Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Phonetics 32 (2004) 277–287

Journal of Phonetics

www.elsevier.com/locate/phonetics

Letter to the editor

Formant frequencies and body size of speaker: a weak relationship in adult humans

Julio González*

Departamento de Psicología Básica, Clínica y Psicobiología, Universidad Jaume I, Campus Riu Sec,
s/n, 12070 Castello, Spain

Received 2 December 2002; received in revised form 9 July 2003; accepted 15 July 2003

Abstract

This paper investigates the relationship between formant frequencies and body size in human adults. In Experiment I, correlation coefficients were obtained between acoustic correlates of the five Spanish vowels uttered by 82 speakers as a function of speakers' heights and weights. In Experiment II correlations were calculated from formant parameters obtained by means of a long-term average analysis of connected speech from 91 speakers. Results of both experiments showed that, in contrast to Fitch's (J. Acoust. Soc. Am. 102 (1997) 1213) findings in macaque vocalizations, the relationship within sex between formant parameters and body size is very weak in human adults. At the same time, it is evident that correlations within the female group are greater than in male group. These results imply that the pattern of individual vocal tract development is relatively free from skeletal size constraints, due to the human descent of larynx from standard mammal position. This disassociation of vocal tract-body size is more important in human males.

1. Introduction

Speaker normalization is a key component in speech perception, allowing listeners to disregard most of the speaker-specific aspects of the speech signal. There is experimental evidence to show that speaker-specific aspects of formants influence phonetic categorization (Ladefoged & Broadbent, 1957; Lieberman, 1984; Pisoni, 1997). More specifically, it is known that the overall formant pattern of speech is used for vocal tract normalization in a given speaker (Ladefoged & Broadbent, 1957).

---

*Tel.: +34-964-72-80-00; fax: +34-964-72-93-50.
E-mail address: gonzalez@psb.uji.es (J. González).

0095-4470/$ - see front matter
doi:10.1016/S0095-4470(03)00049-4

According to the commonly accepted source-filter theory of speech (Müller, 1848; Fant, 1960), the formants, or resonant frequencies of the supralaryngeal vocal tract, are dependent on the length of the vocal tube. The vocal tract is a part of the speakers body and listeners are able to make use of speaker-specific information in the acoustic signal to estimate the speaker weight and height (Lass & Davis, 1976; González & Oliver, submitted). Some authors (Fitch, 1994, 1997, 2000; Fitch & Giedd, 1999) state that the use of formant frequencies as a cue to the body size of a vocalizer has played an important role in the evolution of language. According to Fitch (1997, p. 1220) “if a mechanism for estimating body size from formant dispersion existed in our prelinguistic ancestors, this could have provided a preadaptative basis for vocal tract normalization in humans”.

A frequently cited acoustic parameter as a cue to body size is mean fundamental frequency (f0) of voice (e.g., Darwin, 1871; Morton, 1977; Laver & Trudgill, 1979). In fact, f0 shows a reliable negative correlation with body size across sex and age classes: f0 is higher in children and females and lower in males (on average, with taller and heavier bodies). However, when individual differences are studied by controlling age and sex variables, data has repeatedly shown that there is no correlation within sex between f0 and body size in adult humans (Lass & Brown, 1978; Künzel, 1989; Van Dommelen, 1993; Hollien, Green, & Massey, 1994). Formant dispersion has recently been proposed as a new acoustic cue to body size (Fitch, 1994, 1997). This parameter is defined as the average distance between each adjacent pair of formants. In a study of vocalizations in rhesus macaques, Fitch (1997) found that formant dispersion—defined as the average difference between successive formant frequencies—was closely tied to both vocal tract length (VTL) and body size. Fitch’s data showed two facts: the link between formant dispersion and VTL—explained by the basic assumptions of the source-filter theory applied to the macaque vocalizations, and the link between VTL and body size as a result of a tight anatomical correlation between both elements. The author concludes that unlike f0, formant dispersion could provide a robust cue to body size in most mammals.

Thus far, no study has addressed the degree to which formants and body size are related in human adults. In contrast to Fitch’s (1997) findings in rhesus macaques, the few available studies have only found a very weak link between formants and body size (Van Dommelen & Moxness, 1995; Collins, 2000). The purpose of this paper is to assess the relationship within sex between formant frequencies (and derived measures) and body size (weight/height, and derived measures) of adult speakers. This relationship is of interest for its theoretical and applied (forensic) implications. To this end, two experiments were performed with different types of speech materials: vowels and running speech.

2. Experiment I: vowels

2.1. Method

2.1.1. Subjects

Eighty-two Spanish subjects participated in this experiment, 27 males and 55 females, students at the University Jaume I of Castellón (Spain), with ages ranging from 20 to 30. The height (in cm) and weight (in kg) of each speaker were measured in the Health Service of the University. Means,

Table 1
Mean, standard deviation (S.D.) and range values of the height and weight of male and female subjects from both experiments

	Experiment I			Experiment II		
	Mean	S.D.	Range	Mean	S.D.	Range
Males						
Height (cm)	177.3	6.2	160–189	177.0	5.6	162–189
Weight (kg)	75.5	9.6	62–102	74.3	9.0	62–102
Females						
Height (cm)	163.7	7.4	150–187	163.0	6.6	150–187
Weight (kg)	58.4	8.6	41–80	58.0	7.8	40–80

standard deviations and range values are shown in Table 1. One-sample Kolmogorov–Smirnov tests showed that the distribution of each dataset did not differ significantly from a normal distribution. The means were very close to those presented for the general population. The estimated means of height and weight for Spanish people aged 21 years are: 178 cm and 75 kg for males, and 164 cm and 59 kg for females (SENC, 2000).

2.1.2. Materials

A sustained phonation of each Spanish vowel (/a/, /e/, /i/, /o/, /u/) was recorded from each subject at a comfortable level in a quiet room. The recording (A/D) was performed at 50 kHz with a Shure SM58 microphone at a distance of about 12 cm from the mouth, directly in a CSL-Computerized Speech Lab Model 4300 (Kay Elemetrics Corp.). Then, the samples were downsampled to 12.5 kHz, converted to WAV files and copied in a PC-Pentium for acoustical analysis.

2.1.3. Acoustical analysis

Formant center frequencies were obtained by means of the software package *Praat* v.3.9.2 (Boersma & Weenink, 1996). Both LPC analysis and spectrograms were used to determine the frequency of the first four formants from the first more stable 200 ms of each token—i.e., a portion with flat and steady-state formant contours. The formant tracks were obtained using the default values recommended by the authors of Praat: analysis based on the Burg algorithm of Press, Teukolsky, and Vetterling (1992), 10–12 poles, 25-ms Gaussian window, 10-ms time step between the centers of consecutive windows, and 0.50 pre-emphasis. The Gaussian window was chosen—instead of the more common Hamming or Hanning windows—because according to the authors of Praat it is superior since it gives no sidelobes in the spectrograms. Formant tracks were superimposed on a broadband spectrogram performed with the following parameters: Fourier method (FFT), 5-ms Gaussian window, 2-ms time step, maximum frequency subject to analysis: 5000 Hz, bandwidth: 260 Hz, pre-emphasis: 6 dB/octave. Formant tier and formants from spectrograms were visually compared; when necessary, the number of poles was modified to get an optimal fit. General knowledge about the formant structure of each vowel was kept in mind in deciding the number of poles (e.g., the close proximity of f1 and f2 for the back vowels, or the close proximity of f2 and f3 for the anterior vowels). Once a good adjustment was reached between the formants calculated by LPC analysis and the formants in the spectrogram, f1−f4 were averaged across the most stable 200 ms section of each segment.

The fundamental frequency (f0) of each vowel was extracted from the same 200 ms of signal by means of a robust autocorrelation method described in Boersma (1993).

Reliability of measures. All analysis were repeated with the TF32 software (updated CSpeech, Milenkovic, 1989). Pearson correlations Praat vs TF32 across the averages of formants per vowel were f0: 0.99, f1: 0.97, f2: 0.99, f3: 0.88, f4: 0.79 for males; and f0: 0.99, f1: 0.98, f2: 0.99, f3: 0.80, f4: 0.86 for females.

2.2. Results

Formant and f0 values for males and females corresponded to those presented in previous studies (Quilis, 1981; Bradlow, 1995). Tables 2 and 3 show the Pearson product–moment

Table 2
Male speakers ($N = 27$): Pearson correlations between acoustic parameters (f0, f1−f4) of the five Spanish vowels and speaker height and weight

	Height						Weight					
	f0	f1	f2	f3	f4	*R*	f0	f1	f2	f3	f4	*R*
/a/	−0.21	−0.17	−0.19	−0.20	−0.13	—	−0.15	0.12	−0.14	−0.29	−0.39	—
/e/	−0.26	0.25	−0.57**	−0.34	−0.26	(f1,f2) = 0.65	−0.19	0.06	−0.06	0.15	−0.19	—
/i/	−0.23	−0.10	−0.37	−0.19	−0.19	—	−0.27	0.08	−0.14	−0.17	−0.03	—
/o/	−0.26	−0.08	−0.09	−0.25	−0.04	—	−0.27	0.21	−0.08	−0.18	−0.48*	(f4) = 0.48
/u/	−0.25	−0.04	0.25	−0.20	−0.25	—	−0.22	0.31	−0.08	−0.16	−0.30	—
Means[a]	−0.25	−0.10	−0.40*	−0.32	−0.23	(f2) = 0.40	−0.22	0.20	−0.10	−0.18	−0.33	—

R: Multiple regression from f1 to f4 (stepwise method).

*Significant at the 0.05 level.

**Significant at the 0.01 level.

[a] Coefficients calculated from the parameter means obtained for each subject across the five vowels.

Table 3
Female speakers ($N = 55$)

	Height						Weight					
	f0	f1	f2	f3	f4	*R*	f0	f1	f2	f3	f4	*R*
/a/	0.10	−0.20	−0.36**	−0.07	−0.10	(f2) = 0.36	−0.09	−0.20	−0.29*	0.00	−0.20	(f2) = 0.29
/e/	0.04	0.16	−0.51**	−0.49**	−0.18	(f2,f3) = 0.57	−0.12	0.02	−0.50**	−0.34*	−0.11	(f2) = 0.50
/i/	0.05	−0.04	−0.48**	−0.32*	−0.25	(f2) = 0.48	−0.09	0.02	−0.49**	−0.39**	−0.20	(f2) = 0.49
/o/	0.20	0.16	0.21	−0.18	−0.12	—	−0.01	0.03	−0.09	−0.12	−0.13	—
/u/	0.08	0.15	0.01	−0.07	−0.25	—	−0.08	0.05	−0.03	−0.08	−0.28*	(f4) = 0.28
Means[a]	0.10	−0.08	−0.51**	−0.29*	−0.22	(f2) = 0.51	−0.08	−0.08	−0.47**	−0.24	−0.24	(f2) = 0.47

See the legend of Table 2 for the explanation of symbols.

[a] Coefficients calculated from the parameter means obtained for each subject across the five vowels.

coefficients between the acoustic parameters (f0, f1−f4) extracted from the five Spanish vowels and the heights and weights of male and female speakers. Also, for each subject the mean of each parameter across the five vowels was obtained and the corresponding correlations were calculated.

In agreement with previous findings, no f0 coefficient was statistically significant. Regarding the formant frequencies, only 17 (18%) of the 96 coefficients were significant: 3 for male and 14 for female speakers. All the coefficients were lower than 0.60 (in absolute values), and in all these cases the explained variance was less than 36%. It seems that the most informative parameters for female height and weight were the second and the third formants from the anterior (/i/, /e/) and open-central (/a/) vowels. In male speakers, the second formant of /e/ is strongly correlated with height, and the fourth formant of /o/ with weight. When all the data were pooled (males + females), the percentage of significant coefficients is 93%.

For each vowel in each sex group, a linear multiple regression analysis was performed to predict the height (and weight) from the formant values. Due to possible multicollinearity between independent variables, a stepwise method of selection of variables was chosen. The selection of variables was based in the F-test to consider the increase in R^2 when a variable is entered into a regression equation that already contains the other independent variables. Criteria for adding and removing variables were $F \leqslant 0.05$ and $F \geqslant 0.10$, respectively. R values and selected variables are shown in Tables 2 and 3. In general, the female group presents most significant and higher coefficients. An exception is the /e/ vowel in the male group, which stands out from the other vowel types as it reached a high correlation, mainly due to the second formant.

The following derived measures from formant frequencies were extracted for each subject and each vowel: averages of f1−f2, f1−f3, f1−f4, and f2−f3; differences between consecutive formants; formant dispersions calculated from f1 to f3, and from f1 to f4 (Fitch, 1994, p. 1216). The corresponding correlations with height and weight were obtained on all these parameters. The two highest significant coefficients for male speakers were $r = -0.62$ (f2 − f1 difference in /e/ vs height) and $r = -0.51$ (dispersion of f1−f4 in /o/ vs weight). The corresponding results in female speakers were $r = -0.57$ (average of f1−f3 in /e/ vs height, and average of f2−f3 in /e/ vs height) and $r = -0.53$ (average of f1−f2 in /e/ vs height).

All the correlations were re-calculated with respect to the Log_{10}(Weight), body mass index (BMI $= w/h^2$) and the body surface area (BSA $= (wh)^{1/2}/6$; Mosteller, 1987) where w and h are the weight and height in kg and metres respectively. In general, coefficients were not higher than those obtained separately from the height and weight data.

3. Experiment II: running speech

Correlations obtained in Experiment I varied appreciably depending on the vowel identity. In this experiment we used a long speech segment to disregard the effect of the vowel identity. The analysis technique of a long-term average spectrum (LTAS) was applied to get the characteristic resonances (formants) of each speaker's vocal tract. This type of analysis is useful for speaker identification (Hollien & Majewski, 1977) and recently it has proven to be valuable for vocal tract resonance analysis (Linville & Rens, 2001).

3.1. Method

3.1.1. Subjects

Ninety-one subjects participated in this experiment, 29 males and 62 females, students at the University Jaume I of Castellón (Spain), with ages ranging from 20 to 30. Sixty of these subjects had participated in Experiment I. The new 31 subjects were bilingual since birth—Spanish/Catalan (Valenciano).

The height (in cm) and weight (in kg) of each speaker were measured in the Health Service of the University. Means, standard deviations and range values are shown in Table 1. One-sample Kolmogorov–Smirnov tests showed that the distribution of each dataset did not differ significantly from a normal distribution. As in Experiment I, the means were very close to those presented for the general population.

3.1.2. Materials

Each subject read aloud a paragraph of about 100 words. The paragraph was a Spanish text for 60 subjects and a Catalan (Valenciano) text for the bilingual 31 subjects. Both texts involved a great variety of phonemes in each language. Preliminary comparisons of acoustic parameters extracted from the LTAS did not yield significant differences across languages in f0, f2, f3, and f4. Only f1 showed a significant mean difference (f1 was 50 Hz higher in Catalan presumably because this language has two 'open' vowels that do not occur in Spanish). Therefore, acoustic data from both languages were pooled in a single dataset.

The paragraphs were recorded with a Shure SM58 microphone at a distance of about 12 cm from the mouth, and a Sony-TCD D-8 digital audiotape (DAT) recorder with a sample frequency of 44.1 kHz. Then, the voice signal was digitally transferred to a PC computer and converted to WAV files. Finally, the files were downsampled to 11 025 Hz to perform the LTAS analysis.

3.1.3. Acoustical analysis

In each speech sample an LTAS and an LPC-smoothing analysis were performed using Praat. LTAS was obtained using the default values recommended by the authors of Praat: Fourier method (FFT), 5-ms Gaussian window, 2-ms time step between the centers of consecutive analysis frames, maximum frequency subject to analysis: 5000 Hz, pre-emphasis: 6 dB/octave. The LPC-smoothing was derived from an analysis with 5 peaks for females, 6 peaks for males and a pre-emphasis coefficient of 0.5. By means of edit and draw tools it was verified that LPC profile matched well the average FFT spectrum. In a small number of cases the number of peaks of LPC smoothing was changed to get an optimal fit. The formant (resonance) frequencies were obtained from the first four peaks of the LPC trajectory (see example in Fig. 1).

Reliability of measures: Test–retest reliability was studied in a subsample of 32 tokens (35% of total). Retest measures of formant frequencies were obtained by means of the same procedure. Reliability was high since the acoustical analysis was a semi-automated method. Pearson correlations test–retest were: for f1 0.99, for f2 0.98, for f3 0.96, and for f4 0.90.

3.2. Results

Table 4 shows the correlations between the formant frequencies and the speakers' heights and weights. Dispersions were calculated from the first three and four formants and correlations with

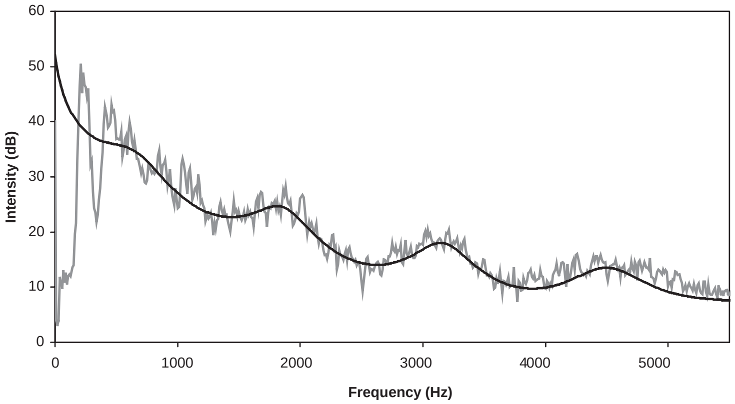

Fig. 1. Long term average spectrum with a superimposed smoothed LPC-spectrum (black line) of speech produced by a female speaker.

Table 4
Pearson correlations after applying an LTAS to running speech

	Formants					F. dispersions	
	f1	f2	f3	f4	*R*	(f1–f3)	(f1–f4)
Males							
Height	−0.15	−0.30	0.27	0.02	—	0.30	0.05
Weight	0.19	0.21	0.06	0.33	—	0.03	0.32
Females							
Height	−0.04	−0.30*	−0.34**	−0.32**	(f3) = 0.34	−0.34**	−0.39**
Weight	−0.27*	−0.34**	−0.16	−0.19	(f2) = 0.34	0.00	−0.06

The coefficients are calculated between acoustic parameters (formants, and formant dispersions) and speaker height and weight in males ($N = 29$), and females ($N = 62$).
R: Multiple regression from f1 to f4 (stepwise method).
* Significant at the 0.05 level.
** Significant at the 0.01 level.

height and weight were obtained. When males and females were pooled in a single dataset, all the coefficients were statistically significant. However, correlations within sex were weaker and all were lower than 0.40 (in absolute values).

As in Experiment I, correlations within females were higher than within males. Coefficients in the male group were not significant partly due to the lower sample size, but also because the values were smaller than in female group, for which seven coefficients reached significance. Results of a stepwise multiple regression analysis performed for each condition are also shown in Table 4. Criteria for adding and removing variables in the equation were the same as in Experiment I.

The recalculated correlation coefficients with respect to the Log_{10}(Weight), BMI, or BSA were not higher than those obtained with respect to the separated height and weight data.

4. General discussion

Unlike macaques and presumably other mammals, adult humans displayed a weak relationship between formant frequencies—and other derived measures, such as formant dispersion—and body size. Since according to the source-filter theory, formant frequencies and formant dispersion are correlated significantly and negatively with the vocal tract length (VTL), it seems that VTL is a weak predictor of the body size in human adults.

It is known that the increase in age from child to adult is marked by a decrease in formant frequencies (Eguchi & Hirch, 1969; Hillenbrand, Getty, Clark, & Wheeler, 1995; Huber, Stathopoulos, Curione, Ash, & Johnson, 1999; Lee, Potamianos, & Narayanan, 1999), and that there are differences in formant frequencies and size of the vocal tract between males and females (Peterson & Barney, 1952; Busby & Plant, 1995; Huber et al., 1999), but little is known of the relationship between formants and body size in adults of the same sex. Two studies have, however, addressed this issue. Van Dommelen and Moxness (1995) studied the ability to judge speaker height and weight from speech samples. In addition to body measures, several acoustic correlates were considered as independent variables. Regression analysis involving f0, formant frequencies, energy below 1 kHz, and speech rate yielded no significant correlations between these parameters and speaker height and weight (the only exception was between male speaker weight and speech rate). Recently, Collins (2000) investigated the relationship between male human vocal characteristics and female judgments about the speaker attractiveness. Body measures (weight, height, hip and shoulder width) and acoustic measures of five Dutch vowels uttered by 34 men (formant frequencies, overall peak and first five harmonic frequencies) were included as independent variables. Results showed that men with voices with low-frequency harmonics were judged as being more attractive, but there was no relationship between any vocal and body characteristic. In addition, regressions on body measures were performed with formant dispersion calculated from all formant frequencies, and from only formant frequencies 3 to 5 (less dependent of the vowel category). None of the regression analyses was significant.

The results of the present study have shown a weak relationship within sex between formant frequencies and body size in human adults. On the other hand, this relationship was greater for female speakers. Fitch (1997) found a strong correlation between formant dispersions and body size in 23 rhesus macaques, thereby demonstrating an acoustic link between vocal tract length (VTL) and body size. Fitch concluded that, unlike voice pitch (f0), formant frequencies—or more specifically, formant dispersions—seem a reliable predictor of body size in macaques and probably other mammal species. In these cases, VTL would be anatomically dependent upon skull size, which is in turn closely correlated with skeletal and body size. In this sense, contemporary human beings would be an exception, most likely because our species has abandoned the structural pattern typical of mammalian vocal tract anatomy.

It is possible to hypothesize two plausible explanations for the weak relationship between formant frequencies and body size in adult humans. (1) Laryngeal motor control is voluntary in humans, and formants can be greatly modified by changing the size and shape of the laryngeal tube. This is due to a direct neural supply from the motor cortex to the brainstem nuclei. On the other hand, in non-human primates, laryngeal control is involuntary and it is regulated by emotional circuitry (Jurgens, 1979). The acoustic consequence of laryngeal tube variation is not yet clear, but it may be responsible for a less strong dependence between formants and body size.

(2) In the absence of any analysis of human speech data, Fitch (1997) comments "the descent of the larynx from the standard mammalian position has the effect of elongating the vocal tract, possibly freeing human tract length from the skeletal size constraints [...], it is possible that the original function of the descending larynx in early hominids was to exaggerate body size. This idea gains support from the observation that there is sexual dimorphism in the degree to which the human larynx descends (a full vertebra lower in males than in females[...]), with no accompanying advantage for males' vowel clarity over that of females". Data of the present study were consistent with this idea, because the correlations between formant measures and body size were greater in female humans. In the vowel study (Experiment I), fourteen Pearson coefficients were significant in the female group, against three in the male group. When the possible effect of vowel identity was discarded (Experiment II), results obtained from an LTAS of connected speech showed clearly the difference between both sex groups: significant correlations emerged only within women. A power analysis showed that this difference was not attributable to a simple difference of sample size, and absolute values were clearly greater in the female coefficients. In men, the vocal tract may be more disassociated from skeletal (and body) size because the larynx has descended to a position deeper in the throat than in women. In fact, two descents of larynx occur throughout the men's life. First, between three months and three years of age, the larynx of all human infants recedes into the throat from the standard mammalian position (Negus, 1949; Lieberman, 1984; Crelin, 1987). Then, male humans undergo a second important descent of larynx during puberty, and they show an additional disproportionate vocal tract lengthening (Fitch & Giedd, 1999). This is likely to result in a pattern of vocal tract growth that is more independent from the skeletal structure. More specifically, the descent of the human larynx is realized by two human-specific morphological developments, one is the separation of the thyroid cartilage from the hyoid bone, and the other is the lowering of the hyoid bone (Nishimura, 2003). Non-human primates have a unified hyoid-larynx complex, resulting in a smaller freedom to alter the VTL. In any case, further research on individual differences within sex in adult humans regarding vocal tract morphology and its relationship with the body size will be necessary in order to obtain definitive conclusions.

Acknowledgements

The author would like to thank the reviewers (Jonathan Harrington, Wim A. van Dommelen, and two anonymous reviewers) for their helpful and valuable comments received on an earlier version of this paper. Specifically, a development of the conclusions presented in the General discussion about the differences between humans and non-human primates was suggested by one of the two anonymous reviewers. This work was supported by *Fundació Caixa Castelló-Bancaixa* and the Universitat Jaume I, Castellón, Spain, Project P1.1A2002-01.

References

Boersma, P. (1993). Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio in a sampled sound. *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam*, *17*, 97–110.

Boersma, P., & Weenink, D. (1996). *Praat: A system for doing phonetics by computer*. Technical Report 132, Institute of Phonetic Science of the University of Amsterdam.

Bradlow, A. R. (1995). A comparative acoustic study of English and Spanish vowels. *Journal of the Acoustical Society of America*, *97*, 1916–1924.

Busby, P. A., & Plant, G. L. (1995). Formant frequency values of vowels produced by preadolescent boys and girls. *Journal of the Acoustical Society of America*, *97*, 2603–2606.

Collins, S. A. (2000). Men's voices and women's choices. *Animal Behaviour*, *60*, 773–780.

Crelin, E. (1987). *The human vocal tract*. New York: Vantage.

Darwin, C. (1871). *The descent of man and selection in relation to sex*. London: Murray.

Eguchi, S., & Hirch, I. J. (1969). Development of speech sounds in children. *Acta Oto-Laryngologica Suppl.*, *257*, 1–51.

Fant, G. (1960). *Acoustic theory of speech production*. The Hague: Mouton.

Fitch, T. W. (1994). *Vocal tract length perception and the evolution of language*. Unpublished Ph.D. thesis, Brown University.

Fitch, T. W. (1997). Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *Journal of the Acoustical Society of America*, *102*, 1213–1222.

Fitch, T. W. (2000). Evolution of speech: A comparative review. *Trends in Cognitive Science*, *4*, 258–267.

Fitch, T. W., & Giedd, J. (1999). Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *Journal of the Acoustical Society of America*, *106*, 1511–1522.

González, J., & Oliver, J. C., Percepción a través de la voz de las características físicas del hablante: identificación de la estatura a partir de una frase o una vocal. *Revista de Psicología General y Aplicada*, submitted.

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America*, *97*, 3099–3111.

Hollien, H., Green, R., & Massey, K. (1994). Longitudinal research on adolescent voice change in males. *Journal of the Acoustical Society of America*, *96*, 2646–2653.

Hollien, H., & Majewski, W. (1977). Speaker identification by long-term spectra under normal and distorted speech conditions. *Journal of the Acoustical Society of America*, *62*, 975–980.

Huber, J. E., Stathopoulos, E. T., Curione, G. M., Ash, T. A., & Johnson, K. (1999). Formants of children, women and men: The effects of vocal intensity variation. *Journal of the Acoustical Society of America*, *106*, 1532–1542.

Jurgens, U. (1979). Neural control of vocalization in non-human primates. In H. D. Steklis, & M. J. Taleight (Eds.), *Neurobiology of social communication in primates: An evolutionary perspective* (pp. 11–44). New York: Academic Press.

Künzel, H. J. (1989). How well does average fundamental frequency correlates with speaker height and weight? *Phonetica*, *46*, 117–125.

Ladefoged, P., & Broadbent, D. E. (1957). Information conveyed by vowels. *Journal of the Acoustical Society of America*, *29*, 98–104.

Lass, N. J., & Brown, W. S. (1978). Correlational study of speakers' heights, weights, body surface areas, and speaking fundamental frequencies. *Journal of the Acoustical Society of America*, *63*, 1218–1220.

Lass, N. J., & Davis, M. (1976). An investigation on speaker height and weight identification. *Journal of the Acoustical Society of America*, *60*, 700–703.

Laver, J., & Trudgill, P. (1979). Phonetic and linguistic markers in speech. In S. Giles (Ed.), *Social markers in speech* (pp. 1–32). Cambridge: Cambridge University Press.

Lee, S., Potamianos, A., & Narayanan, S. (1999). Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *Journal of the Acoustical Society of America*, *105*, 1455–1468.

Lieberman, P. (1984). *The biology and evolution of language*. Cambridge, MA: Harvard University Press.

Linville, S. E., & Rens, J. (2001). Vocal tract resonance analysis of aging voice using long-term average spectra. *Journal of Voice*, *15*, 323–330.

Milenkovic, P. (1989). Cspeech 3.0 (Speech Analysis Software Program, University of Wisconsin, 1415, Johnson Drive, Madison, WI 53706).

Morton, E. S. (1977). On the occurrence and significance of motivation–structural rules in some bird and mammal sounds. *American Naturalist*, *111*, 855–869.

Mosteller, R. D. (1987). Simplified calculation of body surface area. *The New England Journal of Medicine*, *317*, 1098.

Müller, J. (1848). *The physiology of the senses, voice and muscular motion with mental faculties*. London: Walton and Maberly.

Negus, V. E. (1949). *The comparative anatomy and physiology of the larynx*. New York: Hafner.

Nishimura, T. (2003). Comparative morphology of the hyo-laringeal complex in anthropoids: Two steps in the evolution of the descent of the larynx. *Primates*, *44*, 41–49.

Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of vowels. *Journal of the Acoustical Society of America*, *24*, 175–184.

Pisoni, D. (1997). Some thoughts on "Normalization" in speech perception. In K. Johnson, & J. W. Mullennix (Eds.), *Talker variability in speech processing* (pp. 9–32). San Diego, CA: Academic Press.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery B, P. (1992). *Numerical recipes in C: The art of scientific computing*. Cambridge: Cambridge University Press.

Quilis, A. (1981). *Fonética acústica de la lengua española*. Madrid: Gredos.

SENC (2000). Estudio EnKid 1998–2000. Sociedad Española para el estudio de la Obesidad y Nutrición Comunitaria. Madrid.

Van Dommelen, W. A. (1993). Speaker height and weight identification: Reevaluation of some old data. *Journal of Phonetics*, *21*, 337–341.

Van Dommelen, W. A., & Moxness, B. H. (1995). Acoustic parameters in speaker height and weight identification: Sex-specific behaviour. *Language and Speech*, *38*, 267–287.