







Article

Zygoty Effects on Human Voice: Fundamental Frequency Analysis of Brazilian Twins' Speech

Lilian C. Luchesi^{1,2} , Julio C. Cavalcanti^{3,4,5} , Tania K. Lucci² , Vinicius F. David² , Emma Otta²  and Patricia F. Monticelli¹ 

¹Ethology and Bioacoustic Laboratory, Department of Psychology, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, São Paulo, Brazil, ²Psychoethology and Human Ethology Laboratory, Department of Experimental Psychology, Instituto de Psicologia, Universidade de São Paulo, São Paulo, ³Integrated Acoustic Analysis and Cognition Laboratory, Pontifical Catholic University of São Paulo, Rua Ministro de Godoy, São Paulo, Brazil, ⁴Institute of Language Studies, Department of Linguistics, University of Campinas, Campinas, São Paulo, Brazil and ⁵Laboratory of Phonetics, Department of Linguistics, Stockholm University, Stockholm, Sweden

Abstract

Voice production can be influenced by interindividual variations related to genetic, physiological, behavioral, and several environmental factors. Here we examined the effect of zygosity on speaking fundamental frequency (F0) statistical descriptors. Our aims were: (1) to determine whether the genetic similarity between monozygotic (MZ) and dizygotic (DZ) twins affects F0 characteristics, and (2) to quantify the contribution of genetic factors to these characteristics. The study involved 79 same-sex twin pairs of Brazilian Portuguese speakers, comprising 65 MZ and 14 DZ twins, aged 18 to 66 years (31.7 ± 11.6 years), with 21 male and 58 female pairs. Participants were recorded while uttering a greeting phrase and the Brazilian Portuguese version of the 'Happy Birthday to You' song. Speech segments were analyzed using Praat free software, and F0 measures were automatically extracted in both Hertz and semitone scales. Statistical descriptors, including centrality, dispersion, and extreme values of F0 were examined, and the ACE model (i.e., total genetic effects, A; shared environmental influences, C; and nonshared environmental influences, E) was employed to estimate the additive effects of monozygosity. As anticipated, we observed a zygosity effect on several F0 parameters, with more similarity between MZ twins compared to DZ twins. We discuss the genetic influences on F0 parameters and the absence of a monozygosity effect in two of them. Additionally, we briefly address potential biases associated with the selected measurement scale for statistical modeling. Finally, we explore the influence of genetic factors on F0 patterns, as well as environmental, life history and linguistic factors, particularly concerning F0 variation in speech.

Keywords: Intra-sibling similarity; Same-sex twins; Phonetic-acoustic analysis; Voice parameters

(Received 22 May 2024; revise received 27 June 2024; accepted 2 July 2024)

Acoustic Analysis, Assessing Voice Fundamental Frequency

The voice's fundamental frequency (F0) is physically linked to the vibration frequency of the vocal folds. Loudness and pitch constitute fundamental auditory sensations (Oxenham, 2012). Pitch represents the perceptual counterpart of a soundwave's resonant frequency, known as its fundamental frequency (F0), which corresponds to the wave's periodicity or repetition rate. At the same time, loudness is the perceptual counterpart of intensity. From a production-oriented perspective, the voice's frequency is considered the acoustic counterpart of the vibration frequency of the vocal folds (Imamura et al., 2003). The factors underlying its determination have been extensively decomposed and explored by the classical myoelastic-aerodynamic theory of voice production (van den Berg, 1958). The F0 of the voice depends on five interdependent factors related to the vibrating portion of the vocal folds, namely its effective

mass, effective tension, the effective area of the glottis during the cycle, subglottal pressure, and damping (van den Berg, 1958). It represents the lowest frequency in a periodic waveform and plays a pivotal role in determining a voice's pitch. Pitch, conversely, represents the perceptual counterpart of F0. This perception is also influenced by how the listener's auditory system interprets these sounds, allowing us to distinguish between high and low tones and discern a low pitch from a high-pitched voice (Titze, 2000).

The existing literature on voice has predominantly centered on interindividual variations in centrality metrics, which describe the position of F0 along the spectral dimension while ignoring deviations from its average values. To highlight the significance of other aspects of F0 in speech, prosody in conversation fluctuates in response to alterations in frequency contour (Dombrowski & Niebuhr, 2005; Xu, 2012), reflecting abrupt changes in emotional states (da Silva & Barbosa, 2017; Han et al., 2021) and conveying diverse prosodic and linguistic information. This fluctuation in F0, known as intonation, is crucial for effective communication (Hirst & Di Cristo, 1998), directly influencing F0 variability and its standard deviation ($F0_{sd}$), which is statistical indicators together with variation quotients of speech 'liveliness' (Hincks, 2014; Traunmüller & Eriksson, 1994).

Corresponding authors: Patricia Monticelli; Email: monticel@usp.br and Lilian C. Luchesi; Email: lili.luchesi@gmail.com

Cite this article: Luchesi LC, Cavalcanti JC, Lucci TK, David VF, Otta E, Monticelli PF. Zygoty Effects on Human Voice: Fundamental Frequency Analysis of Brazilian Twins' Speech. *Twin Research and Human Genetics* <https://doi.org/10.1017/thg.2024.33>

Moreover, F0 can be characterized by a baseline value, termed F0 baseline ($F0_{\text{base}}$), representing a neutral mode of vibration to which the vocal folds revert after a prosodic shift and reflecting habitual pitch (Lindh & Eriksson, 2007). Additionally, F0 variations span the lowest and highest points during the speech, referred to as F0 minima ($F0_{\text{min}}$) and maxima ($F0_{\text{max}}$) respectively, constituting the F0 dynamic range (Titze, 2006). Furthermore, the pleasantness or tension in voice correlates with the F0 interquartile semi-amplitude ($F0_{\text{SAQ}}$), a measure of dispersion (dos Reis, 2017); individuals with tense voices exhibit elevated F0 measures and increased variability, reflected by higher $F0_{\text{SAQ}}$ values (dos Reis, 2017). Conversely, calmness in speakers is linked to decreased $F0_{\text{med}}$ and $F0_{\text{SAQ}}$ (da Silva & Barbosa, 2017).

The intrinsic geometry of the vocal folds, including factors such as length, depth and thickness, is associated with vibration frequency and F0 variations in response to physiological conditions, age and gender, with children typically exhibiting higher F0 compared to adults, and females generally higher than males (Zhang, 2016). These factors, in turn, are influenced by both intrinsic and extrinsic modifiers, indicating a significant genetic organic impact on an individual's F0, as evidenced by acoustic-phonetic genetic-related speaker studies.

Twin Studies

The association of genetic differences and the proportion of phenotypic variance can be assessed by statistical estimation of heritability measures. The twin approach enables investigation of genetic and environmental influences on human traits, presenting an overall heritability estimate of $h^2 = 0.49$ (Polderman *et al.*, 2015). Among acoustic-phonetic approaches, F0 emerges as one of the most commonly studied parameters in speech and voice analysis, with numerous studies consistently showing a strong correlation between monozygotic (MZ) twins regarding speaking style (Arantes & Eriksson, 2019; Signorello *et al.*, 2020), speaking condition (de Jong *et al.*, 2011), emotional state (Higuchi *et al.*, 1997), and sociocultural factors (Rilliard *et al.*, 2013) on their F0. However, many studies have predominantly focused on a limited set of F0 statistical descriptors, such as mean and standard deviation values, neglecting the genetic impact on other dimensions of the parameter.

For instance, comprehensive research examining the voice quality characteristics of MZ twins across a wide age range (8 to 61 years old) found no influence of sex and age on vocal similarities, with high correlation scores observed for their F0 estimates, supported by auditory perceptual evaluations (van Lierde *et al.*, 2005). Similarly, a case study of a pair of male MZ twins and their age- and sex-matched sibling revealed striking similarities in speech characteristics, particularly in pitch mean, suggesting a degree of family resemblance (Whiteside & Rixon, 2013). F0's potential as a phenotype was examined in twin speakers of American English, revealing an association with age and weight (Przybyla *et al.*, 1992). No disparities were found between MZ and dizygotic (DZ) twins in $F0_{\text{mean}}$ and $F0_{\text{sd}}$, although DZ twins exhibited greater variation in F0 measures (in Hz; Przybyla *et al.*, 1992). Despite these findings, F0 measures may be influenced by factors beyond genetic constitution alone, such as spontaneous and nonspontaneous speech, even among twins.

Examining F0 and its intra-speaker variability during a reading task in adult Dutch speakers, highly similar results were observed between MZ and DZ twins, while no correlation was observed among unrelated peers, indicating that an individual's voice is

influenced by factors beyond genetic constitution alone (Debruyne *et al.*, 2002). Similarly, an analysis of long-term F0 in twin speakers of Australian English revealed that twins tend to exhibit more similar mean long-term F0 values than unrelated pairs, though not always presenting the closest mean F0 values within twin pairs (Loakes, 2006).

Despite the significant contributions of twin studies, they are vastly represented by WEIRD (Western, Educated, Industrialized, Rich and Democratic) societies, particularly by seven countries: the United States, the United Kingdom, Australia, the Netherlands, Sweden, Denmark and Finland (Fernandes *et al.*, 2024). This is evident in the review of human trait heritability in 50 years of twin studies (Polderman *et al.*, 2015), in which they represent more than 90% of the studies. This WEIRD sampling problem involves similar countries in cultural history, social values and standard of living (Uchiyama *et al.*, 2022).

A further issue arises from the disproportionate representativeness of the WEIRD twins, which is identified as a challenge within the portability problem discussed by Matthews and Turkheimer (2022), given that WEIRD societies represent only 2% of humanity. It is thus imperative to acknowledge and evaluate the contributions of twin studies conducted in non-WEIRD populations (Hagenbeek *et al.*, 2022). It is noteworthy that Brazilian twin research remains underrepresented internationally (Fernandes *et al.*, 2024) and, to the best of our knowledge, there are few studies that have specifically examined the F0 speech parameters in twins.

For example, one study focused on the speaker-discriminatory potential of F0 estimates in comparisons between intra-identical twin pairs and across all speakers in Brazilian Portuguese (Cavalcanti *et al.*, 2021a). In another study, the acoustic analyses of connected speech samples and lengthened vowels of adult male twins revealed that $F0_{\text{base}}$, central tendency, and extreme values were mostly discriminatory in both intra-twin pair and cross-pair comparisons, suggesting the influence of speaking style and dialect on dynamic F0 patterns (Cavalcanti *et al.*, 2021b).

In this present study, we aim to assess whether zygosity influences F0 from connected speech using twin intra-sibling similarity by examining statistical descriptors of the parameter in both MZ and DZ twins. Our research focuses on adult Brazilian Portuguese speakers and extends beyond central tendency measures of F0, allowing for a comprehensive analysis of how these variables contribute to differences in F0 across speakers.

Materials and Methods

Participants

The recruitment of twins was conducted online, with each pair receiving a telephone invitation to personally visit the laboratory of the Institute of Psychology at the University of São Paulo (USP) to participate in the research. Peripheral blood samples were collected from each individual's arm for DNA testing to determine zygosity. These samples were sent to a laboratory for genotyping procedures, based on 22 autosomal STR loci, along with amelogenin and DYS391 (cf. Varella *et al.*, *in press*). The study covered recordings of sentences spoken by 88 Brazilian twin pairs, consisting of 72 MZ and 15 DZ pairs, ranging in age from 18 to 66 years (mean age 31.7 ± 11.6 years), who volunteered for the study. Only same-sex twin dyads were included. After addressing all recording issues, such as noise or missing data from one twin within each pair, we analyzed a total of 79 pairs, consisting of 21 male pairs and 58 female pairs, with only one pair residing outside the state of São Paulo.

Table 1. Fundamental frequency (F0) descriptors and their physical or perceptual correlates

F0 parameter	Definition	Refers to
$F0_{\text{mean}}$	The arithmetic means of the F0 values	The average pitch of a speaker's voice
$F0_{\text{base}}$	The estimated base value of F0, represents the speaker's habitual pitch	The fundamental frequency at rest or during normal (habitual) speech
$F0_{\text{sd}}$	The measure of the dispersion of the F0 values around the mean	The F0 variability within a speaker's voice in relation to the mean, related to liveliness in speech
$F0_{\text{min}}$	The lowest F0 value in a given speech chunk	Perceptually, it refers to the lowest pitch of a speaker's voice, as in F0 valleys in connected speech
$F0_{\text{max}}$	The highest F0 value in a given speech chunk	Perceptually, it refers to the highest pitch of a speaker's voice, as in F0 peaks in connected speech
$F0_{\text{SAQ}}$	Half of the difference between the upper and lower quantiles of the F0 distribution	The range of pitch variation within a speech chunk. It relates to the extent of the pitch variation

Note: base, baseline; sd, standard deviation; min, minima; max, maxima; SAQ, interquartile semi-amplitude.

Recording Procedure

Participants were instructed to utter the greeting phrase in Brazilian Portuguese, 'Hi, my name is Pedro' (for male speakers) and 'Hi, my name is Ana' (for female speakers; 'Oi, meu nome é Pedro/Ana'), as well as the spoken version of the Brazilian Portuguese version of 'Happy Birthday to You' song. Recordings were conducted in a soundproof room at the Psychology Institute of the USP using a Zoom H1 Handy Recorder paired with a studio microphone BM8000 positioned approximately 15 cm away from the participant's mouth. Recordings were made in stereo mode at 24-Bit/96kHz resolution and saved in an uncompressed '.wav' format.

Acoustic Analysis

Each recorded voice was segmented into six chunks: the greeting segment 'Hi, my name is Ana/Pedro' comprised two chunks, and the 'Parabéns a você', 'nesta data querida', 'muitas felicidades', 'muitos anos de vida' song was divided into four chunks. From each chunk we extracted various acoustic measures, including mean F0 ($F0_{\text{mean}}$), median F0 ($F0_{\text{med}}$), F0 interquartile semi-amplitude ($F0_{\text{SAQ}}$), F0 baseline ($F0_{\text{base}}$), F0 standard deviation ($F0_{\text{sd}}$), minimum F0 ($F0_{\text{min}}$), and maximum F0 ($F0_{\text{max}}$), expressed in both scales, semitones (st) and Hertz (Hz). Table 1 delineates these measures and their relevance to voice perception. All analyses were conducted using the free software Praat 6.1.32 (Boersma & Weenink, 2020) with the 'Prosody Descriptor Extractor' script (Barbosa, 2021). The F0 floor and ceiling for parameter extraction were set at 60 Hz and 300 Hz respectively, and both Hertz and semitones scales were utilized, with a reference of 1 Hz for the latter.

We employed the Hz scale to represent fundamental frequencies and incorporated a semitone (nonlinear) representation of the same F0 parameters. Production and perceptual differences in human speech that convey emotion can be detected on a logarithmic scale, even without conscious attention (see Bjørklind, 2005; McDermott et al., 2010). By considering the impact of F0 variations on communication, we offer valuable insights for future research into auditory perception of similar voices.

Studies suggest that listeners may discern two tones as distinct if their fundamental frequencies (F0) differ by 3 to 4 semitones (Assmann, 1999; Consoni et al., 2009). Therefore, it is crucial to report intra-twin pair differences surpassing this threshold, as such distinctions could significantly affect future studies on voice perception. By including the semitone scale (st), we aimed to

assess differences between twin siblings potentially perceivable by the auditory system (Vargas et al., 2005) and gender-based disparities, unattainable using the Hz scale and irrespective of individual voice frequency (Costa et al., 2008). Apart from examining variations in acoustic parameters among subjects, we sought to explore the communication implications of such differences, offering insights for future perceptual studies.

Figure 1 represents an illustration of the Praat window and transcription. The upper segment shows the audio signal used for extracting F0 descriptors. The grayscale spectrogram of the Brazilian Portuguese speech segment 'Hi, my name is Ana' is displayed, with F0 and intensity curves overlaid in blue and yellow lines respectively (note the two scales on the spectrogram's vertical axis; yellow and blue lines correspond to values on the right side from 75 to 169.3 Hz). The first tier presents the orthographic transcription of the phrases, while the second tier exhibits the vowel-to-vowel transcription. In our study, F0 parameters were extracted from tier 1, specifically at the phrase level.

Statistical Analysis

We estimated the mean and range values for all F0 parameters in both semitones and Hertz. To compare MZ and DZ dyads, we employed an independent sample Mann-Whitney U test. Covariation among siblings from both MZ and DZ groups was determined using a Wald chi-square mixed model. Sex was included as a factor in the model, with age serving as a covariate variable. Intra-sibling covariance in F0 parameters was applied to estimate heritability (h^2) based on the ACE additive model. This model assumes that all genetic influence arises from additive effects (A), while excluding dominant (D) and epistasis (I) genetic effects. Consequently, the model encompasses total genetic effects (A) along with shared (C) and nonshared (E) environmental influences (Zyphur et al., 2013). Random intercepts and slopes were incorporated into the model, with slopes varying according to zygosity. Heritability was calculated as twice the difference between MZ and DZ covariances. Analyses were conducted using Stata v. 16.0, with the significance level set at $p = .05$.

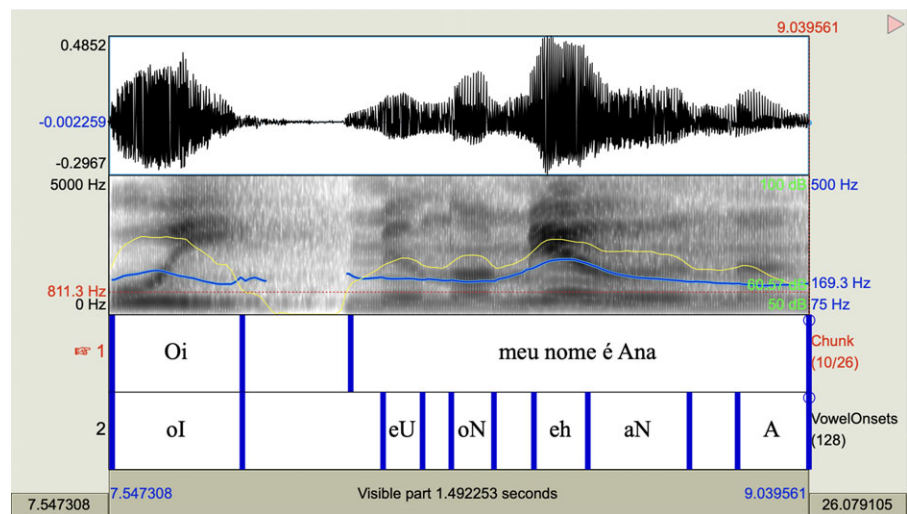
Results

Given the normal distribution of our data and the similarity between $F0_{\text{med}}$ and $F0_{\text{mean}}$ values, we opted to use $F0_{\text{mean}}$ due to its lesser susceptibility to data distribution effects. We first provide a

Table 2. Male and female fundamental frequency (F0) descriptors from 79 pairs of Brazilian twins' speech in hertz (Hz)

	MZ						DZ					
	Male (n = 18)			Female (n = 47)			Male (n = 3)			Female (n = 11)		
	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
F0 _{mean}	120.47	93	158	199.20	130	258	129.50	112	153	200.43	135	270
F0 _{sd}	8.91	2.02	31.07	14.57	0.29	67.53	7.83	2.21	22.42	12.24	1	39.61
F0 _{SAQ}	6.19	0.89	31.92	10.42	0.22	65.89	4.95	1.51	12.80	8.85	0.65	37.46
F0 _{base}	107.49	84	141	175.93	93	239	118.80	89	139	180.54	118	268
F0 _{min}	105.77	80	139	171.77	85	239	117	87	136	177.55	108	267
F0 _{max}	136.18	97	200	220.88	143	338	146.17	119	225	218.14	145	289

Note: sd, standard deviation; SAQ, interquartile semi-amplitude; base, baseline; min, minima; max, maxima.

**Figure 1.** Praat window and transcription of the speech segment 'Hi, my name is Ana' in Brazilian Portuguese.

summarizing overview of the parameters, followed by the outcomes of Hz modeling. Subsequently, we emphasize the distinctions observed when utilizing st for ACE model covariances and effects among twin pairs.

Regarding sex, male MZ and DZ twins exhibited comparable F0 dimensions (Table 2). Five out of six parameters showed differences between male MZ and DZ twins, except for F0_{SAQ}, $U = 1.270$, $df = 1$, $p = .26$, with no disparities found between MZ and DZ females. Besides zygosity, F0_{mean} ranged from 93–158 Hz in males and 130–270 Hz in females, with F0_{sd} varying from 1 to 68 Hz among them and 2–31 Hz in males (F0_{mean} male MZ 120.47 ± 13.46 ; DZ 129.50 ± 10.60 ; female MZ 199.20 ± 25.96 ; DZ 200.43 ± 26.13).

Considering the semitones scale, F0_{mean} ranged from 76 to 88 st in males and 80 to 98 st in females, with higher F0_{sd} observed in females compared to males (Table 3). In general, intrasexual variation was lower than zygosity, except for F0_{SAQ}, which exhibited the most variation both intra-sexual and between MZ and DZ. Analyzing intra-twin differences between MZ and DZ siblings in terms of F0 average parameters, we found a difference only in F0_{base}, $U = 411.5$, $df = 1$, $p = .011$.

For visual representation, the violin plots in Figure 2 show the distribution of specific F0 descriptors among MZ and DZ twins, highlighting differences in F0 distribution across these groups. Outliers were removed based on the interquartile range method. While median values for F0_{mean} and F0_{base} are somewhat

similar between MZ and DZ twins (Table 3 and Figure 2), the distributions skew more towards lower values in MZ. This skewness towards lower values in MZ twins suggests a propensity for certain voices to cluster around specific values more frequently than in DZ twins. Notably, the plots reveal two distinct clusters for MZ twins, one in higher and another in lower F0 regions.

Given the uneven distribution of twin pairs in each group, with higher numbers of MZ than DZ pairs (72 vs. 15), we anticipated a broader range of F0 values for MZ twins in descriptors such as F0_{mean} and F0_{base}, as evidenced by the range (from minimum to maximum values) in the plots (Figure 3). Despite this discrepancy in the number of observations between the twin groups, the distribution for dispersion measures like F0_{sd} and F0_{SAQ} appeared relatively uniform between MZ and DZ.

Intra-Twin Covariation

Twinship exerts an influence on speech parameters, with MZ twins accounting for at least 10% of the variance in all six parameters. A minor effect was observed solely in F0_{SAQ} (<10%). Although we did not conduct statistical analysis, all intra-siblings' variations showed disparities of 2% to 6% (F0_{mean}, F0_{sd}, F0_{base}, and F0_{max}) when comparing values from Hz and st scales (Tables 4 and 5). The environmental component in the ACE model is referred to as the twinning effect.

Table 3. Male and female fundamental frequency (F0) descriptors from 79 pairs of Brazilian twins' speech in semitones (st)

	Monozygotic						Dizygotic					
	Male (n = 18)			Female (n = 47)			Male (n = 3)			Female (n = 11)		
	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
Age	24.2	18	35	32	19	66	21	20	23	29.96	19	53
F0 _{mean}	83	76	88	91	80	98	84	81	87	92	85	97
F0 _{sd}	1.22	0.03	3.85	1.53	0.03	7.06	1	0.06	2.31	1.52	0.06	4.7
F0 _{SAQ}	0.85	0.03	3.83	0.98	0.02	7.1	0.67	0.05	1.61	0.9	0.04	2.88
F0 _{base}	81	75	86	89	78	97	83	78	86	89	79	97
F0 _{min}	81	75	86	88	76	97	83	77	86	88	78	97
F0 _{max}	85	76	93	94	82	101	86	81	94	94	86	100

Note: sd, standard deviation; SAQ, interquartile semi-amplitude; base, baseline; min, minima; max, maxima.

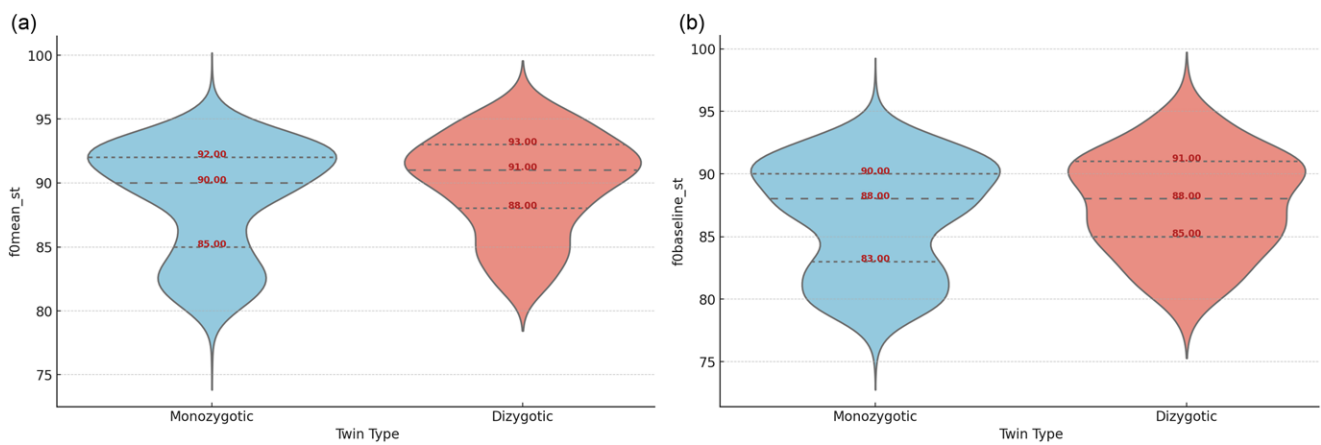


Figure 2. Distribution of mean values of F0_{mean} (a) and F0_{base} (b) within speech chunks between monozygotic and dizygotic twins. Note: base, baseline.

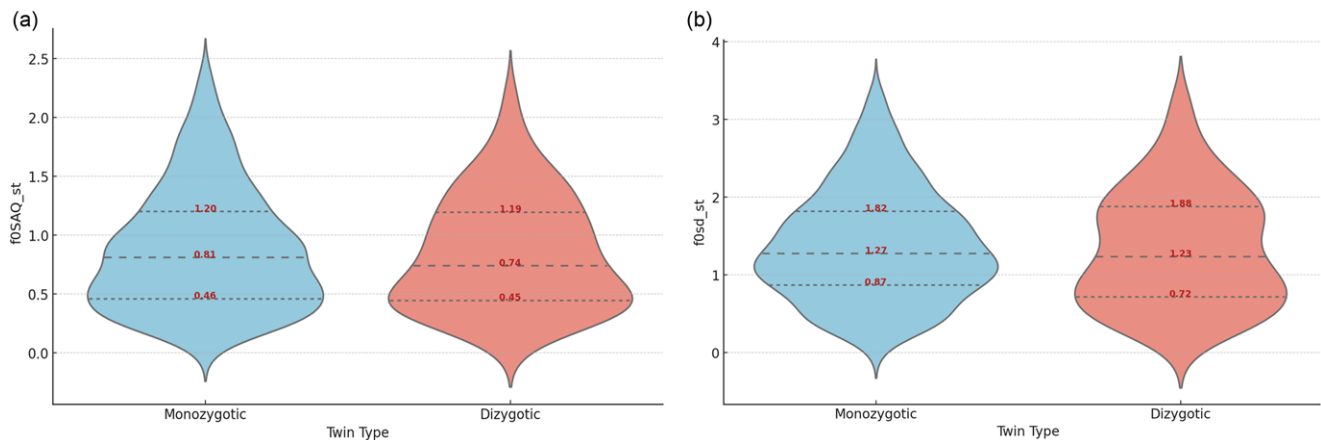


Figure 3. Distribution of mean values of F0_{SAQ} (a) and F0_{sd} (b) within speech chunks between monozygotic and dizygotic twins. Note: SAQ, interquartile semi-amplitude; sd, standard deviation.

Intra-twin covariation in hertz. Table 4 illustrates the covariance between twin siblings across all F0 parameters. Our findings indicate a significant heritability effect on voice, with twinship accounting for 36% of the variance, and monozygosity adding 21.2%, resulting in 57.2% explained variance in F0_{mean}.

Similarly, monozygosity contributed with more 18.3% to intra-sibling variance explaining 46.5% of their similarity in F0_{base}. However, monozygosity did not show an additive effect on F0_{sd} and F0_{min}, with shared environmental factors explaining 11% and 27.6% of the variance, respectively. Sex did not affect two F0

Table 4. Estimated heritability in F0 parameters in hertz considering the effects of twinning (sibling) monozygosity (MZ), sex, and age on intra-siblings' covariance calculated by the ACE model

Parameter	Sibling	MZ	Male		Age	
	(%)	(%)	χ^2	p	χ^2	p
F0 _{mean}	36.0	21.2	8.96	<.001	-0.07	0.001
F0 _{sd}	11.0	0.0	1.68	.093	1.78	0.075
F0 _{SAQ}	8.6	2.4	1.31	.19	2.09	0.037
F0 _{base}	27.8	6.4	3.75	<.001	4.0	<.001
F0 _{min}	27.6	0.0	3.45	.001	3.81	<.001
F0 _{max}	28.2	18.3	4.32	<.001	2.65	0.074

Note: ACE: total genetic effects, A; shared environmental influences, C; nonshared environmental influences, E. sd, standard deviation; SAQ, interquartile semi-amplitude; base, baseline; min, minima; max, maxima.

Table 5. Estimated heritability in F0 parameters in semitones considering the effects of twinning (sibling) monozygosity (MZ), sex, and age on intra-siblings' covariance calculated by the ACE model

Parameter	Sibling	MZ	Male		Age	
	(%)	(%)	χ^2	p	χ^2	p
F0 _{mean}	42.0	18	3.88	<.001	3.47	.001
F0 _{sd}	13.0	0.0	2.41	.016	0.50	.619
F0 _{SAQ}	7.0	5.5	1.97	.049	0.94	.344
F0 _{base}	30.0	5.0	3.80	<.001	4.30	<.001
F0 _{min}	28.0	5.0	3.80	<.001	4.21	.004
F0 _{max}	33.7	12.7	4.04	<.001	2.91	.004

Note: ACE: total genetic effects, A; shared environmental influences, C; nonshared environmental influences, E. sd, standard deviation; SAQ, interquartile semi-amplitude; base, baseline; min, minima; max, maxima.

parameters (F0_{sd}, $\chi^2 = 1.68$, $p = .093$ and F0_{SAQ}, $\chi^2 = 1.31$, $p = .19$), while age influenced four out of six parameters, except F0_{sd}, $\chi^2 = 1.78$, $p = .075$, and F0_{max}, $\chi^2 = 2.65$, $p = .074$.

Intra-twin covariation in semitones. When analyzed in st, we observed slight differences compared to hertz (Table 5). The twinning effect alone explained 42% of the covariance in F0_{mean} and 33.7% in F0_{max}. The additional effect of monozygosity explained 60% of the variance in F0_{mean} and 46.4% in F0_{max}. The genetic similarity negligibly increased 5% the effect on F0_{SAQ} (increased to 12.5%) and F0_{base} (increased to 35%).

In contrast to what was observed when measured in HZ scale (no additive effect, Table 4), F0_{min}, monozygosity added 5% to the twinning effect (28%) totaling 33% of the explained variance between twins. On the other hand, sharing practically the same DNA did not affect the covariance between siblings, when measured in st, with the shared environment alone explaining 13% of the variance in F0_{sd}.

Sex influenced the covariance between twin siblings in all F0 parameters in semitones. Although F0_{SAQ} was just below the significance level threshold, $\chi^2 = 1.97$, $p = .049$, there was no effect of age on two out of six parameters: F0_{sd}, $\chi^2 = 0.5$, $p = .619$ and F0_{SAQ}, $\chi^2 = 0.94$, $p = .344$ (Table 5).

Regarding statistical comparisons intra-twin, only two parameters showed differences exceeding 3 semitones: F0_{mean} and F0_{base}. Specifically, six pairs — five female and one male — exhibited these differences in F0_{mean}, including four pairs with one male MZ

twin. For F0_{base}, seven pairs showed differences, comprised of six female pairs (three of whom were MZ) and one male MZ pair. Among female twins, there was an equal split between MZ and DZ twins.

Comparisons between scales: semitones and hertz. Upon comparing the outcomes in hertz and semitones, we observed that monozygosity contributed to explaining variance across more parameters in semitones, including an additive effect on F0_{min}, which was absent in hertz measurements (Tables 4 and 5). Additionally, sex exerted influence on all parameters, including F0_{SAQ}. However, it is noteworthy that monozygosity did not have an additive effect on F0_{SAQ} in either scale, with the twinning effect remaining the sole relevant factor in this case.

Discussion

In this study, we aimed to assess whether zygosity influences F0 dimensions based on intra-sibling similarity, considering both linear (Hz) and nonlinear (semitones [st]) scales. However, for clarity purposes in our discussion, we will focus solely on st. As anticipated, MZ twins exhibited greater resemblance to each other compared to DZ twins.

Zygosity accounted for variance in five out of six F0 parameters, contributing 18% to F0_{mean} and 12.7% to F0_{max}, with approximately 5% in F0_{SAQ}, F0_{base}, and F0_{min}. The likeness in F0 parameters among MZ siblings is closely tied to similarities in vocal fold geometry and anatomical properties, such as mass, tension area and pressure measures (Zhang, 2016). Our findings reaffirmed expectations that male twins with larger vocal folds would produce lower vibration frequencies (presenting lower F0).

While MZ twins share identical genetic makeup, variations in their environment and experiences can lead to differences in anatomical and physiological characteristics. It is regularly assumed that these differences may become more pronounced as twins age and encounter diverse environments and lifestyles, possibly due to epigenetic changes (Fraga et al., 2005). However, in our sample, we did not observe an age-related effect on intra-twin pairs' F0 variation in a cross-sectional analysis.

The voice variation during speech or 'liveliness' (F0_{sd}) averaged around 1 st in male siblings and 1.5 st in females. In our study, F0_{sd} was not influenced by monozygosity, suggesting that environmental and linguistic factors may play a more significant role in explaining individual variations in this parameter, such as speech flow (Traunmüller & Eriksson, 1994) and variation and modulation estimates, particularly in spontaneous dialogues (Cavalcanti, 2021a, Hincks, 2005).

In addition to assessing differences in acoustic parameters across subjects, we aimed to explore the potential implications of such differences on the communication process, offering insights for future perception studies. Listeners may distinguish two tones if their F0 differences range between 3 st to 4 st (Assmann, 1999; Consoni et al., 2009). Even though we found intra-twin differences in F0_{Mean} and F0_{base} in only six pairs (five females and one male) and F0_{base} in seven pairs, such findings are valuable for understanding how twin voices are perceived, suggesting that twins may differ in the way their pitch is perceived. Other studies exploring different dimensions, such as temporal and spectral characteristics, illustrate how twins can align or diverge in their speech production behavior (cf. Cavalcanti et al., 2021a, 2021b). In other words, different parameters play a role in determining how their voice qualities are perceived as similar or different.

Previous evidence suggests that vocal F0 alone is insufficient to determine zygosity in same-sex twin pairs; instead, a combination of 14 acoustic parameters is required, including $F0_{\text{mean}}$, $F0_{\text{sd}}$, and formant frequencies, among others (Forrai & Gordos, 1983). In our study, we observed no additive effect of zygosity only in $F0_{\text{sd}}$, with $F0_{\text{mean}}$ emerging as the vocal parameter most significantly influenced by monozygosity.

While certain F0 parameters are more influenced by genetic factors (Debruyne et al., 2002; Sataloff, 1995, 1997), temporal patterns (van Lierde et al., 2005) and changes within speech utterances are likely shaped by environmental factors such as accent, dialect, coarticulation pattern and speaking style (Whiteside & Rixon, 2003, 2013) filled with culturally specific codes in prosody (Rilliard et al., 2013). It was expected that MZ twins would exhibit greater similarity in temporal and spectral characteristics compared to DZ twins. However, $F0_{\text{sd}}$ demonstrated the least variance explained by monozygosity, suggesting a lesser genetic influence on this parameter, possibly due to environmental factors that shape ‘pitch variation and intonation’ and thereby neutralizing any genetic effects on this parameter (Debruyne et al., 2002).

Sex exerted influence on all voice parameters in st over the explained variance among MZ and DZ twins, indicating potential differences in genetic and environmental contributions between sexes. However, the imbalance between the MZ and DZ samples must be highlighted. Since zygosity was not known until after data collection, we could not balance the two groups, and statistical tools were used to account for these differences in modeling. In addition, given the limited participation of male pairs in our sample (25% male pairs), caution is warranted in generalizing these findings. The statistical power and possible generalizations are limited by men’s reluctance to participate in research (e.g., friendship — Butera, 2006; or health — Glass et al., 2015).

Future studies should explore correlations and interactions among F0 descriptors and include other acoustic parameters (such as formant frequencies, voice quality metrics and temporal aspects of speech) to provide a more comprehensive understanding of the acoustic phenotype. Also, for clinical applications, other instrumental assessments could be considered to analyze anatomical variables contributing to prosody, which was beyond our study goals. Investigating these parameters together would enable researchers to assess the multivariate effects of genetics, environment and life history on voice characteristics, offering deeper insights into the complex interplay of factors influencing vocal traits in MZ and DZ twins.

In summary, our findings highlight the influence of zygosity and environmental factors on voice, with variations observed between measurement scales (Hz and st). This study offers insights into the intricate determinants of vocal traits, underscoring the nuanced contributions of genetic and environmental factors and the importance of scale selection in investigating hypotheses related to the human voice.

Data availability statement. Due to the nature of the research and ethical restrictions, supporting data is not available.

Acknowledgments. The authors thank the twins who participated in the study, who without their contribution this article would not exist. We are also grateful to the University of Sao Paulo Twin Panel (Painel USP de Gêmeos) colleagues and collaborators who helped the authors during data collection and analysis, especially Bruna Campos Paula, who trained LCL to use the software.

Author contribution. PFM, EO designed the study; EO Funding acquisition; Project administration; TKL recruitment, data acquisition; LCL Investigation,

data preparation and segmentation; LCL, JCC, VFD data analysis; LCL Original Draft manuscript; PFM, TKL, JCC, VFD, EO manuscript review.

Financial support. This work was supported by the Fundação de Amparo à Pesquisa do Estado de São Paulo, FAPESP (E.O. grant number 2014/50282-5), (L.C.L. grant numbers 2020/14250-2, 2022/08063-0), (J.C.C. grant number 2023/11070-1) and by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, CAPES (funding code 001).

Competing interests. None.

Ethical standards. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

References

- Arantes, P., & Eriksson, A. (2019). Quantifying fundamental frequency modulation as a function of language, speaking style and speaker. In *Interspeech 2019* (pp. 1716–1720). The International Speech Communication Association. <https://doi.org/10.21437/Interspeech.2019-2857>
- Assmann, P. F. (1999). Fundamental frequency and the intelligibility of competing voices. In *Proceedings of the 14th International Congress of Phonetic Sciences* (pp. 179–182). University of California Press.
- Barbosa, P. A. (2021). *Prosody descriptor extractor*. Github. Com/Pabarbosa/Prosody-Scripts. <https://github.com/pabarbosa/prosody-scripts.git>
- Björkljnd, A. (2005). Analyses of soprano voices. *The Journal of the Acoustical Society of America*, 33, 575. <https://doi.org/10.1121/1.1908728>
- Boersma, P., & Weenink, D. (2020). *Praat: Doing phonetics by computer*, v 6.1. 21.
- Butera, K. J. (2006). Manhunt: The challenge of enticing men to participate in a study on friendship. *Qualitative Inquiry*, 12, 1262–1282. <https://doi.org/10.1177/1077800406288634>
- Cavalcanti, J. C., Eriksson, A., & Barbosa, P. A. (2021a). Multiparametric analysis of speaking fundamental frequency in genetically related speakers using different speech materials: Some forensic implications. *Journal of Voice*, 38, 1–19. <https://doi.org/10.1016/j.jvoice.2021.08.013>
- Cavalcanti, J. C., Eriksson, A., & Barbosa, P. A. (2021b). Acoustic analysis of vowel formant frequencies in genetically-related and non-genetically related speakers with implications for forensic speaker comparison. *Plos One*, 16, e0246645. <https://doi.org/10.1371/journal.pone.0246645>
- Consoni, F., Peres, D., Lassak, A., Rosa, R., & Ferreira Netto, W. (2009). *Sensitivity to F0 variation in Brazilian Portuguese* [Paper presentation]. 40th Poznan Linguistic Meeting.
- Costa, J. de O., Gama, A. C. C., Oliveira, J. B. de, & de Rezende Neto, A. L. (2008). Avaliação acústica e perceptivo-auditiva da voz nos momentos pré e pós-operatório da cirurgia de implante de pré-fáscia do músculo temporal. *Revista CEFAC*, 10, 76–83. <https://doi.org/10.1590/S1516-18462008000100011>
- da Silva, W., & Barbosa, P. A. (2017). Perception of emotional prosody: investigating the relation between the discrete and dimensional approaches to emotions. *Revista de Estudos da Linguagem*, 25, 1075–1103. <https://doi.org/10.17851/2237-2083.25.3.1075-1103>
- Debruyne, F., Decoster, W., Van Gijssels, A., & Vercammen, J. (2002). Speaking fundamental frequency in monozygotic and dizygotic twins. *Journal of Voice*, 16, 466–471. [https://doi.org/10.1016/s0892-1997\(02\)00121-2](https://doi.org/10.1016/s0892-1997(02)00121-2)
- de Jong, G., Hudson, T., Nolan, F., & McDougall, K. (2011). *The telephone effect on F0* [Paper presentation]. IAFPA 2011 Conference, Vienna, Austria.
- Dombrowski, E., & Niebuhr, O. (2005). Acoustic patterns and communicative functions of phrase-final F0 rises in German: Activating and restricting contours. *Phonetica*, 62, 176–195.
- dos Reis, N. (2017). Medidas Acústicas e Julgamentos Perceptivos de Tensão e de Agradabilidade em Fonação Alaríngea com Prótese Traqueoesofágica. In *Revista Intercâmbio, Especial Expressividade*, XXXVI (pp. 66–85).

- Fernandes, E. de S., Ferreira, I. F., de Felipe, R. P., Segal, N., & Otta, E. (2024). Brazilian twin studies: A scoping review. *Twin Research and Human Genetics*, 27, 105–114. <https://doi.org/10.1017/thg.2024.17>
- Fraga, M. F., Ballestar, E., Paz, M. F., Ropero, S., Setien, F., Ballestar, M. L., Heine-Suñer, D., Cigudosa, J. C., Urioste, M., Benitez, J., Boix-Chornet, M., Sanchez-Aguilera, A., Ling, C., Carlsson, E., Poulsen, P., Vaag, A., Stephan, Z., Spector, T. D., Wu, Y. Z., ... Esteller, M. (2005). Epigenetic differences arise during the lifetime of monozygotic twins. *Proceedings of the National Academy of Sciences*, 102, 10604–10609. <https://doi.org/10.1073/pnas.050039810>
- Forrai G., & Gordos, G. (1983). A new acoustic method for the discrimination of monozygotic and dizygotic twins. *Acta Paediatrica Hungarica*, 24, 315–321.
- Glass, D., Kelsall, H., Slegers, C., Forbes, A. B., Loff, B., Zion, D., & Fritschi, L. (2015). A telephone survey of factors affecting willingness to participate in health research surveys. *BMC Public Health*, 15, 1017. <https://doi.org/10.1186/s12889-015-2350-9>
- Hagenbeek, F. A., van Dongen, J., Pool, R., & Boomsma, D. I. (2022). Twins and omics: The role of twin studies in multi-omics. In A. D. Tarnoki, D. L. Tarnoki, J. Harris, & N. L. Segal (Eds.), *Twin research for everyone: From biology to health, epigenetics, and psychology* (pp. 547–584). Elsevier.
- Han, H. J., Munson, B., & Schlauch, R. S. (2021). Fundamental frequency range and other acoustic factors that might contribute to the clear-speech benefit. *The Journal of the Acoustical Society of America*, 149, 1685–1698. <https://doi.org/10.1121/10.0003564>
- Higuchi, N., Hirai, T., & Sagisaka, Y. (1997). Effect of speaking style on parameters of fundamental frequency contour. In J. P. H. Santen, J. P. Olive, R. W. Sproat, & J. Hirschberg (Eds.), *Progress in speech synthesis* (pp. 417–428). Springer.
- Hincks, R. (2005). Measuring liveliness in presentation speech. In *Proceedings of Interspeech 2005* (pp. 765–768). The International Speech Communication Association. <https://doi.org/10.21437/Interspeech.2005-353>
- Hirst, D., & Di Cristo, A. (Eds.). (1998). A survey of intonation systems. In *Intonation Systems. A survey of twenty languages* (pp. 1–44), Cambridge University Press.
- Imamura, R., Tsuji, D. H., & Sennes, L. U. (2003). *Fisiologia da laringe. Tratado de otorrinolaringologia*.
- Lindh, J., & Eriksson, A. (2007). *Robustness of long time measures of fundamental frequency* [Paper presentation]. Eighth Annual Conference of the International Speech Communication Association, Antwerp, Belgium.
- Loakes, D. (2006). Variation in long-term fundamental frequency: Measurements from vocalic segments in twins' speech. In *Proceedings of the 11th Australian International Conference on Speech Science & Technology* (pp. 205–210).
- Mathews, L. J., & Turkheimer, E. (2022). Three legs of the missing heritability problem. *Studies in History and Philosophy of Science*, 93, 183–191. <https://doi.org/10.1016/j.shpsa.2022.04.004>
- McDermott, J. H., Keebler, M. V., Micheyl, C., & Oxenham, A. J. (2010). Musical intervals and relative pitch: Frequency resolution, not interval resolution, is special. *The Journal of the Acoustical Society of America*, 128, 1943. <https://doi.org/10.1121/1.3478785>
- Oxenham, A. J. (2012). Pitch Perception. *Journal of Neuroscience*, 32, 13335–13338. <https://doi.org/10.1523/JNEUROSCI.3815-12.2012>
- Polderman, T. J., Benyamin, B., De Leeuw, C. A., Sullivan, P. F., Van Bochoven, A., Visscher, P. M., & Posthuma, D. (2015). Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature Genetics*, 47, 702–709. <https://doi.org/10.1038/ng.3285>
- Przybyla, B. D., Horii, Y., & Crawford, M. H. (1992). Vocal fundamental frequency in a twin sample: Looking for a genetic effect. *Journal of Voice*, 3, 261–6.
- Rilliard, A., de Moraes, J. A., Erickson, D., & Shochi, T. (2013). Social affect production and perception across languages and cultures – The role of prosody. *Leitura*, 51, 15–41.
- Sataloff, R. T. (1995). Genetics of the voice. *Journal of Voice*, 9, 16–19. [https://doi.org/10.1016/S0892-1997\(05\)80218-8](https://doi.org/10.1016/S0892-1997(05)80218-8)
- Sataloff, R. T. (1997). *Professional voice: The science and art of clinical care*. Singular.
- Signorello, R., Demolin, D., Henrich Bernardoni, N., Gerratt, B. R., Zhang, Z., & Kreiman, J. (2020). Vocal fundamental frequency and sound pressure level in charismatic speech: A Cross-gender and -language study. *Journal of Voice*, 34, 808.e1–808.e13. <https://doi.org/10.1016/j.jvoice.2019.04.007>
- Titze, I. R. (2000). *Principles of voice production*. National Center for Voice and Speech.
- Titze, I. R. (2006). *The physics of speaking: Exploration in the mechanics of phonation*. National Center for Voice and Speech.
- Trautmüller, H., & Eriksson, A. (1994). The size of F0 excursions in speech production and perception. *Working Papers, Lund University, Department of Linguistics and Phonetics*, 43, 136–139. <https://www.researchgate.net/publication/277868264>
- Uchiyama, R., Spicer, R., & Muthukrishna, M. (2022). Cultural evolution of genetic heritability. *Behavioral and Brain Sciences*, 45, e152. <https://doi.org/10.1017/S0140525X21000893>
- van den Berg, J. (1958). Myoelastic-aerodynamic theory of voice production. *Journal of Speech and Hearing Research*, 1, 227–244. <https://doi.org/10.1044/jshr.0103.227>
- van Lierde, K. M., Vinck, B., De Ley, S., Clement, G., & Van Cauwenberge, P. (2005). Genetics of vocal quality characteristics in monozygotic twins: A multiparameter approach. *Journal of Voice*, 19, 511–518.
- Varella, M. A. C., Fernandes, E. de S., Fridman, C., Lucci, T. K., Defelipe, R. P., Fernandes, L. O., Antonio, A. L. O. G. L., Segal, N., & Otta, E. (in press). Determination of twin zygosity in Brazil: A DNA validation of two short questionnaires. *Estudos de Psicologia*.
- Vargas, A. C., Costa, A. G., & Hanayama, E. M. (2005). Perfil de extensão vocal em indivíduos falantes normais do português brasileiro. *Revista CEFAC*, 7, 108–116. <http://www.redalyc.org/articulo.oa?id=169320490015>
- Whiteside, S. P., & Rixon, E. (2003). Speech characteristics of monozygotic twins and a same-sex sibling: An acoustic case study of coarticulation patterns in read speech. *Phonetica*, 60, 273–297. <https://doi.org/10.1159/000076377>
- Whiteside, S. P., & Rixon, E. (2013). Speech tempo and fundamental frequency patterns: A case study of male monozygotic twins and an age- and sex-matched sibling. *Logopedics Phoniatrics Vocology*, 38, 173–181. <https://doi.org/10.3109/14015439.2012.742562>
- Xu, Y. (2012). Function vs. form in speech prosody—Lessons from experimental research and potential implications for teaching. In J. Romero-Trillo (Ed.), *Pragmatics and prosody in English language teaching* (pp. 61–76). Springer Netherlands.
- Zhang, Z. (2016). Mechanics of human voice production and control. *Journal of Acoustical Society of America*, 140, 2614–2635. <https://doi.org/10.1121/1.4964509>
- Zyphur, M. J., Zhang, Z., Barsky, A. P., & Li, W. D. (2013). An ACE in the hole: Twin family models for applied behavioral genetics research. *The Leadership Quarterly*, 24, 572–594.