
 Received: 29 December 2013, Accepted: 27 May 2014

Edited by: E. Mizraji

Reviewed by: J. Brum, Instituto de Física, Facultad de Ciencias,
Universidad de la República, Montevideo, Uruguay.

Licence: Creative Commons Attribution 3.0

DOI: <http://dx.doi.org/10.4279/PIP.060002>

ISSN 1852-4249

Study of the characteristic parameters of the normal voices of Argentinian speakers

E. V. Bonzi,^{1,2*} G. B. Grad,^{1†} A. M. Maggi,^{3‡} M. R. Muñoz^{3§}

The voice laboratory permits to study the human voices using a method that is objective and noninvasive. In this work, we have studied the parameters of the human voice such as pitch, formant, jitter, shimmer and harmonic-noise ratio of a group of young people. This statistical information of parameters is obtained from Argentinian speakers.

I. Introduction

The voice is a multidimensional phenomenon that must be evaluated using special tools for determining acoustic parameters. These parameters are: the pitch or voice tone, the timbre, considered as the personality of the voice that is particular of each person (determined by fundamental frequency, its harmonics and formants) and the degree of hoarseness.

During sustained vibration, the vocal fold will exhibit variations of fundamental frequency and amplitude; these phenomena are called “frequency perturbation” (jitter) and “amplitude perturbation” (shimmer). They reflect fluctuations in tension and biochemical characteristics of the vocal

folds, as well as variation in their neural control and the physiological properties of the individuals voices.

The acoustic analysis is one of the major advances in the study of voice, increasing the accuracy of diagnosis in this area. Normal values as standards are important and necessary to guide voice professionals.

There are not many studies performed for the Latin languages [1–3]. However, there are several of them for the English language, such as those in Refs. [4–8].

In the same way, the software used for voice therapy is in general designed for other languages than Spanish. A comparison has been made, though, between the two vowel systems of English and Spanish (the variation spoken in Madrid, Spain), which triggered relatively large versus small vowel inventories [9]. That is the reason why we consider it is very important and necessary to produce more results for the Spanish speaking population.

We analyzed 72 audio files of female and male voices from an Argentinian Spanish speaking population to obtain the acoustical parameters using the Praat program [10]. Our data were compared to Bradlow [9], Hualde [11] and Casado Morente *et al.* [12]. The pitches measured were lower than expected and the First formant of the /a/ and /u/

*E-mail: bonzie@famaf.unc.edu.ar

†E-mail: grad@famaf.unc.edu.ar

‡E-mail: alicia.maggi@hotmail.com

§E-mail: eudaimonia13@hotmail.com

¹ Facultad de Matemática, Astronomía y Física, Universidad Nacional de Córdoba, Ciudad Universitaria, 5000 Córdoba, Argentina.

² Instituto de Física Enrique Gaviola (CONICET), Ciudad Universitaria, 5000 Córdoba, Argentina.

³ Escuela de Fonoaudiología, Facultad de Ciencias Médicas, Universidad Nacional de Córdoba, Ciudad Universitaria, 5000 Córdoba, Argentina.

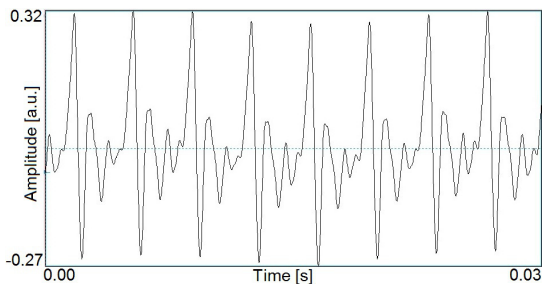


Figure 1: Wave shape of the /a/ sound.

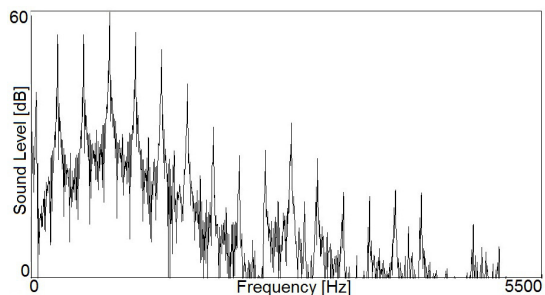


Figure 3: Harmonics of the /a/ vowel.

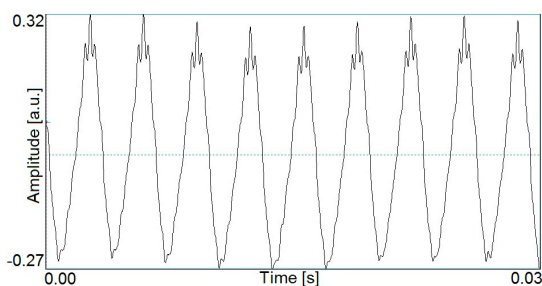


Figure 2: Wave shape of the /i/ sound.

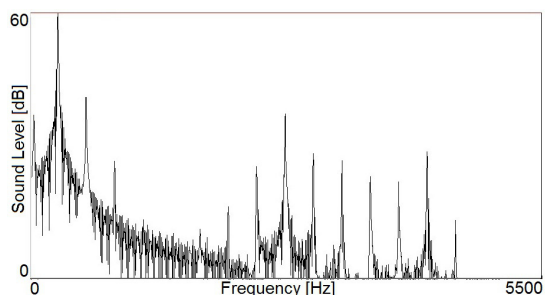


Figure 4: Harmonics of the /i/ vowel.

vowels is higher than the published data. Additionally, the Harmonic to Noise Ratio (HNR) values discriminated per vowel are presented.

II. Measurement methodology

Pitch, First and Second formants, Jitter, Shimmer and Harmonic to Noise Ratio (HNR) are the cornerstones of acoustic measurement of voice signals, and are often regarded as indices of the perceived quality of both normal and pathological voices [13].

In this work, we analyzed the audio files from the five Spanish vowels produced by 72 female and male individuals, in order to study the parameters previously mentioned. The individuals are Argentinian university students whose ages range between 20 and 30, coming from different regions without any special geographical distribution.

The voices were recorded using a Behringer C-1U (USB) cardioid microphone and a notebook.

The microphone was placed at a distance of 10 cm respect to the mouth of the subjects while they were pronouncing the vowels with an intensity and tone that was comfortable in an acoustically treated room. Each sound was sustained for,

at least, five seconds.

The Praat program, commonly used in linguistics for the scientific analysis of the human voice [10], was used to record, analyze the wav files and obtain all the parameters presented in this work. A sample rate of 44100 Hz was used to record the sound file.

The wave shapes of the sounds corresponding to /a/ and /i/ vowels are shown in Figs. 1 and 2. In Figs. 3 and 4, the harmonic components obtained by applying Fourier Transform to the respective vowel signal are shown.

Pitch

The pitch is a perceptual attribute of sound closely related to frequency, being this perception a subjective notion.

In psychoacoustics, the pitch is related to the fundamental frequency of vibration of the vocal cords, allowing the perception of the tone frequency.

Nevertheless, for Praat program [10], the pitch is coincident with the fundamental harmonic of the wave and we used this definition in this work.

This parameter depends on gender, being higher for women and lower for men.

Formants

The voice is created in the vocal cord, shaped as complex sound with harmonics and modified in the vocal tract by the resonating frequencies. Then, the amplitude of harmonics frequencies are enveloped forming a spectrum of energy, the peaks or maximum observed in these spectra are named “formants.” Consequently, a formant is a concentration of acoustic energy around a particular frequency in the speech wave. There are several formants, each one at a different frequency corresponding to a resonance in the vocal tract, and especially the first two are related to the movement of the tongue. The high-low magnitude of the First one (F1) is inversely related to the up-down tongue position and the Second formant (F2) is related to the front tongue position.

Jitter and Shimmer

The naturalness factor of sustained vowels is attributed to a fundamental frequency and the signal amplitude. Still there are unwanted variations in time of the sound signal properties in the voice production.

While jitter indicates the variability or perturbation of fundamental frequency, shimmer refers to the same perturbation but, in this case, related to amplitude of sound wave, or intensity of vocal emission. Jitter is affected mainly by lack of control of vocal fold vibration and shimmer by reduction of glottic resistance and mass lesions in the vocal folds, which are related to the presence of noise at emission and breathiness [10, 14].

Harmonic to Noise Ratio - HNR

The amount of energy conveyed in the fundamental frequency (f_0) and its harmonics, divided by the energy in noise frequencies, is defined as the harmonic-to-noise ratio. Frequencies that are not integer multiples of f_0 are regarded as noise. This parameter is related to the perception of vocal roughness and hoarseness [10].

Normal voices have a low level of noise and high HNR. On the contrary, the degree of hoarseness increases the noise component and decreases HNR.

III. Results and Discussion

The measured data were processed statistically and the results are shown in the Tables 1, 4, 5, 6 and Figs. 5 and 6.

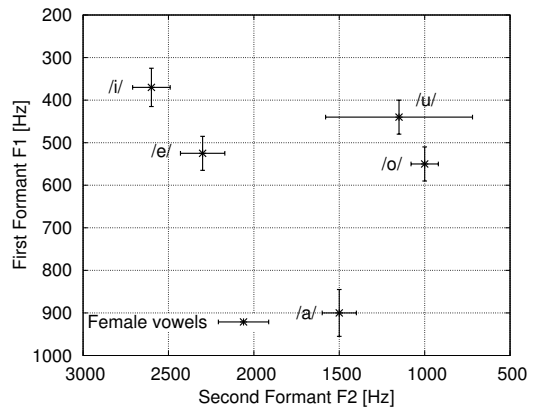


Figure 5: Female formant chart.

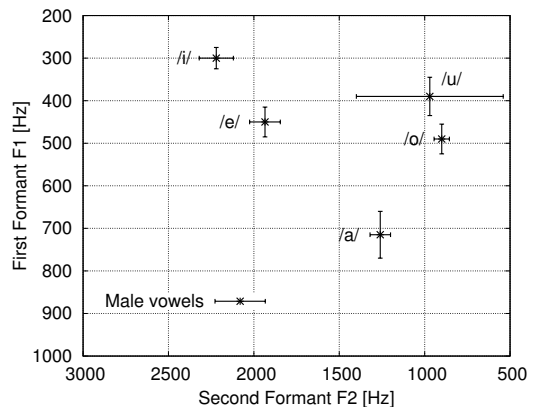


Figure 6: Male formant chart.

The pitches for female and male individuals are shown in Table 1. We used the minimum and maximum values to address the dispersion instead of the standard deviation because the data distribution was not normal. Our values are in general lower for both genders compared to the published data [9, 11, 12].

Tables 2 and 3 show the First and Second formants values and Figs. 5 and 6 show the chart of formants corresponding to female and male populations obtained in this work.

We have compared our male results with formant data of male Spanish speakers published by Bradlow [9].

In general, the First (F1) and Second (F2) formants values are comparable to the published ones.

In particular, the F1 formants for the /a/ and

	Female	Male
Maximum	314	196
Medium	225	128
Minimum	155	85

Table 1: Pitch values of female and male subjects in Hz.

/u/ vowels are higher than the reported ones, 12 and 21 %, respectively.

The Second formant, F2, for the /o/ vowel is lower than Bradlow by 12 %.

On the other hand, we cannot compare our female formant values with published results because we could not find results for female individuals in the literature. Comparing female versus male F1 formants, we observed that most of them are higher by 20 % but in the case of the /o/ vowel the difference is 11 %.

Comparing F2 formants, the female values are higher than the male ones, reaching almost the 25 % for /a/ and /i/ vowels.

Furthermore, the F2 of the /u/ vowel in our samples show an important scatter for both genders, female and male.

In the Tables 4 and 5, the obtained Jitter and Shimmer values for each vowel are shown. They are comparable to the Jitter and Shimmer averages obtained by Casado Morente *et al.* [12] in a study that involves a group of normal people. In our work, we have observed that the Jitter and the Shimmer values of the /a/ vowel are bigger than the corresponding ones of the other vowels.

Finally, the HNR results, see Table 6, are according to the average value presented by Casado Morente *et al.* [12]. However, we could not find in the bibliography the HNR values for each of the five Spanish vowels, so we had to make the comparison with the average of them. In the present work, we have found that the vowels show an increasing HNR value from /a/ to /u/, meaning that /u/ has better signal to noise ratio than the other vowels.

IV. Concluding remarks

The objective of this research was to measure acoustical properties of the Spanish voices of Argentinian speakers.

Vowels	F1 [Hz]	F2 [Hz]
/i/	370 ± 45	2600 ± 110
/e/	525 ± 40	2300 ± 130
/a/	900 ± 55	1500 ± 100
/o/	550 ± 40	1000 ± 80
/u/	440 ± 40	1150 ± 430

Table 2: First and Second formant of female.

Vowels	F1 [Hz]	F2 [Hz]
/i/	300 ± 25	2220 ± 100
/e/	450 ± 35	1935 ± 90
/a/	715 ± 55	1260 ± 60
/o/	490 ± 35	900 ± 45
/u/	390 ± 45	970 ± 430

Table 3: First and Second formant of male.

Vowels	Shimmer Local [%]	Jitter Local [%]
/a/	2.7 ± 1.1	0.31 ± 0.10
/e/	2.1 ± 0.7	0.28 ± 0.08
/i/	2.2 ± 0.6	0.29 ± 0.07
/o/	2.0 ± 0.7	0.26 ± 0.11
/u/	2.1 ± 0.7	0.27 ± 0.09

Table 4: Shimmer and Jitter of female subjects.

Vowels	Shimmer Local [%]	Jitter Local [%]
/a/	3.0 ± 0.9	0.36 ± 0.10
/e/	2.3 ± 0.8	0.33 ± 0.09
/i/	2.3 ± 0.7	0.28 ± 0.08
/o/	2.2 ± 0.8	0.29 ± 0.10
/u/	2.3 ± 0.9	0.25 ± 0.07

Table 5: Shimmer and Jitter of male subjects.

Vowels	Female	Male
/a/	21 ± 3	20 ± 2
/e/	20 ± 2	21 ± 2
/i/	22 ± 3	22 ± 2
/o/	25 ± 3	24 ± 3
/u/	25 ± 4	25 ± 3

Table 6: Harmonic to Noise Ratio of female and male subjects in dB.

These voice parameters are generally assessed subjectively by several authors. This form of perceptual analysis of voice has significant limitations and the subtle interpretative judgments of verbal classifications may not be accurate.

The differences we found in the parameters of the vowels measured in a group of people from Argentina compared to the parameters obtained from Spanish speaking people living in Spain suggests the region of study has an important influence in the results, as expected.

This kind of studies are very useful to compare the properties of normal and pathological voices of people from different regions.

It is necessary to test the same parameters in female Spanish speakers as well.

Such work should be performed in larger quantities and should be extended to other countries or regions of Latin America, especially where different ethnic groups can be found.

-
- [1] W R Rodríguez, O Saz, E Lleida, *Análisis robusto de la voz infantil con aplicación en terapia de voz*, *Areté* **10**, 70 (2010).
- [2] T Cervera, J L Miralles, J González-Álvarez, *Acoustical analysis of Spanish vowels produced by laryngectomized subjects*, *J. Speech Lang. Hear Res.* **44**, 988 (2001).
- [3] J Muñoz, E Mendoza, M D Fresneda, G Carballo, P Lopez, *Acoustic and perceptual indicators of normal and pathological voice*, *Folia Phoniatr. Logop.* **55**, 102 (2003).
- [4] H K Vorperian, R D Kent *Vowel acoustic space development in children: A synthesis of acoustic and anatomic data*, *J. Speech Lang. Hear Res.* **50**, 1510 (2007).
- [5] S P Whiteside, *Sex-specific fundamental and formant frequency patterns in a cross-sectional study*, *J. Acoust. Soc. Am.* **110**, 464 (2001).
- [6] P White, *Formant frequency analysis of childrens spoken and sung vowels using sweeping fundamental frequency production*, *J. Voice*, **13**, 570 (1999).
- [7] R O Coleman, *Male and female voice quality and its relationship to vowel formant frequencies*, *J. Speech Lang. Hear Res.* **14**, 565 (1971).
- [8] S Bennett, *Vowel formant frequency characteristics of preadolescent males and females*, *J. Acoust. Soc. Am.* **69** 231 (1981).
- [9] A R Bradlow *A comparative acoustic study of English and Spanish vowels*, *J. Acoust. Soc. Am.* **97**, 1916 (1995).
- [10] P Boersma, D Weenink, *Praat: doing phonetics by computer [Computer program]*, Version 5.3.51, retrieved 2 June 2013 from <http://www.praat.org/>.
- [11] J I Hualde, *The sounds of Spanish*, Cambridge University Press, Cambridge (2005).
- [12] J C Casado Morente, J A Adrián Torres, M Conde Jiménez, D Piédrola Maroto, V Povedano Rodríguez, E Muñoz Gomariz, E Cantillo Baños, A Jurado Ramos, *Estudio objetivo de la voz en población normal y en la disfonía por nódulos y pólipos vocales*, *Acta Otorrinolaringol. Esp.* **52**, 476 (2001).
- [13] J Kreimana, B R Gerrattb, *Perception of aperiodicity in pathological voice*, *J. Acoust. Soc. Am.* **117**, 2201 (2005).
- [14] H F Wertzner, S Schreiber, L Amaro, *Analysis of fundamental frequency, jitter, shimmer and vocal intensity in children with phonological disorders*, *Rev. Bras. Otorrinolaringol.* **71**, 582 (2005).