

Context-specific acoustic differences between Peruvian and Iberian Spanish vowels

Kateřina Chládková^{a)}

Amsterdam Center for Language and Communication, University of Amsterdam, Spuistraat 210, 1012VT Amsterdam, The Netherlands

Paola Escudero

MARCS Auditory Laboratories, Building 1, University of Western Sydney, Bullecourt Avenue, Milperra, NSW 2214, Australia

Paul Boersma

Amsterdam Center for Language and Communication, University of Amsterdam, Spuistraat 210, 1012VT Amsterdam, The Netherlands

(Received 21 July 2010; revised 28 April 2011; accepted 28 April 2011)

This paper examines four acoustic properties (duration F0, F1, and F2) of the monophthongal vowels of Iberian Spanish (IS) from Madrid and Peruvian Spanish (PS) from Lima in various consonantal contexts (/s/, /f/, /t/, /p/, and /k/) and in various phrasal contexts (in isolated words and sentence-internally). Acoustic measurements on 39 speakers, balanced by dialect and gender, can be generalized to the following differences between the two dialects. The vowel /a/ has a lower first formant in PS than in IS by 6.3%. The vowels /e/ and /o/ have more peripheral second-formant (F2) values in PS than in IS by about 4%. The consonant /s/ causes more centralization of the F2 of neighboring vowels in IS than in PS. No dialectal differences are found for the effect of phrasal context. Next to the between-dialect differences in the vowels, the present study finds that /s/ has a higher spectral center of gravity in PS than in IS by about 10%, that PS speakers speak slower than IS speakers by about 9%, and that Spanish-speaking women speak slower than Spanish-speaking men by about 5% (irrespective of dialect). © 2011 Acoustical Society of America. [DOI: 10.1121/1.3592242]

PACS number(s): 43.70.Fq, 43.70.Kv [AL]

Pages: 416–428

I. INTRODUCTION

This paper presents a detailed acoustic description of vowels of Peruvian Spanish (PS) from Lima and Iberian Spanish (IS) from Madrid in several consonantal and phrasal contexts. The acoustic analyses reported in this paper aim at revealing whether there are differences between these two standard Spanish varieties (henceforth, “dialects”) in the acoustic properties of vowels and, most importantly, whether the cross-dialectal differences are specific for a particular consonantal or phrasal context. We focus on the following acoustic correlates of vowel identity: duration, fundamental frequency (F0), first formant (F1), and second formant (F2).

The phoneme inventories of both Spanish dialects contain the same five monophthongal vowels, namely, /i/, /e/, /a/, /o/, and /u/ (we ignore in this paper sequences such as /je/, /we/, and /ei/, which can be regarded as diphthongs).¹ One expects similarities as well as differences between the two dialects.

As far as similarities between the two dialects are concerned, one expects that there are universal and Spanish-specific differences between the five vowel categories with respect to F1, F2, duration and F0. The five Spanish vowels have been associated with different phonological heights and degrees of backness (Harris, 1969), so one expects them to have different values of F1 (the main acoustic correlate of

vowel height) and/or F2 (the main acoustic correlate of vowel backness). Cross-linguistically, low vowels tend to be produced with longer duration than high vowels of identical phonological length (House and Fairbanks, 1953, p. 111; Lehiste, 1970, p. 18), so one expects this correlation to hold for Spanish, where vowels have no phonological length contrast. For Spanish women and men from Granada, Mendoza *et al.* (2003) indeed found that the low vowel /a/ was produced longer than the high vowels /i/ and /u/. Similarly, Marín Gálvez (1994–1995) found that /a/ was longer than /e/ and /o/, and these longer than /i/ and /u/, for two male speakers of an unspecified variety of Iberian standard Spanish. Although this vowel-intrinsic duration effect can be attributed to the physiology of speech (Lindblom, 1967; Lehiste, 1970, pp. 18 and 19; Solé, 2007, p. 303), it is clear that speakers can control duration independently from F1. A research question of the present study is therefore whether Madrid and Lima Spanish follow this universal tendency, and whether there are between-dialect differences in the strength of the effect.

Likewise, cross-linguistically, low vowels tend to be produced with a lower F0 than high vowels (Lehiste and Peterson, 1961; Whalen and Levitt, 1995). Again, there are physiological explanations for this vowel-intrinsic F0 effect (Ohala and Eukel, 1987), but since speakers can control F0 and F1 independently, it is an open question whether Spanish follows this universal tendency. Albalá *et al.* (2008) found that /e/ had a lower F0 than /i/, /o/, and /a/ (in an unspecified variety of Iberian Spanish), but they advised caution because

^{a)}Author to whom correspondence should be addressed. Electronic mail: k.chladkova@uva.nl

tokens of the vowel /e/ had been recorded in contexts different from the other vowels. Therefore, it is still to be shown whether Spanish follows the universal pattern, and whether IS and PS differ in the extent to which they follow it.

Another expected similarity is the effect of gender on duration, F0, and formants: average male vowels are shorter than average female vowels in many languages (Simpson and Ericsson, 2003) and probably have lower values of F0, F1, and F2 in all languages (American English: Hillenbrand *et al.*, 1995; British English: Deterding, 1995, Portuguese: Escudero *et al.*, 2009; Dutch: Adank *et al.*, 2004; Czech: Chládková *et al.*, 2009). The present paper ascertains these tendencies for Spanish and investigates whether the two dialects might differ in the extent to which they follow them.

In addition to similarities, differences between IS and PS can be expected as well, because the two dialects partially developed separately since the 16th century and can therefore be assumed to be separate linguistic systems. As a result, some of the five vowels may be produced differently between IS and PS, because it has repeatedly been shown that the “same” vowel phoneme (i.e., a vowel phoneme represented in the same way by phonologists) can have different acoustic realizations across languages (Disner, 1983; Bradlow, 1995; Nishi *et al.*, 2008), including across languages that share the same /i, e, a, o, u/ inventory (Greek: Jongman *et al.*, 1989; Hebrew: Most *et al.*, 2000; Spanish: Cervera *et al.*, 2001) and across varieties of a single language (Northern vs Southern Standard Dutch: Adank *et al.*, 2004; Brazilian vs European Portuguese: Escudero *et al.*, 2009). For that reason, several studies on Spanish vowels avoided pooling data of different Spanish dialects and tested speakers of one Spanish dialect only (Argentinean Spanish: Guirao and Borzone de Manrique, 1975; Castilian Spanish: Bradlow, 1995, 1996; Cervera *et al.*, 2001). The differences between the dialects may lie in the formants, in duration (e.g., the degree of stress timing vs syllable timing), and even in F0 (which may be language- and, therefore, dialect-dependent: Yamazawa and Hollien, 1992).

A specific expected difference involves vowels in an /s/ context. The fricative /s/ has been reported to be realized as concave retracted apical alveolar in IS (Navarro Tomás, 1932, pp. 105–107; Harris, 1969, p. 192; Widdison, 1997; Martínez-Celdrán *et al.*, 2003) and is dental in PS (our own observation, which is not unexpected in view of existing literature on the distribution of the varieties of /s/ in Latin America: Canfield, 1981). These different articulations are expected to influence differentially at least the F2 values of the flanking vowels (Gordon *et al.*, 2002; an influence on F3 is also expected but outside the scope of the present study).

Earlier acoustic studies compared the vowels of different dialects of Spanish. Godínez (1978) reported the following differences between Spanish vowels from Argentina, Mexico, and Spain: (1) a smaller vowel space (with respect to F1) in Peninsular Spanish than in the other two dialects and (2) a three-way dialectal difference in the F1 of the two front vowels /i/ and /e/. These findings, however, were from vowels produced by a very low number of only male speakers per dialect (six Mexican, four Argentinean, and six Spanish). Morrison and Escudero (2007) analyzed isolated

vowels in sentence-final position elicited from Spanish speakers from Madrid and Lima, and found differences in duration and fundamental frequency between the two dialects. Morrison and Escudero did test a relatively large number of both female and male participants, but analyzed only sentence-final vowels produced in isolation, which is a context that rarely occurs outside laboratory settings.

Other dialectal studies on Spanish than Morrison and Escudero (2007) did investigate vowels in context, but often disregarded the possible effect of different consonantal or phrasal contexts on vowel production. Thus, Godínez (1978) considered vowels only in the /pVpa/ context, and Bradlow, (1995) only in the /pVta/ and /bVta/ contexts, i.e., without varying place or manner of articulation. Cervera *et al.* (2001) did elicit CVCV words with several different consonant contexts but collapsed all of them in the analysis. An exception to these single- or no-context approaches is Guirao and Borzone de Manrique (1975), who elicited Argentinean Spanish vowels produced in isolation and vowels embedded in two different monosyllables, namely, /bVd/ and /pVs/. They did not collapse the contexts in their analyses and noted that there are some differences in some vowels’ formant values in the different contexts, namely, that the effect of having a context at all (as opposed to isolation) is most apparent in the F1 of /e/ and in the F2 of /u/, and that the /bVd/ context raises the F2 value of /a/. However, their analysis was based solely on visual inspection of vowel formant values plotted in a two-dimensional vowel space, and hence not statistically validated. Like most acoustic studies on Spanish dialects, the majority of studies on the vowels of other languages have not considered contextual variation either (e.g., Jongman *et al.*, 1989; Hillenbrand *et al.*, 1995; Bradlow, 1995; Most *et al.*, 2000; Adank *et al.*, 2004; Escudero *et al.*, 2009; exceptions are, e.g., Strange *et al.* 2005, 2007).

Previous research on coarticulatory effects in speech production has shown that consonants tend to affect the acoustic properties of neighboring vowels: the consonantal place of articulation affects vowel formant transitions (for a general review see Fowler and Saltzman, 1993; for English: Stevens *et al.*, 1966; for Swedish: Lubker and Gay, 1982; for Japanese: Beckman and Shoji, 1984; for Italian: Farnetani and Recasens, 1993; for Dutch: Van Bergem, 1994), and vowel duration varies according to both the phonological voicing and the manner of articulation of the following consonant, e.g., English vowels are longer before voiced than before voiceless obstruents and longer before fricatives than before plosives (Peterson and Lehiste, 1960; Umeda, 1975; Van Santen, 1992). At first sight, it may seem that collapsing a number of consonantal contexts should be viable as long as the number and type of contexts is the same for each of the languages researched. This would hold only if the consonants had the same articulatory realizations and timings across languages. Much is known about the effect of obstruent voicing on the duration of a preceding vowel: when we compare the results of numerous studies, we can conclude that this effect differs greatly across languages (American English: Peterson and Lehiste, 1960; Crystal and House, 1988; Dutch: Slis and Cohen, 1969; Russian: Chen, 1970; Catalan: Dinnsen and Charles-Luce, 1984; Czech: Podlipský, 2009). Much less is

known about the effect of manner of articulation on the duration of a preceding vowel, because this has rarely been studied for languages other than English. Nevertheless, given that fricatives have different realizations across languages, one can assume that the manner effect varies cross-linguistically as well. With respect to consonantal place of articulation, [Strange et al. \(2007\)](#) showed that consonantal context not only affects vowel production in German, English and French in general, but also that the effect of neighboring consonants is different in each of these languages.

Not only consonantal context but also phrasal context has an effect on vowel properties. Previous acoustic analyses have shown that cross-linguistically, phonologically longer phrases are spoken at a higher local rate. That is, segment duration correlates negatively with the number of syllables, moras or segments in a prosodic word or phrase. This dependence of segment duration on the phonological length of higher prosodic units has naturally been found for stress-timed languages such as English ([Lehiste, 1972](#); [Klatt, 1973](#)), Swedish ([Lindblom, 1968](#)) and Dutch ([Nooteboom, 1972](#)), but has also been found for syllable-timed or mora-timed languages such as Italian ([Pickett et al., 1999](#)) or Japanese ([Hirata, 2004](#), Sec. 2.4.4), where the duration of a syllable or mora used to be regarded as independent of the structure of higher prosodic units.

A fast local speaking rate not only yields shorter vowel durations, but might also reduce the size of the vowel space, because *global* speaking rate may do so ([Gay, 1968](#); [Fourakis, 1991](#); [Turner et al., 1995](#)), although this is controversial ([Gay, 1978](#); [Van Bergem, 1993](#)). For such reasons, several cross-dialectal studies compared vowels that were produced in different phrasal contexts: [Guirao and Borzone de Manrique \(1975\)](#) compared isolated vowels to vowels in CVC words and observed that some vowels had different formant values in context than in isolation, and [Strange et al. \(2007\)](#) elicited vowels in isolated words (i.e., phrases with few segments) and in words embedded in sentences (i.e., phrases with more segments) and indeed found that duration differences between vowels were reduced in the sentence context.

To provide a thorough comparison of vowel properties in the two dialects and to uncover any possible differences between the vowels of IS and PS that may otherwise be obscured by collapsing over a number of different contexts, the present study considers (1) the effect of *five different consonantal contexts*, namely, /p/, /t/, /k/, /f/, and /s/, which differ in both place and manner of articulation, and (2) the effect of *phrasal context*, namely, vowels produced in isolated words versus vowels produced in words embedded in a carrier sentence, which are expected to have slower and faster speaking rates, respectively. The chosen consonantal contexts were all voiceless because there are no voiced intervocalic obstruents in these varieties of Spanish (the so-called voiced spirants are actually approximants: [Navarro Tomás, 1932](#); [Martínez-Celdran, 2003](#)). These five consonantal contexts are (at least phonologically) identical to the contexts recorded (but collapsed in the analyses) in [Escudero et al. \(2009\)](#), so that future cross-language comparisons of Spanish and Portuguese vowels may be possible.

The present study adopts from [Escudero et al. \(2009\)](#) the methods of data collection (ten speakers per gender per

dialect) and data analysis. As for the latter, [Escudero et al.](#) developed a new method of vowel formant tracking in which formant analysis settings are adapted to the vowel category and to the speaker at hand. They showed that the *ceiling* of vowel formant analyses, i.e., the maximum analyzable frequency, does not depend on a speaker's gender alone but is, to a large extent, also dependent on vowel category. Since the study of [Escudero et al.](#) revealed acoustic differences between European and Brazilian Portuguese vowels that had not been found by any earlier study, the present paper assumes that their method can reliably uncover formant differences between vowels of the two Spanish dialects considered in the present study. The similarities in the data collection methodology and analysis technique between [Escudero et al.](#) and the present study will allow future comparisons across languages.

The present study first reports a general analysis of Madrid and Lima vowels with collapsed contexts, and then considers how consonantal and phrasal context affects the vowels in the two Spanish dialects.

II. METHOD

A. Participants

Productions were elicited from a total of forty speakers. In order to reliably assess dialect-dependent differences in vowel realizations, an equal number of speakers per dialect were tested, namely, 20 IS and 20 PS speakers. To control for the effect of gender on vowel productions within and between dialects, an equal number of female and male speakers were tested per dialect, i.e., ten female and ten male speakers of each dialect (although one female speaker of PS had to be excluded due to recording problems). We selected young, highly educated, monolingual speakers who had lived all their lives in Lima or Madrid. They were regarded as monolingual if they indicated that they did not have knowledge of any language other than Spanish and English, while their self-estimated proficiency in English, on a scale from 0 to 7, was 2 or below, and that they had been raised by monolingual parents. They were all university students, between 19 and 28 yr of age (mean age was 23.4 for IS and 24.2 for PS), enrolled at Universidad Politécnica or at Universidad Complutense in Madrid and at Pontificia Universidad Católica del Perú in Lima.

B. Data collection

The recordings were made in quiet rooms at the universities in Madrid and Lima using an Edirol R-1 recorder (Roland Corporation, Osaka) and a unidirectional Sony ECM-MS907 condenser microphone (Sony Corporation, Tokyo), with a sample rate of 22050 Hz and 16-bit quantization.

The recording procedure contained 50 trials. In each trial, a speaker read aloud words and sentences that were presented in Spanish orthography on a computer screen. A trial started with a disyllabic nonsense word, presented in isolated position. The shape of the word was CV₁CV₂, where C was one of the consonants /p/, /t/, /k/, /f/, and /s/,² V₁ was the target vowel, i.e., one of the five Spanish monophthongs

/a/, /e/, /i/, /o/, and /u/, and V_2 was meant to be a neutral vowel so as to minimize its influence on V_1 . As Spanish has no schwa, we used for V_2 the two mid vowels /e/ and /o/ to help average the result over front and back environments. In the first half of the trials (i.e., the first 25 trials) the target word was *CVCe*, in the other half (i.e., the remaining 25 trials) the target word was *CVCo*. The isolated target word was immediately followed by a sentence in which the same word was embedded: *En CVCe y CVCo tenemos V*. An example of a trial from the first block and a corresponding trial from the second block would be “Fife. En fife y fifo tenemos i.” (meaning ‘Fife. In fife and fifo we have i.’) and “Fifo. En fife y fifo tenemos i.”, respectively. During the whole recording session, a native Spanish-speaking experimenter (the second author) was present. Before the recording, the speaker practiced reading the sentences in a natural way. In case during the recording a phrase or a part of a trial was judged as unnatural or mispronounced, the experimenter asked the participant to repeat the whole trial, i.e., the isolated target word together with the immediately following sentence.

In each trial, a target vowel with context thus occurred three times, namely, once in the isolated word and twice in the words embedded in the following sentence. The vowels in the sentence-final isolated position are not analyzed in the present study (they form a separate data set, which was analyzed by Morrison and Escudero, 2007). Thus, in the present study five consonantal contexts were combined with five vowel qualities; each subject produced a total of 150 target vowel tokens in context, 50 of which were vowels in the isolated word condition and 100 were vowels in the sentence condition. The data collection procedure yielded a total of 6000 vowel tokens.

C. Data analysis

Since data from one PS female speaker had to be excluded due to erroneous recording, there were a total of 5850 analyzable vowel tokens (3000 IS tokens, 2850 PS tokens). The start and end points of each token were labeled manually in the digitized sound wave, and defined as the zero crossings associated with the first and the last period of the waveform that were judged to have considerable amplitude and a shape resembling that of the central periods of the vowel. The resulting vowel segments were subsequently analyzed for duration, fundamental frequency (F0), first formant (F1) and second formant (F2).

D. Acoustic analysis of duration

The duration of a vowel token was computed as the time between its start point and its end point, as determined via the method of Sec. II C.

E. Acoustic analysis of fundamental frequency

Fundamental frequency was measured in the computer program Praat (Boersma and Weenink, 2009), which determined the F0 curves of all recordings by the cross correlation method. Initially, the pitch range was set to be 60–400

Hz for male speakers, and 120–400 Hz for female speakers. If, for a certain vowel token, Praat failed to determine the F0 with these settings, the token was reanalyzed with different settings depending on the speaker’s gender. The reanalysis of tokens produced by women was done with a lowered pitch floor of 75 Hz, while the reanalysis of tokens produced by men was done with a lowered criterion for voicedness. If the reanalysis failed as well, F0 was determined manually from the waveform. The F0 was measured in steps of 1 ms, and the median F0 value of the central 40% of the vowel was chosen as the token’s representative F0 value, i.e., the value to be used in subsequent analyses.

F. Acoustic analysis of vowel formants

The first and second formants were measured on a single window over the central 40% of the vowel by the Burg algorithm (Anderson, 1978) built into Praat. The window shape was lowered Gaussian (the window is down to 5.0% at the boundaries of the central 40% and physically extends with low tails into the 20% parts before and after the central 40%, at the boundaries of which it reaches 0). Praat had to determine the values of the first five formants in the range between 50 Hz and the value of the *optimal ceiling*. The procedure by which we arrived at the optimal ceiling value is explained below.

To take advantage of having collected a large number of tokens per vowel and speaker (i.e., each of the five vowels was produced 30 times by each speaker), we adopted a method introduced in Escudero *et al.* (2009), which computes the optimal ceiling separately for each vowel category per speaker. The method determines the first five formants 201 times by setting the ceiling in 10 Hz steps between 4500 and 6500 Hz for women and between 4000 and 6000 Hz for men. For each vowel category for each speaker, the ceiling that yields the lowest *within-speaker* variation in F1 and F2 between the 30 tokens (computed as the variance of the thirty logF1 values plus the variance of the thirty logF2 values) is chosen as the optimal ceiling for that vowel of that speaker. Escudero *et al.* (2009) compared vowel formant values measured using the more traditional method of formant analysis with a fixed gender ceiling (i.e., the formant ceiling is the same for all vowels produced by all female or all male speakers) to vowel formants measured with the optimized ceiling method, and showed that the latter method of formant analysis minimizes *between-speaker* variation and thereby provides a more reliable measurement of vowel formants.

III. RESULTS

Sections III A–D report statistical analyses performed in SPSS [version 16.0, release 16.0.1 (November 15, 2007), SPSS, Inc., 1989–2007] and Praat (Boersma and Weenink, 2009). We tested for various effects of dialect, gender, and vowel category on four vowel properties (duration, F0, F1, and F2) and for interactions of dialect, gender and vowel category with phrasal context, consonantal context, and both.

All values were initially measured along linear scales, i.e., ms for duration and Hz for F0 and formants. They were

subsequently transformed into logarithmic values, because duration has been shown to be perceived logarithmically in animals (Gibbon, 1977) as well as humans (Allan and Gibbon, 1991), F0 changes and ranges in women and men are more comparable along logarithmic than along linear axes (Hudson and Holbrook, 1981; Graddol, 1986; Henton, 1989; Tielen, 1992), and the effect of speaker identity (via vocal tract length) on vowel formants tends to be relative (i.e., expressed in terms of ratios, which along logarithmic axes are just differences) rather than absolute (Nearey, 1992). On the log-transformed data, we then applied linear statistical models. When these models yielded, e.g., between-dialect differences, we report these differences in relative terms (with respect to milliseconds and hertz values) rather than in absolute terms (the logarithms themselves, i.e., in octaves, semitones or cents); figures, accordingly, have logarithmic axes with marks in Hz and ms, and reported averages are always geometric averages.

Figures 1 and 2 are detailed graphical visualizations of the median F1 and F2 values per speaker. Figure 1 shows a plot of the median F1 and F2 values of each of the five vowels of each of the ten IS female and nine PS female speakers,

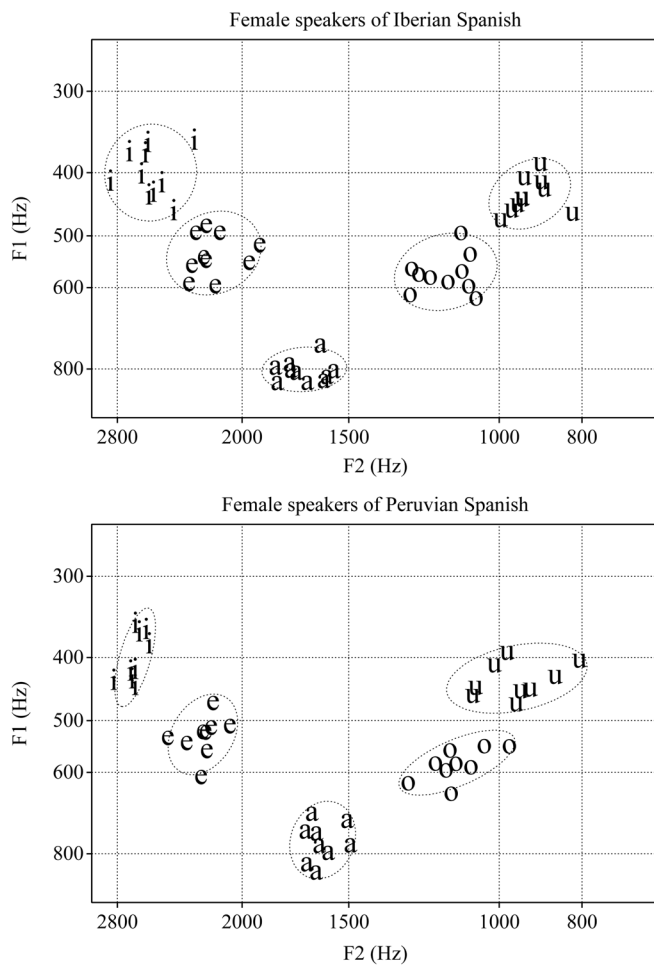


FIG. 1. F1 and F2 values of the vowels of ten female speakers of Iberian Spanish and nine female speakers of Peruvian Spanish. Each vowel symbol represents a median over all 30 tokens produced by one speaker. The ellipses represent two standard deviations from the group mean and thus show the between-speaker variation.

and Fig. 2 does the same for the ten IS male and ten PS male speakers. Table I provides a summary of the average values and standard deviations for all the measured vowel properties (i.e., duration, F0, F1, and F2). Each value in the table is a (geometric) average over ten (PS male, IS male, or IS female) or nine (PS female) speakers. Figures 1 and 2 and Table I report values of vowels as collapsed over all consonantal and phrasal contexts.

A. Universal effects of vowel category and gender

To test whether vowels have different acoustic properties in IS than in PS when all the phrasal and consonantal contexts are collapsed, we ran an exploratory repeated-measures analysis of variance on the duration, F0, F1, and F2 values of the speakers, with vowel category as the within-subject factor and gender and dialect as the between-subject factors (each data point is a median over a speaker's 30 tokens of a certain vowel). In these and all subsequent analyses, if the tests failed to pass Mauchly's test of sphericity, we used Huynh-Feldt's correction, which reduces the number of degrees of freedom by a factor ϵ .

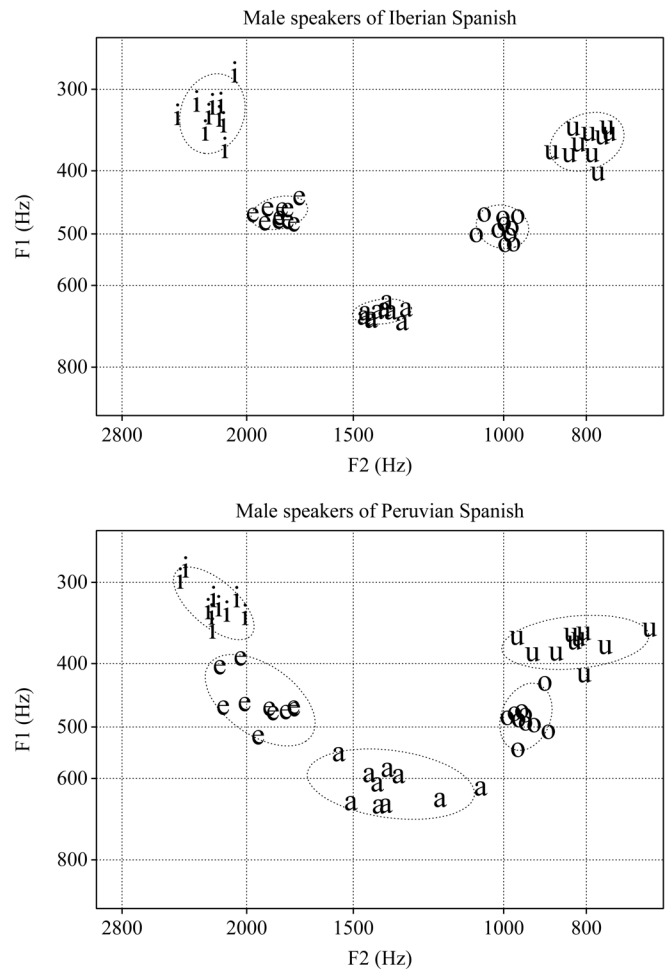


FIG. 2. F1 and F2 values of the vowels of ten male speakers of Iberian Spanish and ten male speakers of Peruvian Spanish. Each vowel symbol represents a median over all 30 tokens produced by one speaker. The ellipses represent two standard deviations from the group mean and thus show the between-speaker variation.

TABLE I. The duration, F0, F1, and F2 of vowels produced by female (F) and male (M) speakers of Iberian Spanish (IS) and Peruvian Spanish (PS). Every cell represents the geometric average over ten speakers (nine speakers in case of PS females); between parentheses appears the between-speaker standard deviation ($N=9$ or 10), expressed as a factor.

Dialect	Measure	Gender	Vowel				
			a	e	i	o	u
IS	duration (ms)	M	77 (1.137)	69 (1.143)	62 (1.147)	70 (1.130)	66 (1.126)
		F	85 (1.101)	76 (1.112)	69 (1.116)	76 (1.130)	70 (1.136)
	F0 (Hz)	M	116 (1.256)	118 (1.240)	125 (1.151)	120 (1.244)	126 (1.161)
		F	216 (1.091)	220 (1.083)	227 (1.081)	222 (1.088)	231 (1.102)
	F1 (Hz)	M	658 (1.022)	464 (1.030)	327 (1.073)	488 (1.039)	361 (1.053)
		F	801 (1.040)	531 (1.077)	400 (1.090)	568 (1.071)	431 (1.064)
F2 (Hz)	M	1389 (1.041)	1832 (1.039)	2195 (1.045)	1003 (1.036)	799 (1.051)	
	F	1691 (1.059)	2159 (1.066)	2560 (1.064)	1155 (1.072)	921 (1.057)	
PS	duration (ms)	M	83 (1.149)	75 (1.157)	67 (1.140)	76 (1.158)	68 (1.161)
		F	87 (1.138)	81 (1.141)	73 (1.189)	81 (1.157)	71 (1.165)
	F0 (Hz)	M	128 (1.123)	131 (1.127)	137 (1.142)	133 (1.134)	139 (1.154)
		F	209 (1.060)	212 (1.061)	220 (1.066)	214 (1.061)	222 (1.062)
	F1 (Hz)	M	612 (1.064)	455 (1.084)	323 (1.066)	483 (1.061)	371 (1.049)
		F	762 (1.071)	525 (1.074)	400 (1.091)	580 (1.059)	430 (1.064)
F2 (Hz)	M	1356 (1.119)	1929 (1.077)	2186 (1.056)	942 (1.036)	824 (1.104)	
	F	1610 (1.045)	2223 (1.048)	2669 (1.028)	1121 (1.081)	954 (1.099)	

The analysis reveals a main effect of vowel category on all the four measures, i.e., duration ($F[4\epsilon, 140\epsilon, \epsilon = 0.910] = 175.833, p = 5.9 \times 10^{-49}$), F0 ($F[4\epsilon, 140\epsilon, \epsilon = 0.352] = 33.170, p = 2.7 \times 10^{-8}$), F1 ($F[4\epsilon, 140\epsilon, \epsilon = 0.642] = 1142.371, p = 7.9 \times 10^{-69}$), and F2 ($F[4\epsilon, 140\epsilon, \epsilon = 0.801] = 2376.632, p = 1.1 \times 10^{-102}$). This finding implies that, as expected, the Spanish population does not realize all vowel categories with the same F1, F2, intrinsic duration, or intrinsic F0. The directions of the effects are as expected (Sec. I, e.g., shorter duration and higher F0 for high vowels), as can be seen in Table I (also Figs. 3, 4 and 5); for discussion see Sec. IV A.

As for the between-subject effects, there is an effect of gender on F0 ($F[1,35] = 184.474, p = 1.6 \times 10^{-15}$), F1 ($F[1,35] = 150.998, p = 3 \times 10^{-14}$), and F2 ($F[1,35] = 160.255, p = 1.3 \times 10^{-14}$), implying that, as expected, vowels produced by Spanish-speaking men differ from vowels

produced by Spanish-speaking women in F0, F1, and F2. The directions of the effects are as expected (higher values for women), as can be seen in Table I (also Figs. 3 and 5). The main effect of gender on duration, with women having longer vowels (Fig. 4), is only marginally significant ($F[1,35] = 3.659, p = 0.064$), which could be due to the low power of the test for duration (caused by the large between-speaker standard deviation for duration that is visible in Table I).

The analysis reveals no main effect of dialect on any of the four measures. However, it does reveal an interaction of vowel category and dialect, which is significant for duration ($F[4\epsilon, 140\epsilon, \epsilon = 0.910] = 2.619; p = 0.043$), for F1 ($F[4\epsilon, 140\epsilon, \epsilon = 0.642] = 3.519; p = 0.024$), and for F2 ($F[4\epsilon, 140\epsilon, \epsilon = 0.801] = 5.394; p = 0.001$). This implies that the duration, F1 and F2 of *some* vowels produced by IS speakers do differ from those of the same vowels produced by PS speakers, and that these differences are not the same for all vowels.

In order to assess which vowels have a different duration, F1 and/or F2 in IS than in PS, we carried out multivariate tests of variance with the duration, F1 and F2 of each vowel category as the dependent variables and gender and dialect as the fixed factors. The analysis reveals a main effect of dialect on the F1 of /a/ ($F[1,35] = 13.900; p = 6.8 \times 10^{-4}$): the vowel /a/ has a higher F1 in IS than in PS by a factor of 1.063 [95% confidence interval (c.i.) = 1.028–1.099]. With respect to the F2, there is a significant effect of dialect for the two mid vowels /o/ ($F[1,35] = 6.466; p = 0.016$) and /e/ ($F[1,35] = 4.781; p = 0.036$); as for the direction of the effect, a comparison of the means shows that these two vowels must be more peripheral in PS than in IS. More specifically, /o/ has a lower F2 in PS than in IS by a factor of 1.048 (95% c.i. = 1.009–1.088), and /e/ has a higher F2 in PS than in IS by a factor of 1.041 (c.i. = 1.003–1.081). Figure 3 allows for visual comparison

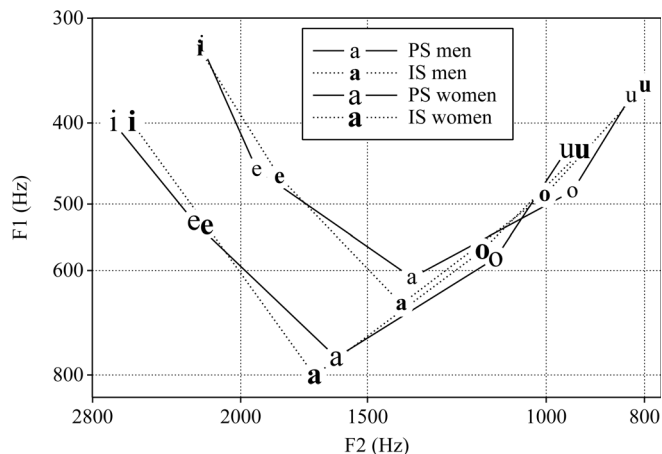


FIG. 3. The vowel spaces of the four groups, collapsed over all contexts. The symbols represent averages of F1 and F2 over the ten (or nine) speakers per group (the same values as in Table I).

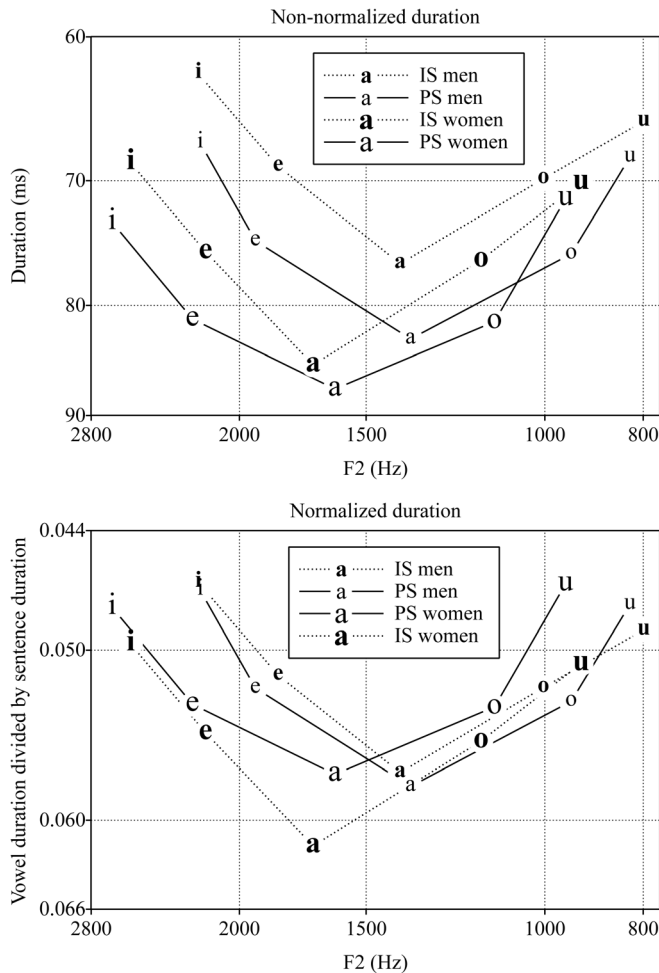


FIG. 4. Top: Average non-normalized duration (values from Table I) as a function of vowel category (expressed by F2). Bottom: Normalized duration (vowel duration divided by each speaker's average sentence duration) as a function of vowel category.

of the male and female vowel spaces in the two dialects: for each vowel category, the figure plots the median F1 and F2 values of Figs. 1 and 2, averaged over the ten (nine) speakers per group. It can be seen that, in line with the statistical results, the mid vowels are slightly more peripheral in PS and the low vowel is higher in PS than in IS, both for women and men. Figures 4 and 5 show plots of duration and F0, respectively.

In the multivariate test for duration, none of the five vowels is shown to be significantly different between the two dialects at the $\alpha=0.05$ level: the most reliable effects of dialect on vowel duration occur for /e/ ($F[1,35]=3.328$; $p=0.077$) and /o/ ($F[1,35]=2.892$; $p=0.098$), which appear to be longer in PS than in IS by a factor of 1.079 (c.i. = 0.991–1.175) and 1.076 (c.i. = 0.986–1.175), respectively. Figure 4 (top) shows a plot of duration as a function of vowel category. Despite the nonsignificant results of the statistical analyses, inspection of the data in the figure and in Table I suggests that all five vowels tend to be slightly longer in PS than in IS, for both genders. If this phenomenon is real (after all, statistical non-significance does not disprove its existence, and the power of the test for duration was low), it may be caused by PS speakers speaking slower than IS

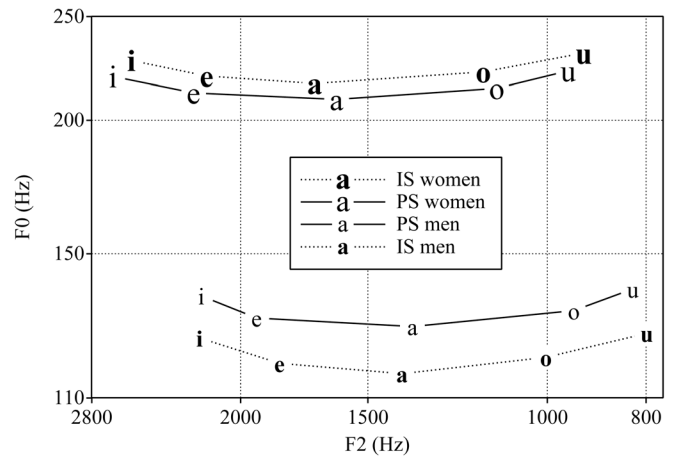


FIG. 5. Average F0 (values from Table I) as a function of vowel category.

speakers. To find out whether any differences in vowel duration are indeed due to differences in global speaking rate, we carried out a univariate analysis of variance on the 39 median sentence durations (i.e. each speaker's median over all vowels, contexts, and replications) with dialect and gender as the fixed factors. The analysis reveals a main effect of dialect ($F[1,35]=15.388$; $p=0.0004$): PS sentences are longer than IS sentences by a factor of 1.088 (c.i. = 1.042–1.137), that is, PS speakers speak slower than IS speakers by a factor of 1.088. The analysis also reveals a main effect of gender on sentence duration ($F[1,35]=5.041$; $p=0.031$): women speak slower than men by a factor of 1.049 (c.i. = 1.005–1.096). No interaction of dialect and gender is seen.

Since we found significant between-dialect and between-gender differences in speaking rate (in terms of sentence duration), we redid the above analysis on vowel duration normalized for global speaking rate, i.e., each vowel duration was divided by the speaker's median sentence duration (Fig. 4, bottom). Recall that the multivariate analysis on non-normalized vowel durations reported in the previous paragraph found a significant effect of dialect at the $\alpha=0.1$ level for two vowels. A similar multivariate analysis on *normalized* vowel durations reveals no effect of dialect (this time the p -values for /e/ and /o/ are above 0.7). The normalization corresponds to a speaker-constant shift in the logarithmic domain, and therefore does not affect the results for the main and interaction effects of vowel category on duration in the repeated-measures analyses reported above. The same independence from normalization goes for most statistical tests on duration in Secs. III B–III D, namely, whenever the factors analyzed involve within-speaker factors (i.e., phrasal context, consonantal context, vowel category). In the following sections, we consider the available phrasal and consonantal contexts to examine whether any dialectal differences have been obscured by collapsing the various contexts, or whether some of the differences that we have located so far apply only in some of the contexts.

B. Phrasal context

The first context-specific analysis was aimed at comparing the effect that the phrasal context alone has on vowel

properties in the two dialects. We performed an exploratory repeated-measures analysis of variance on 390 logarithmic values of each of the four variables, i.e., duration, F0, F1, and F2 in either of the two phrasal contexts (word and sentence). That is, we collapsed the five consonantal contexts, so that 195 values were medians over 10 tokens per vowel category that every speaker produced in the word context, and the other 195 values were medians over 20 tokens per vowel category that every speaker produced in the sentence context. In these repeated-measures analyses, vowel category and phrase were the within-subject factors with five and two levels respectively, and gender and dialect were the between-subject factors.

Apart from the large main effect of vowel category (i.e. similar to the robust effect reported in Sec. III A), these exploratory analyses reveal a main effect of phrasal context on all four variables, i.e., (normalized or non-normalized) duration ($F[1,35] = 209.867$; $p = 2.4 \times 10^{-16}$), F0 ($F[1,35] = 6.300$; $p = 0.017$), F1 ($F[1,35] = 40.975$; $p = 2.3 \times 10^{-7}$), and F2 ($F[1,35] = 4.407$; $p = 0.043$): vowels in word context tend to be longer, have a lower F0, a higher F1, and a lower F2 than vowels in sentence context. In addition, the interaction between vowel category and phrasal context is significant for all four variables, i.e., (normalized or non-normalized) duration ($F[4\epsilon, 140\epsilon, \epsilon = 0.977] = 4.374$; $p = 0.003$), F0 ($F[4\epsilon, 140\epsilon, \epsilon = 0.449] = 4.203$; $p = 0.023$), F1 ($F[4\epsilon, 140\epsilon, \epsilon = 0.891] = 20.217$; $p = 5.8 \times 10^{-12}$), and F2 ($F[4\epsilon, 140\epsilon, \epsilon = 0.798] = 3.698$; $p = 0.012$), implying that different vowels are affected by the phrasal context differently: the effect on F1 (higher in word context) is largest for /a/ and appears to be smallest for /i/ and /u/, and in word context, the F2 of /i/ and /e/ appear to be higher and the F2 of /a/, /o/, and /u/ appear to be lower than in sentence context.

The repeated-measures analysis further reveals a significant interaction effect of phrasal context and gender on (normalized or non-normalized) vowel duration ($F[1,35] = 7.169$; $p = 0.011$): the effect of phrasal context on vowel duration reported above is smaller for men than for women. As for dialectal differences with respect to the effect of phrasal context, the analyses reveal no significant interaction effects on any of the four measures.

C. Consonantal context

The next context-specific analysis was aimed at comparing the effect that the consonantal context alone has on vowel properties in the two dialects. That is, we collapsed the two phrasal contexts and explored whether the factor “consonant” has any effect on vowel duration, F0, F1, or F2. We ran an exploratory repeated-measures analysis of variance on 975 logarithmic values of each of the four variables, i.e., duration, F0, F1, and F2, with the two phrasal contexts collapsed. That is, each of the 975 values is a median over the 6 tokens that each speaker produced for each vowel category in each of the five consonantal contexts, namely, /p/, /t/, /k/, /f/, /s/. In these repeated-measures analyses, vowel category and consonant are the within-subject factors with five levels each, and gender and dialect are the between-subject factors.

As in the context-independent results in Sec. III A, there is again a main effect of vowel category. Now, however, there is also a main effect of consonant on all four measures: i.e., (normalized or non-normalized) duration ($F[4\epsilon, 140\epsilon, \epsilon = 0.907] = 5.812$; $p = 7 \times 10^{-11}$), F0 ($F[4\epsilon, 140\epsilon, \epsilon = 0.951] = 9.556$; $p = 2.6 \times 10^{-20}$), F1 ($F[4\epsilon, 140\epsilon, \epsilon = 0.941] = 25.906$; $p = 5.6 \times 10^{-54}$), and F2 ($F[4\epsilon, 140\epsilon, \epsilon = 0.779] = 271.945$; $p = 1.1 \times 10^{-196}$). As expected (Sec. I), consonantal environment has an impact on the acoustic properties of vowels. The analysis also reveals a significant interaction of vowel category and consonant on (normalized or non-normalized) duration ($F[16,560] = 11.041$; $p = 6.7 \times 10^{-25}$), F1 ($F[16\epsilon, 560\epsilon, \epsilon = 0.791] = 9.902$; $p = 6 \times 10^{-18}$), and F2 ($F[16\epsilon, 560\epsilon, \epsilon = 0.662] = 191.094$; $p = 5 \times 10^{-143}$), which means that different consonants affect different vowels differently. Similarly to the analyses in the previous section that found an interaction effect of phrasal context and gender on duration, the repeated-measures analysis of the present section reveals a significant interaction effect of consonant and gender on (normalized or non-normalized) vowel duration ($F[4\epsilon, 140\epsilon, \epsilon = 0.941] = 5.571$; $p = 4.8 \times 10^{-4}$): a multivariate analysis with 25 dependent variables (the duration of each vowel in each consonant context) reveals that only /i/ in the /s/ context has a significantly different duration in female speakers than in male speakers ($F[1,35] = 5.168$; $p = 0.029$); for female speakers it is longer than for male speakers by a factor of 1.104 (c.i. = 1.011–1.207). The same is seen (with marginal significance) for /e/ in the /s/ context ($F[1,35] = 4.115$; $p = 0.050$) with a factor of 1.093 (c.i. = 1.000–1.196); and for /a/ in the /f/ context ($F[1,35] = 4.062$; $p = 0.052$) with a factor of 1.080 (c.i. = 1.000–1.166).

As for dialect differences in the effect of consonantal context, this analysis uncovers a significant interaction effect of consonant and dialect on two measures, namely, (normalized or non-normalized) duration ($F[4\epsilon, 140\epsilon, \epsilon = 0.907] = 3.561$; $p = 0.011$) and F2 ($F[4\epsilon, 140\epsilon, \epsilon = 0.779] = 5.227$; $p = 0.002$). Apparently, some consonants affect vowel duration and F2 differently in IS than in PS. A significant triple interaction of vowel category, consonant and dialect is seen for F2 ($F[16\epsilon, 560\epsilon, \epsilon = 0.662] = 2.977$; $p = 9.9 \times 10^{-4}$), which suggests that this dialectal consonant-specific effect on F2 applies only to some of the five vowel categories.

Since the exploratory analysis detected several dialect-involving interactions on F2 and normalized duration, we ran a multivariate analysis with the normalized duration and the F2 of every vowel produced in every context as the dependent variables. The aim of this subsequent analysis was to assess where exactly the significant dialectal difference lies and what size and direction it has. The multivariate analysis reveals that three vowels have a reliably different F2 in PS than in IS in the consonantal context of /s/, namely, /o/ ($F[1,35] = 14.878$; $p = 4.7 \times 10^{-4}$), /u/ ($F[1,35] = 4.347$; $p = 0.044$), and /i/ ($F[1,35] = 6.902$; $p = 0.013$); this difference approaches significance for /a/ ($F[1,35] = 3.880$; $p = 0.057$). As for the size of this effect, pairwise comparisons show that in the context of /s/, /o/ has a lower F2 in PS than in IS by a factor of 1.122 (c.i. = 1.057–1.194), /u/ has a lower F2 in PS than in IS by a factor of 1.076 (c.i. = 1.002–1.156), and /i/ has a higher F2 in PS than in IS by a factor of 1.050 (c.i. = 1.011–1.090).

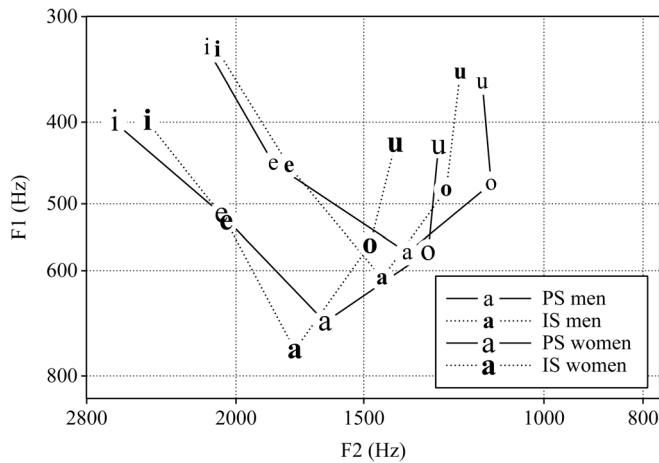


FIG. 6. The vowel spaces of the four groups, in the consonantal context of /s/. The symbols represent averages over the ten (or nine) speakers per group.

The vowel /e/ exhibits a between-dialect difference in F2 in two consonantal contexts, namely, that of /p/ ($F[1,35] = 4.888$; $p = 0.034$) and /t/ ($F[1,35] = 5.007$; $p = 0.032$); pairwise comparisons show that the vowel /e/ has a higher F2 in PS than in IS by a factor of 1.039 in the /t/-context (c.i. = 1.004–1.075) and 1.044 in the /p/-context (c.i. = 1.003–1.085). Figure 6 shows a plot of the vowel spaces in the consonantal context of /s/. In line with the significant effects, the figure shows that /o/, /u/, and /i/ have more peripheral F2 values in PS than in IS. This effect and its possible cause are discussed in Sec. IV C.

For normalized duration, significant cross-dialectal differences are seen only for the vowel /u/ in the /t/-context ($F[1,35] = 4.622$; $p = 0.039$); it is longer in IS than in PS by a factor of 1.116 (c.i. = 1.006–1.240).

To investigate whether the dialectal differences in the /s/ context could be due to a difference in the articulation of /s/ between PS and IS, we measured the spectral center of gravity (CoG) of /s/. This is because CoG in sibilants has been shown to correlate with place of articulation: sibilants articulated further back in the mouth have a lower CoG than sibilants articulated further front in the mouth (Gordon *et al.*, 2002; Zygis and Hamann, 2003). We analyzed the spectral center of gravity (with Praat, using a power of 1 to weigh the amplitude of each frequency) of all 2340 /s/ tokens (39 speakers \times 5 vowels \times 3 words \times 2 positions in the word \times 2 replications) and tabulated each speaker's median center of gravity for each vowel. Figure 7 shows the center of gravity as a function of dialect, speaker and vowel. A repeated-measures analysis of variance with dialect and gender as between-subject factors and vowel as the within-subject factor reveals a main effect of gender ($F[1,35] = 43.466$; $p = 1.3 \times 10^{-7}$) and dialect ($F[1,35] = 11.257$; $p = 0.0019$): women have a higher spectral center of gravity for /s/ than men by a factor of 1.210 (c.i. = 1.141–1.284), and PS /s/ has a higher center of gravity than IS /s/ by a factor of 1.102 (c.i. = 1.039–1.169). No interaction of dialect and gender is seen ($F < 1$). We observe a main effect of vowel ($F[4\epsilon, 140\epsilon, \epsilon = 0.702] = 84.522$; $p = 2.9 \times 10^{-26}$) but no interactions between vowel and dialect or gender. Pair comparisons show that the center of gravity of /s/ next to /o/ and /u/ is different from that next

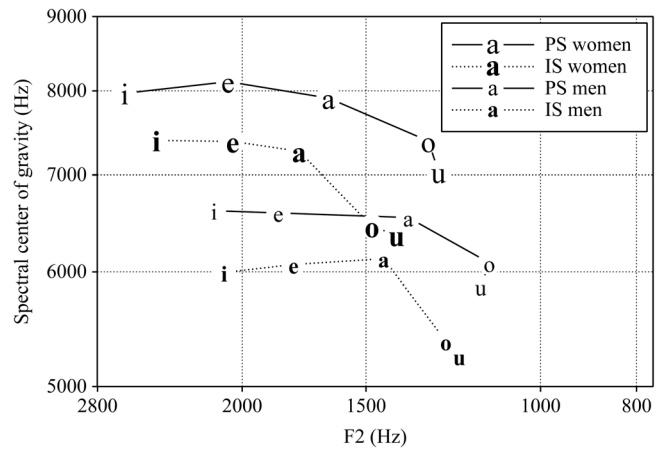


FIG. 7. Average center of gravity of /s/ of the four groups in /sVs/ context, averaged over the prevocalic and postvocalic position and plotted as a function of the vowel category V.

to the other three vowels, and the /o/ and /u/ context also differ from each other (all uncorrected $p < 0.003$).

D. Phrasal and consonantal context

The third and final context-specific analysis was aimed at exploring whether the phrasal and the consonantal context together have dialect-specific effects on vowel properties. We performed a repeated-measures analysis of variance on 1950 logarithmic values of each of the four variables, i.e., duration, F0, F1, and F2. In this analysis, vowel (five levels), phrase (two levels) and consonant (five levels) are the within-subject factors and gender and dialect are the between-subject factors. The analysis reveals a significant interaction effect of phrasal and consonantal context on (normalized or non-normalized) vowel duration ($F[4\epsilon, 140\epsilon, \epsilon = 0.960] = 6.496$; $p = 1.1 \times 10^{-4}$) and F2 ($F[4\epsilon, 140\epsilon, \epsilon = 0.793] = 3.744$; $p = 0.012$), implying that the effect that consonantal context has on vowel duration and F2 differs across phrasal contexts, or that the effect phrasal context has on vowel duration and F2 differs across consonantal contexts, or both. The repeated-measures analyses do not reveal any significant triple or higher interactions involving phrasal context, consonantal context, and the between-subject factors (dialect and/or gender).

IV. DISCUSSION

A. Universal effects

Spanish turns out to exhibit several (near-)universal phenomena that have been attested for other languages

1. Vowel-intrinsic duration and F0

Higher vowels are shorter than lower vowels (Sec. III A, Fig. 4), which corresponds to previous findings for languages that, like Spanish, have no phonological length contrast in vowels, such as French (Rochet and Rochet, 1991), Italian (Esposito, 2002), and Portuguese (Escudero *et al.*, 2009), but also for languages with a binary vowel length contrast such as Swedish (Lindblom, 1967) and Czech (Machač and

Skarnitzl, 2007), where height-dependent duration appears within each length class separately, and even for the “tense” vowels of American English (House and Fairbanks, 1953; House, 1961; Delattre, 1962).

Higher vowels have a higher F0 than lower vowels (Sec. III A, Fig. 5), which corresponds to previous findings for American English (Shadle, 1985), Dutch (Koopmans-van Beinum, 1980), German (Ladd and Silverman, 1984), Italian (Esposito, 2002), Portuguese (Escudero *et al.*, 2009) and a large number of other languages (for a vast survey see Whalen and Levitt, 1995).

2. Gender effects

As entirely expected on the basis of physiology, Spanish-speaking women have higher F0, F1 and F2 than men (Sec. III A). As for duration, women’s *sentences* are reliably longer than men’s by 5% (Sec. III A). Women’s *vowels* were also measured as longer than men’s (by 8%; Fig. 4 top), and although this effect did not reach statistical significance and therefore cannot be reliably generalized to the Spanish-speaking population, the sentence finding and the vowel measurement are compatible with each other and could be causally related.

3. Phrase effects

Vowels are longer in isolated words than in sentences (Sec. III B), which corresponds to what has been found before for stress-timed languages such as English (Lehiste, 1972) and Dutch (Nooteboom, 1972), although Spanish has been described as syllable-timed (Harris, 1969, p. 33) and therefore to exhibit a relative independence of vowel duration from phrase length (see Sec. I); a similar ambiguity as to whether Spanish is properly syllable-timed, in contrast with Japanese, was observed before by Hoequist (1983).

Our findings that in isolated words front vowels have a higher F2, back vowels have a lower F2, and especially /a/ has a higher F1 (Sec. III B) point to the idea that the vowel space is larger in isolated words than in sentence context. Similar findings have been reported for British English (Shearme and Holmes, 1962), Swedish (Lindblom, 1963), and German and French (Gendrot and Adda-Decker, 2005).

B. Spanish-specific findings

There are several respects in which Spanish may be different from other languages. The locations of the five vowels in F1-F2 space (Fig. 3) are Spanish-specific: they are different from those of other languages with five monophthongs, such as Japanese, which has an (articulatorily unrounded and therefore) acoustically frontier /u/, traditionally transcribed as /u/ (Keating and Huffman, 1984), and the short-vowel system of Czech, which has a lower /e/, traditionally transcribed as /e/ (Dankovičová, 1997). Even if one assumes that vowel systems are maximally dispersed (Liljencrants and Lindblom, 1972), such differences are expected on the basis of differences in consonant inventories, differences in auditory cue weighting (Escudero and Boersma, 2004), differences in the featural build-up of the vowels (Boersma and Chládková, 2011), and the differential existence of diphthongs.

Another possible Spanish-specific finding is that F0 is higher in sentence context than in isolated words. This is because our finding is the opposite of what Picheny *et al.* (1986) found for three English speakers who had a higher F0 in clear than in conversational mode. However, these two opposite findings can only be directly compared if one assumes that our isolated words can be associated with a “clear” speaking mode and our sentence context can be associated with a “conversational” speaking mode. Our methodology did not rigorously control for speaking mode because our speakers could freely choose, for instance, between declarative and continuation intonation for the isolated words. Therefore, Spanish may not differ from English in the effect of phrasal context on F0, if data collection methods are held similar across studies.

With respect to consonantal context, the gender × consonant interaction reported in Sec. III C suggests that the voiceless fricatives /f/ and /s/ cause longer vowels in women than in men, even after normalization. Vowel lengthening before fricatives (or shortening before plosives) is a widespread phenomenon (Peterson and Lehiste, 1960). Gender effects on (phonetic) vowel duration differences have been reported before: women often produce larger duration differences between phonologically short and long vowels, and between unstressed and stressed vowels (for a survey see Simpson and Ericsson, 2003).

C. Dialect-specific findings

When collapsing all consonantal and phrasal contexts, the analyses of Sec. III A found three dialect-specific effects. The first effect is that (our specific) sentences tend to be longer in PS than in IS, by 8.8%. We cannot tell whether this is a result of a difference in general speaking rate, or of a dialect difference in the phrasing of these specific sentences.

The second effect is that /a/ has a higher F1 in IS than in PS by 6.3%. This implies that /a/ is realized toward [a] in IS and toward [ɐ] in PS. Given that Secs. III B–III D do not reveal any significant interaction effects involving dialect and context on the F1 of /a/ (or any of the other vowels), the higher /a/ in PS as compared to IS may well be a robust effect found across all consonantal and phrasal contexts.

The third context-independent dialect-specific effect seen in Sec. III A is that the two mid vowels, /e/ and /o/, have more peripheral F2 values in PS than in IS: compared to IS, PS /e/ has a higher F2 by 4.1% and PS /o/ has a lower F2 by 4.8%. In the context of /s/ (Sec. III C), the dialectal difference is slightly more pronounced (8%) and found for a different set of vowels (for /o/, /u/, /i/, and possibly /a/). This dialect-specific effect of /s/ on the F2 of vowels can be attributed to the fact that the place of articulation of the fricative /s/ is different in IS than in PS, and articulatory differences in consonants tend to lead to acoustic differences in the neighboring vowels (see Sec. I). In that respect, Gordon *et al.* (2002) found that fricatives affect both the transition and the mid-point formant values of adjacent vowels. Specifically, in Toda, a Dravidian language spoken in India, retraction in sibilant fricatives leads to lower F3 values at the midpoint of the adjacent vowels as well as to lower F3 and higher F2 transitions to the adjacent

vowels (pp. 165–166). The lower spectral center of gravity of /s/ in IS than in PS reflects its retracted articulation in IS and its dental articulation in PS (see Secs. I and III C). The lower center of gravity for men than for women reflects a perhaps universal gender effect that the literature reports as small and somewhat unstable (English: Jongman *et al.*, 2000; Chickasaw: Gordon *et al.*, 2002; Polish: Zygis and Hamann, 2003) although we find a very strong effect for Spanish (c.i. = 1.141–1.284); the lower center of gravity near /o/ and especially /u/ probably reflects the coarticulation of lip rounding, as was also shown by Mann and Repp (1980, p. 224) and Shadle and Scully (1995). Conversely, the different articulations of /s/ cause the neighboring vowels to behave differently in IS than in PS: given the dialectal differences that we find for vowels in the context of /s/, we conclude that the back vowels /o/ and /u/ are fronted (i.e., [o] and [u]) and the front vowel /i/ is slightly retracted (i.e., [i]) in IS before /s/. Apparently, the well-known finding that back vowels are more centralized after alveolars (Stevens and House, 1963, Fig. 6; Hillenbrand *et al.*, 2001, Fig. 7; Strange *et al.*, 2007, Figs. 3 and 4) is stronger for the apico-alveolar /s/ of Madrid than for the dental /s/ of Lima.

D. Comparison to isolated vowels

Morrison and Escudero (2007) analyzed a complementary data set, namely, the sentence-final isolated vowels of the present recordings. This allows us to compare the situation for vowels in consonantal context (the present study) with the situation for isolated vowels (Morrison and Escudero's study) by studying the degree of overlap of the confidence intervals from the two studies. Regarding duration, Morrison and Escudero found that isolated vowels are shorter in IS than in PS by a factor of 1.339. This difference is also visible in our Fig. 4 for vowels in consonantal context, but is not reliable ($p = 0.155$). Our confidence interval (0.976–1.155) is far below Morrison and Escudero's value (1.339), so if their confidence interval (which they did not report) is comparable to ours, the two confidence intervals do not overlap and we can conclude that the dialect effect is greater for isolated vowels than for vowels in consonantal context.

As for F0, Morrison and Escudero found that isolated vowels have a lower F0 in IS than in PS by a factor of 1.088. The direction and size of the effect of dialect on the F0 of vowels in consonantal context in the present study (Fig. 5), though not reliably different from 1 (c.i. = 0.950–1.119), are compatible with Morrison and Escudero's finding, because their value (1.088) is contained within our confidence interval. With regard to formants, Morrison and Escudero found that isolated vowels have a higher F2 for /o/ in IS than in PS by a factor of 1.108. The direction and size of this effect also compare well with our finding in consonantal context (1.048), because Morrison and Escudero's value is only just outside our confidence interval (1.009–1.088), so if their value (1.108) has a comparable confidence interval, the two confidence intervals overlap to a large degree. Finally, Morrison and Escudero found no difference for the F1 of /a/, while we find a robust context-independent dialectal difference.

V. CONCLUDING REMARKS

The present findings have a number of implications for future cross-dialectal, cross-linguistic, and second-language research with participants from various Spanish-speaking countries. First, given that /a/ has a markedly lower F1 value in PS than in IS (a finding probably not noted in the diachronic or comparative literature), studies investigating cross-language or second-language production and perception with speakers or listeners from various Spanish-speaking countries should relate their results to the participants' background. Many studies with Spanish learners of English, e.g., Flege *et al.* (1997), Escudero and Boersma (2004), and Morrison (2006), did not control for dialectal variation and pooled the data of participants from various Latin American countries and various regions in Spain. Such pooling may not be warranted.

Second, since the /s/-context affects vowels differently in PS than in IS, and this difference can be related to an articulatory and acoustic difference in /s/, one should take into account the possibility that dialectal differences between vowels can be caused by dialectal differences between consonants. Madrid Spanish contrasts an apico-alveolar sibilant [s] with a dental fricative [θ], whereas Lima Spanish only has a dental sibilant [s], which is articulatorily and acoustically in between the two Madrid sounds. We would therefore have been equally justified in using in our experiment the written ⟨z⟩, which is pronounced [θ] in Madrid and [s] in Lima, instead of the written ⟨s⟩, which is pronounced as [s] in Madrid and [s] in Lima.³ This problem stresses the difficulty of identifying phonemes across languages.

Finally, although the dialectal difference in F2 values for /e/ and /o/ is rather small (< 5%), future perception studies will have to show whether native IS and PS listeners perceive these cross-dialectal acoustic differences.

¹We also ignore the possible high-low allophony of the mid vowels /e/ and /o/, which was reported by Navarro Tomás (1932) but could not be confirmed by further research (Martínez-Celdrán, 1984; Morrison, 2004). In any case, our data set would include only what Navarro Tomás identified as the higher allophones.

²In order to follow the methodology of data collection reported in Escudero *et al.* (2009), the context used for /t/ was /tVkV/.

³This ambiguity is analogous to the question of the direction of the historical merger in Latin America, i.e., whether the pronunciation of ⟨z⟩ became that of ⟨s⟩ (the phenomenon of “seseo”) or the pronunciation of ⟨s⟩ became that of ⟨z⟩ (the phenomenon of “ceceo”) (Canfield, 1962; 1981, p. 4; Obaid, 1973).

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.

Albalá, M. J., Battaner, E., Carranza, M., Gil, J., Llisteri, J., Machuca, M. J., Madrigal, N., Marquina, M., Marrero, V., De la Mota, C., Riera, M., and Rios, A. (2008). “VILE: Nuevos datos acústicos sobre vocales del español (VILE: New acoustic data on the vowels of Spanish),” *Language Design: Journal of Experimental and Theoretical Linguistics*, Special Issue 2: New Trends in Experimental Phonetics: Selected Papers From the IV International Conference on Experimental Phonetics, Granada, 11–14 February, Vol. 1, pp. 1–14, available at http://liceu.uab.es/~joaquim/phonetics/VILE/VILE_IVCFE08_GrupoEntonacion.pdf (last viewed 4/27/11).

Allan, L. G., and Gibbon, J. (1991). “Human bisection at the geometric mean,” *Learning Motivation* **22**, 39–58.

Anderson, N. (1978). “On the calculation of filter coefficients for maximum entropy spectral analysis,” in *Modern Spectrum Analysis*, edited by D.G. Childers (IEEE Press, New York), pp. 252–255.

- Beckman, M., and Shoji, A. (1984). "Spectral and perceptual evidence for CV coarticulation in devoiced /s/ and /syu/ in Japanese," *Phonetica* **41**, 61–71.
- Boersma, P., and Chládková, K. (2011). "Asymmetries between speech perception and production reveal phonological structure," in Proceedings of the 17th Congress of Phonetic Sciences, Hong Kong.
- Boersma, P., and Weenink, D. (2009). "Praat: doing phonetics by computer (version 5.1.07)" (computer program), <http://www.praat.org/> (last viewed 5/12/09).
- Bradlow, A. R. (1995). "A comparative acoustic study of English and Spanish vowels," *J. Acoust. Soc. Am.* **97**, 1916–1924.
- Bradlow, A. R. (1996). "A perceptual comparison of the /i/-/e/ and /u/-/o/ contrasts in English and in Spanish: Universal and language specific aspects," *Phonetica* **53**, 55–85.
- Canfield, D. L. (1962). *La pronunciación del español en América: ensayo histórico-descriptivo (The Pronunciation of Spanish in America: Historical-Descriptive Essay)* (Instituto Caro y Cuervo, Bogotá).
- Canfield, D. L. (1981). *Spanish Pronunciation in the Americas* (University of Chicago Press, Chicago).
- 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.
- Chen, M. (1970). "Vowel length variation as a function of the voicing of the consonant environment," *Phonetica* **22**, 129–159.
- Chládková, K., Boersma, P., and Podlipský, V. J. (2009). "On-line formant shifting as a function of F0," in Proceedings of Interspeech 2009, pp. 464–467.
- Crystal, T. H., and House, A. S. (1988). "Segmental durations in connected-speech signals: Current results," *J. Acoust. Soc. Am.* **83**, 1553–1573.
- Dankovičová, J. (1997). "Illustrations of the IPA: Czech," *J. Int. Phonetic Assoc.* **27**, 77–80.
- Delattre, P. (1962). "Some factors of vowel duration and their cross-linguistic validity," *J. Acoust. Soc. Am.* **34**, 1141–1143.
- Deterding, D. (1997). "The formants of monophthong vowels in Standard Southern British English pronunciation," *J. Int. Phonetic Assoc.* **27**, 47–55.
- Dinnsen, D. A., and Charles-Luce, J. (1984). "Phonological neutralization, phonetic implementation and individual differences," *J. Phonetics* **12**, 49–60.
- Disner, S. F. (1983). "Vowel quality: The relation between universal and language specific factors," Ph.D. thesis, UCLA, Los Angeles, available at <http://escholarship.org/uc/item/1wm9n05g> (date last viewed 4/27/11).
- Escudero, P., and Boersma, P. (2004). "Bridging the gap between L2 speech perception research and phonological theory," *Stud. Second Lang. Acquis.* **26**, 551–585.
- Escudero, P., Boersma, P., Rauber, A., and Bion, R. (2009). "A cross-dialect acoustic description of vowels: Brazilian versus European Portuguese," *J. Acoust. Soc. Am.* **126**, 1379–1393.
- Esposito, A. (2002). "On vowel height and consonantal voicing effects: Data from Italian," *Phonetica* **59**, 197–231.
- Farnetani, E., and Recasens, D. (1993). "Anticipatory consonant-to-vowel coarticulation in the production of VCV sequences in Italian," *Lang. Speech* **36**, 279–302.
- Flège, J. E., Bohn, O.-S., and Jang, S. (1997). "Effects of experience on non-native speakers' production and perception of English vowels," *J. Phonetics* **25**, 437–470.
- Fourakis, M. (1991). "Tempo, stress, and vowel reduction in American English," *J. Acoust. Soc. Am.* **90**, 1816–1827.
- Fowler, C. A., and Saltzman, E. (1993). "Coordination and coarticulation in speech production," *Lang. Speech* **36**, 171–195.
- Gay, T. (1968). "Effect of speaking rate on diphthong formant movements," *J. Acoust. Soc. Am.* **44**, 1570–1573.
- Gay, T. (1978). "Effect of speaking rate on vowel formant movements," *J. Acoust. Soc. Am.* **63**, 223–230.
- Gendrot, C., and Adda-Decker, M. (2005). "Impact of duration on F1/F2 formant values of oral vowels: An automatic analysis of large broadcast news corpora in French and German," in Proceedings of Interspeech, Lisbon, pp. 2453–2456.
- Gibbon, J. (1977). "Scalar expectancy theory and Weber's Law in animal timing," *Psych. Rev.* **84**, 279–325.
- Godínez, M., Jr. (1978). "A comparative study of some Romance vowels," in UCLA Working Papers in Phonetics **41**, 3–19, available at <http://escholarship.org/uc/item/5fb7f1kq?page=2> (last viewed 4/27/11).
- Gordon, M., Barthmaier, P., and Sands, K. (2002). "A cross-linguistic acoustic study of voiceless fricatives," *J. Int. Phonetic Assoc.* **32**, 141–174.
- Graddol, D. (1986). "Discourse specific pitch behaviour," in *Intonation in Discourse*, edited by C. Johns-Lewis (College-Hill, San Diego, CA), pp. 221–237.
- Guirao, M., and Borzone de Manrique, A. M. (1975). "Identification of Argentine Spanish vowels," *J. Psycholinguistic Res.* **4**, 17–25.
- Harris, J. (1969). *Spanish Phonology* (MIT, Cambridge, MA).
- Henton, C. G. (1989). "Fact and fiction in the description of female and male pitch," *Lang. Commun.* **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.
- Hillenbrand, J. M., Clark, M. J., and Nearey, T. M. (2001). "Effects of consonant environment on vowel formant patterns," *J. Acoust. Soc. Am.* **109**, 748–763.
- Hirata, Y. (2004). "Effect of speaking rate on the vowel length distinction in Japanese," *J. Phonetics* **32**, 565–589.
- Hoequist, C. (1983). "Syllable duration in stress-, syllable- and mora-timed languages," *Phonetica* **40**, 203–237.
- 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.
- House, A. S. (1961). "On vowel duration in English," *J. Acoust. Soc. Am.* **33**, 1174–1178.
- Hudson, A. I., and Holbrook, A. (1981). "A study of the reading fundamental vocal frequency of young adults," *J. Speech Hearing Res.* **24**, 197–201.
- Jongman, A., Fourakis, M., and Sereno, J. A. (1989). "The acoustic vowel space of Modern Greek and German," *Lang. Speech* **32**, 221–248.
- Jongman, A., Wayland, R., and Wong, S. (2000). "Acoustic characteristics of English fricatives," *J. Acoust. Soc. Am.* **108**, 1252–1263.
- Keating, P. A., and Huffman, M. K. (1984). "Vowel variation in Japanese," *Phonetica* **41**, 191–207.
- Klatt, D. H. (1973). "Interaction between two factors that influence vowel duration," *J. Acoust. Soc. Am.* **54**, 1102–1104.
- 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.
- Ladd, D. R., and Silverman, K. E. A. (1984). "Vowel intrinsic pitch in connected speech," *Phonetica* **41**, 31–40.
- Lehiste, I. (1970). *Suprasegmentals* (MIT Press, Cambridge, MA).
- Lehiste, I. (1972). "The timing of utterances and linguistic boundaries," *J. Acoust. Soc. Am.* **51**, 2018–2024.
- Lehiste, I., and Peterson, G. E. (1961). "Some basic considerations in the analysis of intonation," *J. Acoust. Soc. Am.* **33**, 419–425.
- Liljencrants, J., and Lindblom, B. (1972). "Numerical simulation of vowel quality systems: the role of perceptual contrast," *Language* **48**, 839–862.
- Lindblom, B. (1963). "Spectrographic study of vowel reduction," *J. Acoust. Soc. Am.* **35**, 1773–1781.
- Lindblom, B. (1967). "Vowel duration and a model of lip-mandible coordination," *Speech Transmission Laboratory Quarterly Progress and Status Report*, Vol. 8, issue 4, pp. 1–29, available at http://www.speech.kth.se/prod/publications/files/qpsr/1967/1967_8_4_001-029.pdf (last viewed 4/27/11).
- Lindblom, B. (1968). "Temporal organization of syllable production," in *Speech Transmission Laboratory Quarterly Progress and Status Report* Vol. 9, issue 2-3, pp. 1–5, available at http://www.speech.kth.se/prod/publications/files/qpsr/1968/1968_9_2-3_001-005.pdf (last viewed 4/27/11).
- Lubker, J. F., and Gay, T. (1982). "Anticipatory labial coarticulation: Experimental, biological and linguistic variables," *J. Acoust. Soc. Am.* **71**, 437–448.
- Machač, P., and Skamitzl, R. (2007). "Temporal compensation in Czech?," in Proceedings of the 16th Congress of Phonetic Sciences, Saarbrücken, pp. 537–540.
- Mann, V. A., and Repp, B. H. (1980). "Influence of vocalic context on perception of the [j]-[s] distinction," *Perception Psychophys.* **28**, 213–228.
- Marín Gálvez, R. (1994). "La duración vocálica en español (Vowel duration in Spanish)," *Estudios Lingüística. Universidad Alicante* **10**, 213–226.
- Martínez-Celdrán, E. (1984). "Cantidad e intensidad en los sonidos obstruyentes del castellano: Hacia una caracterización acústica de los sonidos aproximantes (Quantity and intensity in Castilian obstruents: Towards an acoustic characterization of approximants)," *Estudios de Fonética Experimental* Vol. 1, pp. 73–129, available at <http://www.raco.cat/index.php/EFE/article/viewFile/144191/264165> (last viewed 4/27/11).
- Martínez-Celdrán, E., Fernández-Planas, A. M., and Carrera-Sabaté, J. (2003). "Illustrations of the IPA: Castilian Spanish," *J. Int. Phonetic Assoc.* **33**, 255–259.

- Mendoza, E. Carballo, G. Cruz, A. Fresneda, M. D. Muñoz, J., and Marrero, V. (2003). "Temporal variability in speech segments of Spanish: context and speaker related differences," *Speech Comm.* **40**, 431–447.
- Morrison, G.S. (2004). "An acoustic and statistical analysis of Spanish mid-vowel allophones," *Estudios de Fonética Experimental* Vol. 13, pp. 11–37, available at <http://www.ub.edu/labfon/XIII-5.pdf> (last viewed 4/27/11).
- Morrison, G. S. (2006). "L1 & L2 production and perception of English and Spanish vowels: A statistical modeling approach," Ph.D. thesis, University of Alberta, Edmonton.
- 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.
- Navarro Tomás, T. (1932). *Manual de pronunciación española (Manual of Spanish Pronunciation)* (Centro de Estudios Históricos, Madrid).
- Nearey, T. M. (1992). "Applications of generalized linear modeling to vowel data," in *Proceedings of ICSLP-1992*, pp. 583–586.
- Nishi, K., Strange, W., Akahane-Yamada, R., Kubo, R., and Trent-Brown, S. A. (2008). "Acoustic and perceptual similarity of Japanese and American English vowels," *J. Acoust. Soc. Am.* **124**, 576–588.
- Nootboom, S. G. (1972). "Production and perception of vowel duration: A study of durational properties of vowels in Dutch," Ph.D. thesis, Utrecht University.
- Obaid, A. H. (1973). "The vagaries of the Spanish 's,'" *Hispania* **56**, 60–67.
- 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 Lehiste, I. (1960). "Duration of syllable nuclei in English," *J. Acoust. Soc. Am.* **32**, 693–703.
- Picheny, M. A., Durlach, N. I., and Braida, L. D. (1986). "Speaking clearly for the hard of hearing II: Acoustic characteristics of clear and conversational speech," *J. Speech Lang. Hearing Res.* **29**, 434–446.
- Pickett, E. R., Blumstein, S. E., and Burton, M. W. (1999). "Effects of speaking rate on the singleton/geminate consonant contrast in Italian," *Phonetica* **56**, 135–157.
- Podlipský, V. J. (2009). "Reevaluating perceptual cues: Native and non-native perception of Czech vowel quantity," Ph.D. thesis, Palacký University, Olomouc, available at http://www.anglistika.upol.cz/jonas/pub/Podlipsky_2009_-_Reevaluating_perceptual_cues_-_Native_and_non-native_perception_of_Czech_vowel_quantity.pdf (last viewed 4/27/11).
- 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.
- Shadle, C. H. (1985). "Intrinsic fundamental frequency of vowels in sentence context," *J. Acoust. Soc. Am.* **78**, 1562–1567.
- Shadle, C. H., and Scully, C. (1995). "An articulatory-acoustic-aerodynamic analysis of [s] in VCV sequences," *J. Phonetics* **23**, 53–66.
- Shearme, J. N., and Holmes, J. N. (1962). "An experimental study of the classification of sounds in continuous speech according to their distribution in the formant 1–formant 2 plane," in *Proceedings of the Fourth International Congress of Phonetic Sciences*, Helsinki, pp. 234–244.
- Simpson, A. P., and Ericsson, C. (2003). "Sex-specific durational differences in English and Swedish," in *Proceedings of the 15th Congress of Phonetic Sciences*, Barcelona, pp. 1113–1116.
- Slis, I. H., and Cohen, A. (1969). "On the complex regulating the voiced-voiceless distinction I," *Lang. Speech* **12**, 80–102.
- 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. N., and House, A. S. (1963). "Perturbation of vowel articulations by consonantal context: An acoustics study," *J. Speech Hearing Res.* **6**, 111–128.
- Stevens, K. N., House, A. S., and Paul, A. P. (1966). "Acoustical description of syllabic nuclei: An interpretation in terms of a dynamic model of articulation," *J. Acoust. Soc. Am.* **40**, 123–132.
- Strange, W., Bohn, O-S., and Nishi, K. (2005). "Contextual variation in the acoustic and perceptual similarity of North German and American English vowels," *J. Acoust. Soc. Am.* **118**, 1751–1762.
- 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.
- Turner, G. S., Tjaden, K., and Weismer, G. (1995). "The influence of speaking rate on vowel space and speech intelligibility for individuals with amyotrophic lateral sclerosis," *J. Speech Hearing Res.* **38**, 1001–1013.
- Umeda, N. (1975). "Vowel duration in American English," *J. Acoust. Soc. Am.* **58**, 434–435.
- Van Bergem, D. R. (1993). "Acoustic vowel duration as a function of sentence accent, word stress, and word class," *Speech Comm.* **12**, 1–23.
- Van Bergem, D. R. (1994). "A model of coarticulatory effects on the schwa," *Speech Commun.* **14**, 143–162.
- Van Santen, J. P. H. (1992). "Contextual effects on vowel duration," *Speech Commun.* **11**, 513–546.
- Whalen, D. H., and Levitt, A. G. (1995). "The universality of intrinsic F0 of vowels," *J. Phonetics* **23**, 349–366.
- Widdison, K. A. (1997). "Phonetic explanations for sibilant patterns in Spanish," *Lingua* **102**, 253–264.
- Yamazawa, H., and Hollien, H. (1992). "Speaking fundamental frequency patterns of Japanese women," *Phonetica* **49**, 128–140.
- Zygis, M., and Hamann, S. (2003). "Perceptual and acoustic cues of Polish coronal fricatives," in *Proceedings of the 15th Congress of Phonetic Sciences*, Barcelona, pp. 395–398.