was approximately 15% for each variety in relation to the initial experiment design. To ensure that the lexical items used in the experiment were universally understood by the speakers, we took the precaution of having a STPP speaker recognize and validate the images and vocabulary used during the test preparation phase.

While selecting the words for our study, our criterion was the presence of one of the 7 vowels in a stressed syllable. It is important to note that this work represents a pioneering effort, and the elicited corpus primarily serves as a general documentation resource. Consequently, we were unable to exercise control over the specific segmental and syllabic contexts that surround and encompass the target vowels. As a

result, there is inherent variation in the phonetic context within the corpus, and this variation has the potential to influence the acoustic measurements conducted in this research.

As shown in Table 1, most of the analyzed vowels originate from 2 primary syllabic structures: CVC (50/74, 67.6%) and CV (23/74, 31.1%). In the case of CVC syllables, the onset consonant can encompass a range of possibilities, including stops (both voiceless and voiced) such as /p, b, t, d, k, g/, fricatives (both voiceless and voiced) like /f, v, s, z/, nasals /m, n/, and liquids /r, ʎ/. Within the analyzed dataset, the coda of these CVC syllables predominantly featured either /r/ or /l/ – in addition to these liquids, we identified, as proposed by Balduino (2022), /N/ and /S/ occupying the coda position in STPP. It is worth noting that liquids are frequent subjects of phonological processes in STPP (XXX). Regarding rhotics in the onset position, they can manifest as a tap [ɾ], an alveolar trill [r], a uvular trill [ʀ], or even a uvular fricative [ʁ]. In the coda position, we observed the same variations along with instances of rhotic deletion. As for laterals in the coda, they can undergo deletion, velarization [ɫ], or vocalization [w].[5] In Table 1, our focus is primarily on the segmental contexts of CVC and CV.C syllables. There was only one occurrence of the VC syllable (1/74, 1.35%), represented by the word *urso* [ˈuɾ.sʊ], 'bear' (as seen in Table 1). In this case, the coda is occupied by a rhotic.[6]

Table 1. Segmental Context of the Stressed Vowels in the Corpus

CVC (67.6%)				**CV.C (31.1%)**		
/t, f, v, m/	i	/r, l/		/k, m, n/	i	/t, g, n, l/
/d, v, m, r/	e	/r, l/		/s, v, m, l/	e	/t, g, l, ʎ/
/d, p, s, m, ʎ/	ɛ	/r, l/		/t, n/	ɛ	/k, r, l/
/d, b, k, g, s, n, m/	a	/r, l/		/t, d, k, f, l/	a	/t, d, b, s, v/
/t, p, b, k, s, f, r/	ɔ	/r, l/		/b, v/	ɔ	/b, z/
/t, d, p, b/	o	/r, l/		/b, g, s/	o	/l, z, r/
/d, p, k, z/	u	/r, l/		/t, s, n, r/	u	/k, g, v, ʒ, n/

The segmental contexts, as well as the phonological processes that impact the consonants specified in Table 1, particularly in the coda position, are noteworthy as they highlight the significant variability concerning the manner and point of articulation of consonants coarticulated with stressed vowels. This variability has the potential to influence the measured formant frequencies in Hz. To mitigate such interference, we employed the Linear Prediction (LPC) technique. Subsequently, we

[5] As noted by one of the reviewers of this article, considering the syllabic structure, it is well-established that liquids, which are more sonorous than obstruents, tend to exhibit a preference for the nucleus. This is because liquids often function as syllabic consonants. Therefore, it is preferable to avoid including liquids (as well as rhotics) in the syllable under analysis for two main reasons: (i) the ordering of consonants in onset and coda clusters is typically asymmetrical, with liquids often positioned closer to the nucleus, and (ii) their inclusion could affect the analysis of the data.

[6] A more comprehensive discussion of syllables in São Tomé and Príncipe Portuguese (STPP) can be found in Balduino (2020), Balduino (2022) and Balduino (2023b).

conducted formant measurements exclusively within the central spectral portion of the vowel, where the formant frequencies remained stable and stationary.

3.2. Participants

The recordings were conducted with twelve participants – 6 men and 6 women. All of them had medium to high level of education,[7] and were within the age range of 18 to 23 years old. They are from the cities of São Tomé, in São Tomé Island, and Santo António, located on Príncipe Island. These participants are speakers of PST (6 participants, 3 men and 3 women) and PP (6 participants, 3 men and 3 women). In terms of their proficiency in a second language (L2), we observed varying degrees of bilingualism, which is a common occurrence in an environment of rich linguistic contact like STPP.[8] Consequently, in addition to their proficiency in the Portuguese language (the mother tongue of all participants), 4 of them were also proficient in other languages spoken in the archipelago, such as Santome/Forro and Kabuverdianu.

Table 2. Participants – Controlled Speech

Info.	Age	Sex	Schooling	Another spoken language	Variety
I	18	F	High	Kabuverdianu	PST
II	23	M	Medium	Kabuverdianu	PST
III	19	M	High	Santome	PST
IV	22	F	Medium	---	PST
V	18	M	High	---	PST
VI	18	F	High	---	PST
VII	18	M	High	---	PP
VIII	18	F	High	---	PP
IX	18	F	High	---	PP
X	18	M	High	---	PP
XI	19	F	High	Kabuverdianu	PP
XII	20	M	High	---	PP

The four bilingual speakers reported that they learned the Creole language alongside Portuguese in a home environment. Thus, Portuguese is reported by the speakers as their mother tongue alongside Kabuverdianu or Santome. All the bilingual speakers reported using the Creole languages with family members, especially older relatives such as parents and mainly grandparents. Portuguese is also used in these family circumstances, in addition to being used daily in public contexts and among friends (younger generation). In the specific case of Kabuverdianu, the speakers also reported using the language within the Kabuverdianu community beyond the family

[7] The level of education was divided as follows: medium (8-9 years) and high (10-12 years).

[8] Please note we relied on the participants' self-reported proficiency in another language; proficiency tests were not administered. Therefore, in future data collection efforts, it is possible to devise questions and tests to assess the proficiency in the various languages spoken by the participants in order to analyze the interference of one language in another in more detail.

circle. Finally, the research participants also stated that they listen to music written and sung in the Creole languages and in Portuguese and have frequent contact with both languages.

Fant (1966), Takefuta et al. (1972), Johnson (2005), and Pépiot (2009; 2012), among others, have highlighted that distinctions between male and female voices are influenced by a convergence of factors encompassing social, anatomical, and, consequently, articulatory, and acoustic aspects. Recognizing the potential for acoustic variations arising from differences in participants' voices, we conducted separate measurements in the STPP dataset where we factored in participants' sex.

3.3. Analysis instruments

To analyze the data, we manually extracted the first (F1) and second (F2) formants. This extraction was performed using Praat software (Boersma & Weenink 2020). Given the participants involved in this research and the distinct acoustic properties inherent to the various surveyed groups, we conducted separate analyses of the data produced by PST and PP speakers. Subsequently, we proposed a combined examination of the STPP dataset. We normalized the raw values of F1 and F2 using the Lobanov normalization method,[9] which effectively excludes physiological differences in formant values while preserving sociolinguistic differences. Additionally, for duration normalization, we employed the z-score method in RStudio.

When considering the methodological procedures employed in this study, it is important to underscore the challenges and limitations inherent to the process of acoustic analysis based on fieldwork data. While the recordings were conducted within enclosed environments, such as the Príncipe radio station and the Brazilian Embassy in São Tomé, they were not carried out in acoustically controlled cabins designed for precise acoustic studies. Consequently, there may be some ambient noise present that could potentially restrict the scope of our analysis. As previously mentioned, segments with compromised spectral quality due to audio issues were excluded from our analysis. This revised version maintains the original content while enhancing clarity and readability.

4. Acoustic characterization of stressed vowels in STPP: first and second formants (F1 e F2)

Regarding vowel formants, we present the normalized values extracted in table 3 in different varieties spoken in STP. In this table, the average of the first formants (F1 and F2) for each of the 7 stressed vowels is the benchmark for the acoustic definition of stressed vowels in the macrovarieties of Portuguese spoken in STP. This table serve as a reference, wherein the average of the first formants (F1 and F2) for each of the 7 stressed vowels plays a pivotal role in defining the acoustic characteristics of stressed vowels across the macrovarieties of Portuguese spoken in STP.

[9] The Lobanov method presents the following formula for normalization: Fn[v]N = (Fn[v] - mean(n)) / SD(n), where Fn[v]N is the normalized value for Fn[v] (i.e., for formant n of vowel v), mean(n) corresponds to the average value for formant n for the speaker in question, and finally, SD(n) corresponds to the standard deviation for formant n of the speaker (http://lingtools.uoregon.edu/norm/norm1.php)

Table 3. Normalized Average Values (in Hertz) of F1 and F2 for Stressed Vowels - PST and PP - Men and Women

Male Speakers						Female Speakers				
Vowel	PST		PP			Vowel	PST		PP	
	F1	F2	F1	F2			F1	F2	F1	F2
i	303	2153	253	2136		i	315	2091	289	2061
e	386	2016	366	1958		e	380	2037	351	1897
ɛ	504	1810	501	1886		ɛ	498	1926	573	1866
ə	530	1389	521	1466		ə	536	1503	---	---
a	625	1350	619	1393		a	740	1387	733	1317
ɔ	537	919	562	901		ɔ	547	934	592	1162
o	413	952	397	889		o	407	957	354	940
u	339	937	299	928		u	310	761	303	814

Upon examining the average formant values produced by the speakers, we observed that the dispersion areas of F1 and F2 for each vowel are constrained, effectively maintaining the segmental distinctions between the analyzed vowels. These findings are further clarified by the graphical representations in Figures 1 and 2.

Figure 1. Stressed Vowels in PST

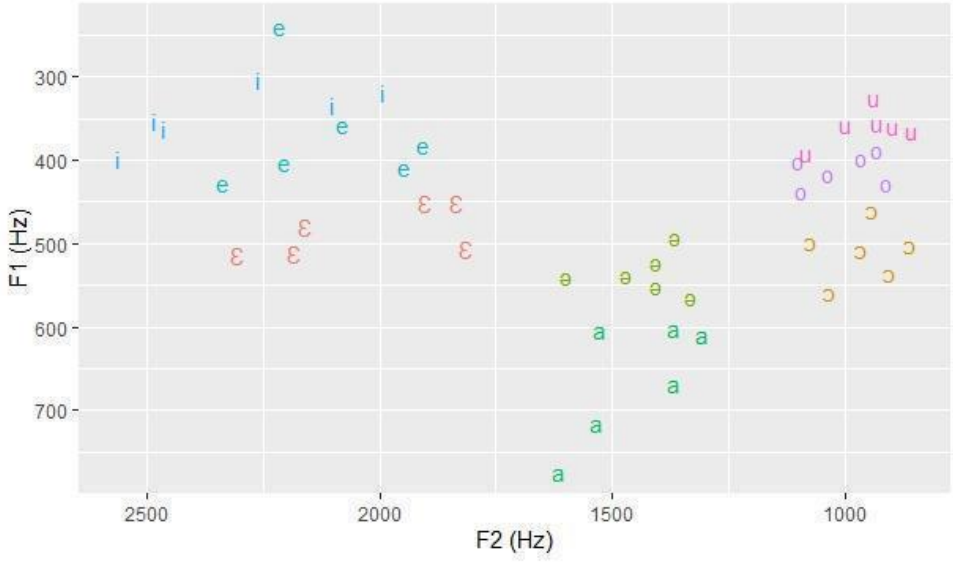
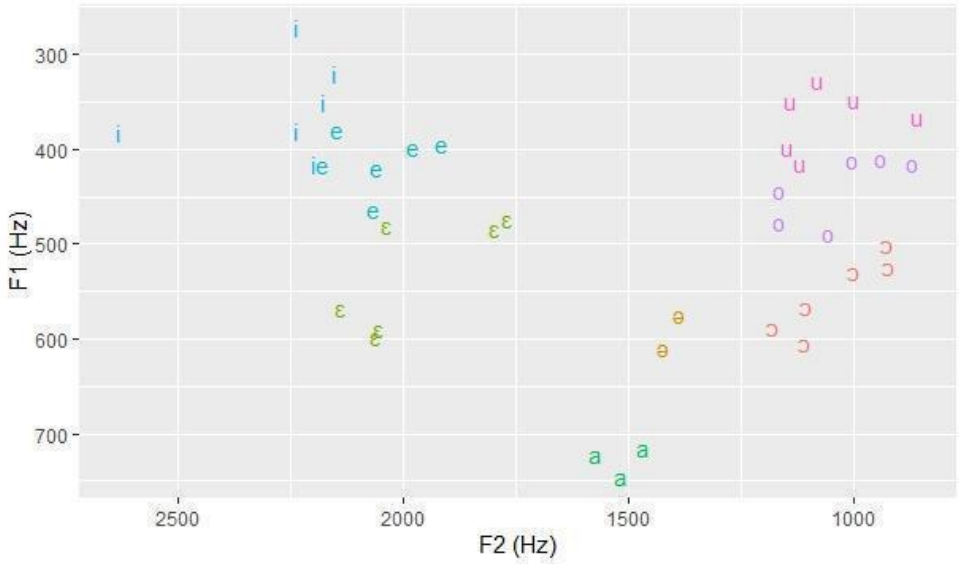


Figure 2. Stressed Vowels in PP



As illustrated in the plots (refer to Figures 1 and 2), F1 is represented on the vertical axis, while F2 is displayed on the horizontal axis. Even when considering the unique production system of each participant during speech, it remains possible to establish average values for both F1 and F2. Additionally, it is possible to identify the vowel trapezoid formed by the frequencies of the examined vowels. In this study, we were able to observe the realization of [ə] (here understood as a central vowel allophone of [a])[10] in stressed syllables both in PST and PP. This finding differs from findings in Balduino (2022), Christofoletti and Araujo (2018), and Santiago (2019).[11]

To compare the vowels [a] and [ə], we performed a statistical modeling of the data using RStudio. To do this, we conducted two linear regression with mixed effects, one for F1 and other for F2. While the data for F2 were not significant, the tests indicated that [a] and [ə] had a significant effect for F1 ($p > 0.001$) as shown in Table 4.

[10] For a discussion on the phonetic and phonological status of the segment(s) reported as schwa in the literature, see Veloso (2010). In this work, the author distinguishes between central vowels, epenthetic vowels, and unmarked vowels. In this paper, however, [ə] is understood as a specific vowel quality, the canonical [ə], but this vowel seems to have no special status in stressed positions.

[11] The sound [ə] was observed in 2 primary syllabic structures: (i) CV and (ii) CVC. In the first data group, the consonant "C" encompassed a range of sounds, including stops [b, t, d, k, g], fricatives [s, f, ʁ], and liquids [l, r, ɾ]. The sole instance where the initial consonant was a nasal [m] occurred in the word mar - 'sea.' However, Balduino (2018) has shown that, when in coda or following an onset, nasal consonants may exert some influence on F1 of coarticulated vowels, a phenomenon not applicable to the segmental contexts examined in this study. In the second data group, the onset consonant could be any of the aforementioned consonants, while the coda consistently featured either a lateral or a rhotic. Lastly, the schwa sound was observed in monophthongs (e.g., *sal* – 'salt') as well as in initial positions (e.g., *garfo* – 'fork'), medial positions (e.g., *alface* – 'lettuce'), and final positions (e.g., *sofá* – 'sofa') within words. However, there was insufficient data available to evaluate its occurrence at the absolute beginning of words in syllables without an onset, such as in the pattern V.CV.CV.

Table 4. Linear Regression with Mixed Effects - F1.

	Estimate	Standard Error	t value	p-value
Intercept	607.215	14.715	15.536	< .0001
[a]	39.687	3.933	209.494	< .0001

Note: random effects = (vowel| speaker) + (vowel| word). R2:0.699.

After providing an overview of the data plots for the 2 macrovarieties under study and testing the difference between [a] and [ə], we organized the identified dispersion areas into average values. This was made based on the data collected from men (indicated in dark blue) and women (indicated in light blue). The result allowed us to generate acoustic trapezoidal representations, which are illustrated in Figures 3 and 4. Upon careful examination of these plots, we noticed that in both PST and PP, women exhibited higher formant values in Hz for stressed vowels – a fact influenced by anatomical factors.

Figure 3. Plot of General Stressed Vowels in PST – Men (dark blue) and Women (light blue)

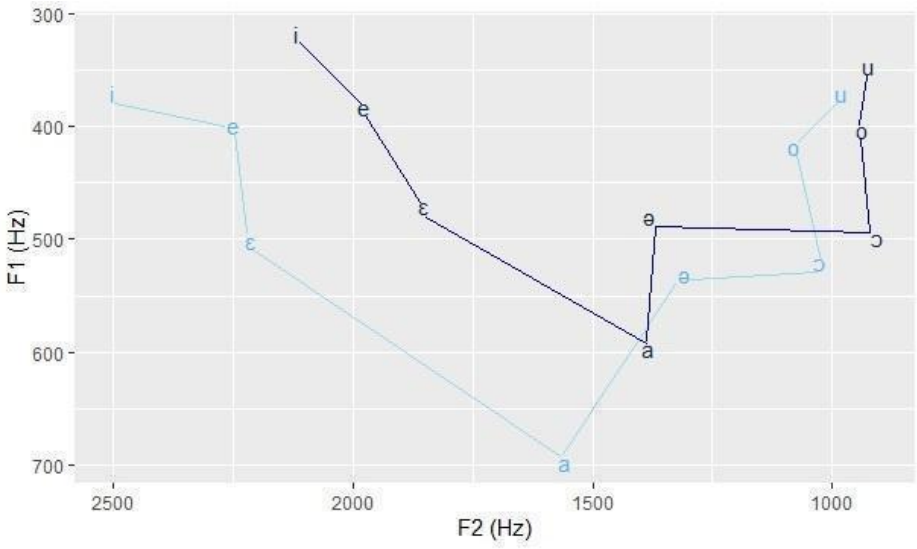
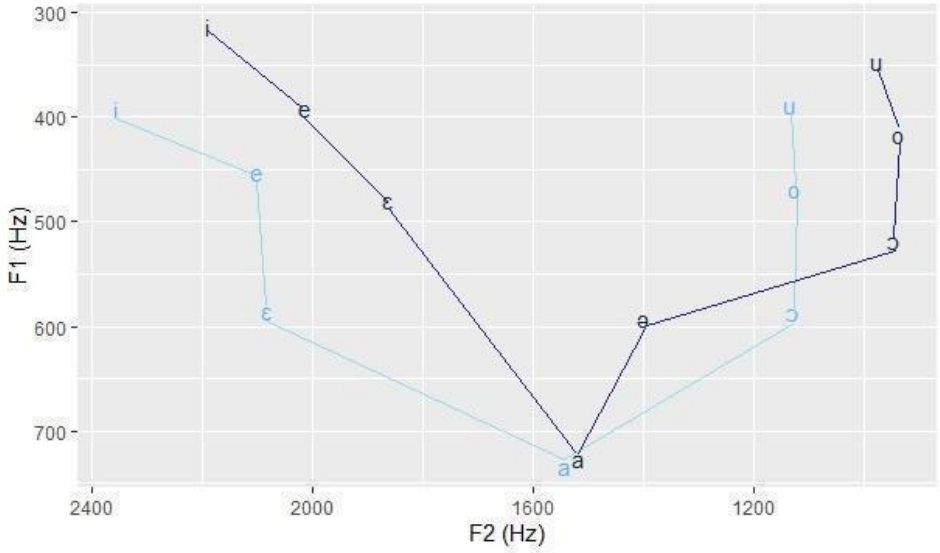


Figure 4. Plot of General Stressed Vowels in PP – Men (dark blue) and Women (light blue)



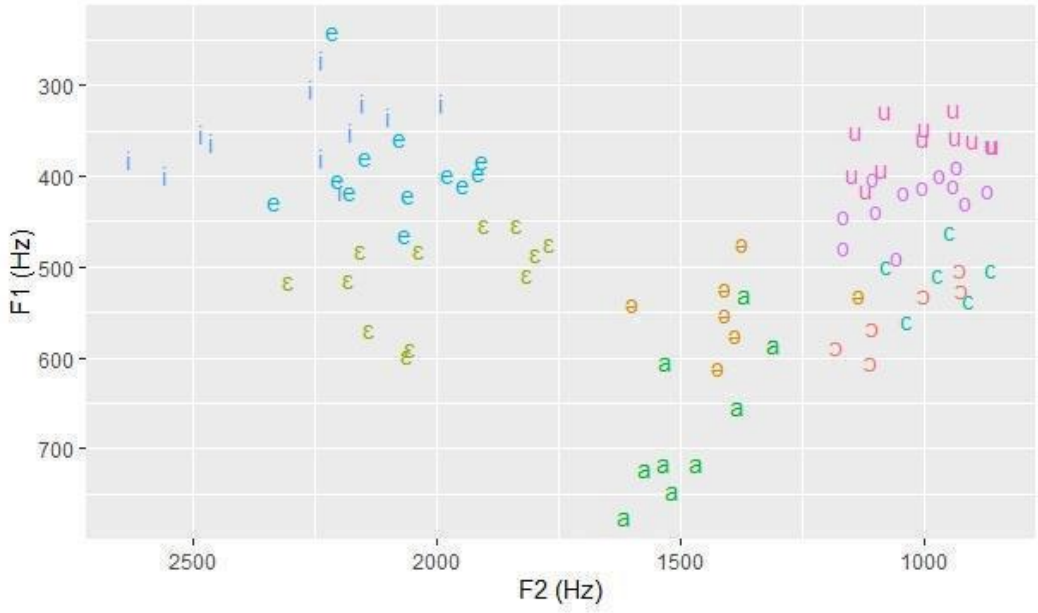
Considering the above and STPP in its entirety, Table 5 presents the average of the first formants (F1 and F2) for both men and women.

Table 5. Normalized Average Values (in Hertz) of F1 and F2 for Stressed Vowels - STPP - Men and Women

STPP					
Vowel	**Men**		**Vowel**	**Women**	
	F1	**F2**		**F1**	**F2**
i	278	2144.5	i	302	2076
e	376	1987	e	365.5	1967
ɛ	502.5	1898	ɛ	535.5	1896
ə	525.5	1427.5	ə	536	1503
a	622	1371.5	a	736.5	1352
ɔ	549.5	910	ɔ	569.5	1048
o	405	920.5	o	381.5	948.5
u	319	932.5	u	319.5	910

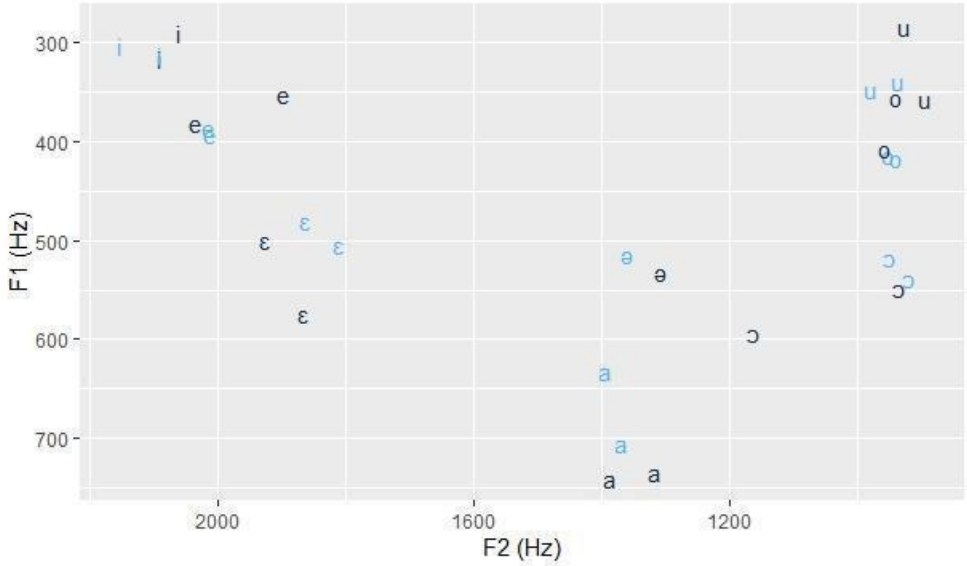
Figure 5 illustrates the plot of general frequency of the first 2 formants, considering male and female data. Despite observing a different distribution for the 8 vowels identified in STPP, there are areas with token contact points, as indicated by a/ə; or; o/ɔ; e/i.

Figure 5. Plot of General Stressed Vowels in STPP



Plotting the average values of these data results in the acoustic space shown in figure 6, in which we distinguish between male (dark blue) and female (light blue) data.

Figure 6. Plot of General Stressed Vowels in STPP – Men (dark blue) and Women (light blue)



Aligned with the average values of F1 and F2 as presented in Table 5, the plots indicate an overall resemblance between F1 and F2 in the production of stressed oral vowels in both varieties spoken in STP. This similarity is reinforced when comparing the average values of the first formants in PST and PP with urban varieties from BP and EP – data collected, more specifically, in the cities of São Paulo and Lisbon (cf. Escudero et al. 2009) (tables 6 and 7).

Table 6. Normalized Average Values (in Hertz) of F1 and F2 for Stressed Vowels - PST, PP, BP and EP - Men

Vowel	Men							
	PST		PP		BP		EP	
	F1	F2	F1	F2	F1	F2	F1	F2
i	303	2153	253	2189	285	2198	284	2161
e	386	2016	366	2014	357	2028	355	2028
ɛ	504	1850	501	1886	518	1831	455	1836
ə	530	1389	521	1466	–	–	–	–
a	625	1350	619	1393	683	1389	661	1365
ɔ	537	919	562	901	532	927	491	934
o	413	952	397	889	372	804	363	843
u	339	937	299	928	310	761	303	814

Source: Authors (PST and PP values) and Escudero et al. (2009) (BP and EP values).

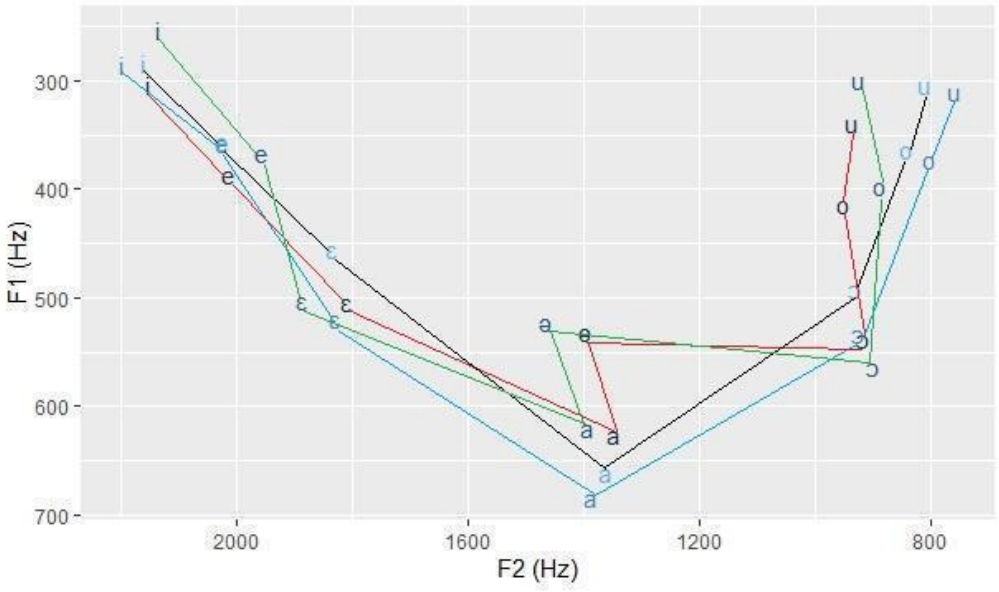
Table 7. Normalized Average Values (in Hertz) of F1 and F2 for Stressed Vowels - PST, PP, BP and EP - Women

Vowel	Women							
	PST		PP		BP		EP	
	F1	F2	F1	F2	F1	F2	F1	F2
i	315	2091	289	2061	307	2676	313	2760
e	380	2037	351	1897	425	2468	402	2508
ɛ	498	1926	573	1866	646	2271	511	2360
ə	536	1503	–	–	–	–	–	–
a	740	1387	733	1317	910	1627	781	1662
ɔ	547	934	592	1162	681	927	592	1118
o	407	957	354	940	442	893	422	921
u	356	894	283	926	337	812	335	862

Source: Authors (PST and PP values) and Escudero et al. (2009) (BP and EP values).

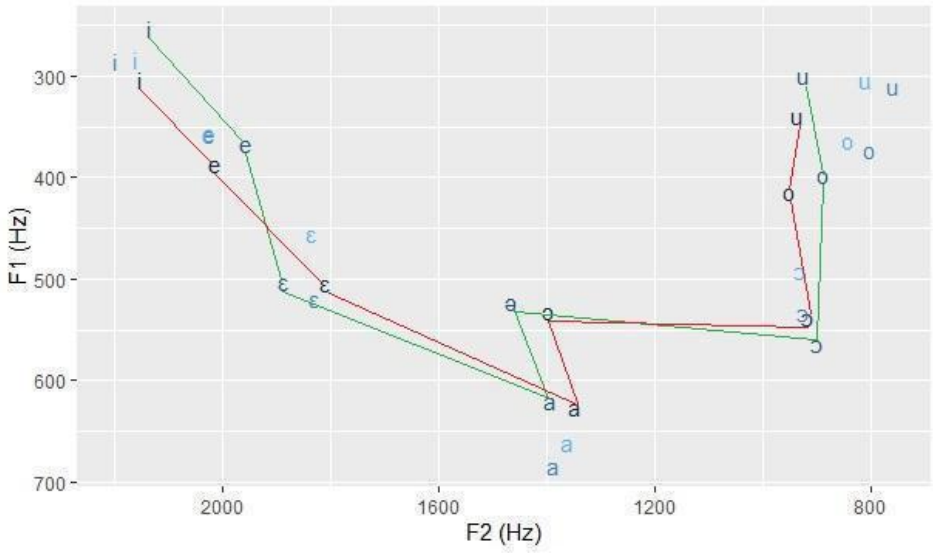
In Figures 7, 8, and 9, we plot stressed vowels for 4 varieties: PST (red), PP (green), BP (blue), and EP (black). We maintain the distinction between PST and PP, refraining from analyzing macrovarieties as STPP. This methodological choice was made to contrast the acoustic space of these varieties. In Figure 7, specifically, we present the plot corresponding to average values produced by male participants. Despite methodological differences between this study and that of Escudero et al. (2009), a noteworthy proximity is observed in the values obtained, with PST and PP emerging as the varieties with the most limited acoustic space.

Figure 7. Plot of General Stressed Vowels in PST (red), PP (green), BP (blue) and EP (black) - Men



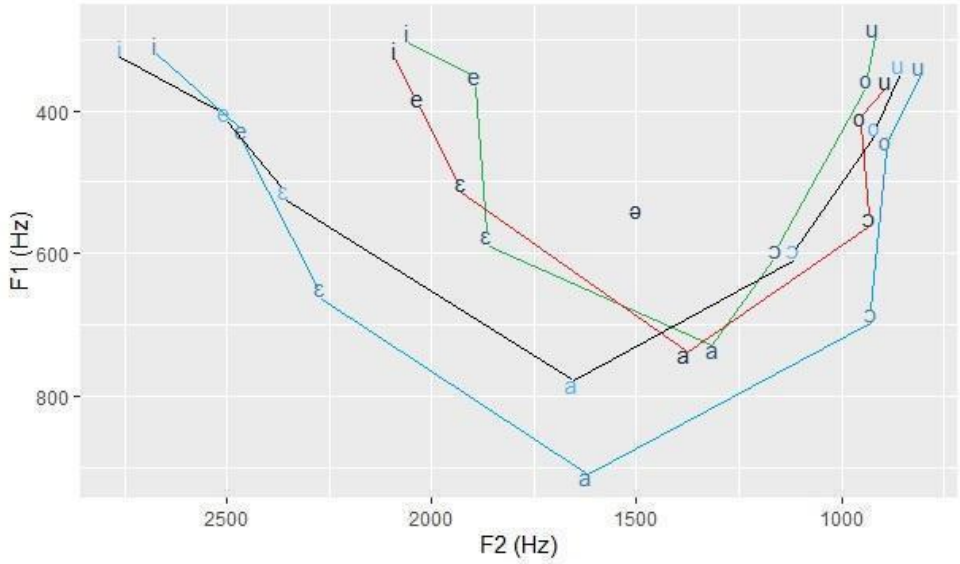
When analyzing data from PST and PP, a noticeable trend emerges wherein the acoustic spaces of these 2 varieties appear closer in comparison to the acoustic spaces of EP and BP, as depicted in Figure 7. Despite this proximity, the data produced by PST men exists within a slightly more compressed space. High vowels have a higher F1 in relation to the F1 of these same vowels in PP (F1 [i/u] PST > F1 [i/u] PP), which makes the arrangement of [i] and [u] in the plot in figure 7 closer to [e] and [o] in PST than in PP. Conversely, the F1 values of [a] and [ə] of PST are slightly larger compared to the F1 of [a] and [ə] of PP (F1 [a/ə] PST > F1 [a/ə] PP). This could lead to the expansion of the acoustic space of the PST, as [a] and [ə] occupy lower portions of the plot in figure 7 and 8. However, the difference between the F1 of [a] and [ə] across varieties is minimal, which keeps the y-axis space more compressed for PST compared to PP (cf. figure 7 and 8). Turning to F2, mid vowels and high vowels in PP exhibit a smaller F2 (excluding [ɛ]) and, therefore, are more posteriorized. In contrast, PST vowels tend to have a larger F2 and, therefore, are more anterior.

Figure 8. Plot of General Stressed Vowels in PST (red) and PP (green) - Men



This result, however, is partially evidenced in data produced by women, as represented in Figure 9.

Figure 9. Plot of General Stressed Vowels in PST (red), PP (green), BP (blue), and EP (black) - Women



When examining female data, contrasting the 4 varieties reveals a distinct pattern: PST and PP (urban varieties) exhibit a comparatively smaller acoustic space in comparison to dialects such as São Paulo (BP) and Lisbon (EP). Notably, both front vowels display a reduced F2 when compared to the Brazilian and European varieties considered by Escudero et al. (2009). Furthermore, the low vowel [a] in STP varieties is higher and posteriorized. In the comparison between PST and PP, a noteworthy similarity emerges: Drawing a parallel with male data represented in Figure 8, these varieties share a more analogous acoustic space with each other than with BP and EP.

By analyzing the normalized mean values (measured in Hertz) of F1 and F2 for stressed vowels in descending order within each variety, the following scales emerges:

Table 8. Normalized Average Values (in Hertz) of F1 and F2 for Stressed Vowels in Descending Order Within Each Urban variety - Men

	F1				**F2**			
Vowel	**Variety**				**Variety**			
i	PST >	BP >	EP >	PP	BP >	PP >	EP >	PST
e	**PST >**	**PP >**	BP >	EP	BP =	EP >	**PST >**	**PP**
ɛ	BP >	**PST >**	**PP >**	EP	**PP >**	**PST >**	EP >	BP
ə	----	----	**PST >**	**PP**	----	----	**PP >**	**PST**
a	BP >	EP >	**PST >**	**PP**	PP >	BP >	EP >	PST
ɔ	**PP >**	**PST >**	BP >	EP	EP >	BP >	**PST >**	**PP**
o	**PST >**	**PP >**	BP >	EP	**PST >**	**PP >**	EP >	BP
u	PST >	BP >	EP >	PP	**PST >**	**PP >**	EP >	BP

Table 9. Normalized Average Values (in Hertz) of F1 and F2 for Stressed Vowels in Descending Order Within Each Urban Variety - Women

	F1				**F2**			
Vowel	**Variety**				**Variety**			
i	PST >	EP >	BP >	PP	EP >	BP >	**PST >**	**PP**
e	BP >	EP>	**PST >**	**PP**	EP >	BP >	**PST >**	**PP**
ɛ	BP >	PP >	EP >	PST	EP >	BP >	**PST >**	**PP**
ə	----	----	----	PST	----	----	----	PST
a	BP>	EP >	**PST >**	**PP**	EP >	BP >	**PST >**	**PP**
ɔ	BP >	PP =	EP >	PST	EP >	**PP >**	**PST >**	BP
o	BP >	EP >	**PST >**	**PP**	**PST >**	**PP >**	EP >	BP
u	PST >	BP >	EP >	PP	**PP >**	**PST >**	EP >	BP

The analysis reveals that PST and PP are only separated into 7 segments, especially concerning F1 values. This finding suggests that, in fact, PST and PP form a macrovariety, as defended by Balduino (2022). Also, BP and EP are generally close (except for 6 cases) and present different behaviors regarding F1 and F2: Overall, BP precedes EP in terms of F1, while EP precedes BP in terms of F2. This finding suggests a prevalent trend where EP presents more anterior values of F2 while BP tends to have more posterior values. One could not perform a direct comparison of [ə] values among the varieties, as Escudero et al. (2009) do not report this variation in São Paulo (BP) and Lisbon (EP) varieties.

In this regard, it is worth noting that other studies dedicated to the analysis of the acoustic space of stressed vowels in different dialects of BP, such as Madruga, Hamann, and Abaurre (2020), also do not confirm the presence of [ə] in the stressed vowel inventory. When investigating stressed vowels in varieties of the North, Northeast, and Southeast of Brazil, the authors report the existence of 7 phonetic vowels [i, e, ɛ, a, ɔ, o, u] for each of the Brazilian dialects examined. The same is observed in Sandalo, Abaurre, and Madruga (2013), who, upon analyzing the dialects spoken in Porto Alegre and Salvador (South and Northeast of Brazil), also attests the

phonetic occurrence of the same 7 vowels evidenced by Escudero et al. (2009) and Madruga, Hamann, and Abaurre (2020). Therefore, the oral vowel /a/ is not raised in stressed syllables in BP.

The literature dedicated to EP, on the other hand, suggests distinct behaviors among Portuguese dialects. Although the study by Escudero et al. (2009) – used here for comparative purposes – supports the presence of 7 vowels in EP, different studies have pointed to the allophone between [a] and [ɐ] in the stressed syllable in dialects of the central coast of Portugal, specifically in the Lisbon norm (Mateus & d'Andrade 2000). Broadly speaking, in this dialect, [ɐ] tends to occur in a stressed syllable when preceded by a nasal segment, as in câmara [ˈkɐmɐɾɐ] 'chamber', as well as to distinguish the 1st person plural of the present indicative compramos [kõˈpɾɐmuʃ] 'we buy' and the 1st person plural of the past perfect indicative compramos [kõˈpɾamuʃ] 'we bought' (cf. Varanda; Barroso; Rato 2016).[12] In other dialects, such as from Braga (northern Portugal), [a] is the preferred variant, even in the highlighted segmental and morphosyntactic contexts (Varanda, Barroso & Rato 2016). In this regard and based on a perceptual discrimination test, Horn, Rinke, and Flores (2020) also demonstrate that speakers from the 2 dialectal groups (Lisbon and Braga) differ in the perception of the contrast between the 2 central vowels under investigation: Speakers of the Northern variety tend to differentiate less between [a] and [ɐ] compared to speakers from Lisbon. Therefore, we were able to note that /a/ raising is also reported in some varieties of EP, albeit marked in different linguistic contexts (we have not yet identified a specific segmental context for the production of [ə] in STPP) and producing different results ([ɐ] for EP and [ə] for STPP).

Additionally, Candeias, Lopes, and Perdigão (2011) analyzing the realization of [ə] (segment reported by the authors as schwa) in European Portuguese (EP), indicate the existence of different phonetic realizations for [ə], such as [ə], [i], and its deletion. However, the study focuses on unstressed syllables and concludes that [ə], [Ø], and [i] are allophonic realizations of /ə/, with this vowel characterized as a weak vowel that is regularly reduced and often deleted – a result already reported in the literature in studies such as Andrade (1996), Andrade and Mascarenhas (1995), and Veloso (2007), which also analyze vowel reduction in unstressed syllables in European Portuguese (EP).

The study of Candeias, Lopes, and Perdigão (2011) for EP differs from ours because, in addition to considering the production of [ə] in unstressed syllables, there is a theoretical distinction between the segment that the authors call schwa and the central vowel observed in stressed syllables in STPP. In this regard, Veloso (2010) points out that in many languages, central vowels act as epenthetic vowels, which has led to a terminological equivalence between labels such as central vowel, schwa, epenthetic vowel, default vowel, and unmarked vowel. The author argues that it is necessary to establish some terminological (and ontological) distinctions among these different terms and concepts. This also seems to be the case for STPP.

Even though in European Portuguese, similar to other languages such as French, 'schwa' is used to refer to a reduced vowel or a vowel with special

[12] As pointed out by one of the reviewers of this article, in the Northern varieties of Portuguese spoken in Portugal, the distinction between the 1st person plural forms of the present indicative and the past perfect indicative, which is typically marked by the binomial ɐ/a, does not exist.

phonological status, here it identifies a specific vowel quality, the canonical [ə]. However, this vowel appears to have no special status in stressed positions.

In addition to this and despite the F1 and F2 values provided in this paper, it is necessary to set forth some considerations that might impact our results along with the proposed comparisons relative to BP and EP.

Firstly, there are substantial methodological differences between this paper and that of Escudero et al. (2009) (and even among other aforementioned studies related to EP and BP). While both investigations involve data from men and women under 30, Escudero et al.'s (2009) work does not include bilingual speakers. In contrast, in our study, given the multilingual context in which STPP has emerged and is spoken, 4 participants self-identify as speakers of 2 creole languages spoken in STP, Santome and Kabuverdianu.

Specifically addressing Kabuverdianu,[13] Freitas (2022) carries out an acoustic analysis of the vowels of the Príncipe variety, revealing that the following central vowels are found in tonic position: [a], [ɐ] and [ə].[14] Even though the open vowel appears in the majority of data (65%), the presence of other variants, not commonly found in stressed positions, cannot be disregarded. Comparing these findings with those of this study, it is noteworthy that [ə] is identified in both PST and PP. Although Freitas' work focused on Príncipe's data, the potential for similarities between Kabuverdianu varieties from São Tomé and Príncipe (as observed in the case of Portuguese) raises the intriguing possibility of Kabuverdianu influencing the vowel acoustic space in PST and PP. This aspect needs to be examined in more detail in future studies. An interesting approach would be comparing the acoustic spaces of bilingual participants who are proficient in Portuguese and Kabuverdianu with those of monolingual Portuguese speakers; and, after that, contrasting this result with the acoustic space of Kabuverdianu.

In addition to the intricacies of language contact, another factor that may influence our result is the fact that Escudero et al. (2009) perform greater control over the segmental contexts used to extract the target vowels. The authors work with the first vowel of a disyllabic sequence (CV.CV), where the consonants (C) are, necessarily, 2 voiceless stops or 2 identical voiceless fricatives, specifically /p, t, k, f, s/, as exemplified in /pepo/ (Escudero et al. 2009, p. 3). Acknowledging that surrounding consonants can affect formants of both tautosyllabic and heterosyllabic vowels, it is likely that this factor influence our findings, since we analyzed data with a broad segmental context relative to the phones surrounding the target vowels.

The influence of anatomical differences and articulatory habits on participants' acoustic productions is a crucial consideration. Escudero et al. (2009) work with a broader production, comprising 20 participants, whereas the present study considered 12 participants. Therefore, factors related to the presence of bilingual speakers, the composition and size of our corpus, as well as articulatory differences within and also between varieties, may impact our results. Aligned with Barbosa and Madureira (2015, p. 307), we believe that studies with a greater number of individuals in line with the

[13] Regarding Santome, we are not aware of acoustic studies that we can use for comparative purposes.

[14] Regarding Kabuverdianu (in Cape Verde), there are studies discussing the varieties spoken on the different islands, such as Santiago (which brings together most of the studies), Fogo, Maio, Santo Antão (Quint 2000; Lang 2002; 2014; Rodrigues 2007; Moreira 2020).

"assessment of tract sizes" among Brazilians, Portuguese and Santomeans are interesting and necessary for more robust conclusions.

In any case, albeit preliminary, the acoustic analysis proposed in this article offers an initial and interesting foundation for future phonetic and phonological examinations regarding STPP. Understanding PST and PP formants, alongside the STPP, helps "the adequate search for frequency bands where we can expect to find formants and avoid gross errors in obtaining formant frequency values" (Barbosa; Madureira 2015, p. 307). Furthermore, by establishing the frequency ranges of PST and PP, this study highlighted the presence of [ə] in the phonetic inventory of these varieties, which may provide interesting clues about phonological processes of STPP. The production of [ə] in a tonic context increases the number of contrasts within STPP; however, even so, it does not lead to a substantial expansion of the acoustic space of this variety. In comparison to BP and EP, STPP's acoustic space is more compressed, especially for the F1 values: [ə], with average F1 values of 525.5 Hz - men and 536 Hz - women, is close to [ɛ], a vocoid whose F1 is 502.5 Hz - men and 535.5 Hz - women. This similarity places these segments close in the acoustic space, generating a cost to the system in terms of distinctiveness of F1 (Flemming 2004).

In addition to reinforcing the interpretation that STPP exhibits a smaller acoustic space compared to São Paulo (BP) and Lisbon (EP) varieties, the inclusion of [ə] may also reveal phonological considerations linked to syllabic positional prominence. In this regard, linguistic literature reports that languages typologically present less vowel contrast in unstressed syllables than in stressed ones (Crosswhite 2004; Flemming 2004; Kingston 2007). In the context of STPP, the relevance of positional prominence was attested in processes such as devoicing and deletion of unstressed vowels, raising and lowering of pretonic vowels, and hetero- and tautosyllabic vowel nasalization (Balduino 2018; 2022). Given this fact, Balduino (2022) and Balduino and Freitas (2022) hypothesize that, in STPP, vowel segments lacking word stress may undergo lenition and deletion to maximize the prominence contrasts within the word, maximizing, for example, the prominence contrast between the segments in stressed and unstressed syllables (cf. Crosswhite 2004; Nevins 2012) and even among constituent segments of vowel sequences, potentially preserving the diphthong against processes such as monophthongization (Balduino 2023). However, the presence of the vowel reduction of /a/ to [ə] in a highly prominent position such as the stressed syllable conflicts with this hypothesis. This issue deserves to be examined in more detail by future studies.

5. Conclusions

This article presented an acoustic analysis of STPP considering the first formant (F1) and the second formant (F2) of stressed vowels. Initially, we separately discussed the acoustic space of stressed vowels in PST and PP and then the STPP macrovariety was analyzed as a whole. The results were compared with those found for BP and EP, more specifically for the dialects spoken in São Paulo and Lisbon (Escudero et al. 2009).

Regarding the plot of acoustic spaces from F1 and F2, among the 4 varieties, PST and PP exhibit a more compressed acoustic space and are closer together, which argues in favor of their forming a unique variety – STPP, as defended by Balduino (2022) and others. BP displays the largest acoustic space and, generally, is closer to

PE; additionally, BP vowels are more posterior. Another noteworthy result is the presence of [ə] in STPP and its urban microvarieties, i.e., both in PST and PP. This has not been observed in São Paulo and Lisbon varieties either, as reported by Escudero et al. (2009) and Madruga, Hamann and Abaurre's research (2020), which evaluates F1 and F2 of stressed vowels produced by speakers from the northeast, north, and southeast regions of Brazil. In EP, /a/ also has variable realizations in the tonic position with alternation between [a] and [ɐ] in some dialects, but not between [a] and [ə] (Horn; Rinke; Flores 2020). This suggests that, among the documented varieties of Portuguese, [ə] in stressed syllables seems to be uniquely found in STPP and its microvarieties.

As mentioned, STPP has not been extensively studied and has not yet been the focus of any acoustic study. Therefore, this paper aims to provide an initial description of the acoustic space of stressed vowels and set forth some hypotheses based on its results. However, some questions remain open, demanding a more detailed experimental examination to investigate issues like: (i) Are vowel phenomena, such as the raising of /a/, responsible for triggering vowel neutralization when combined with certain cases of syllable duration and/or syllabic prominence? (ii) Are these same processes related to the acoustic dispersion of vocoids and their respective distinctiveness? (iii) How do neighboring consonants (front and back) influence the realization of vowels in stressed syllables? (iv) How does the current scenario of contact between languages in STP influence the Portuguese spoken in the archipelago?

We hope this study will serve as a starting point for future research focused on these topics.

References

Agostinho, Ana Lívia. 2015. *Fonologia e Método Pedagógico do lung'Ie*. Ph.D thesis, University of São Paulo.

Agostinho, Ana Lívia, Bandeira, Manuele, & Gabriel Antunes Araujo. 2016. O lung'ie na educação escolar de São Tomé e Príncipe. *Trabalhos em linguística aplicada* 55(3): 591–618. https://doi.org/10.1590/010318135164183401

Agostinho, Ana Lívia, & Maiara Mendes. 2020. A grafia dos róticos intervocálicos no português da Ilha do Príncipe: fusão fonológica e ensino. *Veredas – Revista de Estudos Linguísticos* 24(3): 54–176.

Andrade, Amália. 1996. Reflexões sobre o 'e mudo' em Português europeu. In I. Duarte, & I. Leiria (eds), *Actas do Congresso Internacional sobre o Português*, vol. II, 303–344. Lisboa: Colibri/APL.

Andrade, Amália, & Mascarenhas, I. 1995. Sobre a variação fonética de /i/ – uma primeira abordagem. In: *Actas do X Encontro Nacional da Associação Portuguesa de Linguística*, 25-43. Lisbon: Colibri/APL.

Araujo, Gabriel. 2020. Há uma política linguística para o português em São Tomé e Príncipe? In: S. Souza, & F. Calvo Olmo (eds), *Línguas em português – A Lusofonia numa visão Crítica*, 173–197. Porto: Universidade do Porto Press.

Balduino, Amanda Macedo. 2018. A nasalidade vocálica no português falado em São Tomé e Príncipe. Masters Dissertation, University of São Paulo.

Balduino, Amanda Macedo. 2020. Templates silábicos no Português do Príncipe: processos em coda e 'prevalência' de sílabas CV. *Filologia e Linguística Portuguesa* 22(2): 247–270. https://doi.org/10.11606/issn.2176-9419.v22i2p247-270

Balduino, Amanda Macedo. 2022. Fonologia do Português de São Tomé e Príncipe. Ph.D. thesis, University of São Paulo.

Balduino, Amanda Macedo. 2023a. Os ditongos no português de São Tomé e Príncipe: um estudo sobre a monotongação, a vocalização e a ditongação. Postdoctoral Report, University of São Paulo.

Balduino, Amanda Macedo. 2023b. Padrões silábicos no Português de São Tomé e Príncipe: relações complexas. *Working Papers em Linguística* 24(1): 132–159. https://doi.org/10.5007/1984-8420.2023.e92297

Balduino, Amanda Macedo, Bandeira, Manuele, & Shirley Freitas. 2022. A língua portuguesa de São Tomé e Príncipe/ The Portuguese language of São Tomé and a Príncipe. *Fórum Linguístico* 19(3): 8125–8143. https://doi.org/10.5007/1984-8412.2022.e79197

Balduino, Amanda Macedo, & Shirley Freitas. 2022. Vogais pretônicas do português urbano de São Tomé. *Lingüística* 38(2): 119–138. https://doi.org/10.5935/2079-312x.20220017

Bandeira, Manuele. 2017. *Reconstrução fonológica e lexical do protocrioulo do Golfo da Guiné*. Ph.D thesis, Universidade de São Paulo.

Barbosa, Plínio Almeida, & Sandra Madureira. 2015. *Manual de fonética acústica experimental: aplicações a dados do português*. São Paulo: Cortez.

Baxter, Alan. 2018. O Português dos Tongas de São Tomé. In M. Santos Duarte Oliveira, & G. Antunes Araujo (eds), *O Português na África Atlântica*, 297–324. São Paulo: Humanitas/ Fapesp.

Boersma, Paul, & Weenick, David. (1992–2022). *Praat: doing phonetics by computer* [Computer program]. Version 6.2.06, restaurada em 23 de janeiro de 2022 de http://www.praat.org/.

Bouchard, Marie-Eve. 2017. *Linguistic Variation and change in the Portuguese of São Tomé*. Ph.D. Thesis, New York University.

Bouchard, Marie-Eve. 2018. Subject Pronoun Expression in Santomean Portuguese. *Journal of Portuguese Linguistics* 17(5): 1–29. https://doi.org/10.5334/jpl.191

Bouchard, Marie-Eve. 2019. Language shift from Forro to Portuguese: Language ideologies and the symbolic power of Portuguese on São Tomé Island. *Lingua* 228: 1–16. https://doi.org/10.1016/j.lingua.2019.06.013

Bowern, Claire. 2008. *Linguistic fieldwork: a practical guide*. New York: Palgrave Macmillan.

Braga, Gabriela. 2018. *A prosódia do português de São Tomé: a entoação do contorno neutro*. Masters Dissertation, University of São Paulo.

Brandão, Silvia Figueiredo, Pessanha, Davi Bretas, Pontes, Stefany, & Monique Corrêa. 2017. Róticos na variedade urbana do Português de São Tomé. *PAPIA* 27(2): 293–315.

Brickell, Timothy. 2018. Linguistic fieldwork: perception, preparation, and practice. In P. Austin, & L. Gawne (eds), *Language Documentation and Description*, vol. 15, 179–207. London: EL Publishing.

Candeias, Sara, Veiga, Arlindo, Lopes, Carla, & Fernando Perdigão. 2011. A Realização do Schwa no Português Europeu. In *Anais da II Jornada de Descrição do Português*, Cuiabá, 26–31.

Christofoletti, Alfredo, & Gabriel Araujo. 2018. Vogais e Ditongos no Português Vernacular de São Tomé e Príncipe. In M. Santos Duarte Oliveira, & G. Antunes Araujo (eds), *O Português na África Atlântica*, 258–296. São Paulo: Humanitas/ Fapesp.

Crosswhite, Katherine. 2004. Vowel Reduction. In B. Hayes, R. Kirchner, & D. Steriade (eds), *Phonetically based Phonology*, 191–231. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511486401.007

David, Cristina. 2018. *A língua portuguesa e as línguas de contacto: o impacto da política linguística em São Tomé e Príncipe*. Ph.D. Thesis, University of Lisbon.

Escudero, Paola, Boersma, Paul, Schurt, Andréia Rauber, & Ricardo Bion. 2009. A cross-dialect acoustic description of vowels: Brazilian and European Portuguese. *The Journal of the Acoustical Society of America* 126(3): 1379–1393.

Fant, Gunnar. 1966. A note on vocal tract size factors and non-uniform F-pattern scaling. *Speech Transmission Laboratory, Quarterly Progress and Status Report* 7: 22–30

Figueiredo, Carlos. 2009. A configuração do SN Plural do Português Reestruturado da Comunidade de Almoxarife - São Tomé. *Journal of Ibero-Romance Creoles* 1(1): 28–55.

Figueiredo, Carlos. 2010. *A concordância plural variável no sintagma nominal do português reestruturado da comunidade de Almoxarife, São Tomé*. Ph.D. Thesis, University of Macau.

Flemming, Edward. 2004. Contrast and perceptual distinctness. In B. Hayes, R. Kirchner, & D. Steriade (eds), *Phonetically based Phonology*, 232–276, Cambridge: Cambridge University Press.

Freitas, Shirley. 2022. O sistema vocálico do kabuverdianu do Príncipe. Postdoctoral Report, Federal University of Uberlândia.

Gonçalves, Rita. 2010. *Propriedade de Subcategorização verbal no português de S. Tomé*. Masters Dissertation, University of Lisbon.

Horn, Valerie, Rinke, Esther, & Cristina Flores. 2020. Dialectal Variation in European Portuguese Central Vowel Perception. *Journal of Portuguese Linguistics* 19(1): 1. https://doi.org/10.5334/jpl.232

INE – Instituto Nacional de Estatística. 2012 *População Segundo Línguas Faladas*, RGPH 2012. Available in: <https://www.ine.st/phocadownloa/userupload/Documentos/Atlas/Caracter%C3%ADsticas%20Educacionais/5.%20Popula%C3%A7%C3%A3o%20seg.%20L%C3%ADnguas%20Faladas%20-%2054.pdf>. Accessed on: 10 sept. 2021.

Johnson, Keith. 2005. Speaker normalization in speech perception. In D. Pisoni, & R. Remez (eds), *The Handbook of Speech Perception*, 363–389. Blackwell Publishers.

Kingston, John. 2007. The phonetics-phonology interface. In P. de Lacy (ed.), *The Cambridge Handbook of Phonology*, 401–434. Cambridge: Cambridge University Press.

Lang, Jürgen. 2002. *Gramática do Crioulo da Ilha de Santiago (Cabo Verde). Capítulo 1*. Available in: <http://iilp.wordpress.com/2012/11/14/umagramatica-do-crioulo/>. Accessed on: 12 nov. 2013.

Lang, Jürgen (ed.). 2014. *A variação geográfica do crioulo caboverdiano*. Erlangen: FAU University Press.

Madruga, Magnun Rochel, Hamann, Silke, & Maria Bernadete Abaurre. 2020. Gradient and categorical assimilation of pretonic vowels in Brazilian Portuguese. *Journal of Portuguese Linguistics* 19: 1–21. https://doi.org/10.5334/jpl.234

Mateus, Maria Helena, & Ernesto D'Andrade. 2000. *The Phonology of Portuguese*. Oxford: Oxford University Press.

Moreira, Ana Karina Tavares. 2020. *Documentação e descrição gramatical e lexical do crioulo afro-português da ilha do Fogo (República de Cabo Verde, Africa*

Ocidental). Ph.D. Thesis, Institut National des Langues et Civilisations Orientales (INALCO), Langage, langues et cultures d'Afrique (LLACAN), National Centre for Scientific Research (CNRS).

Nevins, Andrew. 2012. Vowel Lenition and Fortification in Brazilian Portuguese. *Letras de Hoje* 47(3): 228–233.

Pépiot, Erwan. 2009. *The making of French vocalic triangles: the case of a woman's voice versus a man's voice*. Ph.D. Thesis, University of Paris.

Pépiot, Erwan. 2012. Voice, speech and gender: male-female acoustic differences and cross-language variation in English and French speakers. In *XVèmes Rencontres Jeunes Chercheurs de l'ED 268*, Jun 2012, Paris, France.

Pereira, Rodrigo, Hagemeijer, Tjerk, & Maria João Freitas. 2018. Consoantes róticas e variação no português de São Tomé. *Revista da Associação Portuguesa de Linguística* 4: 206–224.

Quint, Nicolas. 2000. *Grammaire de la Langue Cap-Verdienne*. Paris: L'Harmattan.

Rodrigues, Ulisdete. 2007. *Fonologia do caboverdiano: das variedades insulares à unidade nacional*. Ph.D. Thesis, University of Brasilia.

Sândalo, Filomena, Abaurre, Maria Bernadete, Madruga, Magnum. 2013. Dispersão e Harmonia Vocálica em dialetos em Português Brasileiro. *Organon* 28(54): 13–30.

Santiago, Ana Maria. 2019. *As vogais do português do Príncipe*. Undergraduate Monograph, Federal University of Santa Catarina.

Santiago, Ana Maria, Balduino, Amanda, Agostinho, Ana Lívia, & Eduardo Correa Soares. 2022. As Vogais no Português do Príncipe. In T. Hagemeijer, M. Santos Duarte de Oliveira, & C. Figueiredo (eds), *O Português na África Atlântica*, vol 2. São Paulo: Humanitas/ Fapesp.

Takefuta, Yukio et al. 1972 A statistical analysis of melody curves in the intonation of American English. In *Proceedings of the 7th International Congress of Phonetic Sciences*, Montreal, 1035–1039.

Varanda, Patrícia, Barroso, Henrique, & Anabela Rato. 2016. Estudo acústico de /a/ acentuado na fala bracarense [An acoustic study of stressed /a/ in the Bracarense speech]. *Revista Portuguesa de Humanidades/Estudos Linguísticos* 20(1): 101–135.

Vaux, Bert, & Justin Cooper. 2003. *Introduction to Linguistic Field Methods*. München: Lincom Europa.

Veloso, João. 2007. Schwa in European Portuguese: the Phonological Status of [ɨ]. In O. Crouzet, & J.-P. Angoujard (eds), *Actes des/Proceedings of JEL'2007. Schwa(s). 5.èmes Journées d'Etudes Linguistiques*, Nantes, 55–60.

Veloso, João. 2010. Central, epenthetic, unmarked vowels and schwas: a brief outline of some essential differences. *Linguística – Revista de Estudos Linguísticos da Universidade do Porto* 5(1): 193–213.