

An acoustic description of Brazilian Portuguese oral vowels¹

ANDRÉIA SCHURT RAUBER
(Universidade do Minho)

Resumo

Este estudo objetivou investigar a produção das sete vogais orais do português brasileiro (PB) por mulheres e homens monolíngües dos estados brasileiros do Rio Grande do Sul, Santa Catarina e Paraná. Além de fornecer uma descrição acústica das vogais, com dados de duração, F0 e dos três primeiros formantes, o estudo também investigou em quais pistas acústicas (duração e/ou qualidade espectral) os participantes se basearam para pronunciar os sons vocálicos. Os dados foram obtidos através da leitura de frases contendo palavras com as vogais-alvo em sílaba tônica. Os resultados mostram que a qualidade espectral é a principal pista acústica usada pelos participantes deste estudo para fazer distinções entre as vogais do PB.

1. Introduction

This study provides a description of the acoustic characteristics (duration, fundamental frequency (F0), and the first three formant values (F1, F2, and F3)) of the seven Brazilian Portuguese (BP) oral vowels in stressed position (/i, e, ε, a, ɔ, o, u/). A limited number of studies provided acoustic descriptions of BP oral vowels (Faveri, 1991;

¹ Acknowledgements: This work is part of the study which investigates the perception and production of English vowels by native speakers of Brazilian Portuguese and was supported by a CAPES (Brazilian Ministry of Education) grant. I would like to thank Professor Dr. Paul Boersma, from the University of Amsterdam, for the scripts that allowed the vowel formants and duration to be measured automatically.

Lima, 1991; Moraes *et al.*, 1996; Pereira, 2001; Seara, 2000), and all these studies investigated the productions of male participants who had lived in the capitals of several Brazilian states. The participants in Faveri (1991), Lima (1991), Pereira (2001), and Seara (2000) were from Florianópolis (South); and the participants in Moraes *et al.* (1996) were from three regions: Recife and Salvador (Northeast), Rio de Janeiro and São Paulo (Southeast), and Porto Alegre (South). The vowels measured in the present study were produced by both male and female participants from non-capital cities in the South of Brazil. Thus, this study investigates a different variety and goes further in analyzing how the participants make use of duration and spectral quality to produce BP vowels.

2. Method

2.1. Participants

The productions of twelve BP monolingual speakers (6 women and 6 men) were recorded, four speakers (two women and two men) from each of the Brazilian southern states Rio Grande do Sul (RS), Santa Catarina (SC) and Paraná (PR). The women's ages ranged from 20 to 31 years (mean = 27 years), and the men's ages ranged from 20 to 36 years (mean = 26 years). The speakers were from the following cities: Rio Grande-RS, Santa Maria-RS, Chapecó-SC, and Cascavel-PR. None of the participants spoke any foreign language and all affirmed to have had no contact with speakers of other foreign languages in their daily routines.

2.2. Corpus

The corpus consisted of 70 words and nonce words, 10 for each of the BP vowels (/i, e, ε, a, ɔ, o, u/) of the following 5 phonological structures: pV.pV, tV.kV, kV.kV, fV.fV, sV.sV. The words and nonce words were contextualized in a carrier sentence preceded by one of the words: *CVCe/o. Em CVCe e CVCo temos V* (C = consonant, V = vowel). All the words were stressed on the first syllable. Thus, the participants would read sentences like: *Pêpe. Em pêpe e pêpo temos ê* (Pêpe. In pêpe and pêpo we have ê).

The phonological contexts were chosen by taking into account the place of articulation and the absence of voicing. Since each sentence contained two target words formed by the same vowel and consonantal context, the second syllable of the first and second words ended in the graphic vowels “e” and “o”, respectively, but were always read as [ɪ] and [ʊ] due to a production BP process of raising of the vowels /e/ and /o/ in word-final unstressed position. These two vowels were chosen because they have similar F1 values. For words containing the vowels /e/, /o/, /ɛ/ or /ɔ/ in the stressed syllable, the vowels were written as ê, ô, é or ó, respectively, to help the participants identify the vowel to be produced. All literate Brazilians are familiar with these diacritics. Table 1 shows the target BP words read by the participants.

Table 1: Target words read by the BP monolinguals

Vowel	pVpV	tVkV	kVkV	fVfV	sVsV
[i]	pipe/pipo	tike/tiko	kike/kiko	fife/fifo	sisse/sisso
[e]	pêpe/pêpo	têque/têco	quêque/quêco	fêfe/fêfo	sêsse/sêsson
[ɛ]	pépe/pépo	téque/téco	quéque/quéco	fêfe/féfo	sêsse/sésson
[a]	pape/papo	taque/taco	caque/caco	fafe/fafo	sasse/sasso
[ɔ]	pópe/pópo	tóque/tóco	cóque/cóco	fôfe/fófo	sósse/sósson
[o]	pôpe/pôpo	tôque/tôco	côque/côco	fôfe/fôfo	sósse/sôsson
[u]	pupe/pupo	tuque/tuco	cuque/cuco	fufe/fufo	susse/susson

The isolated word was not considered for analysis. Each BP monolingual read the 35 target sentences twice, resulting in 140 tokens (7 vowels x 2 tokens per sentence x 5 contexts x 2 repetitions). The total of BP tokens analyzed was 1680 (140 tokens x 12 participants).

2.3. Procedure

The recordings were made in a quiet room at the participants' homes in Brazil, and in case there was some background noise during the recordings, the participants were asked to reread the affected item(s). Each participant was asked to read the target sentences at normal speed. To help maintain a falling intonation each sentence was shown on a different card. The participants were allowed to read the item on the next card only if the recording was satisfactory in terms of intonation and vowel height. Even though the speakers were informed

that the two target words in the sentence should rhyme with the word in isolation, and even with the signaling in orthography (ê, é, ô, ó), in the sentences containing the vowels [e]-[ɛ] and [o]-[ɔ] many participants tended to mix vowels, pronouncing an open and a closed vowel in the same sentence. In the instances when this type of confusion occurred, the participants were asked to repeat the sentence until all the vowels were pronounced with a similar quality.

2.4. Recording equipment

All the productions analyzed in the present study were recorded with a Sony MZ-NHF800 minidisk recorder, with a Sony ECM-MS907 condenser microphone at 22 kHz, with 16-bit accuracy.

2.5. Data analysis

2.5.1. Duration

Before running a script that can automatically and reliably measure formants, each vowel was manually segmented and labeled in the digitized sound wave by using the program Praat, version 4.4.23 (Boersma/Weenink, 2006). Either the beginning or the end of the selection was close to a zero crossing, that is, when the wave crosses zero amplitude. The start and end points were considered the first and last periodic pulses on the waveform that had considerable amplitude and resembled the vowel period. The choice for voiceless consonantal contexts was exactly to facilitate duration measurements, since these consonants allow a more precise identification of the first and last constant periodic pulses of the vowel.

2.5.2. Fundamental frequency (F0)

In order to measure the fundamental frequency, the central 40% of the target vowels was measured automatically with Praat using autocorrelation analysis. The pitch floor was set at 60 Hz for men and 120 Hz for women, and the pitch ceiling was set at 400 Hz for both. After the pitch values of the center of the vowel were found, their median was calculated.

2.5.3. *The first three formants*

The formant measurements for F1, F2 and F3 of each vowel were made by applying the burg algorithm (Anderson, 1978) built in Praat to calculate the Linear Predictive Coding (LPC) spectra with the number of formants per frame defined as 5. Due to the great differences between the vocal tract shapes, the formant ceiling was determined for each type of vowel produced by each participant. In order to find the optimal ceiling, all values from 4500 to 6500 Hz in 10-Hz steps were calculated for women, and all values from 4000 to 6000 Hz in 10-Hz steps were calculated for men. The optimal ceiling for each vowel for each speaker was the one which yielded the lowest standard deviation of F2 over the measured values, which were a total of 20 for each vowel. The window length was set to 50 ms, and the time steps were defined as 25 ms.

3. Results

As described in Section 2.1., the BP monolinguals are from three different Brazilian states: Rio Grande do Sul, Santa Catarina and Paraná. An analysis of variance (ANOVA) of the independent variables F1 and F2 revealed that there are significant differences among participants within the group of men and within the group of women. Tukey post-hoc tests showed that there is no homogeneous tendency that allows the participants to be grouped by dialect, because, for instance, one participant from one language variety did not have any statistically significant formant value difference from participants from the other two varieties, but for some vowels this participant's formant values differed significantly from the other participant from the same variety. This lack of consistency in the results can be explained by the limited number of participants from each region (only 2 per gender per variety). Thus, henceforth the participants will be considered to pertain to only one variety: the BP spoken in non-capital cities of the southern Brazilian states. The participants will only be grouped by gender.

The mean, median and standard deviation (SD) of the formant values of all the vowels produced by the BP female and male participants in the five phonological contexts are shown in Tables 2 and 3, respectively. The mean values are plotted in Figures 1 and 2.

**Table 2: BP women's duration, F0, F1, F2 and F3 values
(N = number of vowels)**

		/i/	/e/	/ɛ/	/a/	/ɔ/	/o/	/u/
	N	120	120	120	120	120	120	120
Dur.	Mean	92	111	127	127	123	111	93
	Median	94	111	128	127	124	112	94
	SD	19	24	26	24	23	23	19
F0	Mean	241	222	206	202	206	221	245
	Median	234	214	202	200	204	212	240
	SD	37	34	31	32	31	23	40
F1	Mean	298	414	606	890	631	422	326
	Median	286	408	611	896	636	416	320
	SD	41	36	51	87	74	38	49
F2	Mean	2710	2540	2282	1667	1091	908	880
	Median	2694	2558	2283	1682	1098	904	825
	SD	151	192	152	143	150	128	235
F3	Mean	3200	3021	2912	2580	2693	2880	2875
	Median	3248	3058	2964	2627	2676	2902	2904
	SD	328	279	286	321	217	184	211

**Table 3: BP men's duration, F0, F1, F2 and F3 values
(N = number of vowels)**

		/i/	/e/	/ɛ/	/a/	/ɔ/	/o/	/u/
	N	120	120	120	120	120	120	120
Dur.	Mean	95	114	130	130	130	113	96
	Median	95	116	134	132	131	114	98
	SD	17	19	21	17	20	18	16
F0	Mean	151	146	136	136	135	146	157
	Median	153	143	136	134	134	146	154
	SD	41	26	30	28	30	35	35
F1	Mean	292	344	494	651	542	379	308
	Median	293	339	497	686	574	390	308
	SD	23	35	64	109	69	37	21
F2	Mean	2212	2080	1908	1405	971	874	834
	Median	2199	2061	1888	1383	967	849	790
	SD	130	170	133	157	105	128	174
F3	Mean	2950	2755	2614	2346	2334	2466	2526
	Median	2972	2734	2620	2314	2346	2452	2468
	SD	214	200	156	224	266	186	261

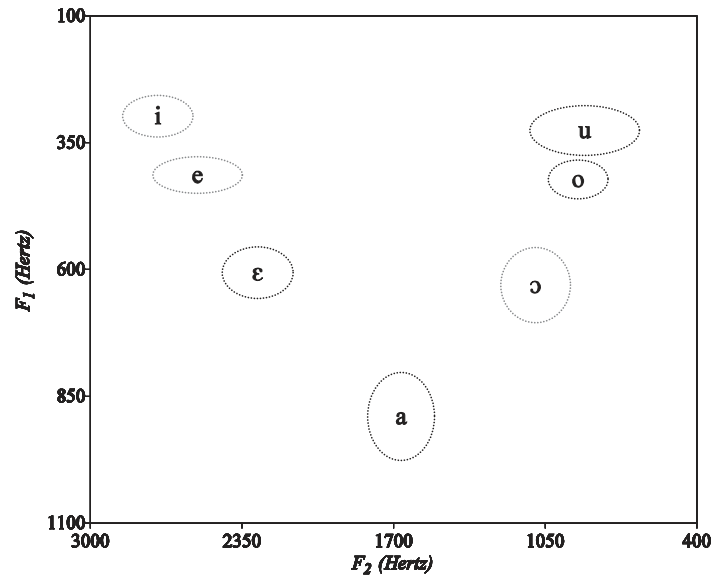


Figure 1: BP female participants' mean and SD (in ellipses) of the F1 and F2 values in Hz

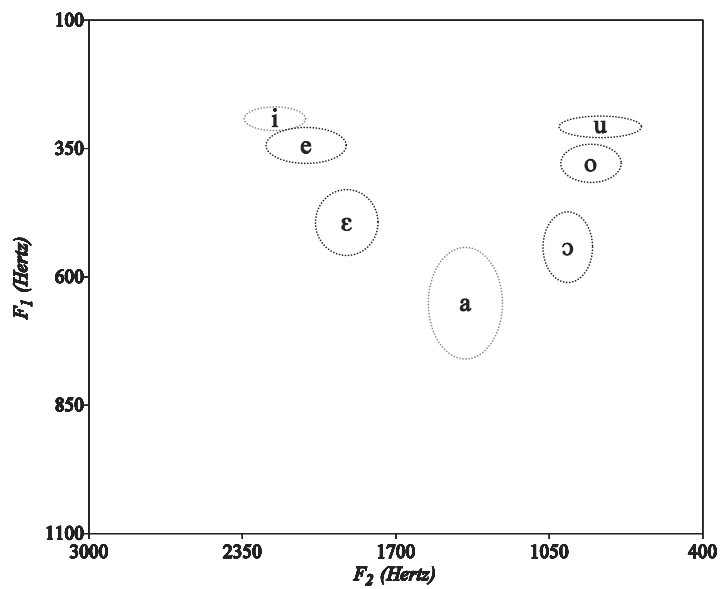


Figure 2: BP male participants' mean and SD (in ellipses) of the F1 and F2 values in Hz

The vowels plotted in Figure 2 show that the women's vowel system is slightly more symmetrical than the men's if the series of front and back vowels are taken into account. Accordingly, paired-samples t-tests confirm the visual impression from the vowel plots. The results show that the men from the Brazilian Southern states have higher front than back vowels, but none of the women from the Brazilian Southern states has the front vowels significantly higher than their back vowel counterparts, as can be seen in Table 4.

Table 4: Results of t-tests comparing F1 values within the groups of BP female and male monolinguals

Vowels	Women	Men
/i/ - /u/	t(5) = -2.611, p = .05	t(5) = -4.503, p = .0003
/e/ - /o/	t(5) = -0.653, p = .54	t(5) = -3.661, p = .007
/ɛ/ - /ɔ/	t(5) = -1.200, p = .15	t(5) = 5.726, p = .0003

The plots also show that the men's vowel system is more compact, as would be expected, the distance between the high vowels and the low central vowel being significantly smaller for men than for women, as revealed by the two-tailed independent-samples t-tests of the Euclidean distances between the vowels /i/-/a/-/u/ pronounced by the two gender groups:

- 1) /i/-/a/: t(10) = 5.353, p < .0001;
- 2) /u/-/a/: t(10) = 4.567, p < .0001; and
- 3) /i/-/u/: t(10) = 4.738, p < .0001.

The results concerning duration differences between the target vowels confirmed that the lower the vowel the greater its duration, which is a typical intrinsic vowel characteristic, since the more the jaw needs to open to articulate a vowel, the longer its duration. As regards the women's and men's productions, no statistically significant length difference was found between the duration values between women and men for (a) the low vowels /ɛ/, /a/ and /ɔ/; (b) the mid vowels /e/ and /o/; or (c) the high vowels /i/ and /u/. An ANOVA revealed a nonsignificant duration vs. gender interaction (F = .457, p = .841). Figure 3 shows the mean values for each vowel produced by each gender in all the phonological contexts.

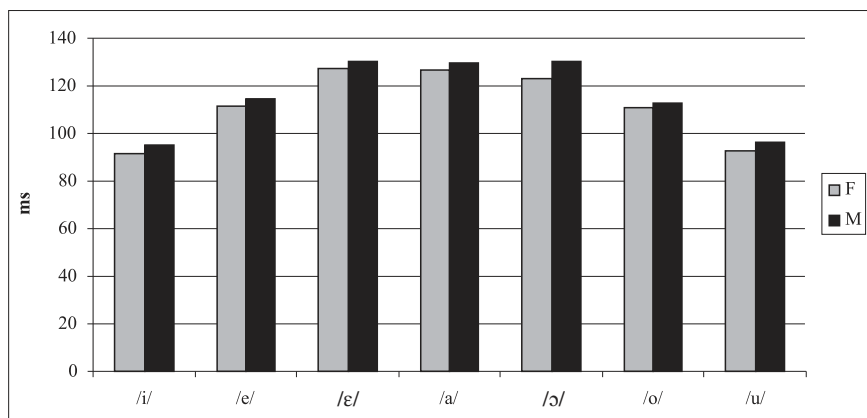


Figure 3: Vowel duration values by monolingual BP speakers

Thus, the duration results obtained with the participants of the present study indicate that BP vowels cannot be divided into a group of short and a group of long vowels. Low vowels differ significantly from mid and high vowels; however, this difference is not made to maximize contrasts, but simply because lower vowels need some more time to be articulated than do mid and high vowels.

4. Conclusion

The purpose of this article was to provide the acoustic characteristics of the BP oral vowels produced by Brazilian monolinguals from non-capital cities of Rio Grande do Sul, Paraná and Santa Catarina. The figures corroborated the results of previous studies (Faveri, 1991; Seara, 2000) in that duration correlates with F1, since low vowels are significantly longer than mid vowels, which are significantly longer than high vowels. Although there are duration differences between BP vowels, these differences do not imply that there are groups of long and short vowels arranged to maximize contrasts. As expected, the duration differences are simply an intrinsic vowel property, since the lower the vowel the more time it needs to be articulated. Therefore, although duration could be a cue used to define vowel height, the first formant is still the main cue to distinguish between vowels, possibly because of the reliability of this cue, which does not vary considerably according to speech rate.

Works cited

- ANDERSON, N. (1978), "On the calculation of filter coefficients for maximum entropy spectral analysis", *Modern Spectral Analysis*, IEEE Press, pp. 252-255.
- BOERSMA, Paul / Weenink, David (2006), *PRAAT: doing phonetics by computer (Version 4.4.23)* (Computer program), retrieved 12 June 2006, from <http://www.praat.org>.
- FAVERI, Cláudia Borges de (1991), *Análise da duração das vogais orais do português de Florianópolis-Santa Catarina*. Unpublished master's thesis. Universidade Federal de Santa Catarina.
- LIMA, Ronaldo (1991), *Análise acústica das vogais orais do português de Florianópolis*. Unpublished master's thesis. Universidade Federal de Santa Catarina.
- MORAES, João Antonio / CALLOU, Dinah / LEITE, Yonne (1996), "O sistema vocálico do português do Brasil: caracterização acústica", in M. Kato (ed.), *Gramática do Português Falado 5*, Campinas: Editora da Unicamp, pp. 33-53.
- PEREIRA, Ana Luiza Dias (2001), *Caracterização acústica do sistema vocálico tônico oral florianopolitano: alguns indícios de mudança*. Unpublished master's thesis. Universidade Federal de Santa Catarina.
- SEARA, Izabel Christine (2000), *Estudo acústico-perceptual da nasalidade das vogais do português brasileiro*. Unpublished doctoral dissertation. Universidade Federal de Santa Catarina.