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

Paola Escudero[a)] 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 A. H. Bion
Department of Psychology, Stanford University, Jordan Hall, Building 420, 450 Serra Mall, Stanford, California 94305

(Received 18 July 2008; revised 22 June 2009; accepted 24 June 2009)

This paper examines four acoustic correlates of vowel identity in Brazilian Portuguese (BP) and European Portuguese (EP): first formant (F1), second formant (F2), duration, and fundamental frequency (F0). Both varieties of Portuguese display some cross-linguistically common phenomena: vowel-intrinsic duration, vowel-intrinsic pitch, gender-dependent size of the vowel space, gender-dependent duration, and a skewed symmetry in F1 between front and back vowels. Also, the average difference between the vocal tract sizes associated with /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 in both varieties of Portuguese the vowel-intrinsic duration effect is larger than in many other languages. Differences between BP and EP are found in duration (BP has longer stressed vowels than EP), in F1 (the lower-mid front vowel approaches its higher-mid counterpart more closely in EP than in BP), and in the size of the intrinsic pitch effect (larger for BP than for EP). © *2009 Acoustical Society of America.* [DOI: 10.1121/1.3180321]

PACS number(s): 43.70.Fq, 43.70.Kv, 43.72.Ar [AL] Pages: 1379–1393

I. INTRODUCTION

The aim of this article is to investigate the acoustic characteristics of the seven oral vowels that Brazilian Portuguese (BP) and European Portuguese (EP) have in common in stressed position, namely, the vowels /i, e, ɛ, a, ɔ, o, u/, and thereby to find out what aspects of the Portuguese vowel inventory are universal, Portuguese-specific, or dialect-specific.

Studies that described Portuguese vowels in phonological or impressionistic articulatory terms (e.g., Câmara, 1970; Mateus, 1990; Bisol, 1996; Mateus and d'Andrade, 1998, 2000; Barroso, 1999; Moraes, 1999; Cristófaro Silva, 2002; Barbosa and Albano, 2004; Mateus *et al.*, 2005) agree that the Portuguese vowel inventory has an internal symmetry: apart from the central low vowel /a/, there are three unrounded front vowels (i, e, ɛ) and three rounded back vowels (u, o, ɔ) between which we can identify three pairings, namely, two high vowels (i-u), two higher-mid vowels (e-o) and two lower-mid vowels (ɛ-ɔ).[1] Because of the general relation between vowel height and the first formant (F1), we expect that the members of each pair have almost identical F1 values, and one research question is whether this is true for Portuguese. In fact, languages with large symmetric vowel inventories have been reported to have slightly higher F1 values for each back vowel as compared to its corresponding front vowel: American English (Peterson and Barney, 1952; Clopper *et al.*, 2005; Strange *et al.*, 2007), Parisian French (Strange *et al.*, 2007), Northern German (Strange *et al.*, 2007), Dutch (Koopmans-van Beinum, 1980),[2] and BP (Moraes *et al.*, 1996, p. 35; Seara, 2000, pp. 80, 91, 102, 112, and 141); one research question is whether this holds for both varieties of Portuguese.

Portuguese has been reported to have no phonological length distinctions in vowels (Falé, 1998, p. 257; Mateus *et al.*, 2005, p. 140). For such languages, it has been reported that low vowels tend to have a longer duration than high vowels (e.g., for French: Rochet and Rochet, 1991, p. 57, Fig. 7b). The effect can even be seen in languages that do have phonological length, such as English (House and Fairbanks, 1953, p. 111). 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 Lindblom (1967) and Lehiste (1970, pp. 18 and 19) in regarding it as a universal physiological property of speech production: open vowels require more jaw lowering, hence more time, than closed vowels. Since speakers can in principle control duration and F1 independently, it is, however, an open question whether Portuguese follows this cross-linguistic tendency or not. If Portuguese does follow the tendency, it is relevant to know the extent to which Portuguese does this; if this extent is larger than in other languages, it would be evidence for an exaggeration of the use of duration as a cue to vowel height.

[a)]Author to whom correspondence should be addressed. Electronic mail: paola.escudero@uva.nl

Portuguese has never been reported to have phonological tone. For such languages, it has been reported that low vowels tend to have a lower F0 than high vowels (for a long list of languages, see Whalen and Levitt, 1995). Lehiste and Peterson (1961) call the effect *intrinsic fundamental frequency*. Again, articulatory explanations have been proposed, mainly in terms of a pull of the tongue on the larynx (Ohala and Eukel, 1987), but speakers can also control F0 and F1 independently, so it is an open question whether Portuguese follows this universal tendency or not, and if so, whether it does so to a larger extent than other languages, i.e., whether it exaggerates F0 differences as a cue to vowel height.

Several Romance languages with a comparable symmetric seven-vowel inventory as Portuguese show signs that the lower-mid vowels are merging with the higher-mid vowels in some regional varieties: Italian (Maiden, 1997, p. 8), French (Landick, 1995), and Catalan (Recasens and Espinosa, 2009). One of our research questions is whether any signs of future merger can be observed in either of the two Portuguese varieties under scrutiny.

As for differences between female and male speakers, we expect Portuguese to exhibit the following near-universal effects. First, females have generally higher F0 and formants than males. Second, women tend to have a larger vowel space than men, even along logarithmic scales, i.e., in terms of a ratio of the F1 values of /a/ versus /i, u/; the cause of this 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 easier to discriminate by listeners than female F1 values (Goldstein, 1980; Ryalls and Lieberman, 1982; Diehl *et al.*, 1996). Third, women have longer vowel durations than men (Simpson and Ericsdotter, 2003); the source of this effect has been sought in the physiology (Simpson, 2001, 2002) as well as in the idea that women put more effort in trying to speak clearly (Byrd, 1992; Whiteside, 1996). As for differences between BP and EP, Moraes *et al.* (1996) report, comparing their BP results with the EP results of Delgado-Martins (1973), that /i/ and /u/ have a higher F1 in BP than in EP; the question is whether this result will still hold when comparing BP and EP with identical measurement methods.

Answering these research questions on the basis of earlier acoustic descriptions of Portuguese vowels (Delgado-Martins, 1973, 2002, pp. 41–52; Callou *et al.*, 1996; Moraes *et al.*, 1996; Seara, 2000) is difficult, because none of these studies provided direct cross-dialectal comparisons, investigated a sufficient number of speakers, included female speakers, or reported all four acoustic characteristics of all vowels; also, the results of multiple studies can hardly be combined, as a result of differences in measurement methods. The methodology employed in the present study is designed to answer the research questions with more confidence: (1) it compares 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) 40 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; and (4) acoustic analyses were made of vowel duration, fundamental frequency, and the first two vowel formants. This methodology allows us to address all of the research questions mentioned above, as well as to explore any unpredicted differences between females and males or between BP and EP.

Finally, the present paper aims at providing reliable values for duration by measuring vowels only between voiceless consonants, and at providing typical formant values by measuring vowels only between stops and fricatives. Elicitation of multiple tokens per speaker allows us to automatically define the formant ceiling of the LPC analysis on the basis of within-speaker and within-vowel variation, thus allowing more reliable automatic formant measurements. This methodology is explained in detail so that it can be used as a reference for future studies on vowel formant analyses.

II. METHOD

A. Participants

In order to obtain relatively homogeneous and comparable groups of BP and EP 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 ("I don't understand a word") to 7 ("I understand like a native speaker"), 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" (more precisely: "age-, social-economic-status-, and region-dependent variety of the standard language") 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), and the males' mean age was 18.7 years (s.d. 0.8).

B. Data collection procedure

All 40 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 quantization. The 20 BP recordings were made at the Escola Superior de Propaganda e Marketing (ESPM) in São Paulo, and the 20 EP recordings were made 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 /pepo/ and

/saso/ (*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 in two phrasal positions, namely, 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/, and 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 six tokens of each vowel embedded in each consonant context. From these six tokens, we chose the two isolated words (i.e. one with final *e*, and one with final *o*) and the two best exemplars of the tokens embedded in the carrier sentence (one with final *e*, and one with final *o*). Two native speakers of Portuguese chose these best exemplars on the basis of their recording quality, i.e., the tokens with no background noise or hesitation during the production of the whole sentence. The final isolated vowels were not considered in the analysis. Thus, 20 productions (2 phrasal positions × 2 word-final vowels × 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 × 7 vowels × 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, 2008) was used to measure the F0 curves of all 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 center 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 six of the 2800 tokens, which were creaky), 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, which was noisy), 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 values measured in steps of 1 ms in the central 40% of the vowel: ignoring the first and last 30% 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.

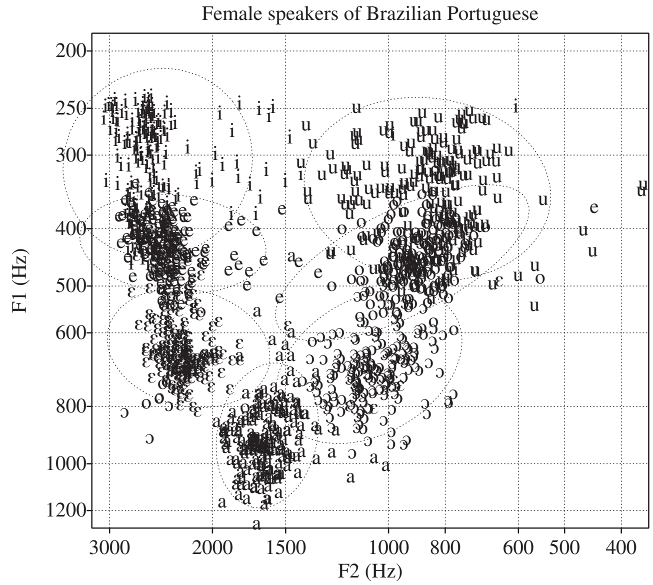

FIG. 1. The first and second formants of the 1400 vowel tokens of the Brazilian women, measured with a fixed (gender-specific) formant ceiling of 5500 Hz. The ellipses show two estimated standard deviations and have been designed to cover 86.5% of the data points (for normally distributed data).

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% of the vowel.[3] 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). These gender-specific *formant ceilings* of 5000 and 5500 Hz reflect the different average vocal tract lengths of men versus women (since looking for five formants entails that the ceiling is meant to lie between F5 and F6, one can estimate the vocal tract length as $5c/(2\cdot\text{ceiling})$, where c is the speed of sound). The 1400 F1-F2 pairs thus measured for the Brazilian women are plotted in Fig. 1.

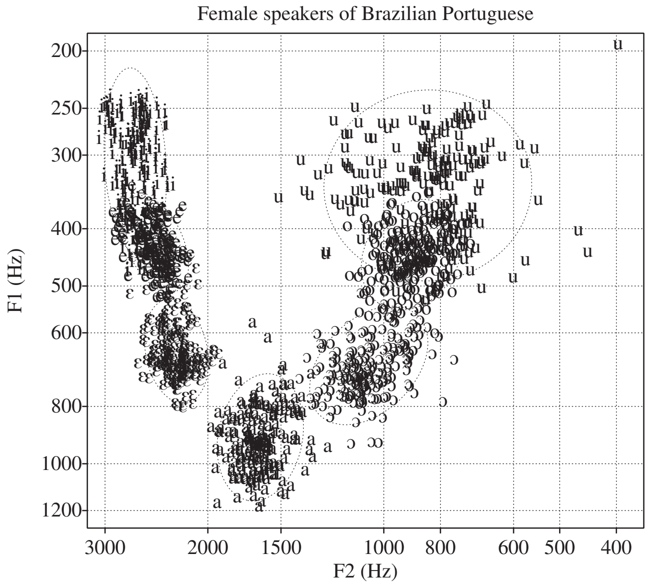

FIG. 2. The first and second formants of the 1400 vowel tokens of the Brazilian women, measured with optimized (speaker- and vowel-specific) formant ceilings.

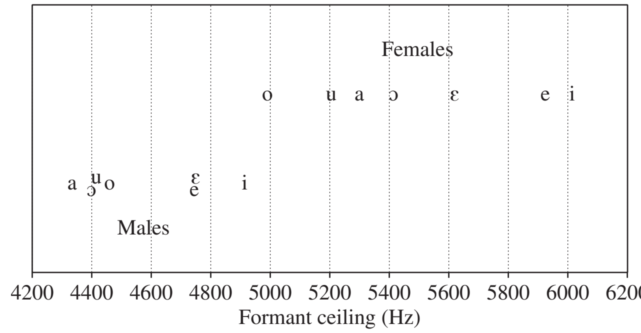

FIG. 3. Median optimal ceilings for each gender-vowel combination.

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

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 *et al.*, 2002: smallest bandwidths, continuity in time, correlation between original and LPC-generated spectrogram; also described by Adank, 2003, and used by Adank *et al.*, 2004), 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 20 measured F1-F2 pairs. This variation is computed along the same logarithmic scales as seen in Fig. 1, namely, as the variance of the 20 log(F1) values plus the variance of the 20 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.

Figure 2 shows that the variation between the vowel tokens has decreased appreciably: almost all outliers have gone, and although only the variation in 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 category on the optimal formant ceiling. Each vowel symbol in that figure represents the median of 20 optimal ceilings (because there are 20 speakers of each gender and the two dialects are pooled).

Figure 3 shows that both gender and vowel category 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 category 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 upper and lower lip (rounded and protruded for /u/, and spread and retracted for /i/) and probably a difference in the height of the larynx (lowered for /u/: Ewan and Krones, 1974; Riordan, 1977). Generally, the three spread vowels /i/, /e/, and /ɛ/ come with shorter vocal tracts than the three rounded vowels /u/, /o/, and /ɔ/, and this must be reflected in the values of the higher formants (Kent and Read, 2002, p. 32); as the formant ceiling lies between F5 and F6, the formant ceiling will on average be higher for the spread than for the rounded vowels. Since a correct formant ceiling influences the reliability of the measurements of *all* formants, including F1 and F2, this result 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–VI present the detailed results of the acoustic measurements and statistical analyses aimed at answering

TABLE I. Geometric averages of vowel duration, F0, F1, F2, F3, and formant ceilings for female (F) and male (M) speakers of BP and EP. Between parentheses: the standard deviations, converted back to ratios of ms and Hz. Every cell represents ten speakers.

			/i/	/e/	/ɛ/	/a/	/ɔ/	/o/	/u/
BP	Duration (ms)	F	99 (1.210)	122 (1.195)	141 (1.192)	144 (1.173)	139 (1.145)	123 (1.151)	100 (1.201)
		M	95 (1.216)	109 (1.200)	123 (1.232)	127 (1.186)	123 (1.209)	110 (1.189)	100 (1.205)
	F0 (Hz)	F	242 (1.096)	219 (1.098)	210 (1.092)	209 (1.088)	211 (1.093)	225 (1.098)	252 (1.087)
		M	137 (1.199)	131 (1.186)	124 (1.183)	122 (1.199)	122 (1.178)	132 (1.194)	140 (1.223)
	F1 (Hz)	F	307 (1.198)	425 (1.082)	646 (1.076)	910 (1.078)	681 (1.087)	442 (1.094)	337 (1.192)
		M	285 (1.077)	357 (1.077)	518 (1.089)	683 (1.095)	532 (1.160)	372 (1.100)	310 (1.070)
	F2 (Hz)	F	2676 (1.056)	2468 (1.061)	2271 (1.051)	1627 (1.062)	1054 (1.099)	893 (1.054)	812 (1.054)
		M	2198 (1.078)	2028 (1.076)	1831 (1.072)	1329 (1.088)	927 (1.108)	804 (1.092)	761 (1.100)
	F3 (Hz)	F	3296 (1.073)	3074 (1.048)	2897 (1.077)	2625 (1.119)	2653 (1.114)	2627 (1.158)	2691 (1.123)
		M	2952 (1.066)	2719 (1.077)	2572 (1.050)	2324 (1.084)	2335 (1.069)	2380 (1.060)	2309 (1.078)
	Ceiling (Hz)	F	6001 (1.086)	5933 (1.094)	5463 (1.166)	5577 (1.076)	5260 (1.137)	4938 (1.113)	5090 (1.095)
		M	5230 (1.155)	5063 (1.181)	5010 (1.137)	4463 (1.105)	4436 (1.077)	4522 (1.068)	4458 (1.064)
EP	Duration (ms)	F	92 (1.154)	106 (1.151)	115 (1.137)	122 (1.144)	118 (1.141)	110 (1.158)	94 (1.208)
		M	84 (1.142)	97 (1.147)	106 (1.162)	108 (1.183)	104 (1.149)	99 (1.144)	83 (1.151)
	F0 (Hz)	F	216 (1.084)	211 (1.082)	204 (1.075)	201 (1.086)	204 (1.076)	211 (1.084)	222 (1.092)
		M	126 (1.177)	122 (1.165)	117 (1.156)	115 (1.151)	117 (1.151)	123 (1.171)	127 (1.187)
	F1 (Hz)	F	313 (1.243)	402 (1.125)	511 (1.154)	781 (1.186)	592 (1.270)	422 (1.150)	335 (1.230)
		M	284 (1.085)	355 (1.090)	455 (1.131)	661 (1.075)	491 (1.111)	363 (1.107)	303 (1.085)
	F2 (Hz)	F	2760 (1.033)	2508 (1.040)	2360 (1.031)	1662 (1.078)	1118 (1.091)	921 (1.184)	862 (1.144)
		M	2161 (1.048)	1987 (1.058)	1836 (1.068)	1365 (1.060)	934 (1.078)	843 (1.090)	814 (1.127)
	F3 (Hz)	F	3283 (1.054)	3007 (1.043)	2943 (1.042)	2535 (1.170)	2729 (1.086)	2636 (1.188)	2458 (1.204)
		M	2774 (1.057)	2559 (1.057)	2475 (1.049)	2333 (1.041)	2414 (1.077)	2429 (1.072)	2315 (1.041)
	Ceiling (Hz)	F	5875 (1.090)	5734 (1.087)	5662 (1.096)	5278 (1.085)	5259 (1.132)	5165 (1.123)	5066 (1.119)
		M	4570 (1.153)	4733 (1.148)	4792 (1.098)	4523 (1.120)	4537 (1.137)	4512 (1.108)	4366 (1.065)

the specific research questions mentioned in the Introduction and finding differences between the two dialects and between the two genders. These sections report the effects of vowel category, gender and dialect on formants, duration, and fundamental frequency. Table I summarizes the average values for all these quantities (also shown in Figs. 6–8); each number in the table is a geometric average over ten speaker values, each of which is a median over 20 tokens (2 phrasal positions × 2 word-final vowels × 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).

Since duration, F0, and formants are by definition positive quantities, they are expected to be normally distributed along logarithmic scales, and all statistical investigations in this and the following sections are therefore performed on log-transformed values; this decision is also inspired by the fact that duration is perceived and represented logarithmically (Gibbon, 1977; Allan and Gibbon, 1991), that F0 ranges are comparable for men and women only along a logarithmic scale (Henton, 1989; Tielen, 1992), and that the influence of a specific articulation on the height of formants (in hertz) must be expressed as a *ratio* (rather than as a difference) that is independent of the vocal tract size (if the vocal tract shape is constant). For readability, all averages of logarithmic values are transformed back to milliseconds or hertz, so that the reported averages are in effect geometric averages over the original values in milliseconds or hertz, as in Table I. Also, observed differences between groups in the log domain are reported as ratios between groups, and an observed reliable difference between groups in the log domain is reported as a (duration, F0, F1, or F2) ratio between groups that is reliably different from 1. Another consequence is that all figures use logarithmic axes. In Table I, the standard deviations in the log domain are expressed as ratios in the milliseconds or hertz domains; for example, if a certain average is 400 Hz and the corresponding standard deviation is 1.100, then one standard deviation up from the average is 440 Hz, two standard deviations up is 484 Hz, and one standard deviation down is 363.636 Hz.

Table I does not express what kind of variation the seven standard deviations in a row are due to; do the standard deviations of F0, for instance, reflect the fact that every speaker comes with a different small pitch range, or do they reflect the fact that every speaker randomly determines which vowel has what F0? To thus separate main speaker effects from speaker-vowel interaction effects, and to evaluate the differences between the dialects and between the genders, each of the statistical investigations into duration, F0, F1, and F2 (Secs. IV B, IV F, V, and VI) starts out with an exploratory repeated-measures analysis of variance (conducted with SPSS) on 280 logarithmic values (40 speakers × 7 vowels), which are the median values of the 20 tokens of each of the 7 vowels produced by the 40 speakers. In every repeated-measures analysis, dialect and gender act as between-subjects factors and vowel category acts 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, we decided to use Huynh–Feldt's correction, which multiplies the number of 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. After each exploratory analysis we perform tests that directly address a specific research question raised in the Introduction, by investigating the behavior of a within-speaker measure specifically designed for the purpose.

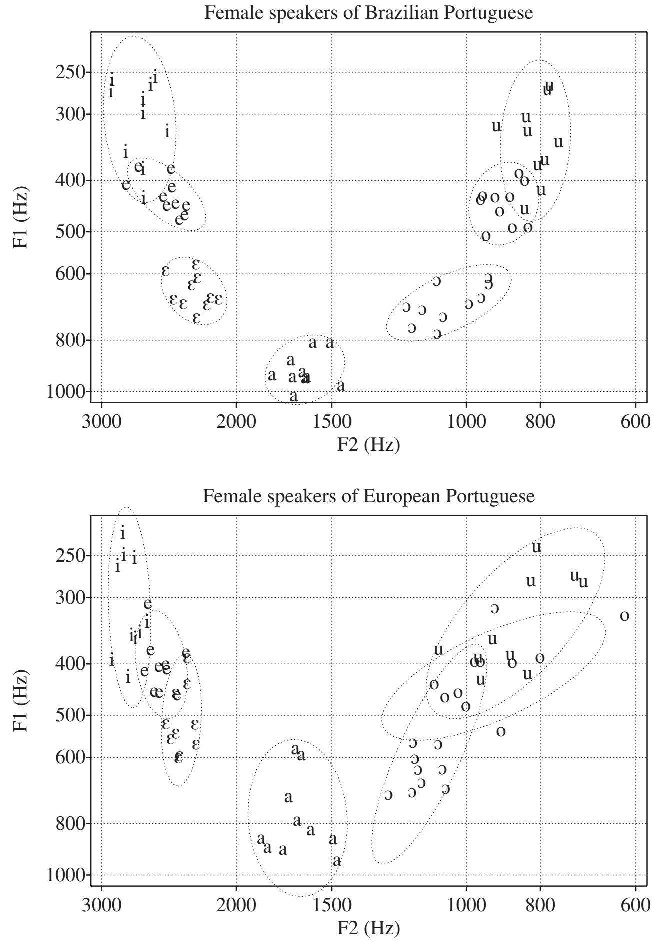

FIG. 4. First and second formants of ten BP and ten EP women.

IV. RESULTS FOR FORMANTS

A. The speakers' median formants

Figures 4 and 5 show the median F1 and F2 values for the ten female and ten male speakers of each dialect. In each of the four figures, each vowel occurs ten times because there were ten 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 were all measured with the same formant ceiling, namely, the formant ceiling that minimizes the variation among the 20 F1 and F2 values (Sec. II E).

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

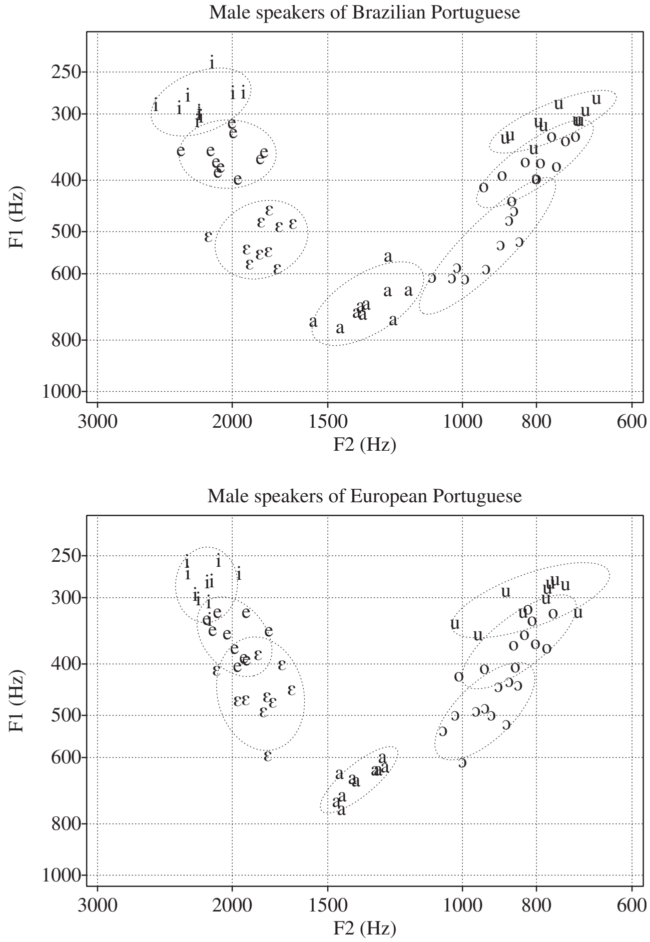

FIG. 5. First and second formants of ten BP and ten EP men.

B. Exploratory analysis of F1

The exploratory repeated-measures analysis of variance reveals a large main effect of vowel category on F1 (η_p^2 $=0.950$; $F[6\varepsilon, 216\varepsilon, \varepsilon=0.609]=684.926$; $p=9\times10^{-85}$). As expected from the Introduction, and clearly visible in Fig. 6, the main determiner of F1 is the phonological vowel height: coarsely speaking, 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 subtler effect (of vowel place) is investigated in Sec. IV C.

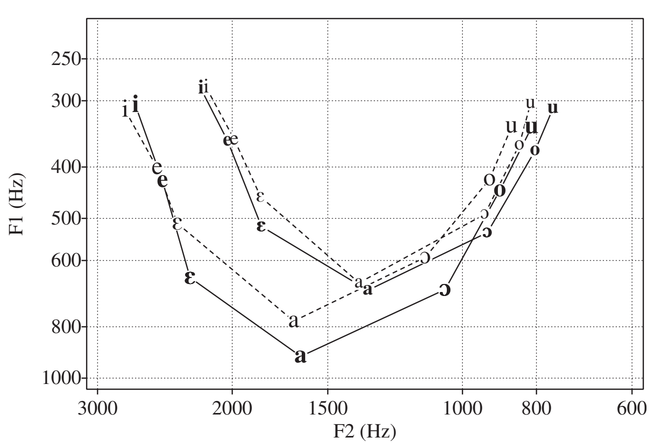

FIG. 6. The vowel spaces of the four groups. Solid lines and bold symbols=BP; dashed lines=EP. Large font: women; small font: men.

As expected, the analysis also reveals a large main effect of gender on F1 ($\eta_p^2=0.394$; $F[1,36]=23.430$; $p=2.4\times10^{-5}$): Portuguese-speaking women tend to have higher F1 values (geometric average: 478 Hz; the 95% confidence interval runs from 456 to 501 Hz) than Portuguese-speaking men (409 Hz; c.i.=390–429 Hz). The gender effect on F1 is therefore a ratio of 1.170 (c.i.=1.095–1.249), 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 gender effect on F1 may have to be viewed in relation to interaction effects. Since the interaction of gender and dialect is not reliably different from zero ($F[1,36]=0.492$; $p=0.488$), and neither is the triple interaction of gender, dialect, and vowel ($F[6\varepsilon, 216\varepsilon, \varepsilon=0.609]=1.219$; $p=0.306$), it remains to consider the interaction of gender and vowel, which is indeed reliable ($\eta_p^2=0.113$; $F[6\varepsilon, 216\varepsilon, \varepsilon=0.609]=4.604$; $p=0.0023$). Figure 6 suggests that this is because women take up a greater part of the F1 continuum than men. This is investigated in detail in Sec. IV D.

Finally, the analysis reveals a nearly significant 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 ($\eta_p^2=0.158$; $F[6\varepsilon, 216\varepsilon, \varepsilon=0.609]=6.777$; $p=9.5\times10^{-5}$). Apparently, some vowels have different heights in (São Paulo) BP than in (Lisbon) EP. This is investigated in detail in Sec. IV E.

C. The effect of vowel place on F1

One of the research questions in the Introduction is whether Portuguese follows the cross-linguistic trend that (rounded) back vowels tend to have higher F1 values than the corresponding (unrounded) front vowels. Figure 6 does show that for all four groups of speakers each back vowel has a higher average F1 than its front counterpart, but the figure does not show that this can be generalized to the Portuguese-speaking population. The exploratory analysis of Sec. IV A does yield an answer by reporting within-subjects comparisons. That is, a speaker's F1 of /u/ is higher than that of his or her /i/ by a factor of 1.082, the F1 of /o/ is higher than that of /e/ by a factor of 1.039, and the F1 of /ɔ/ is higher than that of /ɛ/ by a factor of 1.078. All three factors are reliably greater than 1 (uncorrected two-tailed $p=9.1\times10^{-12}$, 5.6×10^{-5}, and 7.1×10^{-5}, respectively): their 98.30% confidence intervals (i.e., Šidák-corrected for three planned comparisons) are 1.060–1.103, 1.017–1.061, and 1.034–1.125, respectively. The conclusion is that in the Portuguese-speaking population, each back vowel has a higher mean F1 than its corresponding front vowel. A multivariate analysis of variance with dialect and gender as factors and the three front-back differences as dependents reveals no influence of dialect, gender, or dialect × gender on the front-back differences.

Simple sign counting reveals that this correlation between F1 and backness holds for a majority of individual speakers: for 38 of the 40 speakers, the F1 of /u/ is higher than the F1 of the same speaker's /i/. Likewise, the /o/-/e/ difference is positive for 32 of the 40 speakers, and the /ɔ/-/ɛ/ difference for 35 of the 40 speakers (the 15 exceptions happen to be maximally evenly distributed over the four groups, and maximally randomly distributed over the speakers). By not labeling the vowel symbols for speaker, Figs. 4 and 5 obscure this consistent effect (for instance, the four EP 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/). Sign counting therefore confirms again that there is a consistent correlation between F1 and phonological backness.

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

One of the research questions in the Introduction is whether Portuguese-speaking females have larger vowel spaces (along logarithmic axes) than Portuguese-speaking males. To answer this, we define a speaker's *F1 space size* as the ratio of the F1 of his or her low vowel /a/ and the (geometric) average F1 of his or her high vowels /i/ and /u/. We thus compute 40 F1 space sizes and subject these to a two-way analysis of variance with dialect and gender as factors. Since an interaction between gender and dialect was not found ($F[1,36]=2.395$, $p=0.130$), we report here only the two main effects.

The average F1 space size of the 20 women turns out to be 2.613, and that of the 20 men only 2.276. The female F1 space is therefore $2.613/2.276=1.148$ times (0.199 octaves) larger than the male F1 space (c.i. $=1.046-1.260$; the ratio is reliably different from 1 with $F[1,36]=9.052$, $p=0.0048$). 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. 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, and that of the Europeans 2.331. For the combined population of men and women, the Brazilian F1 space is therefore 1.095 times larger than the European F1 space (c.i. $=0.998-1.201$). This is not very reliably different from 1 ($F[1,36]=3.895$, $p=0.056$).

E. Vowel height differences between the two dialects

One of the research questions in the Introduction is which vowels are different in the two dialects. We first investigate this by a multivariate analysis of variance on the seven F1 values, with dialect and gender as factors. Since the dialect-gender interaction is not significant (Wilks' $\Lambda[7,30]=0.837$, $p=0.566$), we focus on the main effect of dialect. The vowel /ɛ/ turns out to be very reliably lower (higher F1) in BP than in EP ($F[1,36]=27.468$, $p=7.1\times 10^{-6}$). A difference in the same direction is found for its back counterpart /ɔ/ ($F[1,36]=4.973$, $p=0.032$) and for the vowel /a/ ($F[1,36]=7.162$, $p=0.011$), although these differences are not very reliable (regarding the multiple comparisons). The hypothesis by Moraes *et al.* (1996) mentioned in the Introduction is not confirmed: for the 40 speakers, /u/ has indeed a higher F1 in BP than in EP (ratio 1.013), but /i/ has a lower F1 in BP than in EP (ratio 0.992); neither of these ratios generalize reliably to the populations (they are different from 1 with $p=0.779$ and 0.866); in fact, the upper bounds of the confidence intervals (0.923–1.112 and 0.900–1.093) show that the extent of any lowering of the high vowels cannot be greater than 11.2%.

From the mere fact that we found that /ɛ/ is lower in BP than in EP whereas we found no difference for /e/, we cannot yet conclude that in BP /ɛ/ is lowered more than /e/ (from differences in p values no inferences can be made about the relative sizes of an effect), and we cannot therefore answer yet our research question about the difference between the /ɛ/-/e/ distances in BP and EP. Both of these problems are addressed in the remainder of this section.

In order to establish any dialectal difference in /ɛ/-/e/ distance, one can take advantage of the fact that all seven vowels have been spoken by the same 40 speakers, i.e., we have information about the internal structure of each speaker's vowel space. Thus, the log(F1) differences between every speaker's /ɛ/ and /e/ were computed, as well as those between every speaker's /ɔ/ and /o/. A multivariate analysis of variance with dialect and gender as factors was performed on the two sets of 40 values. The only significant effect is that of dialect ($\Lambda[2,35]=0.451$, $p=8.8\times 10^{-7}$), and it turns out that the F1 ratio of /ɛ/ and /e/ is very reliably greater in BP (observed average 1.485; uncorrected 95% c.i. $=1.437-1.535$) than in EP (1.276; c.i. $=1.235-1.319$): the ratio of these ratios is $1.485/1.276=1.164$ (c.i. $=1.111-1.219$), which is reliably different from 1 ($F[1,36]=43.391$, $p=1.1\times 10^{-7}$). Likewise, the F1 ratio of /ɔ/ and /o/ is greater for the 20 Brazilians (1.482; c.i. $=1.409-1.559$) than for the 20 Europeans (1.377; c.i. $=1.309-1.449$); the ratio of these ratios is 1.076 (c.i. $=1.002-1.156$), which is reliably different from 1 at the $\alpha=0.05$ level ($F[1,36]=4.326$, $p=0.045$). We conclude that the acoustic distance between lower-mid and higher-mid vowels is indeed larger in BP than in EP.

We subsequently address the other question, namely, what is behind these observed differences in the acoustic mid-vowel distances: are these differences due to /ɛ/ and /ɔ/ being lower in BP than in EP or due to /e/ and /o/ being higher in BP than in EP? Table I and Fig. 6 indicate that the latter possibility is unlikely: for both women and men, the mean BP /e/ and /o/ are *lower* than the mean EP /e/ and /o/. The next hypothesis to consider is 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 D). In that case, the lowness of /ɛ/ and /ɔ/ should disappear if the F1 values are normalized for the F1 space size. To assess whether this is the case, we compute the *relative heights* of the four mid vowels for each speaker. For instance, the relative height of /ɛ/ within the front vowel space can be defined as $(\log \mathrm{F1(a)} - \log \mathrm{F1(\varepsilon)})/(\log \mathrm{F1(a)} - \log \mathrm{F1(i)})$, and the relative height of /o/ within the back vowel space can be defined as $(\log \mathrm{F1(a)} - \log \mathrm{F1(o)})/(\log \mathrm{F1(a)} - \log \mathrm{F1(u)})$.

A multivariate (four vowels) two-way (dialect, gender) analysis of variance reveals no effect of gender on relative height ($\Lambda[4,33]=0.883$, $p=0.376$) and no interaction of dialect and gender ($\Lambda[4,33]=0.961$, $p=0.855$). We therefore only report on the main effect of dialect ($\Lambda[4,33]=0.423$, $p=1.0\times10^{-5}$). 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 0.315 (c.i.=0.275–0.355), but the average EP /ɛ/ has a relative height of 0.455 (c.i.=0.415–0.496), i.e., it lies close to the center of the F1 dimension; the difference between the dialects is highly reliable ($F[1,36]=25.022$; $p=3.0\times10^{-5}$). For /ɔ/, the difference between BP and EP is in the same direction (0.303 versus 0.353), but is not significant ($F[1,36]=1.250$; $p=0.271$). The higher-mid vowels seem to have very similar relative heights in the two dialects: /e/ has 0.730 for BP and 0.737 for EP, and /o/ has 0.752 for BP and 0.748 for EP. We conclude that the lower BP /ɛ/ remains even after normalizing for BP's larger F1 space.

The results of the previous paragraph suggest that the cause of the smaller /ɛ/-/e/ distance in EP could lie in a lower F1 for /ɛ/, but to be absolutely statistically certain (again, different degrees of statistical significance do not entail different effect sizes) one has to investigate whether the dialectal difference in the relative height of /ɛ/ is greater than that of /e/. This can be determined by subjecting the 40 *average mid vowel heights*, namely, $(\log F1(a)-(\log F1(\varepsilon)+\log F1(e))/2)/(\log F1(a)-\log F1(i))$, to a two-way analysis of variance. The effect of dialect on this measure is indeed significant ($F[1,36]=6.450$; $p=0.016$). We conclude that the smaller /ɛ/-/e/ distance in EP as compared to BP is due more to a raised /ɛ/ than to a lowered /e/ (within a normalized F1 space). For a discussion of the implications see Sec. VII A.

F. Effects on F2

As expected, the repeated-measures analysis of the variance of F2 reveals a large main effect of gender ($F[1,36]=120.857$; $p=4.7\times10^{-13}$): 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. The EP speakers turn out to have higher F2 values than the BP speakers, but this difference cannot be reliably generalized to their populations ($F[1,36]=3.009$; $p=0.091$). An interaction of dialect and gender is not found ($F[1,36]<1$).

As for the within-subject effects, the analysis reveals the expected main effect of vowel category on F2 ($F[6\varepsilon,216\varepsilon,\varepsilon=0.423]=1826.704$; $p=1.6\times10^{-78}$), as well as a reliable interaction between vowel and gender ($F[6\varepsilon,216\varepsilon,\varepsilon=0.423]=9.339$; $p=5.5\times10^{-5}$). From Fig. 6, the cause of the latter appears to be that the size of the F2 space (the /u/-/i/ distance) is larger for females than for males; this is investigated in detail below. The analysis reveals no interaction between vowel and dialect ($F<1$) and no triple interaction between vowel, dialect, and gender ($F<1$).

A multivariate analysis of variance on the F2 values of the seven vowels reveals neither a main effect of dialect[4] nor an effect of the interaction of dialect and gender; the main effect of gender ($\Lambda[7,30]=0.143$, $p=5.0\times10^{-11}$) is that /a,ɛ,e,i,ɔ/ have a very reliably higher F2 for women than for men ($F[1,36]\geqslant28.953$, $p\leqslant4.7\times10^{-6}$); for /u/ ($F[1,36]=3.329$; $p=0.076$) and /o/ ($F[1,36]=8.125$; $p=0.0072$), the observed average effect is in the same direction but in itself less reliably generalizable to the population (given the multiplicity of the tests). The hypothesis that all vowels simultaneously have a higher F2 for women than for men is nevertheless confirmed at the $\alpha=0.10$ level (in the case of such an inclusive hypothesis, the multiplicity of tests also raises the chance of a type II error, so that one is allowed to use a higher α than usual: Winer, 1962, p. 13).

Analogously to the F1 space size of Sec. IV C, we define a speaker's *F2 space size* as the ratio of the F2 of his or her /i/ and the F2 of his or her /u/. When we subject the 40 sizes to a two-way analysis of variance, we find no effect of dialect ($F[1,36]=2.076$, $p=0.158$) or of dialect×gender ($F<1$), and the main effect of gender ($F[1,36]=16.504$, $p=2.5\times10^{-4}$) is that for the 20 men, the average ratio is 2.768 (c.i.=2.616–2.929), and for the 20 women it is 3.249 (c.i.=3.070–3.437); the ratio of these ratios is 1.174 (c.i.=1.083–1.271). We conclude that the size of the F2 space is greater for Portuguese-speaking women than for men, i.e., that the gender difference in F2 is larger for /i/ than for /u/.

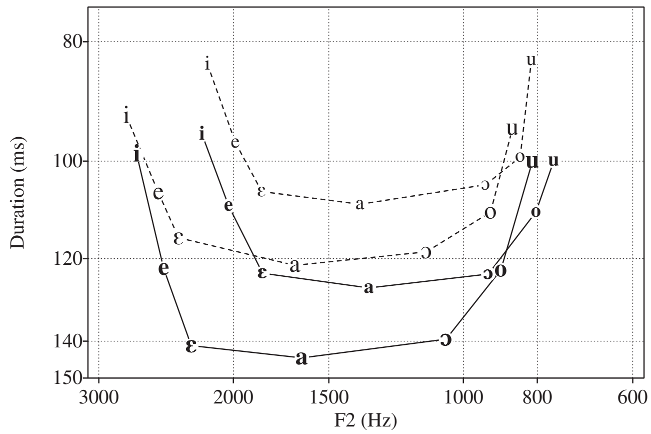

FIG. 7. Mean duration as a function of vowel category. The purpose of the inclusion of the F2 axis and the reversal of the vertical axis is to provide vowel space shapes that are similar in orientation and extent as the more usual ones in Fig. 6. Solid lines and bold symbols=BP; dashed lines=EP. Large font: women; small font: men.

V. RESULTS FOR DURATION

The fact, mentioned in the Introduction, that the Portuguese vowel system does not use vowel length as a phonological feature does not preclude 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 ten speakers.

A. Exploratory analyses

The repeated-measures analysis of the variance of duration reveals that the main effect of vowel category is very reliable ($F[6\varepsilon,216\varepsilon,\varepsilon=0.811]=243.358$, $p=5\times10^{-76}$); this issue is investigated in detail in Sec. V B. The duration of the vowels is influenced by dialect ($\eta_p^2=0.180$; $F[1,36]=7.915$, $p=0.008$): vowels are longer in BP than in EP by a factor of 1.148 (c.i.=1.039–1.269); this is investigated further in Sec. V C. The expected main effect of gender (see Introduction) is barely significant ($\eta_p^2=0.103$; $F[1,36]=4.125$, $p=0.050$): women's vowels are longer than men's vowels by a ratio of 1.105 (c.i.=1.0001–1.221); this is discussed in Sec. V C as well. The analysis does not reveal an interaction between gender and dialect ($F<1$), 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 are reliable, at least under the somewhat forgiving Huynh–Feldt correction ($F[6\varepsilon,216\varepsilon,\varepsilon=0.811]$ =2.426, 3.829, 3.671; $p=0.039$, 0.0028, 0.0038); Fig. 7 suggests, for instance, that specifically /u/ is shortened specifically by EP men.

A multivariate analysis of variance on all vowel durations shows that at the $\alpha=0.10$ level, all seven vowels are longer in BP than in EP (/a,ɛ,ɔ/: $F[1,36]\geqslant10.770$, $p\leqslant0.0023$; /e/: $F=6.480$, $p=0.015$; /u/: $F=5.020$, $p=0.031$; /o/: $F=4.981$, $p=0.032$; /i/: $F=3.648$, $p=0.064$).

B. Vowel-intrinsic duration

From the Introduction, one can expect an effect of vowel height on duration, and Fig. 7 confirms this expectation. In fact, for 39 of the 40 speakers, the median of his or her 20 measured /i/ tokens is shorter than the median of his or her 20 measured /e/ tokens. Within the analysis of Sec. V A, pairwise comparisons between the seven vowels yield the following results for vowels of adjacent phonological heights: /i,u/ are shorter than /e,o/ (all four uncorrected two-tailed $p<3\times10^{-13}$), /e,o/ shorter than /ɛ,ɔ/ (all four $p<2\times10^{-10}$), /ɛ/ shorter than /a/ ($p=0.0072$), and /o/ shorter than /a/ ($p=0.000\,34$). We conclude with confidence that lower vowels are longer than higher vowels in Portuguese.

Given the structure of the phonological vowel space, a second potential effect may be worth investigating, namely, whether duration depends on the front-back distinction. The result of the three relevant pairwise comparisons is that /i/ is shorter than /u/ ($p=0.036$) and /e/ is shorter than /o/ ($p=0.029$); the difference between /ɛ/ and /ɔ/ is not significant ($p=0.940$). This subject is not pursued further here (a possible explanation is given in Sec. VII C), and the focus below is solely on the traditional vowel-intrinsic duration effect, which is the relation between duration and height.

To investigate the size (rather than just the existence) of the vowel-intrinsic duration effect (for cross-linguistic comparison), we define for each speaker the *vowel-intrinsic duration ratio* as the ratio between the duration of his or her /a/ and the average duration of his or her /i/ and /u/. We subject the 40 values thus obtained to a two-way analysis of variance. The average vowel-intrinsic duration ratio of the 40 speakers is 1.339 (c.i.=1.304–1.374). The ratio is comparably slightly influenced by dialect ($\eta_p^2=0.100$; $F[1,36]=3.988$, $p=0.053$), gender ($\eta_p^2=0.118$; $F[1,36]=4.794$, $p=0.035$), and an interaction of dialect and gender ($\eta_p^2=0.110$; $F[1,36]=4.454$, $p=0.042$); a one-way analysis of variance with the four speaker groups as the levels of the single factor confirms that the BP females have a larger vowel-intrinsic duration ratio than any of the other three groups (Tukey's "honestly significant difference" *post hoc* test: all three $p\leqslant0.030$), which do not differ significantly among themselves (all three $p\geqslant0.999$). Comparisons with other languages, and their implications, are discussed in Sec. VII C.

C. Dialect and gender differences in duration: Results of speaking rate?

The observed differences in vowel duration between the groups might potentially arise from between-group differences in speaking rate. To investigate whether such differences exist, we perform three between-group analyses of speaking rate.

For the first analysis we measured the durations of the utterance parts "em susse e susso," "em sasse e sasso," and so on, for all seven vowels but only for the consonant /s/; averaging over the seven vowels yields one typical sentence duration per speaker. When we subject the 40 values to a two-way analysis of variance, we find no reliable effect of dialect, gender, or dialect×gender (all three $p\geqslant0.142$). Hence, no difference in speaking rate is detected here.

For the second analysis we measured the durations of the /s/ *before* the target vowel, i.e., the initial consonants "s" of the words "susse," "sasse," and so on, for all seven vowels; averaging over the seven vowels yields one typical initial /s/ duration per speaker. A two-way analysis of variance again finds no reliable effect of dialect, gender, or dialect ×gender (all three $p\geqslant0.219$). So again no difference is found between the dialects.

For the third analysis we measured the durations of the /s/ *after* the target vowel, i.e., the medial consonants "ss" of the words susse, sasse, and so on, for all seven vowels; averaging over the seven vowels yields one typical medial /s/ duration per speaker. A two-way analysis of variance reveals an effect of dialect alone ($p=0.012$; the other two $p\geqslant0.205$): the postvocalic /s/ is *shorter* in BP than in EP, opposite to the difference in vowel durations. Hence, it looks as if the Brazilians compensate for their longer stressed vowels by shortening the following consonant. This suggests that the duration difference in the stressed vowels is not caused by a difference in speech rate between the dialects.

VI. RESULTS FOR FUNDAMENTAL FREQUENCY

The fact, mentioned in the Introduction, that the Portuguese vowel system does not use tone as a phonological feature does not preclude that different vowels may have quite different fundamental frequencies, and that fundamental frequencies may differ between dialects (as they are expected to do between genders). Figure 8 shows the dependence of F0 on vowel, dialect, and gender. Each symbol represents a value of F0 (and F2) averaged over the median F0 (and F2) values of ten speakers.

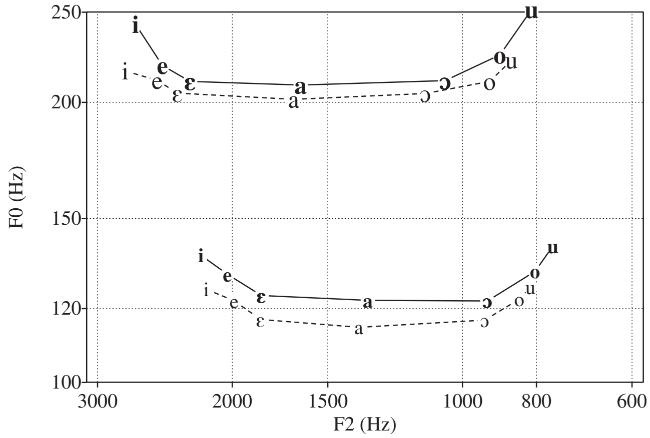

FIG. 8. Mean F0 as a function of vowel category. Solid lines and bold symbols=BP; dashed lines=EP. Top: women; bottom: men.

A. Exploratory analysis

The exploratory analysis of variance of F0 finds the expected large main effect of gender ($\eta_p^2=0.833$; $F[1,36]=179.793$, $p=1.4\times10^{-15}$): the 20 women have a (geometric) average F0 of 216.60 Hz, the 20 men one of 125.07 Hz; the F0 of Portuguese-speaking women is therefore a factor of 1.732 higher than that of Portuguese-speaking men (c.i. =1.567–1.913). We find no reliable main effect of dialect ($F[1,36]=0.007$, $p=0.932$). Within speakers we find a main effect of vowel category ($F[6\varepsilon,216\varepsilon,\varepsilon=0.492]=136.121$, $p=5.3\times10^{-36}$) and an interaction of vowel and dialect ($F=11.224$, $p=2.1\times10^{-6}$), both of which can be observed in Fig. 8 and are discussed in Sec. VI B. We find no reliable interaction of vowel and gender ($F=2.499$; $p=0.064$) or triple interaction of vowel, gender, and dialect ($F=2.276$; $p=0.085$).

B. Vowel-intrinsic F0

From the Introduction, one can expect an effect of vowel height on F0, and Fig. 8 confirms this expectation. In fact, for all 40 speakers, both /i/ and /u/ have a higher F0 than /a/. Within the analysis of Sec. VI A, pairwise comparisons between the seven vowels yield the following results for vowels of adjacent phonological heights: /i,u/ have a higher F0 than /e,o/ (all four $p<2\times10^{-9}$), /e,o/ higher than /ɛ,ɔ/ (all four $p<4\times10^{-11}$), and /ɛ,ɔ/ higher than /a/ ($p=0.000\,55$ and 0.0040). We conclude with confidence that lower vowels have a lower F0 than higher vowels in Portuguese. The fundamental frequency also seems to depend on place: /u/ has a higher F0 than /i/ ($p=0.000\,22$) and /o/ than /e/ ($p=0.049$); the difference between /ɔ/ and /ɛ/ is less than one standard error (and in the wrong direction; $p=0.334$).

To investigate the size of the vowel-intrinsic F0 effect, we define for each speaker the *vowel-intrinsic F0 ratio* as the ratio between the average F0 of the high vowels /i/ and /u/ and the F0 of the low vowel /a/. When we subject the 40 values thus obtained to a two-way analysis of variance, we find a reliable main effect of dialect ($F[1,36]=12.301$, $p=0.0012$): the average ratios are 1.158 for the 20 Brazilians and 1.095 for the 20 Europeans. The ratio is therefore greater for BP than for EP by a factor of 1.057 (c.i.=1.024–1.092; $p=0.000\,62$). Neither a main effect of gender ($F[1,36]=0.987$, $p=0.327$) nor an interaction between gender and dialect ($F[1,36]=4.454$, $p=0.079$) is reliably detected.

VII. DISCUSSION

This section compares the results of Secs. IV–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 has found that the four-way phonological vowel height contrast of Portuguese is a strong determiner of F1. That is, the seven vowels divide up into four F1 regions, where each back vowel has an F1 similar to its corresponding front vowel. This is an unsurprising observation given the phonological discussions in the Introduction and given the fact that most languages with large vowel inventories exhibit this kind of symmetry. Section IV B has also found that 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. Section IV C finds that back vowels consistently have slightly higher F1 values than their front counterparts. We speculate that a universal principle might be involved, because this effect has been found for several languages with large vowel inventories (mentioned in the Introduction), and even for five-vowel inventories the relation still seems to apply to the /i/-/u/ contrast: Iberian Spanish (the control subjects of Cervera *et al.*, 2001), Japanese (Nishi *et al.*, 2008), Czech (Chládková *et al.*, 2009), and Hebrew (Most *et al.*, 2000).

According to Sec. IV D, the BP F1 space size is 1.201 times larger for females than for males, and for the EP speakers this *female-to-male F1 space size ratio* is 1.097. In order to assess the universality of these gender differences, 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 languages and even closely related varieties in this respect.

Both Portuguese values happen to fall in between the two Dutch ones.

The combined evidence of Sec. IV E leads to the conclusion that /ɛ/ is higher (less open, having a lower absolute and relative F1) in EP from Lisbon than in BP from São Paulo. None of the studies on Portuguese vowels mentioned in the Introduction reported this dialectal difference. Regarding the ideas in the Introduction, and the location of /ɛ/ near the center of the F1 continuum, we might well be watching an impending merger (in EP) of /ɛ/ into /e/, as is also happening in Italian, French, and Catalan (see Introduction).

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

Section IV F makes four observations. First, phonological front- and backness is a strong determiner of F2 in Portuguese. This is an unsurprising observation given that Portuguese, as most languages, uses vowel place to distinguish between vowel categories. 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/ might be more fronted in EP than in BP.[4] This could have been seen by comparing earlier publications on BP (Callou *et al.*, 1996) 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 F, together with the presumably equally large uncertainties in the values reported for other languages, do not allow firm 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 (Sec. V A). 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, e.g., men tend to have stiffer articulators than women (as speculated by Simpson, 2001, 2002, but not confirmed by Simpson 2003).[5]

Second, vowels are longer in BP than in EP (Sec. V A). A comparable difference has been found in the Spanish-speaking neighbors: Morrison and Escudero (2007) found that Peruvian Spanish vowels (from Lima) were 34% longer than European Spanish vowels (from Madrid). Causation by dialectal differences in speaking rate can probably be ruled out (Sec. V C).

Third, lower vowels are longer than higher vowels (Sec. V B). In Portuguese, this vowel-intrinsic duration effect turns out to be strong: the duration ratio of low and high vowels is 1.339. The effect is stronger than in most other languages without a phonological length contrast, such as Iberian Spanish (the control subjects of 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). This language-dependence suggests that in Portuguese the effect is not solely of an automatic articulatory nature: it seems that Portuguese has turned duration into a language-specific (minor) 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, both in production (Heffner, 1937; House and Fairbanks, 1953; Luce and Charles-Luce, 1985) and in perception (Denes, 1955; Raphael, 1972).

Fourth, back vowels might be longer than their front counterparts (Sec. V B). For the high vowels, this was also found by Seara (2000). This effect may be epiphenomenal: back vowels have higher F1's than front vowels (Sec. VII A), and since F1 covaries with duration (see previous paragraph), back vowels are expected to have longer durations than front vowels.

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

Section VI identifies three influences on F0. First, the ratio by which Portuguese-speaking women have a higher average F0 than men is 1.732 (Sec. VI A). It 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; Most *et al.* (2005) report a ratio of 1.518 for Hebrew. All these ratios are much smaller than the ratio 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 might therefore well be zero, so that the effect could just be physiologically determined. However, comparing the gender-dependence of F0 across studies may be less than reliable, because the F0 difference between men and women tends to be largest at the age of our subjects (young adults) and tends to fall at later ages (Baken, 2005).

Second, high vowels have a higher F0 than low vowels, with a ratio of 1.158 for the Brazilians and a reliably smaller ratio of 1.095 for the Europeans (Sec. VI B). This vowel-intrinsic F0 effect is comparable to those reported 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). In Portuguese, the dialect-dependence suggests that the intrinsic F0 is not an automatic consequence of articulation. However, this dependence might be caused by the dialect-dependence of duration, but the literature has never identified a universal negative correlation between F0 and duration (for vowels with a constant F1), so such a cause does not seem likely.

Third, back vowels seem to have a higher F0 than front vowels in Portuguese (Sec. VI B). 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.

VIII. CONCLUSION

The present study finds several general properties of Portuguese vowels that they have in common with vowels in many other languages: they exhibit intrinsic F0 (Secs. VI B and VII D) and intrinsic duration (Secs. V B and VII C), the sizes of the F1 and F2 spaces are larger for women than for men (Secs. IV D, IV F, VII A, and VII B), F0 and formant values are higher for females than for males (Secs. IV A, IV F, VI A, VII A, VII B, and VII D), females' vowels are longer than those of males (Secs. V A and VII C), and the structure of the vowel inventory is basically symmetric (Secs. IV B and VII A) although back vowels have slightly higher F1 values than their front counterparts (Secs. IV C and VII A).

A Portuguese-specific 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 (Secs. V B and VII C); just as happened with the voicing-dependent vowel lengthening in English, one can predict that Portuguese *listeners* use this cue to a greater extent than listeners of other languages. Future research will have to verify this prediction.

There are three reliably established dialect-specific findings. One is that BP vowels are longer than EP vowels (Secs. V A, V C, and VII C). Another is that the vowel-intrinsic F0 effect is greater in BP than in EP (Secs. VI B and VII D). The third is that the lower-mid vowel /ɛ/ is higher in EP than in BP, and that it is closer to /e/ in EP than in BP (Secs. V B and VII C), a situation which might signal a future merger. To establish whether we are really witnessing a sound change in progress, a larger investigation with more age groups, social-economic strata, and regional varieties is called for. Such a more comprehensive study could also address some other questions that we had to leave open, such as the possible lowering of high vowels and the degree of articulatory automaticity of the intrinsic duration and intrinsic F0 effects.

At the methodological level, the proposed formant ceiling optimization method found that the average difference of the vocal tract lengths associated with /i/ and /u/ is comparable to the average difference of the female and male vocal tract lengths. Future investigations involving automatic formant measurements could benefit from this observation.

ACKNOWLEDGMENTS

This research was supported by NWO (Netherlands Organization for Scientific Research) Grant No. 016.024.018 to P.B. and by a CAPES (Committee for Postgraduate Courses in Higher Education, Brazilian Ministry of Education) grant to A.S.R. We 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.

[1]Some of the authors (Mateus *et al.*, 2005, p. 79) group /ɛ/ and /ɔ/ with /a/ by calling them "low vowels;" there seems to be no reason for this move other than minimizing the number of phonological features.

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

[3]A technical detail: the Gaussian-like shape of the window requires tails that capture another 20% of the vowel duration on each side of the central 40%.

[4]One could look specifically into the degree of fronting of /u/, knowing that /u/ was historically fronted (auditorily) in several European languages (dates approximate): 1st-century BC Greek (Sihler, 1995, p. 37), 5th-century Slavic (Stieber, 1979, p. 23), Old Dutch (Schönfeld, 1932, p. 82), 9th-century French (Meyer-Lübke, 1908, p. 53), 15th-century Swedish (Kock, 1911, p. 191), 20th-century southern British English (Harrington *et al.*, 2008). The European speakers indeed have a higher F2 than the Brazilians, but this cannot at this point be reliably generalized to the populations ($F[1,36]=3.676$; $p=0.063$).

[5]If vowel duration is related to speaking rate, identical utterances should be longer when spoken by women than when spoken by men. Whiteside (1996) did find this, but Simpson (2001) did not. Our Portuguese data can neither confirm nor disconfirm such gender differences in speaking rate (Sec. V C).

Adank, P. (**2003**). "Vowel normalization: A perceptual-acoustic study of Dutch vowels," Ph.D. thesis, University of Nijmegen.

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.

Allan, L. G., and Gibbon, J. (**1991**). "Human bisection at the geometric mean," Learn Motiv **22**, 39–58.

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

Baken, R. J. (**2005**). "The aged voice: A new hypothesis," J. Voice **19**, 317–325.

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

Barroso, H. (**1999**). *Forma e substância de expressão da língua portuguesa (Form and substance of the Portuguese language expression)* (Almedina, Coimbra).

Bisol, L. (**1996**). *Introdução a estudos de fonologia do português brasileiro (Introduction to studies on the phonology of Brazilian Portuguese)* (Editora Universitária da Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre).

Boersma, P., and Weenink, D. (**2008**). "Praat: doing phonetics by computer (Version 5.0.43)" [Computer program], retrieved 9 December 2008 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 (The vocalism of the Portuguese of Brazil)," Letras de Hoje (Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre) **31**(2), 27–40.

Câmara, J. M., Jr. (**1970**). *Estrutura da língua portuguesa (Structure of the Portuguese Language)* (Vozes, Petrópolis).

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

Chládková, K., Boersma, P., and Podlipský, V. J. (**2009**). "On-line formant shifting as a function of F0," in Proceedings of Interspeech 2009.

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 (The Phonetics and Phonology of Portuguese)* (Contexto, São Paulo).
Delgado-Martins, M. R. (**1973**). "Análise acústica das vogais orais tônicas em português (Acoustic analysis of the stressed oral vowels in Portuguese)," Boletim de Filologia (University of Lisbon) **22**, 303–314.
Delgado-Martins, M. R. (**2002**). *Fonética do português: trinta anos de investigação (The Phonetics of Portuguese: Thirty Years of Research)* (Caminho, Lisbon).
Denes, P. (**1955**). "Effect of duration on the perception of voicing," J. Acoust. Soc. Am. **27**, 761–764.
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 thyroumbrometer," J. Phonetics **2**, 327–335.
Falé, I. (**1998**). "Duração das vogais tónicas e fronteiras prosódicas: uma análise em estruturas coordenadas (Duration of stressed vowels and prosodic boundaries: An analysis on coordinated structures)," Actas do XIII Encontro Nacional da Associação Portuguesa de Linguística (Colibri, Lisbon), pp. 255–269.
Gibbon, J. (**1977**). "Scalar expectancy theory and Weber's Law in animal timing," Psychol. Rev. **84**, 279–325.
Goldstein, U. (**1980**). "An articulatory model for the vocal tracts of growing children," Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.
Hagiwara, R. (**1997**). "Dialect variation and formant frequency: The American English vowels revisited," J. Acoust. Soc. Am. **102**, 655–658.
Harrington, J., Kleber, F., and Reubold, U. (**2008**). "Compensation for coarticulation, /u/-fronting, and sound change in standard southern British: An acoustic and perceptual study," J. Acoust. Soc. Am. **123**, 2825–2835.
Heffner, R.-M. (**1937**). "Notes on the length of vowels," Am. Speech **12**, 128–134.
Henton, C. G. (**1989**). "Fact and fiction in the description of female and male pitch," Language & Communication **9**, 299–311.
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.
Kent, R. D., and Read, C. (**2002**). *The Acoustic Analysis of Speech*, 2nd ed. (Singular, San Diego).
Kock, A. (**1911**). *Svensk ljudhistoria (Swedish Sound History)* (Gleerup, Lund), Vol. **2**.
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).
Landick, M. (**1995**). "The mid-vowels in figures: hard facts," The French Review **69**, 88–102.
Lehiste, I. (**1970**). *Suprasegmentals* (MIT, Cambridge, MA).
Lehiste, I., and Peterson, G. E. (**1961**). "Some basic considerations in the analysis of intonation," J. Acoust. Soc. Am. **33**, 419–425.
Lindblom, B. (**1967**). "Vowel duration and a model of lip-mandible coordination," Speech Transm. Lab. Q. Prog. Status Rep. **4**, 1–29.
Luce, P. A., and Charles-Luce, J. (**1985**). "Contextual effects on vowel duration, closure duration, and the consonant/vowel ratio in speech production," J. Acoust. Soc. Am. **78**, 1949–1957.
Maiden, M. (**1997**). "Vowel systems," in *The Dialects of Italy*, edited by M. Maiden and M. Parry (Routledge, London), pp. 7–14.
Mateus, M. H. M. (**1990**). *Fonética, fonologia e morfologia do português (The Phonetics, Phonology, and Morphology of Portuguese)* (Universidade Aberta, Lisbon).
Mateus, M. H. M., and d'Andrade, E. (**1998**). "The syllable structure in European Portuguese," DELTA [Documentação de Estudos em Linguística Teórica e Aplicada] (Pontifícia Universidade Católica de São Paulo, São Paulo) **14**, 13–32.
Mateus, M. H. M., and d'Andrade, E. (**2000**). *The Phonology of Portuguese* (Oxford University Press, Oxford).
Mateus, M. H. M., Falé, I., and Freitas, M. (**2005**). *Fonética e fonologia do português (Portuguese Phonetics and Phonology)* (Universidade Aberta, Lisbon).
Meyer-Lübke, W. (**1908**). *Historische Grammatik der französischen Sprache. 1. Laut- und Flexionslehre (Historical Grammar of the French Language. 1. Phonology and Inflectional Morphology)* (Carl Winter, Heidelberg).
Moraes, J. A. (**1999**). "Um algoritmo para a correção/simulação da duração dos segmentos vocálicos em português (An algorithm to correct/simulate duration in Portuguese vocalic segments)," in *Estudos da prosódia (Prosody Studies)*, edited by E. Scarpa (Editora da Unicamp, Campinas), pp. 69–84.
Moraes, J. A., Callou, D., and Leite, Y. (**1996**). "O sistema vocálico do português do Brasil: caracterização acústica (The vocalic system of the Portuguese of Brazil: Acoustic characterization)," in *Gramática do português falado (The Grammar of Spoken Portuguese)*, 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.
Most, T., Amir, O., and Tobin, Y. (**2000**). "The Hebrew vowel system: Raw and normalized acoustic data," Lang Speech **43**, 295–308.
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.
Nishi, K., Strange, W., Akahane-Yamada, R., Kubo, R., and Trent-Brown, S. (**2008**). "Acoustic and perceptual similarity of Japanese and American English vowels," J. Acoust. Soc. Am. **124**, 576–588.
Ohala, J. J., and Eukel, B. (**1987**). "Explaining the intrinsic pitch of vowels," in *In Honor of Ilse Lehiste*, edited by R. Channon and L. Shockey (Foris, Dordrecht), pp. 207–215.
Peterson, G. E., and Barney, H. L. (**1952**). "Control methods used in a study of vowels," J. Acoust. Soc. Am. **24**, 175–184.
Raphael, L. J. (**1972**). "Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English," J. Acoust. Soc. Am. **51**, 1296–1303.
Recasens, D., and Espinosa, A. (**2009**). "Dispersion and variability in Catalan five and six peripheral vowel systems," Speech Commun. **51**, 240–258.
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**, pp. 54–57.
Ryalls, J. H., and Lieberman, P. (**1982**). "Fundamental frequency and vowel perception," J. Acoust. Soc. Am. **72**, 1631–1634.
Schönfeld, M. (**1932**). *Historiese grammatika van het Nederlands (Historical Grammar of Dutch)* (Thieme, Zutphen).
Seara, I. C. (**2000**). "Estudo acústico-perceptual da nasalidade das vogais do português brasileiro (Acoustical-perceptual study on the nasality of the vowels of Brazilian Portuguese)," Ph.D. thesis, Universidade Federal de Santa Catarina, Florianópolis.
Sihler, A. L. (**1995**). *New Comparative Grammar of Greek and Latin* (Oxford University Press, New York).
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. (**2003**). "Possible articulatory reasons for sex-specific differences in vowel duration," in Proceedings of the sixth International Seminar on Speech Production, Sydney, pp. 261–266.
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.
Stieber, Z. (**1979**). *Zarys gramatyki prorównawczej języków slowiańskich (An Outline of the Comparative Grammar of the Slavic Languages)* (Państwowe Wydawnictwo Naukowe, Warsaw).
Stevens, K. (**1998**). *Acoustic Phonetics* (MIT, 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.
Tielen, M. T. J. (**1992**). "Male and female speech: An experimental study of sex-related voice and pronunciation characteristics," Ph.D. thesis, University of Amsterdam.
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. Phonetic Assoc. **26**, 23–40.
Winer, B. J. (**1962**). *Statistical Principles in Experimental Design* (McGraw-Hill, New York).
Yamazawa, H., and Hollien, H. (**1992**). "Speaking fundamental frequency patterns of Japanese women," Phonetica **49**, 128–140.