Perceptual and Motor Skills, 2007, 105, 215-220.

CORRELATIONS BETWEEN SPEAKERS' BODY SIZE AND ACOUSTIC PARAMETERS OF VOICE[1,2]

JULIO GONZÁLEZ

University Jaume I

Summary.—Generally, there is a significant relationship between some acoustic measures (F_0 and formant parameters) and the body size of speakers; however, data become less clear when age and sex variables are controlled. To date, no other vocal parameter apart from F_0 has been studied in relation to body size. In the present study, correlations between a set of 27 parameters of the Multi-dimensional Voice Program (Kay Elemetrics Corp.) and 4 body measures were obtained from 134 speakers of both sexes belonging to one age group (20–29 years). Correlations within sex groups were null or very weak, and all significant coefficients were below .35.

As humans grow, their vocal folds also increase in length and mass (Titze, 1989), which results in a lowering of the fundamental frequency (F0) of speech. Vocal tract length also increases (Fitch & Giedd, 1999), and the formant frequencies of the vowels decrease (Huber, Stathopoulus, Curione, Ash, & Johnson, 1999) as a result. Conversely and on the average, women are shorter and have smaller vocal folds and vocal tracts than men (Fitch & Giedd, 1999). Consequently, in general there are relationships between some acoustic measures of speech and the body sizes of speakers.

Data become less clear, however, when individual differences are studied and controlled by age and sex. Contrary to a common belief, the correlation within sex between F_0 and body size is null or very weak for human adults (Lass & Brown, 1978; Künzel, 1989; van Dommelen, 1993; Hollien, Green, & Massey, 1994; González, 2004; Rendall, Kollias, Ney, & Lloyd, 2005). In contrast to Fitch's findings (1997) in macaque vocalizations, the relationship within sex between formant parameters and body size is very weak in human adults (González, 2004; Rendall, *et al.*, 2005).

No other vocal parameters, such as jitter, shimmer, noise-to-harmonic ratio, etc. apart from F_0 have been studied with respect to their relationship with speakers' body size. Here, correlations are presented for a multidimensional set of 27 voice parameters and four body parameters (height, weight, Body Mass Index, body surface area) within a wide sample of people of the same age.

[1]Address correspondence to Julio González, Ph.D., Departamento de Psicología Básica, Clínica y Psicobiología, Universidad Jaume I, 12070 Castellón, Spain or e-mail (gonzalez@psb.uji.es).
[2]This study was partly supported by the Ministry of Science and Technology of Spain (I+D+I, Project BSO2003-01002/PSCE).

DOI 10.2466/PMS.105.1.215-220

Method

Speakers

Participants were 134 young adults of both sexes at the University of Jaume I of Castellón, of whom 53 were men (ages 20–29 years, $M = 22.0$, $SD = 2.2$) and 81 were women (ages 20–27 years, $M = 21.6$, $SD = 1.5$). Heights ranged between 162 and 197 cm ($M = 177.2$, $SD = 6.4$) for men, and 150 and 187 cm ($M = 163.6$, $SD = 6.3$) for women. Weights ranged between 52 and 102 kg ($M = 74.6$, $SD = 9.5$) for men, and 40 and 80 kg ($M = 57.2$, $SD = 7.1$) for women. The means were very close to the estimated means of height and weight for Spanish people ages 21 years (SENC, 2000) (178 cm and 75 kg for men; 164 cm and 59 kg for women). For each speaker, Body Mass Index ($BMI = w/h^2$) and body surface area (Mosteller, 1987) [$BSA = (wh)^{1.2}/6$] were calculated, where w and h are the weight and height in kilograms and meters, respectively.

Voice Samples and Analysis

The recording of voice samples was performed with a Shure SM58 microphone at an approximate distance of 15 cm from the mouth. The voice parameters were extracted with the Multi-dimensional Voice Program™ Model 4305 of Kay Elemetrics Corporation implemented in a Computerized Speech Lab (CSL Model 4300B, Kay Elemetrics Corp., Pine Brook, NJ, USA). This program was chosen because the prior results have shown its robustness and reliability (Kent, Vorperian, & Duffy, 1999; González, Cervera, & Miralles, 2002) and its value as a tool in both research and clinical applications (van As, Hilgers, Verdonck-de Leeuw, & Koopmans-van Beinum, 1998; Xue & Fucci, 2000; González, Cervera, & Llau, 2003; Kent, Vorperian, Kent, & Duffy, 2003; Butha, Patrick, & Garnett, 2004; Campisi, Low, Papsin, Mount, & Harrison, 2006).

Following the MDVP operations manual (Kay Elemetrics, 1993), the speakers were asked to produce a sustained phonation of the /a/ vowel for 3 sec. at a comfortable pitch and loudness. Samples were recorded at 50 kHz and stored directly in the host computer in a soundproof room at the University Laboratories.

Voice Parameters

By means of the MDVP software, the following parameters were obtained (Kay Elemetrics, 1993) for each participant.

Fundamental frequency parameters

F_0 (Hz): Average fundamental frequency of the voice sample
Fhi (Hz): Highest fundamental frequency of the voice sample
Fl_o (Hz): Lowest fundamental frequency of the voice sample
STD: Standard Deviation of the fundamental frequency
PRF: Phonatory F_0-Range in semitones

Frequency perturbation parameters

Jita (us): Absolute Jitter, or the period-to-period variability of the pitch in microseconds
Jitt (%): Jitter percent
RAP (%): Relative Average Perturbation, or the variability of the pitch with smoothing factor of 3 periods
PPQ (%): Pitch Perturbation Quotient, or the variability of the pitch with smoothing factor of 5 periods
sPPQ (%): Smoothed Pitch Perturbation Quotient, or the variability of the pitch with smoothing factor of 55 periods
vF_0 (%): The relative standard deviation of the fundamental frequency

Amplitude perturbation parameters

ShdB (dB): Shimmer in dB, or the period-to-period variability of the peak-to-peak amplitude
Shim (%): Shimmer percent
APQ (%): Amplitude Perturbation Quotient, or the variability of the peak-to-peak amplitude at smoothing of 11 periods
sAPQ (%): Smoothed Amplitude Perturbation Quotient, or the variability of the peak-to-peak amplitude at smoothing of 55 periods
vAm (%): Peak-Amplitude Variation, or the relative standard deviation of peak-to-peak amplitude

Noise parameters

NHR: Noise to Harmonic Ratio
VTI: Voice Turbulence Index, or the relative energy of high frequency noise
SPI: Soft Phonation Index, or ratio of the harmonic energy in the range 70–1600 Hz to the harmonic energy in the range 1600–4500 Hz

Tremor Parameters

Fftr (Hz): F_0-Tremor Frequency, or the frequency of the most intensive F_0-modulating component of tremor
Fatr (Hz): Amplitude Tremor Frequency, or the frequency of the most intensive amplitude-modulating component of tremor
FTRI (%): Frequency Tremor Intensity Index, or ratio of the F_0-tremor to the total frequency magnitude of the voice sample
ATRI (%): Amplitude Tremor Intensity Index, or ratio of the amplitude tremor to the total amplitude of the voice sample

Parameters of Subharmonic components

NSH: Number of Subharmonic Segments
DSH (%): Degree of Subharmonics, or relative evaluation of subharmonic to F_0 components in the voice sample

Parameters of Voice irregularities

NUV: Number of Unvoiced Segments
DUV (%): Degree of Voiceless, or relative evaluation of nonharmonic areas in the voice sample

Results and Discussion

Table 1 shows Pearson correlations for the 27 voice parameters with the four body parameters (height, weight, Body Mass Index, and body surface area) for all speakers and for both sexes separately. The strong negative correlations obtained across all subjects for F_0 and other fundamental frequency

TABLE 1
PEARSON CORRELATIONS AMONG VOICE PARAMETERS AND BODY MEASURES OF BODY MASS INDEX (BMI) AND BODY SURFACE AREA (BSA)

Measure[a]	All ($N=143$)				Men ($n=53$)				Women ($n=86$)			
	H[a]	W[a]	BMI	BSA	H	W	BMI	BSA	H	W	BMI	BSA
Fundamental Frequency Parameters												
F_0 (Hz)	–.65†	–.69†	–.42†	–.72†	–.25	–.14	.00	–.19	.06	–.17	–.22	–.13
Fhi (Hz)	–.64†	–.66†	–.39†	–.69†	–.26	–.16	–.02	–.21	.08	–.05	–.10	–.02
Flo (Hz)	–.62†	–.66†	–.41†	–.69†	–.24	–.09	.04	–.14	.00	–.20	–.22*	–.17
STD (Hz)	–.09	–.11	–.07	–.11	–.07	–.02	.02	–.04	.07	.05	.00	.06
PFR (semit.)	–.14	–.13	–.05	–.14	–.10	–.16	–.11	–.17	.07	.13	.10	.13
Frequency Perturbation Parameters												
Jita (us)	.07	.16*	.18*	.15*	.02	.09	.09	.08	.04	.24*	.24*	.22
Jitt (%)	–.18*	–.10	.03	–.13	–.03	.04	.06	.02	.06	.25*	.24*	.23*
RAP (%)	–.20*	–.11	.02	–.14	–.04	.03	.06	.01	.06	.25*	.23*	.23*
PPQ (%)	–.19*	–.10	.03	–.14	–.03	.03	.04	.01	.03	.23*	.24*	.20
sPPQ (%)	–.07	.02	.09	–.01	–.04	.17	.20	.14	.07	.19	.17	.18
vF_0 (%)	–.04	–.05	–.03	–.05	–.01	.04	.05	.03	.07	.05	.01	.06
Amplitude Perturbation Parameters												
ShdB (dB)	–.04	–.01	.03	–.03	–.15	–.06	.03	–.10	.15	.19	.11	.21
Shim (%)	–.03	.00	.04	–.01	–.15	–.06	.03	–.09	.15	.21	.13	.23*
APQ (%)	.06	.09	.10	.09	–.09	–.01	.05	–.04	.16	.22*	.14	.24*
sAPQ (%)	–.02	.09	.15*	.06	–.13	.13	.21	.08	.19	.31†	.21	.32†
sAm (%)	–.10	–.15*	–.12	–.15*	–.01	.04	.05	.03	.26*	.14	–.03	.19
Noise Parameters												
NHR	.25†	.27†	.17*	.28†	.04	.17	.15	.16	.11	.09	.02	.11
VTI	.21†	.21†	.10	.22†	.11	.27	.21	.27	.11	–.05	–.14	–.02
SPI	.01	–.03	–.05	–.02	–.08	–.29*	–.23	–.29*	–.03	–.01	.00	–.01
Tremor Parameters												
Fftr (Hz)	–.03	–.03	–.02	–.03	–.01	.06	.06	.05	.01	–.04	–.05	–.03
Fatr (Hz)	.03	–.05	–.09	–.03	.09	–.04	–.10	–.01	.03	–.05	–.08	–.03
FTRI (%)	–.02	–.02	–.01	–.03	–.16	.09	.18	.03	.07	.05	.01	.06
ATRI (%)	–.11	–.14	–.10	–.14	–.05	.05	.08	.03	.19	.08	–.05	.12
Parameters of Subharmonic Components												
DSH (%)	–.02	–.09	–.10	–.07	–.06	–.21	–.18	–.21	.20	.10	–.03	.15
NSH	–.01	–.08	–.09	–.07	–.06	–.21	–.18	–.21	.20	.10	–.03	.15
Parameters of Voice Irregularities												
DUV (%)	.04	.10	.10	.09	–.08	.04	.08	.01	–.06	.01	.06	–.01
NUV	.04	.10	.10	.09	–.08	.04	.08	.01	–.06	.01	.06	–.01

[a]See text; H = Height, W = Weight. *$p < .05$ (t test, two tails). †$p < .01$ (t test, two tails).

parameters were almost exclusively due to variations between the sexes. In agreement with previous literature, the correlations within sex between F_0 and body size were quite weak, ranging from 0 to –.26. F_0 (and Fhi, Flo) was related with height for men but not for women. On the contrary, F_0 and Flo were related with weight parameters (Weight, Body Mass Index) only for women.

Overall, all significant correlations within sex were weak. Stronger correlations were found for men between SPI and VTI parameters and Weight and body surface area. The Soft Phonation Index (SPI) is related to the spectral slope, being the ratio between the lower-frequency harmonic energy (70–1600 Hz) to the higher-frequency harmonic energy (1600–4500 Hz). Male speakers with higher values of Weight, Body Mass Index, and body surface area tend to have a less "soft" phonation with a more flat spectral slope (a lower SPI). The Voice Turbulence Index (VTI) measures the relative energy of high-frequency noise and is related with "breathiness" in phonation. Heavier males and those with larger body surface area values could have a voice with more breathiness (a higher VTI).

Among the women, a significant positive correlation was found between some Frequency Perturbation parameters (Jitta, Jitt, RAP, sPPQ) and their Weight, Body Mass Index, and body surface area values. Heavier women and those with a higher Body Mass Index and body surface area tend to present more irregularities in their voice pitch. In addition, some significant correlations were observed within the Amplitude Perturbation domain, particularly a relatively stronger relation between sAPQ and the Weight and body surface area variables. The smoothed Amplitude Perturbation Quotient (sAPQ) measures the long-term variability of the peak-to-peak amplitude with a smoothing factor of 55 periods. This parameter could be related with amplitude voice tremors (Deliyski, Orlikoff, & Kahane, 1991; Kay Elemetrics, 1993), and according to our data, heavier and larger women would have higher sAPQ values than slimmer and shorter women. Finally, a significant correlation was found between Subharmonic parameters (DSH, NSH) and Height. Taller women would be more likely to have subharmonic components that are typical of diplophonic voices and voices with a glottal fry than shorter women.

REFERENCES

Butha, T., Patrick, L., & Garnett, J. D. (2004) Perceptual evaluation of voice quality and its correlation with acoustic measurements. *Journal of Voice*, 18, 299-304.

Campisi, P., Low, A. J., Papsin, B. C., Mount, R. J., & Harrison, R. V. (2006) Multidimensional voice program analysis in profoundly deaf children: quantifying frequency and amplitude control. *Perceptual and Motor Skills*, 103, 40-50.

Deliyski, D., Orlikoff, R., & Kahane, J. (1991) Multi-dimensional acoustic analysis of spasmodic dysphonia. *Abstracts of ASHA Convention*, Atlanta, GA.

Fitch, W. T. (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, W. T., & 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. (2004) Formant frequencies and body size of speaker: a weak relationship in adult humans. *Journal of Phonetics*, 32, 277-287.

González, J., Cervera, T., & Llau, M. J. (2003) Acoustic analysis of pathological voices compressed with MPEG system. *Journal of Voice*, 17, 126-139.

González, J., Cervera, T., & Miralles, J. L. (2002) [Acoustic voice analysis: reliability of a set of multi-dimensional parameters]. *Acta Otorrinolaringológica Española*, 53, 256-268. [in Spanish]

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.

Huber, J. E., Stathopoulus, 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.

Kay Elemetrics. (1993) *Multi-dimensional Voice Program (MDVP)*. Pine Brook, NJ: Author. [Computer program]

Kent, R. D., Vorperian, H. K., & Duffy, J. R. (1999) Reliability of the Multi-dimensional Voice Program for the analysis of voice samples of subjects with dysarthria. *American Journal of Speech-Language Pathology*, 8, 129-136.

Kent, R. D., Vorperian, H. K., Kent, J. F., & Duffy, J. R. (2003) Voice dysfunction in dysarthria: application of the Multi-dimensional Voice Program. *Journal of Communication Disorders*, 36, 281-306.

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

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.

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

Rendall, D., Kollias, S., Ney, C., & Lloyd, P. (2005) Pitch (F_0) and formant profiles of human vowels and vowel-like baboon grunts: the role of vocalizer body size and voice-acoustic allometry. *Journal of the Acoustical Society of America*, 117, 944-955.

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

Titze, I. R. (1989) Physiologic and acoustic differences between male and female voices. *Journal of the Acoustical Society of America*, 85, 1699-1707.

van As, C. J., Hilgers, F. G., Verdonck-de Leeuw, I. M., & Koopmans-van Beinum, F. (1998) Acoustical analysis and perceptual evaluation of tracheoesophageal prosthetic voice. *Journal of Voice*, 12, 239-248.

van Dommelen, W. A. (1993) Speaker height and weight identification: re-evaluation of some old data. *Journal of Phonetics*, 21, 337-341.

Xue, S., & Fucci, D. (2000) Effects of race and sex on acoustic features of voice analysis. *Perceptual and Motor Skills*, 91, 951-958.

Accepted June 20, 2007.