A cross-dialect acoustic description of vowels: Brazilian and European Portuguese

Paola Escudero and Paul Boersma
Amsterdam Center for Language and Communication, University of Amsterdam, Spuistraat 210, 1012VT Amsterdam, The Netherlands

Andréia Schurt Rauber
Center for Studies in the Humanities, University of Minho Campus de Gualtar, 4710-057 Braga, Portugal

Ricardo Bion
International School for Advanced Studies, Cognitive Neuroscience Sector Via Stock, 2/2, 34135 Trieste, Italy

(Version 17 July 2008)

This paper examines four acoustic correlates of vowel identity in Brazilian and European Portuguese: first formant (F1), second formant (F2), duration, and fundamental frequency (F0). Both varieties of Portuguese display some universal phenomena: vowel-intrinsic duration, vowel-intrinsic pitch, gender-dependent size of the vowel space, gender-dependent duration, and an asymmetry in F1 between front and back vowels. Also, the average difference between the vocal tract sizes for /i/ and /u/, as measured from formant analyses, is comparable to the average difference between male and female vocal tract sizes. A language-specific phenomenon is that both varieties of Portuguese have a large vowel-intrinsic duration effect. Differences between Brazilian Portuguese (BP) and European Portuguese (EP) are found in duration (BP has longer stressed vowels than EP), in F1 (the lower mid vowels and possibly the low vowel are more open in BP than in EP), and in the size of the intrinsic pitch effect (larger for BP than for EP).

PACS number(s): 43.70.Fq, 43.70.Kv, 43.72.Ar

I. INTRODUCTION

The aim of this article is to describe and compare the acoustic characteristics of the seven oral vowels that Brazilian Portuguese (BP) and European Portuguese (EP) have in common in stressed position, i.e. the vowels /i, e, ɛ, a, ɔ, o, u/. Most earlier studies described either BP or EP vowels in impressionistic articulatory terms (e.g., Câmara, 1970; Mateus, 1990; Bisol, 1996; Mateus and d’Andrade, 1998, 2000; Cristófaro Silva, 2002; Barbosa and Albano, 2004). The few studies that have provided acoustic descriptions of BP or EP vowels (Delgado-Martins, 1973; Faveri, 1991; Lima, 1991; Callou *et al.*, 1996; Pereira, 2001) had several limitations that the present study overcomes.

Specifically, the earlier acoustic studies share five limitations: (1) none of them considered a comparison between dialects, and none of them followed the same procedure as any other, so that cross-dialectal comparisons cannot be made; (2) in all of them, data were collected from a small number of speakers (a minimum of 3 and a maximum of 8 per dialect); (3) none of them tested female subjects; (4) none of them controlled for the voicing of neighboring consonants despite its influence on vowel duration (Peterson and Lehiste, 1960), and the only study on EP (Delgado-Martins, 1973) had liquids and nasals following the target vowels, making it difficult to precisely identify vowel boundaries and control for vowel nasalization; and (5) none of these studies reported on the fundamental frequency of either variety, or on both the spectral quality and the duration of BP vowels.[1]

The present study aims at overcoming these limitations in the following ways: (1) it provides the first comparison between the acoustic properties of BP and EP vowels, and follows as closely as possible the methods of data collection reported in Adank *et al.* (2004) in order to allow future comparisons across experiments and languages; (2) forty speakers, 20 BP and 20 EP, produced a total of 5600 vowel tokens; (3) half of the speakers in each dialect were male and half were female; (4) the vowel tokens were always produced between voiceless stops or fricatives to allow easy and accurate formant measures; and (5) acoustic analyses were made of vowel duration, fundamental frequency and the first two vowel formants. In addition, five different consonantal contexts were used, which perhaps yields a less biased sample than the ones reported in many previous studies.

Additionally, the present study innovates in the methodology for formant analysis. It uses a new procedure to automatically define the formant ceiling of the LPC analysis based on within-speaker and within-vowel variation, thus allowing more accurate automatic formant measurements.

II. METHOD

A. Participants

In order to obtain relatively homogeneous and comparable groups of Brazilian and European Portuguese participants, all participants were chosen to be highly educated young adults from the largest metropolitan area in each country. They were selected from groups of volunteers that completed a background questionnaire: if they met three requirements, they could be enlisted as speakers for the present study. The requirements were that they had lived in either São Paulo or Lisbon throughout their lives, that they did not speak any foreign language with a proficiency of 3 or more on a scale from 0 to 7, and that they were undergraduate students under 30 years of age. In this way, 20 BP speakers from São Paulo and 20 EP speakers from Lisbon were selected. For each dialect there were equal numbers of men and women, so that the gender-dependence of the vowels could be investigated as easily as the dialect-dependence. For BP, the females' mean age was 23.2 years (standard deviation 4.3 years) and the males' mean age was 22.5 years (s.d. 4.7); for EP speakers, the females' mean age was 19.8 years (s.d. 1.5), the males' 18.7 years (s.d. 0.8).

[1] Exceptionally, Seara (2000) did control the consonantal environment and did measure F0, but being interested mainly in nasalized vowels, she considered only five of the seven oral vowels.

B. Procedure

The recordings were made in a quiet room with a Sony MZ-NHF800 minidisk recorder and a Sony ECM-MS907 condenser microphone, with a sample rate of 22 kHz and 16-bit accuracy. The BP data were collected at the Escola Superior de Propaganda e Marketing (ESPM) in São Paulo, and the EP data were collected at the Instituto de Engenharia de Sistemas e Computadores (INESC) and at the University of Lisbon, both in Lisbon.

The target vowels /i, e, ɛ, a, ɔ, o, u/ were orthographically presented to the speakers as *i*, *ê*, *é*, *a*, *ó*, *ô*, and *u*, respectively, embedded in a sentence written on a computer screen. Each vowel was produced as the first vowel in a disyllabic CVCV sequence (C = consonant, V = vowel), where the two consonants were two identical voiceless stops or fricatives; this yielded nonce words such as /pe̲po/ and /sa̲so/ (*pêpo* and *sasso*) where the underlined vowel is the target vowel. The consonants were always voiceless so as to allow easy measurement of duration; the analysis was restricted to the five consonants /p, t, k, f, s/, i.e. the voiceless consonants that Portuguese shares with Spanish, in order to allow future cross-language comparisons. The speakers always stressed the first syllable of the nonce word, helped by the orthographic conventions of Portuguese. In the final unstressed syllable, where Portuguese has only three vowels, the participants only read the vowels /e/ and /o/, which are usually pronounced as [ɪ] and [ʊ] in BP (Cristófaro-Silva, 2002, p. 86) and (if audible at all) as [ɨ] and [u] in EP (Mateus and d’Andrade, 2000, p. 18).

The disyllabic nonce words were read both in isolation and embedded in an immediately following carrier sentence similar to the one used in Adank *et al.* (2004). The sentences were read twice in two blocks; in the first block the isolated word had a final /e/, in the second block it had a final /o/. An example of an isolated word with sentence in block 1 was therefore “*Pêpe*. Em *pêpe* e *pêpo* temos *ê*”, which means ‘*Pêpe*. In *pêpe* and *pêpo* we have *ê*.’ The corresponding example from block 2 would be “*Pêpo*. Em *pêpe* e *pêpo* temos *ê*”.

The words and sentences were presented on a computer screen. In case the participants misread a word or sentence, they were asked to repeat it before the next word or sentence was presented.

Each participant thus produced eight tokens of each vowel embedded in each consonant context, from which only four were selected for analysis: the two words produced in isolation (i.e. one with final *e*, one with final *o*) and the two best exemplars (judged by the clarity of the recordings) of the words produced embedded in the carrier sentence (one with final *e*, one with final *o*); the isolated vowel at the end of the carrier sentence was not used in the analysis. Thus, 20 productions (2 phrasal positions x 2 word-final vowels x 5 consonantal contexts) were analyzed for each of the 7 vowels of each participant. This yielded a total of 2800 vowel tokens per dialect (20 productions x 7 vowels x 20 speakers).

C. Acoustic analysis: duration

For duration measurements the start and end points of each of the 5600 vowel tokens were labeled manually in the digitized sound wave. Because all flanking consonants were voiceless and unaspirated, the start and end points of the vowel could be determined relatively easily by finding the first and last periods that had considerable amplitude and whose shape resembled that of more central periods, with both points of the selection chosen to be at a zero crossing of the waveform.

D. Acoustic analysis: fundamental frequency

In order to determine the F0 of each of the 5600 vowel tokens, the computer program Praat (Boersma and Weenink, 1992–2007) was used to measure the F0 curves of all 40 recordings by the cross-correlation method, which is especially suitable for measuring short vowels. The pitch range for the analysis was set to 60–400 Hz for men and 120–400 Hz for women. If the analysis failed on any of the speaker's vowel tokens, i.e., if Praat considered the entire vowel centre voiceless, the analysis for that token was redone in a way depending on the speaker's gender: if the analysis failed for a woman (which happened for five of the 2800 tokens), the analysis was retried with a pitch floor of 75 Hz, and if it failed for a man (which happened for 1 of the 2800 tokens), the analysis was retried with a lower criterion for voicedness. In this way, all 5600 vowel tokens eventually yielded F0 values. To get a robust measure of the F0 of the vowel, the median F0 value was taken of the central 40 percent of the vowel: ignoring the first and last 30 percent of the vowel reduces the effect of the flanking consonants, and taking the median rather than the mean reduces the effect of F0 measurement errors.

E. Acoustic analysis: optimized formant ceilings

For each of the 5600 vowel tokens, F1 and F2 were determined with the Burg algorithm (Anderson, 1978), as built into the Praat program. The analysis was done on a single window that consisted of the central 40 percent of the vowel.[2] As an initial approximation, Praat was made to search for five formants in the range from 50 Hz to 5500 Hz (for female speakers) or 5000 Hz (for male speakers). The 1400 F1-F2 pairs for the Brazilian women are plotted in Fig. 1.

[2] A technical detail: the Gaussianlike shape of the window requires tails that capture another 20 percent of the vowel duration on each side of the central 40 percent.

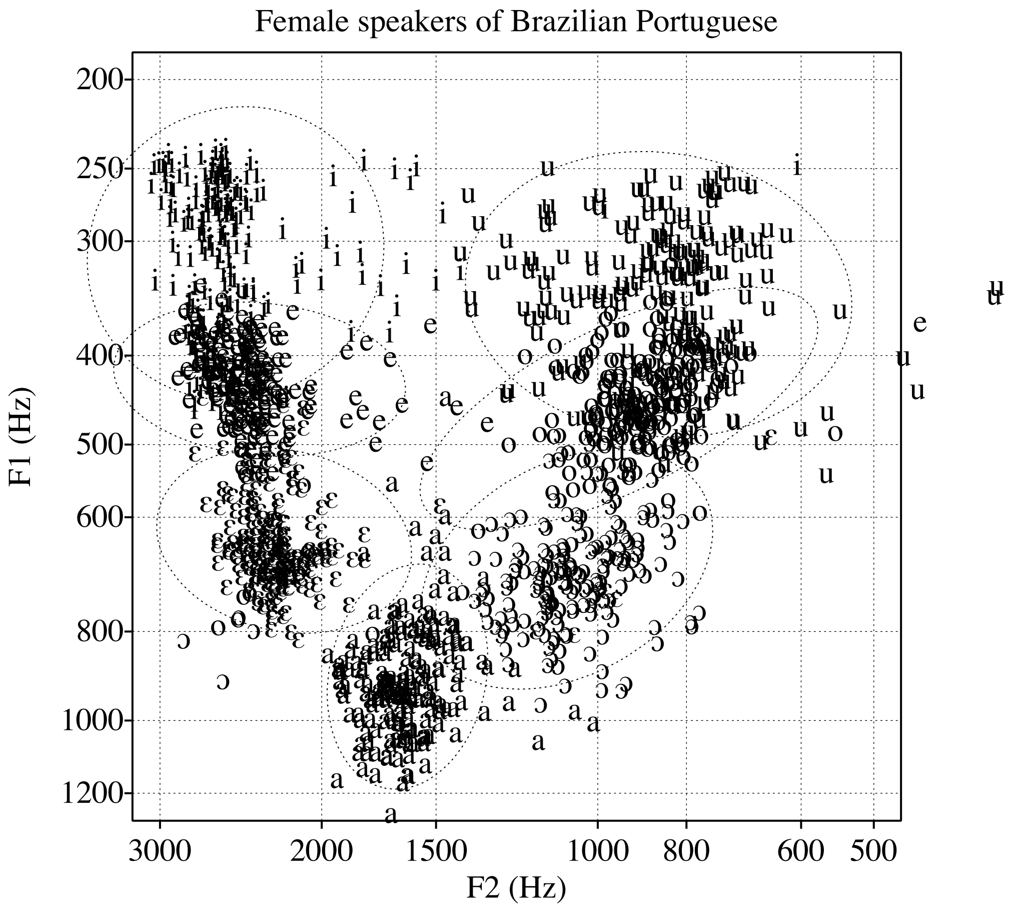

FIG. 1. The first and second formants of the 1400 vowel tokens of the Brazilian women, measured with a fixed (gender-specific) formant ceiling of 5500 Hz.

Figure 1 shows several unlikely values for some formants: for several back vowels the F2 has been analysed as nearly identical to F1, there are /ɔ/ and /o/ tokens in the lower left whose F2 has been analysed as an F1, and the second tracheal resonance of /i/, between 1500 and 2000 Hz (Stevens, 1998, p. 300), has often been analysed as an F2. Figure 1 shows the large 2σ ellipses that these outliers cause (for normally distributed formants, such ellipses are designed to cover 86.5 percent of the data points). Such shifts in the numbering of formants indicate that the fixed gender-specific formant ceilings of 5000 and 5500 Hz could be problematic.

Although the manner of visualization in Fig. 1 overrepresents the outliers, a method was designed to adapt the formant ceilings to the speaker and the vowel at hand. This could be done by some general method that optimizes a formant track by a number of criteria (e.g. Nearey, Assmann and Hillenbrand, 2002: smallest bandwidths; continuity in time; correlation between original and LPC-generated spectrogram), but the present paper instead takes advantage of the fortunate circumstance that each vowel was produced 20 times by each speaker.

The procedure to optimize the formant ceiling for a certain vowel of a certain speaker runs as follows. For all 20 tokens the first two formants are determined 201 times, namely for all ceilings between 4500 and 6500 Hz in steps of 10 Hz (for women) or for all ceilings between 4000 and 6000 Hz in steps of 10 Hz (for men). From the 201 ceilings, the ‘optimal ceiling' is chosen as the one that yields the lowest variation in the twenty measured F1-F2 pairs. This variation is computed along the same logarithmic scales as seen in Fig. 1, namely as the variance of the twenty log(F1) values plus the variance of the twenty log(F2) values. Thus, the procedure ends up with 280 optimal ceilings, one for each vowel of each speaker. With the 70 speaker-vowel-dependent ceilings for Brazilian women, Fig. 1 turns into Fig. 2.

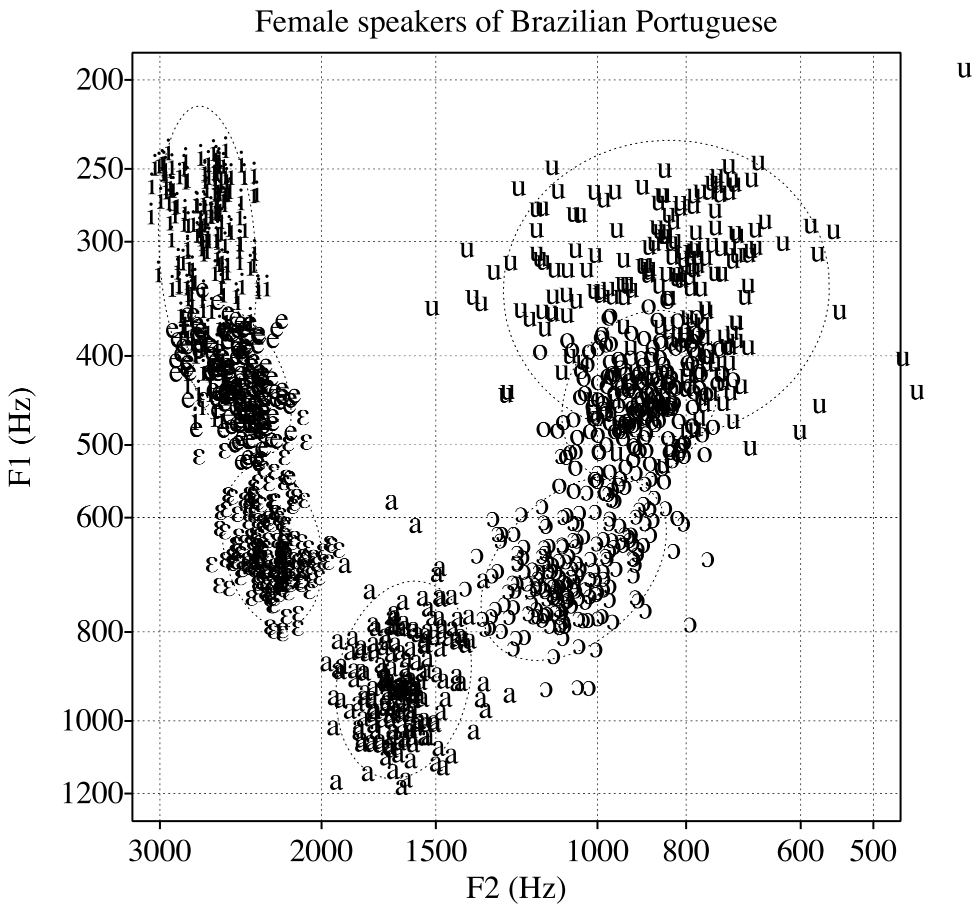

FIG. 2. The first and second formants of the 1400 vowel tokens of the Brazilian women, measured with optimized (speaker- and vowel-specific) formant ceilings.

Figure 2 shows that the variation between the vowel tokens has decreased appreciably: almost all outliers have gone, and although only the variation of the formant values of a vowel *within* a speaker (not that *between* speakers) has been explicitly minimized, the 2σ ellipses have shrunk, especially in the F2 direction.

To illustrate that the ceiling optimization method does something sensible, Fig. 3 shows the effects of gender and vowel on the optimal formant ceiling. Each vowel symbol in that table represents the median of 20 optimal ceilings (because there are 20 speakers of each gender and the two dialects are pooled).

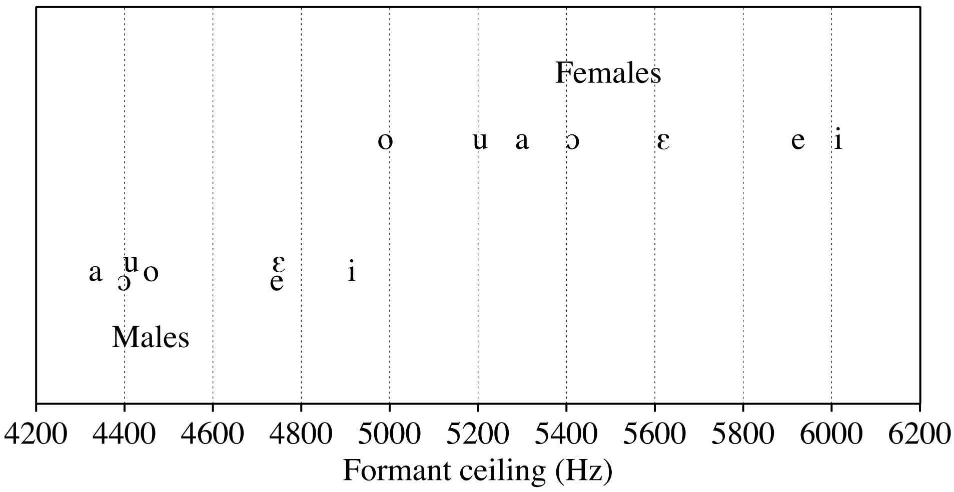

FIG. 3. Median optimal ceilings for each gender-vowel combination.

Figure 3 shows that both gender and vowel have strong effects on what the optimal ceiling is. The median of the 140 optimal ceilings for the women is 5450 Hz, and the median of the 140 optimal ceilings for the men is 4595 Hz, which is a factor of 1.186 lower. This difference must reflect the difference in vocal tract lengths between men and women; it constitutes a justification for the use of different formant ceilings for men and women in computer analyses for formant frequencies. Interestingly, however, the effect of vowel is of comparable size as the effect of gender: the median of the 40 optimal ceilings for /u/ is 4600 Hz, and the median of the 40 optimal ceilings for /i/ is 5625 Hz, which is a factor of 1.223 higher. This difference must reflect a difference in the length of the channel between the lips (rounded and protruded for /u/, spread and retracted for /i/) and probably a difference in the height of the larynx (lowered for /u/: Ewan and Krones, 1974; Riordan, 1977). This difference suggests that automated formant measurement methods should take into account vowel-related vocal tract lengths to a larger extent than they usually do.

III. SUMMARY OF RESULTS

Sections IV through VI present the detailed results of the acoustic measurements and statistical analyses aimed at investigating general properties of Portuguese and finding differences between the two dialects. These sections report the effects of vowel, gender and dialect on formants, duration, and fundamental frequency. Table I summarizes the average values for all these quantities (also shown in Figs. 6, 7, and 8); each number is a geometric average over 10 speaker values, each of which is a median over 20 tokens (2 phrasal positions x 2 word-final vowels x 5 consonant environments, see Sec. II B; using the median minimizes the influence of occasional measurement errors). Following much existing cross-dialectal work (Hagiwara, 1997; Adank *et al.*, 2004; Clopper *et al.*, 2005), the table has been split not only for dialect but also for gender, because males may speak differently as a group from females, and sound change (which is a likely source of any difference between BP and EP) may proceed with a different speed for males than for females (Labov, 1994, p. 156).

TABLE I. Geometric averages of vowel duration, F0, F1 and F2 for female (F) and male (M) speakers of Brazilian Portuguese (BP) and European Portuguese (EP).

			/i/	/e/	/ɛ/	/a/	/ɔ/	/o/	/u/
BP	Duration (ms)	F	99	122	141	144	139	123	100
		M	95	109	123	127	123	110	100
	F0 (Hz)	F	242	219	211	209	211	225	252
		M	137	131	124	122	122	132	140
	F1 (Hz)	F	307	425	646	910	681	442	337
		M	285	357	518	683	532	372	310
	F2 (Hz)	F	2676	2468	2271	1627	1054	893	812
		M	2198	2028	1831	1329	927	804	761
EP	Duration (ms)	F	92	106	115	122	118	110	94
		M	84	97	106	108	104	99	83
	F0 (Hz)	F	216	211	205	202	204	211	222
		M	126	122	117	115	117	123	127
	F1 (Hz)	F	313	402	511	781	592	422	335
		M	284	355	455	661	491	363	303
	F2 (Hz)	F	2760	2508	2360	1662	1118	921	862
		M	2161	1987	1836	1365	934	843	814

Since duration, F0, F1, and F2 are by definition positive quantities, all statistical investigations in the following sections are performed on log-transformed values. As a result, all figures use logarithmic axes, and all averages reported are geometric averages over the original values in milliseconds or Hertz, as in Table I; the same goes for any reported dimensionless factors.[3]

Each of the statistical investigations into duration, F0, F1, and F2 (Secs. IV B, IV E, V, VI) starts out with an exploratory repeated-measures analysis of variance (conducted with SPSS) on 280 logarithmic values (40 speakers x 7 vowels), with dialect and gender as between-subjects factors and vowel as a within-subjects factor. For all four acoustical dimensions, Mauchly’s sphericity test suggests that the numbers of degrees of freedom for the vowel effects have to be reduced. Accordingly, it was decided to use Huynh-Feldt’s correction, which multiplies the degrees of freedom (6 for the numerator, 216 for the denominator in the F-test) by a factor ε, which tends to be around 0.5.

IV. RESULTS FOR FORMANTS

A. The speakers’ median formants

Figures 4 and 5 show the median F1 and F2 values for the 10 female and 10 male speakers of each dialect. In each of the four figures, each vowel occurs ten times because there were 10 speakers of that gender and dialect. Each vowel symbol’s vertical position represents the median of the speaker’s 20 F1 values, and its horizontal position represents the median of the speaker’s 20 F2 values. The 20 F1-F2 pairs that lie behind each vowel symbol

[3] A geometric average is computed by the exponentiation of the arithmetic average of log-transformed values.

were all measured with the same formant ceiling, namely the formant ceiling that minimizes the variance within the 20 F1 and F2 values (Sec. II E).

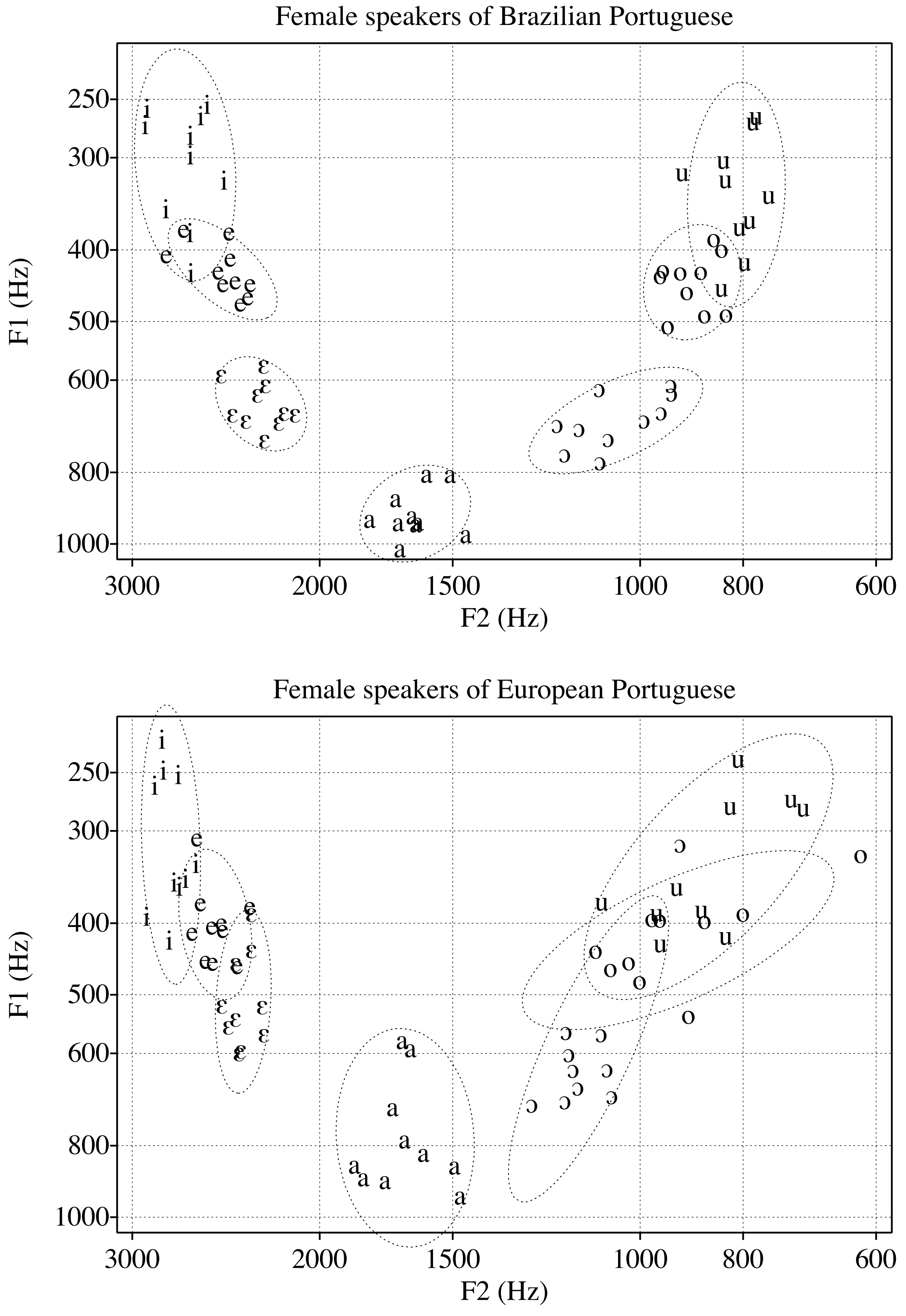

FIG. 4. First and second formants of ten Brazilian and ten European Portuguese women.

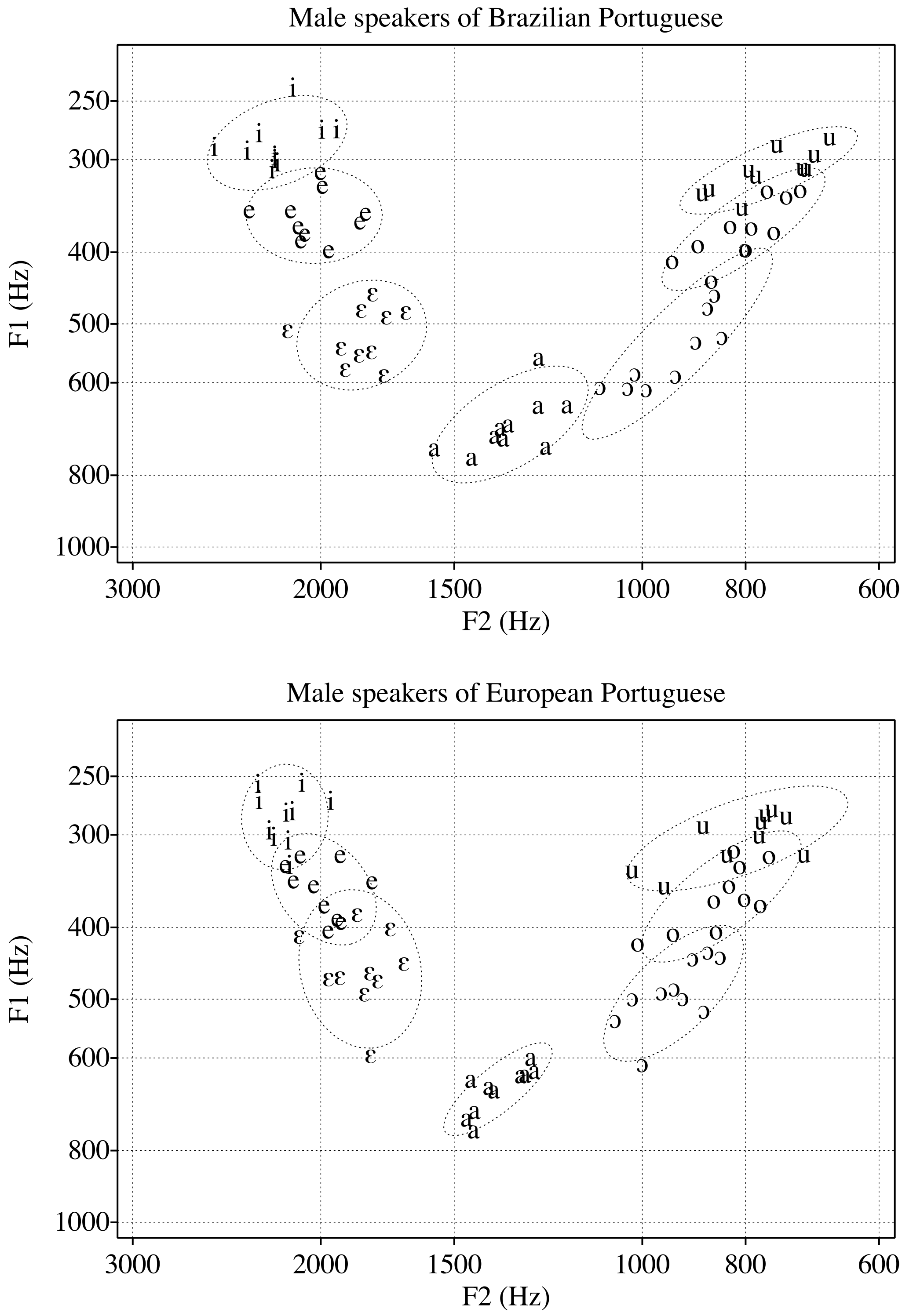

FIG. 5. First and second formants of ten Brazilian and ten European Portuguese men.

Figure 6 shows the mean F1 and F2 values for the seven vowels for the four groups. Each symbol represents a geometric mean of 10 speakers' median F1 and F2 values. The following sections consider F1 and F2 separately.

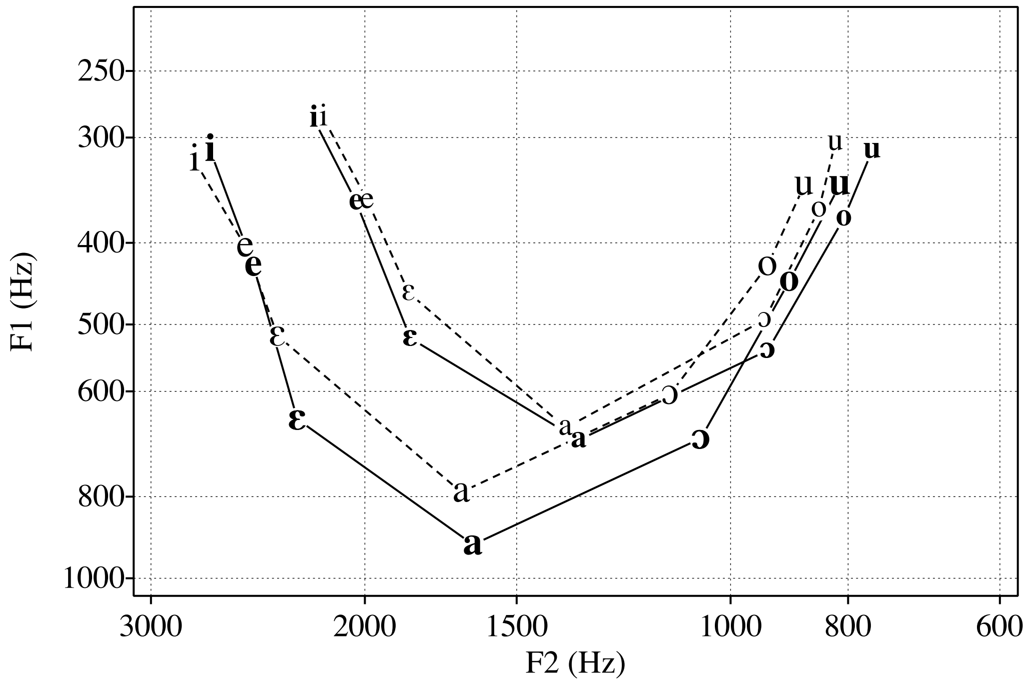

FIG. 6. The vowel spaces of the four groups.
Solid lines and bold symbols = Brazilian Portuguese; dashed lines = European Portuguese.
Large font: women; small font: men.

B. The effect of vowel and gender on F1

The exploratory analysis of variance reveals a main effect of vowel on F1 ($F(6\varepsilon,216\varepsilon,\varepsilon{=}0.609) = 684.926$; $p = 9{\cdot}10^{-85}$): different vowels tend to have different F1 values. The main determiner of F1 is the phonological vowel height: the low vowel /a/ has the highest F1, followed by the lower mid vowels /ɛ/ and /ɔ/, then the higher mid vowels /e/ and /o/, and finally the high vowels /i/ and /u/ which have the lowest F1. A consistent secondary effect on F1 is caused by the front-back distinction. Here it is convenient that all vowels were measured with the same speakers. For 38 of the 40 speakers, the F1 of /u/ is higher than the F1 of the same speaker's /i/. Averaged over all 40 speakers, the F1 of /u/ is higher than that of /i/ by a factor of 1.082; although this front-back effect is small, the factor is very reliably different from 1 ($t(39) = 10.00$; one-tailed $p = 1.3{\cdot}10^{-12}$), and its 95% confidence interval runs from 1.065 to 1.099. Likewise, the F1 of /o/ is higher than that of /e/ for 32 of the 40 speakers, who have an average ratio of 1.039 ($t(39) = 4.67$; $p = 1.8{\cdot}10^{-5}$; confidence interval from 1.022 to 1.056). Finally, the F1 of /ɔ/ is higher than that of /ɛ/ for 35 of the 40 speakers, who have an average ratio of 1.078 ($t(39) = 4.27$; $p = 6.1{\cdot}10^{-5}$; confidence interval from 1.040 to 1.118). Thus, there is a consistent correlation between F1 and phonological backness. By not labelling the vowel symbols for speaker, Figs. 4 and 5 obscure this consistent effect (for instance, the four European Portuguese speakers with the conspicuously low F1 values for /i/ in Fig. 4 are the same as those with the conspicuously low F1 values for /u/). Figure 6 does show that each back vowel symbol is lower (has a higher F1) than its front counterpart, but it does not show that this holds for the individual speakers.

The next observation to be made is that the analyses indicate that the interaction of gender and dialect is not reliably different from zero (p = 0.488), and neither is the triple interaction of gender, dialect and vowel (p = 0.298). This is important because it restricts the number of explanations that have to be considered for the effects that *are* found, thereby facilitating the analyses in the next sections.

The analyses reveal a main effect of gender on F1 ($F(1,36) = 23.430$; $p = 2.4\cdot10^{-5}$): women tend to have higher F1 values (geometric average: 478 Hz) than men (409 Hz). The gender effect on F1 is therefore a ratio of 1.170, which compares well (as it should) with the female-male ratio of 1.186 found for the optimal formant ceilings in Sec. II E.

It is possible that the effect of gender just found may have to be viewed in relation with interaction effects. Since two of the possible interactions (i.e. those involving dialect and gender) were ruled out above, one only has to consider the interaction effect of gender and vowel on F1, which is indeed reliable ($F(6\varepsilon,216\varepsilon,\varepsilon{=}0.609) = 4.604$; $p = 0.0023$). Figure 6 suggests that this is because women take up a greater part of the vowel height continuum than men. This is investigated in the next section.

C. The effect of gender and dialect on the size of the F1 space

Each speaker's *F1 space size* can be defined as the ratio of the F1 of their low vowel /a/ and the (geometric) average F1 of their high vowels /i/ and /u/. Thus computed, the average F1 space size of the 20 women turns out to be 2.613, that of the 20 men only 2.276. For the combined population of BP and EP speakers, the female F1 space is therefore 2.613/2.276 = 1.148 times (0.199 octaves) bigger than the male F1 space (the 95% confidence interval of this ratio is 1.047..1.259; the ratio is different from 1 with $F(1,36) = 9.052$, $p = 0.005$). As suggested at the end of Sec. IV B, therefore, Portuguese-speaking women indeed take up a larger part of the F1 space than men, even along a logarithmic F1 scale. For a comparison with other languages see Sec. VII A.

The F1 space size may also depend on the dialect. The average F1 space size of the 20 Brazilians is 2.552, that of the Europeans 2.331. For the combined population of men and women, the Brazilian F1 space is therefore 1.095 times bigger than the European F1 space (c.i. = 0.998..1.202). This may not be very reliably different from 1 ($F(1,36) = 3.895$, $p = 0.056$), but it should be taken into account as a possible explanation for later findings (Sec. IV D).

An interaction between gender and dialect was not found ($F(1,36) = 2.395$, $p = 0.130$).

D. Comparing F1 across dialects with gender normalization

The analysis of variance suggests a main effect of dialect on F1 ($F(1,36) = 4.052$; $p = 0.052$), but the cause of this is probably the reliable interaction effect of dialect and vowel on F1 ($F(6\varepsilon,216\varepsilon,\varepsilon{=}0.609) = 6.777$; $p = 9.5\cdot10^{-5}$). Apparently, some vowels have different heights in (São Paulo) Brazilian than in (Lisbon) European Portuguese.

By performing a multivariate analysis of variance on the seven F1 values, with dialect and gender as factors, one can determine which vowels are different in the two dialects (the dialect-gender interaction term is included in the model but not significant for any of the 7 vowels; its exclusion from the model turns out not to change the results). The vowel /ɛ/ turns out to be very reliably lower (higher F1) in Brazilian than in European Portuguese

($F(1,36) = 27.468$, $p = 7.1 \cdot 10^{-6}$). Its back counterpart /ɔ/ ($F(1,36) = 4.973$, $p = 0.032$) and the vowel /a/ ($F(1,36) = 7.162$, $p = 0.011$) are probably lower as well.

These results apply if one regards the vowel tokens of a certain category as independent from the vowel tokens of any other category. However, all 7 vowels have been spoken by the same 40 speakers. Distances within the speaker's vowel inventories may shed light on the cause behind the finding that the lower mid vowels /ɛ/ and /ɔ/ have a higher F1 in Brazilian than in European Portuguese. Thus, the log(F1) differences between every speaker's /e/ and /ɛ/ were computed. It turns out that the Brazilian F1 ratio of /ɛ/ and /e/ (estimated as 1.485; 95% c.i. = 1.452..1.519) is very reliably greater than the European /ɛ/–/e/ F1 ratio (estimated as 1.276; 95% c.i. = 1.223..1.332): the ratio of these ratios is estimated as 1.485/1.276 = 1.164, with a 95% confidence interval of 1.110..1.219, which is different from 1 with $t(38) = 6.564$ and a two-tailed $p = 9.6 \cdot 10^{-8}$ (genders pooled) or with $F(1,36) = 43.391$ and a two-tailed $p = 1.1 \cdot 10^{-7}$ (analysis of variance with dialect and gender as factors). Likewise, the Brazilian F1 ratio of /ɔ/ and /o/ (estimated as 1.482; c.i. = 1.416..1.552) is reliably greater than the European /ɔ/–/o/ F1 ratio (1.377; c.i. = 1.298..1.461): the ratio of these ratios is estimated as 1.076 (c.i. = 1.001..1.157), which is different from 1 with $t(38) = 2.057$, two-tailed $p = 0.047$.

Is this difference due to /ɛ/ and /ɔ/ being lower in Brazilian than in European Portuguese, or due to /e/ and /o/ being higher in Brazilian than in European Portuguese? Table I and Fig. 6 indicate that the latter possibility is unlikely: for both women and men, the median Brazilian /e/ and /o/ are *lower* than the median Portuguese /e/ and /o/. It is more likely that the relative openness of the lower mid vowels in BP is due to the larger F1 space that BP speakers may be using (Sec. IV C). To assess this hypothesis, the *relative heights* of all vowels within the log(F1) space are computed for each speaker.

For that purpose, the relative height of each speaker's /a/ is defined as zero, and the relative height of the average height of her /i/ and /u/ is defined as 1. Linear interpolation yields 40 relative height values for each vowel (except /a/). A two-way analysis of variance reveals no effect of gender on the relative height of any of the six vowels (all six p values are 0.259 or more) and no interaction of dialect and gender (all six p values are 0.539 or more). The two genders can therefore be pooled, and the 20 relative heights of the Brazilian Portuguese can simply be compared with those of the European Portuguese.

If all vowels were equally spaced along the log(F1) dimension, the lower mid vowels would have a relative height of 0.333. The average Brazilian /ɛ/ indeed has a relative height of 0.329 (c.i. = 0.298..0.361), but the average EP /ɛ/ has a relative height of 0.473 (c.i. = 0.425..0.521), i.e. it lies close to the centre of the F1 dimension; the difference between the dialects is highly reliable ($t(38) = 5.253$; $p = 6.0 \cdot 10^{-6}$). For /ɔ/, the difference between BP and EP is in the same direction (0.289 versus 0.340), but the statistical reliability has apparently been lost in normalization ($p = 0.241$).

The higher mid vowels seem to have very similar relative heights in the two dialects: /e/ has 0.764 for BP and 0.766 for EP, and /o/ has 0.716 for BP and 0.718 for EP. As expected, the high vowels have relative heights around 1: /i/ has 1.048 for BP and 1.041 for EP, and /u/ has 0.952 for BP and 0.959 for EP. Finally, /a/ has of course a relative height of 0 for everybody.

E. The effect of gender and dialect on F2

The analysis of variance reveals a large main effect of gender on F2 ($F(1,36) = 120.857$; $p = 4.7 \cdot 10^{-13}$): women's F2 values are higher than those of men by an average factor of 1.183, which compares well with the values found for the formant ceiling in Sec. II E and for F1 in Sec. IV B. A main effect of dialect ($F(1,36) = 3.009$; $p = 0.091$) cannot be excluded. An interaction of dialect and gender is not found ($F(1,36) < 1$).

The analysis also reveals a main effect of vowel ($F(6\varepsilon, 216\varepsilon, \varepsilon{=}0.423) = 1826.704$; $p = 1.6 \cdot 10^{-78}$): not all vowels have the same F2. A reliable interaction between vowel and gender is found ($F(6\varepsilon, 216\varepsilon, \varepsilon{=}0.423) = 9.339$; $p = 0.00006$); apparently, the size of the F2 space is larger for females than for males: although the F2s of /u/ are comparable for the two genders, those of /i/ are very different. As with F1, the analysis reveals no interaction between vowel and dialect ($F<1$) and no triple interaction between vowel, dialect, and gender ($F<1$).

The analyses of variance on the F2 values of the 7 vowels separately reveal that all vowels, except perhaps /u/ ($F(1,36) = 3.329$; $p = 0.076$) and /o/ ($F(1,36) = 8.125$; $p = 0.007$), have a very reliably higher F2 for women than for men (all five remaining F values are 28.953 or higher, i.e. $p \leq 4.7 \cdot 10^{-6}$). The only barely detectable effect of dialect on F2 may be that /u/ is more fronted in European than in Brazilian Portuguese ($F(1,36) = 3.676$; $p = 0.063$). No reliable interaction between dialect and gender is seen for any of the vowels.

As with F1, the size of the F2 space is greater for Portuguese-speaking women than for men. For each speaker, this size is computed as the ratio of the F2 of her /i/ and the F2 of her /u/. For the 20 men, the average ratio is 2.768 (c.i. = 2.606..2.940), for the 20 women it is 3.249 (c.i. = 3.068..3.440); the ratio of these ratios is 1.174 (c.i. = 1.083..1.272).

V. RESULTS FOR DURATION

No literature on the Portuguese vowel system suggests that vowel length could be a phonological feature. This does not preclude, however, that different vowels may have quite different phonetic durations, and that vowel durations may differ between dialects and between genders. Figure 7 shows the dependence of duration on vowel, dialect and gender. Each symbol represents a value of duration (and F2) averaged over the median duration (and F2) values of 10 speakers.

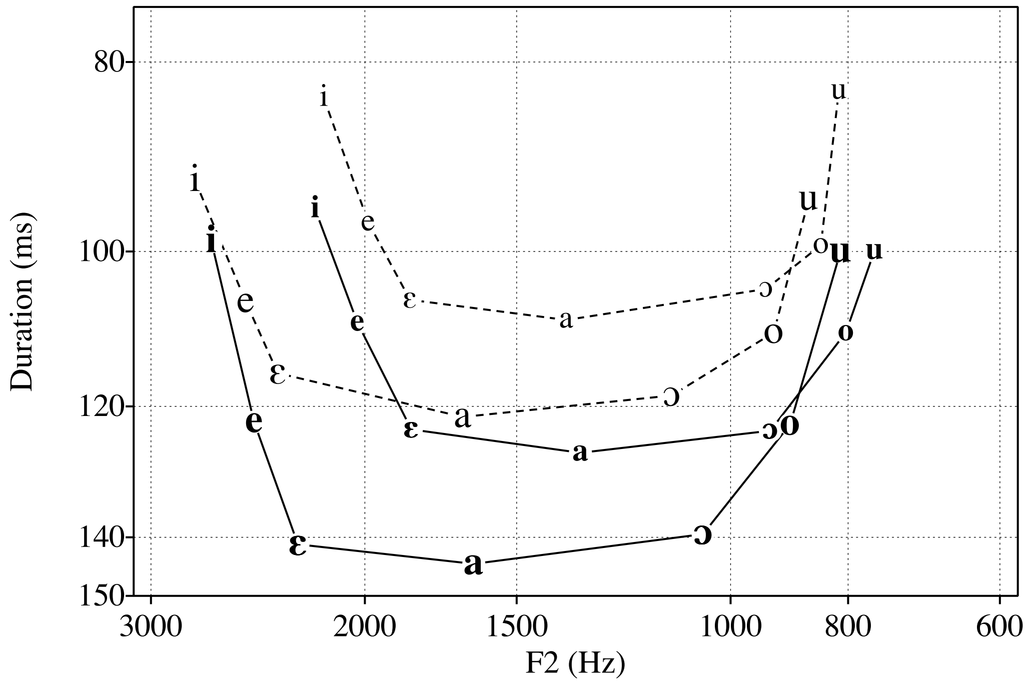

FIG. 7. Mean duration as a function of vowel.
Solid lines and bold symbols = Brazilian Portuguese; dashed lines = European Portuguese.
Large font: women; small font: men.

An exploratory analysis of variance reveals that the duration of the vowels is influenced by vowel ($F(6\varepsilon,216\varepsilon,\varepsilon=0.811) = 243.358$, $p = 5\cdot10^{-76}$), gender ($F(1,36) = 4.125$, $p = 0.050$) and dialect ($F(1,36) = 7.915$, $p = 0.008$). The analysis does not reveal an interaction between gender and dialect ($F(1,36) < 1$), i.e. the difference between the two solid curves in Fig. 7 is not reliably different from the difference between the two dashed curves. The two-way interactions between gender and vowel and between dialect and vowel, and the three-way interaction between gender, dialect, and vowel have low F-values (2.426, 3.829, and 3.671, respectively) that are nevertheless moderately reliable, at least under the somewhat forgiving Huynh-Feldt correction ($df1 = 6\varepsilon$, $df2 = 216\varepsilon$, $\varepsilon=0.811$: $p = 0.039$, 0.0028, 0.0038). The reader can consult Fig. 7 for possible causes of these potential interactions; the following paragraphs describe only the main effects.

The main effects can be described as follows. First, all vowels are longer for women than for men: the average factor is 1.105 (the 95% confidence interval is 1.053..1.160). Second, all vowels are longer in Brazilian Portuguese than in European Portuguese: the average factor is 1.148 (c.i. = 1.096..1.203). Third, the duration depends somewhat on the front-back distinction: /i/ is shorter than /u/ ($p = 0.021$), and /e/ is shorter than /o/ ($p = 0.014$); the difference between /ɛ/ and /ɔ/ is not significant. Fourth, the duration depends strongly on vowel height; this is investigated in detail in the following paragraphs.

There is a large effect of vowel height on duration, as can be seen in Fig. 7. For 39 of the 40 speakers, the median of her 20 measured /i/ tokens is shorter than the median of her 20 measured /e/ tokens. Reliable differences between vowels of different phonological heights are: /i,u/ are shorter than /e,o/ (all four comparisons yield $p < 10^{-12}$), /e,o/ are shorter than /ɛ,ɔ/ (all four $p < 10^{-9}$), and /ɛ,ɔ/ are shorter than /a/ (one-tailed p = 0.0029 and 0.00010).

One can again utilize the fact that all vowels have been spoken by the same speakers. For each speaker, the *vowel-intrinsic duration ratio* is computed as the ratio between the duration of their /a/ and the average duration of their /i/ and /u/. Thus computed, the average vowel-intrinsic duration ratio of the 40 speakers is 1.339 (c.i. = 1.300..1.379). The ratio is influenced by dialect ($F(1,36) = 3.988$, $p = 0.053$), gender ($F(1,36) = 4.794$, $p = 0.035$), and an interaction of dialect and gender ($F(1,36) = 4.454$, $p = 0.042$): it seems that the BP females have a larger vowel-intrinsic duration ratio than the other groups.

To investigate whether duration could be a direct consequence of F1, one can look at the extent to which duration and F1 covary *within* vowels (rather than *between* vowels as above). For this, the original 5600 measurements of F1 and duration have to be considered again. Since phrasal position is likely to influence both F1 and duration strongly, we consider here the seven isolated-word vowels and the seven sentence-internal vowels separately. Each of the 14 vowels, then, comes with 400 tokens (40 speakers x 5 following consonants x 2 word-final vowels), each of which can be normalized for speaker (and therefore for dialect and gender and their interaction) and normalized for the main effects (averaged over all 40 speakers) of the five following consonants (p, t, k, f, s) and the two word-final vowels (e, o). With speaker, vowel, phrasal position, following consonant, and word-final vowel factored out (and under the assumption that factors like gender and following consonant do not interact), only two of the 14 vowels show a reliable correlation between F1 and duration: the vowel /e/ in isolated words ($r = -0.15$, $p = 0.001$) and the vowel /ɛ/ in sentence-internal position ($r = +0.22$, $p < 0.00001$). The only straightforward generalization one could formulate about these two pieces of data is that longer vowel tokens tend to be more peripheral, i.e. lengthening causes closed vowels to become more closed and open vowels to become more open. What the data do *not* show, specifically, is a monotonic relation between duration and F1 within vowels; if the vowel-intrinsic duration found in the previous paragraphs had been due to such a relation, all reliable correlation coefficients would have been expected to be positive.

VI. RESULTS FOR FUNDAMENTAL FREQUENCY

The exploratory analysis of variance of F0 finds a main effect of gender ($F(1,36) = 179.842$, $p = 1.4 \cdot 10^{-15}$), a possible but nonreliable main effect of dialect ($F(1,36)=2.641$, $p = 0.113$), and no interaction between gender and dialect ($F(1,36) = 0.009$, $p = 0.924$). Within speakers there is a main effect of vowel ($F(6\varepsilon,216\varepsilon,\varepsilon=0.494)=136.031$, $p = 4.0 \cdot 10^{-36}$) and an interaction of vowel and dialect ($F = 11.314$, $p = 1.9 \cdot 10^{-6}$); an interaction of vowel and gender ($F = 2.542$) and a triple interaction of vowel, gender, and dialect ($F = 2.284$) might exist but have not been reliably detected ($p = 0.061$ and $p = 0.084$, respectively).

Since the exploratory analysis did not find any reliable interactions between gender and dialect or between gender and vowel, one can pool the BP and EP speakers when measuring the influence of gender on F0, and one can average over the seven vowels. The effect of gender on F0 turns out to be large: the 20 women have an average F0 of 216.63 Hz, the 20 men one of 125.07 Hz. The F0 of Portuguese-speaking women is therefore a factor of 1.732 higher than the F0 of Portuguese-speaking men, with a 95% confidence interval of 1.593..1.883 ($t(38) = 13.30$).

The exploratory analysis found a reliable interaction effect of vowel and dialect on F0, i.e., the effect of vowel on F0 depends on the dialect. As with duration, there are reliable differences between vowels of different phonological heights (Fig. 8): within-speaker comparisons reveal that /i,u/ have a higher F0 than /e,o/ (all four $p < 10^{-8}$), /e,o/ higher than /ɛ,ɔ,a/ (all four $p < 10^{-10}$), and /ɛ,ɔ/ higher than /a/ (one-tailed $p = 0.00019$ and 0.0029). Each speaker's *vowel-intrinsic F0 ratio* is computed as the ratio between the average F0 of the high vowels /i/ and /u/ and the F0 of the low vowel /a/. For all 40 speakers, this ratio turns out to be greater than 1. There is a reliable main effect of dialect ($F(1,36) = 12.497$, $p = 0.0011$): the average ratio is 1.157 for the 20 Brazilians and 1.094 for the 20 Europeans. The ratio is therefore greater for BP than for EP by a factor of 1.057 (c.i. = 1.023..1.093; $t(38) = 3.43$). Neither a main effect of gender ($F(1,36) = 1.042$, $p = 0.314$) nor an interaction between gender and dialect ($F(1,36) = 3.434$, $p = 0.072$) is reliably detected.

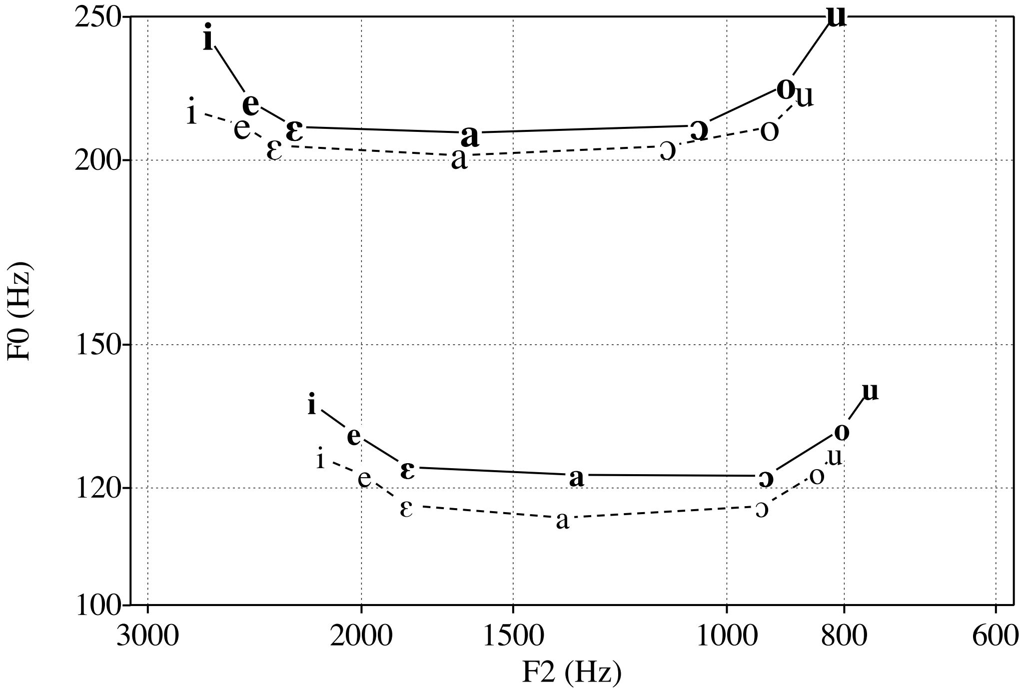

FIG. 8. Mean F0 as a function of vowel.
Solid lines and bold symbols = Brazilian Portuguese; dashed lines = European Portuguese.
Top: women; bottom: men.

The cause of the vowel-intrinsic F0 (the anti-correlation of F0 and F1 between vowels) can be investigated in the same way as the cause of the intrinsic duration in Sec. V, namely by computing within-vowel correlations between F0 and F1, with the main effects of speaker, following consonant, and word-final vowel factored out. Three of the 14 vowels show reliable positive correlations between F1 and F0: the vowel /ɛ/ in isolated words ($r = +0.15$, $p = 0.001$), sentence-internal /e/ ($r = +0.28$, $p < 10^{-8}$), and sentence-internal /o/ ($r = +0.26$, $p < 10^{-6}$). These correlations are opposite from what one would expect if the relation between F1 and F0 were an automatic result of the articulation.

VII. DISCUSSION

This section compares the results of Secs. IV to VI to earlier findings in the literature, and tries to find explanations for the phenomena observed. Universal aspects, Portuguese-specific aspects, and dialect-specific aspects are identified.

A. First formant: universal, Portuguese-specific, dialect-specific

Section IV B makes three observations. First, phonological vowel height is a strong determiner of F1 in Portuguese. This is an unsurprising observation given that Portuguese, as all languages, uses vowel height to distinguish between vowel categories. Second, back vowels have consistently higher F1 values than their front counterparts. This was found before for Brazilian Portuguese by Moraes *et al.* (1996, p. 35) and Seara (2000, pp. 80, 91, 102, 112, 141), and may reflect a universal principle, because it was also found for American English (Peterson and Barney, 1952; Clopper *et al.*, 2005; Strange *et al.*, 2007), Iberian Spanish (Cervera *et al.*, 2001), Parisian French (Strange *et al.*, 2007), Northern German (Strange *et al.*, 2007), and Dutch (Koopmans-van Beinum, 1980)[4]. Third, women tend to have higher F1 values than men. This is an unsurprising observation reported abundantly in the previous literature (e.g. Peterson and Barney, 1952), and well understood in terms of the differences in vocal tract length between women and men. The gender effect on F1 is a ratio of 1.170.

According to Sec. IV C, the Brazilian Portuguese F1 space size is 1.201 times larger for females than for males, and for the European Portuguese speakers this *female-to-male F1 space size ratio* is 1.097. In order to assess the universality of this gender difference, one can compare these ratios to those of other languages. It is difficult to compare F1 values between studies because of the different data collection methods (speaking rate, speaking style) and different formant analysis methods (formant ceilings, number of formants measured, pre-emphasis). One can hope, however, that most of these issues have little influence on the female-male F1 ratio that one can extract from any specific study. For the American English speakers of Peterson and Barney (1952), then, the ratio is 0.978. For the American English speakers of Hillenbrand *et al.* (1995), the ratio is also 0.978. This suggests that American English women have a vowel space that may be shifted with respect to that of American English men, but is not larger (along a logarithmic scale). For the Northern Standard Dutch speakers of Adank *et al.* (2004), the ratio is 1.260, and for the Southern Standard Dutch speakers in that study the ratio is 1.032; apparently, there can be large differences between closely related varieties in this respect. Both Portuguese values fall in between the two Dutch ones. The cause of the effect has been sought in the physiology (Simpson, 2001) as well as in the idea that males reduce their F1 space size because their F1 values are better to discriminate by listeners than female F1 values (Goldstein, 1980; Ryalls and Lieberman, 1982; Diehl, Lindblom, Hoemeke, and Fahey, 1996).

The combined evidence of Sec. III D leads to the conclusion that /ɛ/, and possibly /ɔ/ as well, are lower (more open, having a higher absolute and, at least for /ɛ/, relative F1) in Brazilian Portuguese from São Paulo than in European Portuguese from Lisbon. To the knowledge of the present authors, this observation has not been reported before.

[4] Adank *et al.* (2004) do not confirm this result for either of the two regional standard varieties of Dutch that they investigate.

B. Second formant: universal, Portuguese-specific, dialect-specific

Section IV E makes four observations. First, not all vowels have the same F2. This corresponds to the undisputable fact that Portuguese contrasts front and back vowels. Second, women have higher F2 values than men. As with F1, the well-understood explanation lies in the differences between the vocal tract sizes (the gender effect on F2 is a ratio of 1.183, which is comparable to the effect on F1). Third, /u/ may be more fronted in European than in Brazilian Portuguese. This could have been seen by comparing earlier publications on BP (Faveri, 1991; Lima, 1991; Callou *et al.*, 1996; Pereira, 2001) and EP (Delgado-Martins, 1973).

Fourth, Portuguese-speaking women not only have larger F1 space sizes than men, they also have larger F2 space sizes. The average Portuguese *female-to-male F2 space size ratio* is 1.174. For the American English speakers of Peterson and Barney (1952), the ratio is 1.116; for those of Hillenbrand *et al.* (1995), it is 1.089. For the Northern Dutch speakers of Adank *et al.* (2004), the ratio is 1.002, for the Southerners it is 1.166 (when compared with the F1 case, it is now the opposite group that exhibits large gender differences). The Portuguese ratio seems to be larger than that of English and Dutch. However, the large confidence interval reported in Sec. IV E, together with the presumably equally large uncertainties in the values reported for other languages, do not allow big conclusions to be drawn.

C. Duration: universal, Portuguese-specific, dialect-specific

Section V identifies four influences on duration in Portuguese.

First, vowels are longer for women than for men. This influence of gender on duration is not specific to Portuguese. Simpson and Ericsdotter (2003) report on many studies which find that female speakers produce longer vowels than male speakers in many Indo-European languages, such as English, German, Jamaican Creoles, French and Swedish, but also in non-Indo-European languages, such as Creek. This gender effect may have a socio-phonetic origin (Byrd, 1992; Whiteside, 1996), e.g. women tend to speak more clearly than men, or a physiological one (Simpson, 2001, 2002).

Second, vowels are longer in Brazilian Portuguese than in European Portuguese. A comparable difference has been found in the Spanish-speaking neighbours: Morrison and Escudero (2007) found that Peruvian Spanish vowels (from Lima) were 34% longer than European Spanish vowels (from Madrid).

Third, each vowel comes with its own typical duration. A height-dependence of duration has been found before for several languages, such as English (House and Fairbanks, 1953, p. 111) and French (Rochet and Rochet, 1991, p. 57, Fig. 7b). In fact, the effect is so widespread that Lehiste (1970, p. 18) calls it *intrinsic vowel duration*. As for the cause of the effect, a recent review on controlled and mechanical properties of speech (Solé, 2007, p. 303) follows Lehiste (1970, p. 18–19) in regarding it as a universal physiological property of speech production: open vowels require more jaw lowering, hence more time, than closed vowels. In Portuguese, the effect turns out to be strong: the duration ratio of low and high vowels is 1.339. This is more than in most other languages without a phonological length contrast, such as Iberian Spanish (Cervera *et al.*, 2001: a ratio of 1.14; Morrison and Escudero, 2007: 1.04), Peruvian Spanish (Morrison and Escudero, 2007: 0.94), or European French (Rochet and Rochet, 1991: a ratio of 1.13; Strange *et al.*, 2007: 1.11). Although Sec. V did not find an automatic effect of F1 on duration (or the reverse), this does not rule out the existence of such an automatic effect completely. However, since in Portuguese the vowel-intrinsic duration effect is stronger than in other languages, one can well conceive that Portuguese must have turned duration into a language-specific cue for phonological vowel identity, analogously to how e.g. English vowel duration has become a cue for the phonological voicing of a following obstruent.

Fourth, the back vowels seem to be longer than their front counterparts. For the high vowels, this was also found by Seara (2000).

D. Fundamental frequency: universal, Portuguese-specific, dialect-specific

Section VI identifies three influences on F0.

First, Portuguese-speaking women have a higher average F0 than men. While the direction of the effect is well understood in terms of differences in vocal fold length and weight, the size of the effect may be language-specific. The ratio of 1.732 found here can be compared to the ratios of 1.687 and 1.690 found for American English by Peterson and Barney (1952) and Hillenbrand *et al.* (1995), respectively. The data of Adank *et al.* (2004) reveal ratios of 1.497 for Northern Dutch and 1.730 for Southern Dutch. The ratio is much smaller than that found for Japanese (Yamazawa and Hollien, 1992), where the gender difference in F0 is apparently culturally influenced. Since Portuguese joins in with the majority of languages, it can be concluded that the cultural influence of gender on F0 in Portuguese is the same as that in this majority of languages, and may therefore well be zero, so that the effect is just physiologically determined.

Second, high vowels have a higher F0 than low vowels, with a ratio of 1.157 for the Brazilians and a reliably smaller ratio of 1.094 for the Europeans. This effect has been reported widely before, with comparable ratios, e.g. for American English (House and Fairbanks, 1953: a ratio of 1.092) and Dutch (Koopmans-van Beinum, 1980: 1.098; Adank *et al.*, 2004: 1.222); for a long list of languages, see Whalen and Levitt (1995). Lehiste and Peterson (1961) call the effect *intrinsic fundamental frequency*. In Portuguese, both the dialect-dependence and the fact that F0 does not covary with F1 *within* vowels suggest that the intrinsic F0 is not an automatic consequence of F1. As intrinsic duration, intrinsic F0 may have turned into a vowel-specific cue. However, the smallness of the effect seems to contradict this, so it is hard to give a conclusive verdict on this problem.

Third, back vowels have a higher F0 than front vowels in Portuguese. This was also reliably found for English in a meta-analysis by Whalen and Levitt (1995). No causes for the effect seem to be known.

E. Speaker normalization

Figures 4 and 5 show large degrees of overlap between the clouds of vowels. If listeners, when trying to identify the correct vowel category, had to rely solely on the F1 and F2 values of any incoming vowel token, they would make many errors. Fortunately, listeners are capable of normalizing away some of the differences between speakers. For instance, they could normalize a speaker's log(F1) values by subtracting the average log(F1) of the speaker's vowel space (Nearey, 1978). This has been done in Fig. 9, for F1 and F2 separately. For this figure, each speaker's average log(F1) is subtracted from each of his or her 7 log(F1) values, and then the average log(F1) of the 10 speakers is added back in (this is the method used by Morrison and Nearey, 2006). The clouds of vowels are seen to be well-separated.

It has to be noted that the statistical tests of Secs. IV A and IV E cannot be performed with the speaker-normalized data of Fig. 9: the normalization method reduces the variation *within* each of the four groups while keeping the variation *between* the groups constant; hence, any statistical method that relies on comparing within-group variation with between-group variation (such as analyses of variance, *t*-tests, and Wilcoxon tests) would overestimate the statistical significance of the differences between the two dialects or between the two genders.

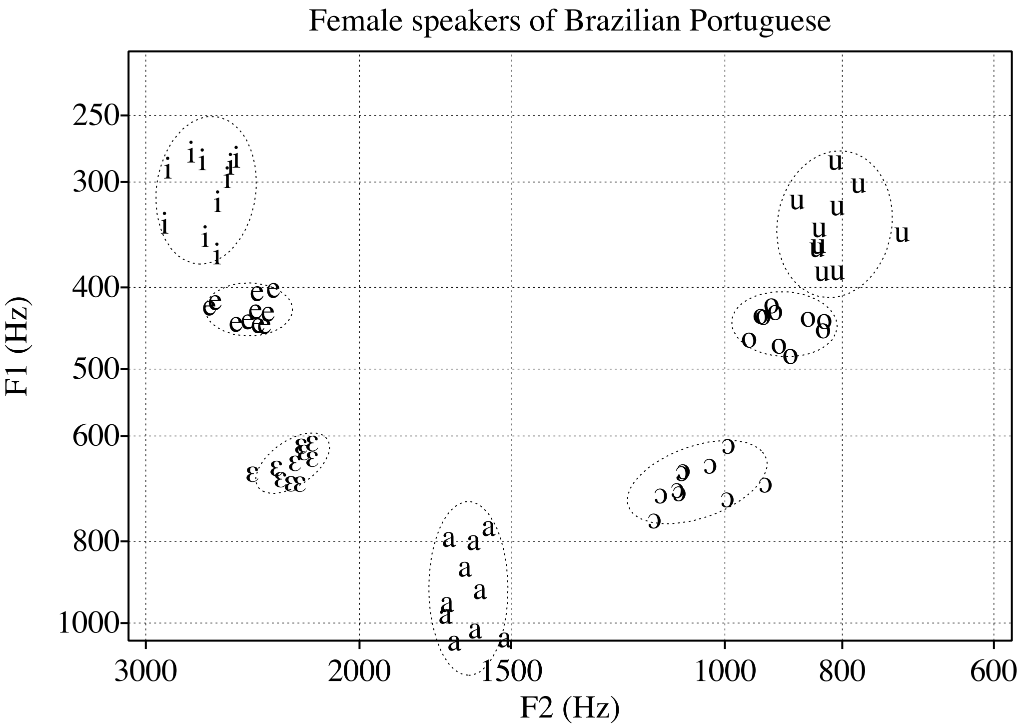

FIG. 9. Normalized vowel spaces for Brazilian women.

VIII. CONCLUSION

The present study finds several general properties of Portuguese vowels, which they have in common with vowels in many other languages: they can be kept apart by listeners (Sec. VII E), they exhibit intrinsic F0 (Sec. VI) and intrinsic duration (Sec. V), and the sizes of the F1 and F2 spaces are larger for women than for men (Secs. IV C, E). A less universal finding is that Portuguese speakers seem to have turned vowel duration into a cue for vowel identity, to an extent that goes beyond the automatic lengthening of open vowels; one can predict that Portuguese *listeners* use this cue to a greater extent than listeners of other languages.

The study also finds several differences between Brazilian Portuguese (BP) and European Portuguese (EP). The lower mid vowels /ɛ/ and /ɔ/ are lower in BP than in EP, and their distance to /e/ and /o/ is greater. Since there is no indication that their distance to /a/ is smaller in BP than in EP, it may be the case that the lower half of the vowel space is lower in BP than in EP. This difference has a number of implications for L2 and cross-language perception. First, the challenges that listeners of other seven-vowel languages meet with when learning Portuguese will not just depend on the L1 language (e.g. Italian or Catalan) but also on the target dialect (EP or BP). Second, listeners of a language with a larger vowel inventory, such as English or Dutch, will have different problems with the two Portuguese dialects; for instance, the Dutch vowel contrast /ɪ/-/ɛ/ may match better with the /e/-/ɛ/ contrast of BP than with that of EP. Finally, listeners of a language with smaller vowel inventories, such as Spanish or Greek, will have different problems with the two varieties as well; for instance, Spanish listeners may find it easier to learn the contrasts /e/-/ɛ/ and /o/-/ɔ/ in BP than in EP. Further research may be able to test these predictions.

ACKNOWLEDGMENTS

This research was supported by NWO (Netherlands Organization for Scientific Research) grant 016.024.018 to Boersma and by a CAPES (Committee for Postgraduate Courses in Higher Education, Brazilian Ministry of Education) grant to Rauber. The authors would like to acknowledge the contribution of Denize Nobre Oliveira on the testing of participants and manual vowel segmentation, and of Ton Wempe for technical support and preliminary analyses.

References

Adank, P., Van Hout, R., and Smits, R. (**2004**). "An acoustic description of the vowels of Northern and Southern standard Dutch," J. Acoust. Soc. Am. **116**, 1729–1738.

Anderson, N. (**1978**). "On the calculation of filter coefficients for maximum entropy spectral analysis," in *Modern Spectral Analysis* (IEEE Press, New York).

Barbosa, P. A., and Albano, E. C. (**2004**). "Brazilian Portuguese: Illustrations of the IPA," J. Int. Phon. Assoc. **34**, 227–232.

Bisol, L. (**1996**). *Introdução a estudos de fonologia do português brasileiro*. (EDIPUCRS, Porto Alegre).

Boersma, P., and Weenink, D. (**1992–2007**). Praat: doing phonetics by computer (Version 4.6.15) [Computer program]. Retrieved August 21, 2007, from http://www.praat.org/

Byrd, D. (**1992**). "Preliminary results on speaker-dependent variation in the TIMIT database," J. Acoust. Soc. Am. **92**, 593–596.

Callou, D., Moraes, J., and Leite, Y. (**1996**). "O vocalismo do português do Brasil," Letras de Hoje **31(2)**, 27–40.

Câmara Jr., J. M. (**1970**). *Estrutura da língua portuguesa* (Vozes, Petrópolis).

Cervera, T., Miralles, J. L., and González-Álvarez, J. (**2001**). "Acoustical analysis of Spanish vowels produced by laryngectomized subjects," J. Speech Lang. Hearing Res. **44**, 988–996.

Clopper, C. G., Pisoni, D. B., and De Jong, K. (**2005**). "Acoustic characteristics of the vowel systems of six regional varieties of American English," J. Acoust. Soc. Am. **118**, 1661–1676.

Cristófaro Silva, T. (**2002**). *Fonética e fonologia do português* (Contexto, São Paulo).

Delgado-Martins, M. R. (**1973**). "Análise acústica das vogais orais tônicas em português," Boletim de Filologia **22**, 303–314. [republished 2002 in Delgado-Martins, M. R., *Fonética do Português: Trinta anos de investigação* (pp. 41-52). Lisbon: Caminho]

Diehl, R. L., Lindblom, B., Hoemeke, K. A., and Fahey, R. P. (**1996**). On explaining certain male-female differences in the phonetic realization of vowel categories," J. Phonetics **24**, 187–208.

Ewan, W., and Krones, R. (**1974**). "Measuring Larynx Movement Using the Thyreumbrometer," J. Phonetics **2**, 327–335.

Faveri, C. B. (**1991**). "Análise da duração das vogais orais do português de Florianópolis-Santa Catarina," M.A. thesis, Universidade Federal de Santa Catarina.

Goldstein, U. (**1980**). "An articulatory model for the vocal tracts of growing children," Ph.D. thesis, M.I.T., Cambridge, MA.

Hagiwara, R. (**1997**). "Dialect variation and formant frequency: The American English vowels revisited," J. Acoust. Soc. Am. **102**, 655–658.

Hillenbrand, J., Getty, L. A., Clark, M. J., and Wheeler, K. (**1995**). "Acoustic characteristics of American English vowels," J. Acoust. Soc. Am. **97**, 3099–3111.

House, A. S., and Fairbanks, G. (**1953**). "The influence of consonant environment upon the secondary acoustical characteristics of vowels," J. Acoust. Soc. Am. **25**, 105–113.

Koopmans-van Beinum, F. J. (**1980**). "Vowel contrast reduction. An acoustic and perceptual study of Dutch vowels in various speech conditions," Ph.D. thesis, University of Amsterdam.

Labov, W. (**1994**). *Principles of Linguistic Change. Volume I: Internal Factors* (Blackwell, Oxford).

Lehiste, I. (**1970**). *Suprasegmentals* (MIT Press, Cambridge, Mass.)

Lehiste, I., and Peterson, G. E. (**1961**). "Some basic considerations in the analysis of intonation," J. Acoust. Soc. Am. **33**, 419–425.

Lima, R. (**1991**). "Análise acústica das vogais orais do português de Florianópolis-Santa Catarina," M.A. thesis, Universidade Federal de Santa Catarina.

Mateus, M. H. M. (**1990**). *Fonética, fonologia e morfologia do português* (Universidade Aberta, Lisbon).

Mateus, M. H. M., and d'Andrade, E. (**1998**). "The syllable structure in European Portuguese," Documentação de Estudos em Linguística Teórica e Aplicada (DELTA) **14**, 13–32.

Mateus, M. H. M., and d'Andrade, E. (**2000**). *The phonology of Portuguese* (Oxford University Press, Oxford).

Moraes, J. A., Callou, D., and Leite, Y. (**1996**). "O sistema vocálico do português do Brasil: caracterização acústica," in *Gramática do Português falado*, 5, edited by M. Kato (Editora da Unicamp, Campinas), pp. 33–53.

Morrison, G. S., and Escudero, P. (**2007**). "A cross-dialect comparison of Peninsular- and Peruvian-Spanish vowels," in Proceedings of the 16th Congress of Phonetic Sciences, Saarbrücken, pp. 1505–1508.

Morrison, G. S., and Nearey, T. M. (**2006**). "A cross-language vowel normalisation procedure," Canadian Acoustics **34**, 94–95.

Nearey, T. M. (**1978**). *Phonetic Feature Systems for Vowels* (Indiana University Linguistics Club, Bloomington).

Nearey, T. M., Assmann, P. F., and Hillenbrand, J. M. (**2002**). "Evaluation of a strategy for automatic formant tracking," J. Acoust. Soc. Am. **112**, 2323.

Pereira, A. L. D. (**2001**). "Caracterização acústica do sistema vocálico tônico oral florianopolitano: alguns indícios de mudança," M.A. thesis, Universidade Federal de Santa Catarina.

Peterson, G. E., and Barney, H. L. (**1952**). "Control methods used in a study of vowels," J. Acoust. Soc. Am. **24**, 175–184.

Peterson, G. E., and Lehiste, I. (**1960**). "Duration of syllable nuclei in English," J. Acoust. Soc. Am. **32**, 693–703.

Riordan, C. J. (**1977**). "Control of vocal-tract length in speech," J. Acoust. Soc. Am. **62**, 998–1002.

Rochet, A. P. and Rochet, B. L. (**1991**). "The effect of vowel height on patterns of assimilation nasality in French and English," In Proceedings of the 12th International Congress of Phonetic Sciences, Aix, Vol. **3**, 54–57.

Ryalls, J. H. and Lieberman, P. (**1982**). "Fundamental frequency and vowel perception," J. Acoust. Soc. Am. **72**, 1631–1634.

Seara, I. C. (**2000**). "Estudo acústico-perceptual da nasalidade das vogais do português brasileiro," Ph.D. thesis, Universidade Federal de Santa Catarina.

Simpson, A. P. (**2001**). "Dynamic consequences of differences in male and female vocal tract dimensions," J. Acoust. Soc. Am. **109**, 2153–2164.

Simpson, A. P. (**2002**). "Gender-specific articulatory-acoustic relations in vowel sequences," J. Phonetics **30**, 417–435.

Simpson, A. P., and Ericsdotter, C. (**2003**). "Sex-specific durational differences in English and Swedish," in Proceedings of the 15th Congress of Phonetic Sciences, Barcelona, pp. 1113–1116.

Solé, M. J. (**2007**). "Controlled and mechanical properties in speech: a review of the literature," in *Experimental Approaches to Phonology*, edited by M.J. Solé, P. Beddor and M. Ohala (Oxford University Press, Oxford), pp. 302–321.

Stevens, K. (**1998**). *Acoustic Phonetics* (MIT Press, Cambridge, MA).

Strange, W., Weber, A., Levy, E. S., Shafiro, V., Hisagi, M., and Nishi, K. (**2007**). "Acoustic variability within and across German, French, and American English vowels: Phonetic context effects," J. Acoust. Soc. Am. **122**, 1111–1129.

Whalen, D. H., and Levitt, A. G. (**1995**). "The universality of intrinsic F_0 of vowels," J. Phonetics **23**, 349–366.

Whiteside, S. P. (**1996**). "Temporal-based acoustic-phonetic patterns in read speech: Some evidence for speaker sex differences," J. Int. Phon. Assoc. **26**, 23–40.

Yamazawa, H., and Hollien, H. (**1992**). "Speaking fundamental frequency patterns of Japanese women," Phonetica **49**, 128–140.