Can blind persons accurately assess body size from the voice?

Article (Accepted Version)

BIOLOGY LETTERS

rsbl.royalsocietypublishing.org

Research

CrossMark click for updates

Cite this article: Pisanski K, Oleszkiewicz A, Sorokowska A. 2016 Can blind persons accurately assess body size from the voice? *Biol. Lett.* 20160063. http://dx.doi.org/10.1098/rsbl.2016.0063

Received: 22 January 2016
Accepted: 21 March 2016

Subject Areas:
behaviour, cognition, evolution, biophysics

Keywords:
vocal communication, size normalization, visual impairment, congenital blindness, formant, height

Author for correspondence:
Katarzyna Pisanski
e-mail: k.pisanski@sussex.ac.uk

Electronic supplementary material is available at http://dx.doi.org/10.1098/rsbl.2016.0063 or via http://rsbl.royalsocietypublishing.org.

THE ROYAL SOCIETY PUBLISHING

Animal behaviour

Can blind persons accurately assess body size from the voice?

Katarzyna Pisanski[1,2], Anna Oleszkiewicz[1] and Agnieszka Sorokowska[1,3]

[1]Institute of Psychology, University of Wroclaw, Wroclaw, Poland
[2]Mammal Vocal Communication and Cognition Research Group, School of Psychology, University of Sussex, Sussex, UK
[3]Smell and Taste Clinic, Department of Otorhinolaryngology, TU Dresden, Germany

KP, 0000-0003-0992-2477

Vocal tract resonances provide reliable information about a speaker's body size that human listeners use for biosocial judgements as well as speech recognition. Although humans can accurately assess men's relative body size from the voice alone, how this ability is acquired remains unknown. In this study, we test the prediction that accurate voice-based size estimation is possible without prior audiovisual experience linking low frequencies to large bodies. Ninety-one healthy congenitally or early blind, late blind and sighted adults (aged 20–65) participated in the study. On the basis of vowel sounds alone, participants assessed the relative body sizes of male pairs of varying heights. Accuracy of voice-based body size assessments significantly exceeded chance and did not differ among participants who were sighted, congenitally blind or lost their sight later in life. Accuracy increased significantly with relative differences in physical height between men, suggesting that both blind and sighted participants used reliable vocal cues to size (i.e. vocal tract resonances). Our findings demonstrate that prior visual experience is not necessary for accurate body size estimation. This capacity, integral to both nonverbal communication and speech perception, may be present at birth or may generalize from broader cross-modal correspondences.

1. Introduction Q1

The human voice can reliably communicate a host of ecologically relevant information about the speaker, including the speaker's body size. In particular, larger individuals with longer vocal tracts produce lower and more closely spaced formant frequencies (vocal tract resonances) [1], and as a result, formants reliably indicate body size in a number of mammalian species [2] including humans [3,4]. Several other voice parameters tied to sex hormone levels, including fundamental frequency (perceived as voice pitch), have been identified as potential indicators of human height, weight or body shape, particularly among men [5–7]. Indeed, vocal communication of body size may have been most relevant for our male ancestors, for whom largeness and physical dominance likely brought higher social and reproductive success [8].

Several studies have demonstrated that sighted human listeners can accurately assess men's relative body size from the voice alone, typically associating lower fundamental and formant frequencies with larger size [3,9,10]. However, how we acquire this ability remains unknown. One parsimonious possibility is that this ability is acquired through learning, following repeated audiovisual pairings of low voice frequencies with large bodies. However, this possibility is necessarily weakened by evidence that the human voice explains only a fraction of the variance in body size when sex and age are controlled [4], and that listeners, while fairly accurate, often use erroneous voice cues to judge body size at this level [3,10,11]. A second possibility is that listeners generalize broader sound–size relationships, such that large objects produce lower resonances, to the voice and body [3]. Similarly, systematic stereotypes linking low frequencies to masculinity, dominance and threat [8,12] may link these same vocal parameters to physical largeness [13,14]. These latter possibilities suggest that humans' ability to accurately assess body size from the voice may in fact be acquired without the need for visual input or is present at birth. This study, the first to examine voice-based body size estimation in a sample of blind persons, was designed to test this prediction.

2. Methods

(a) Participants

Ninety-one healthy adults (50 men, 41 women) participated in the study, including 28 congenitally or early blind (aged 24–65, mean $= 38.2 \pm 11.8$ years) and 40 late blind adults (aged 23–65, mean $= 48.7 \pm 10.7$ years). Following previous classifications of early and late blindness [15], early blindness was defined as a complete loss of vision before 2 years of age, i.e. before completion of visual development [16]. Blind participants had no residual vision, light perception or neurological impairments (for descriptive statistics detailing causes of vision loss in the late blind adults see electronic supplementary material, table S1). Twenty-three sighted adults participated as controls (aged 20–65, mean $= 39.2 \pm 14.3$ years). Blind and sighted participants were closely matched by age and sex. All participants reported normal hearing, provided written informed consent and were compensated for their participation.

(b) Voice stimuli

Thirty adult men were recorded speaking the monophthong vowels /a/ /i/ /ɛ/ /o/ and /u/ in a sound-controlled booth using a Sennheiser condenser microphone with cardioid pick-up pattern. Audio was digitally encoded at a sampling rate of 96 kHz and 32-bit amplitude quantization and stored onto a computer as WAV files. Voice stimuli were amplitude normalized to 70 dB RMS SPL in Praat [17] and then randomly paired to form 60 unique voice pairs, divided into four groups (15 voice pairs per group). The differences in height between men in the voice pairs ranged from 0 to 21 cm (mean $= 7.4 \pm 5.6$ cm) and did not differ across the four groups of voice stimuli (one-way ANOVA: $F_{3,56} = 0.067$, $p = 0.997$). In 70% of voice pairs, the taller man had lower and more closely spaced formants than did the shorter man.

(c) Experimental procedure

Following a standardized interview in which we collected personal and demographic information and confirmed the absence of injuries and disorders, participants were randomly assigned to assess the relative body size of one of four groups of voice stimuli. Participants completed the experiment in individual sessions wherein voices were presented via a custom computer interface and through Sennheiser HD 201 professional headphones. Each participant completed a total of 15 trials; the presentation order of trials and voices within each pair was randomized. On each trial, participants were presented with two men's voices and were asked to select which of the two voices belonged to the larger man. The experimenter executed the interface and inputted participants' verbal responses into the programme, which automatically loaded the next trial. To create identical testing conditions, sighted participants were asked to close their eyes during the experiment, and all participants were seated with their backs to the computer.

3. Results

A generalized linear model fitted with maximum-likelihood estimation was used to examine the proportion of accurate body size assessments (i.e. correctly identifying the taller of two men). Sight (sighted, late blind, congenitally or early blind) sex of listener (male, female) and stimulus group (1–4) were included as factors, and age of listener as a covariate. The model revealed no significant differences in the accuracy of body size assessments among participants who were sighted or blind (Wald $\chi^2_2 = 0.46$, $p = 0.79$; figure 1*a*). Listener sex $\chi^2_1 = 0.33$, $p = 0.56$), listener age ($\chi^2_1 = 1.02$, $p = 0.31$) and stimulus group ($\chi^2_3 = 1.8$, $p = 0.62$) did not affect performance, and removing these variables from the omnibus model did not change the pattern of results (i.e. no effect of sight: $\chi^2_2 = 1.6$, $p = 0.92$). Models including two-way (all $\chi^2 < 2.0$, all $p > 0.16$) and three-way relationships (all $\chi^2 < 2.8$, all $p > 0.83$) showed no interactions among any of the factors. Mean accuracy of body size assessments significantly exceeded chance (0.5) for sighted ($p = 0.01$), late blind ($p = 0.002$) and congenitally or early blind participants ($p = 0.035$), as indicated by two-way non-parametric binomial tests (figure 1*a*).

A logit model was used to regress counts of accurate size assessments against the relative difference in height between men in each given voice pair (log transformed and excluding negligible height differences less than or equal to 0.5 cm), with sight included as a factor (goodness-of-fit, likelihood ratio $\chi^2 = 234.38$, d.f. $= 142$, $p < 0.001$). The logistic regression indicated that accuracy of size assessments increased significantly with relative differences in body size ($Z = 2.2$, $p = 0.037$, 95% CI: 0.10–0.91; figure 1*b*), and that sight had no effect on this relationship ($Z = -0.75$, $p = 0.46$, 95% CI: -0.93 to 0.42). Mean size assessment accuracy reached 87.8% correct (83% for sighted, 80% for late blind and 100% for congenitally or early blind participants) on trials in which the difference in height between men was maximal (21 cm).

4. Discussion

We demonstrate that blind men and women can accurately estimate relative differences in men's body size from the voice alone, with the same degree of accuracy as sighted adults. Listener's size assessment accuracy increased with the relative difference in height between the men whose voices were assessed. This finding indicates that both blind and sighted participants were using reliable vocal cues to size (i.e. formants/vocal tract resonances [1,4]). Prior visual experience is therefore not a necessary prerequisite for accurate body size estimation. The ability to judge body size from the voice may be learned through general correspondences linking low-frequency sounds to large size (e.g. in animal vocalizations or in the resonances produced by inanimate objects ([3,18] for discussion), may be acquired through non-visual cross-modal correspondences (e.g. pairing the sound of a person's voice with the height from which that voice is projected), and/or may have a strong innate component.

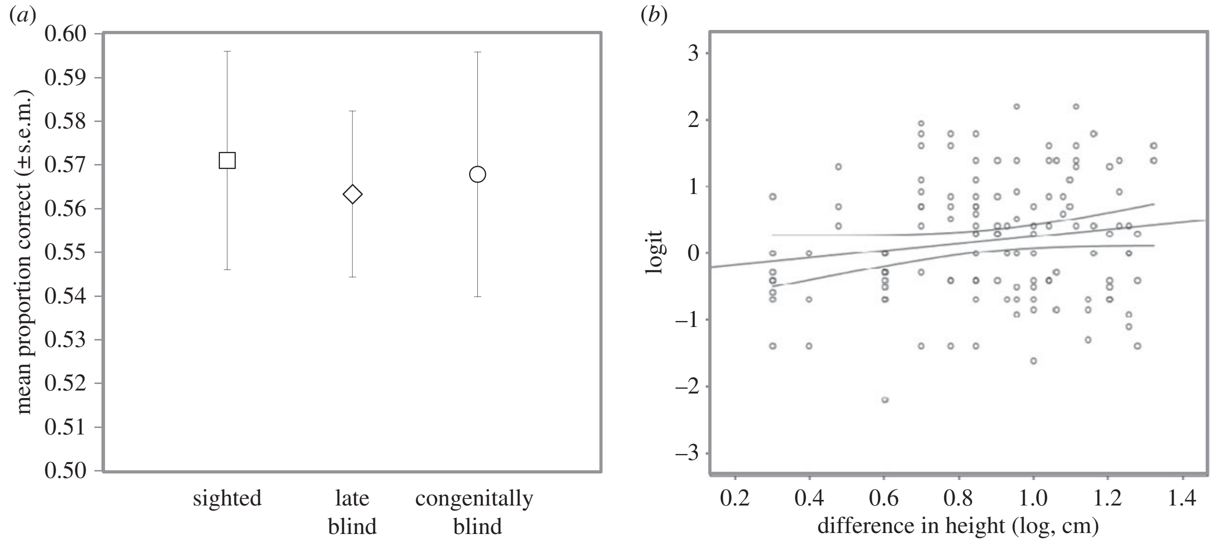

Figure 1. (*a*) Accuracy in voice-based assessments of relative body size (mean proportion correct size assessments) did not differ among sighted, late blind, and congenitally or early blind adults and exceeded chance (0.5) in all three groups; (*b*) for both sighted and blind participants, accuracy increased significantly as the relative difference in height between men in the voice pairs increased. Q5

Given a lack of visual information on which to rely, as well as subsequent structural reorganization of the auditory cortex following blindness [19], one might predict that blind persons will rely more strongly on vocal information during social communication compared with sighted persons, and may even show an advantage in voice perception tasks. Indeed, in the absence of direct visual cues, vocal estimates of body size are important for developing a mental representation of another person. Our results indicate that blind persons do not show an advantage in voice-based body size assessments of men. Similarly, previous studies suggest that although blind adults outperform their sighted counterparts in low-level auditory tasks testing spatial localization or pitch discrimination, blind persons generally do not show a significant advantage in voice recognition tasks (see [19] for review).

Voice-based estimation of body size has an important function not only for social communication, but also for speech recognition [1,20]. In addition to indicating body size [4], and other social characteristics such as dominance [8], changes in formant spacing produce different vowel sounds. To accurately segregate body size information from speech content produced by speakers with diverse vocal tract lengths, human listeners must first perform speaker 'size normalization' (see [21] for review). Size normalization occurs at an early stage in the auditory processing of speech and other sounds, indicative of a highly general, automatic and low-level mechanism [22,23]. Indeed, infants as young as four months of age are able to infer size-related information from vowel sounds [24].

This study is the first, to the best of our knowledge, to examine voice-based size estimation in blind persons as well as in an older, i.e. non-student sample of sighted or blind adults. Our results corroborate those reported for sighted student samples in which the accuracy of relative size assessments exceeded chance and increased with the magnitude of the height difference between speakers [3,10]. Our results show that this ability does not deteriorate with age. Previous studies report equivocal findings as to whether male listeners process size information differently than do female listeners [3,9,10]. In our study, listener sex had no effect. Sex differences in harmonic spacing [9,10] may, however, make it easier for listeners to estimate body size from women's than men's voices [3]. Thus, the authors are presently testing whether blind adults show any advantage or disadvantage when estimating women's body size from the voice.

Ethics. The study was performed in accordance with the American Psychological Association's ethical standards in the treatment of human participants and was approved by the Ethical Committee of the Institute of Psychology, University of Wroclaw (project no. 2013/11/B/HS6/01522).

Data accessibility. The datasets supporting this article have been uploaded as electronic supplementary material.

Authors' contributions. All authors contributed to the conception and design of the experiment; K.P. programmed the experiment; A.S. and A.O. collected the data. The paper was drafted by K.P. and critically reviewed and approved by all authors, who agree to be accountable for the work.

Competing interests. The authors report no competing interests.

Funding. We received no funding for this study.

Acknowledgements. We thank Malgorzata Szagdaj, Anna Szagdaj, Katarzyna Gwozdziewicz, Natalia Wernecka, Anna Trzepizur and Joanna Widomska for assisting in participant recruitment and data collection. This work was supported by a scholarship from the Polish Ministry of Science and Higher Education (2013–2016) and National Science Center OPUS grant no. (2013/11/B/HS6/01522) to A.S.

References

1. Fitch WT. 2000 The evolution of speech: a comparative review. *Trends Cogn. Sci.* **4**, 258–267. (doi:10.1016/S1364-6613(00)01494-7)
2. Taylor AM, Reby D. 2010 The contribution of source-filter theory to mammal vocal communication research. *J. Zool.* **280**, 221–236. (doi:10.1111/j.1469-7998.2009.00661.x)
3. Rendall D, Vokey JR, Nemeth C. 2007 Lifting the curtain on the wizard of Oz: biased voice-based impressions of speaker size. *J. Exp. Psychol. Hum. Percept. Perform.* **33**,

1208–1219. (doi:10.1037/0096-1523.33. Q2 5.1208)

4. Pisanski K *et al.* 2014 Vocal indicators of body size in men and women: a meta-analysis. *Anim. Behav.* **95**, 89–99. (doi:10.1016/j.anbehav.2014.06.011)
5. Pisanski K, Jones BC, Fink B, O'Connor JJM, DeBruine L, Roder S, Feinberg DR. 2015 Voice parameters predict sex-specific body morphology in men and women. *Anim. Behav.* **112**, 13–22. (doi:10.1016/j.anbehav.2015.11.008)
6. Evans S, Neave N, Wakelin D, Hamilton C. 2008 The Q4 relationship between testosterone and vocal frequencies in human males. *Physiol. Behav.* **93**, 783–788. (doi:10.1016/j.physbeh.2007.11.033)
7. Q3 Hughes SM, Harrison MA, Gallup GGJr. 2009 Sex-specific body configurations can be estimated from voice samples. *J. Soc. Evol. Cult. Psychol.* **3**, 343. (doi:10.1037/h0099311)
8. Puts D, Apicella CL, Cardenas RA. 2012 Masculine voices signal men's threat potential in forager and industrial societies. *Proc. R. Soc. B* **279**, 601–609. (doi:10.1098/rspb.2011.0829)
9. Charlton BD, Taylor AM, Reby D. 2013 Are men better than women at acoustic size judgements? *Biol. Lett.* **9**, 20130270. (doi:10.1098/rsbl.2013.0270)
10. Pisanski K, Fraccaro PJ, Tigue CC, O'Connor JJM, Feinberg DR. 2014 Return to Oz: voice pitch facilitates assessments of men's body size. *J. Exp. Psychol. Hum. Percept. Perform.* **40**, 1316–1331. (doi:10.1037/a0036956)
11. Bruckert L, Lienard J-S, Lacroix A, Kreutzer M, Leboucher G. 2006 Women use voice parameters to assess men's characteristics. *Proc. R. Soc. B* **273**, 83–89. (doi:10.1098/rspb.2005.3265)
12. Morton ES. 1977 On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *Am. Nat.* **111**, 855–869. (doi:10.1086/283219)
13. Pisanski K, Mishra S, Rendall D. 2012 The evolved psychology of voice: evaluating interrelationships in listeners' assessments of the size, masculinity, and attractiveness of unseen speakers. *Evol. Hum. Behav.* **33**, 509–519. (doi:10.1016/j.evolhumbehav.2012.01.004)
14. Ohala JJ. 1984 An ethological perspective on common cross-language utilization of F0 of voice. *Phonetica* **41**, 1–16. (doi:10.1159/000261706)
15. Rombaux P, Huart C, De Volder AG, Cuevas I, Renier L, Duprez T, Grandin C. 2010 Increased olfactory bulb volume and olfactory function in early blind subjects. *Neuroreport* **21**, 1069–1073. (doi:10.1097/WNR.0b013e32833fcb8a)
16. Wiesel TN. 1982 The postnatal development of the visual cortex and the influence of environment. *Biosci. Rep.* **2**, 351–377. (doi:10.1007/BF01119299)
17. Boersma P, Weenink D. 2015 *Praat: doing phonetics by computer*.
18. Spence C. 2011 Crossmodal correspondences: a tutorial review. *Atten. Percept. Psychophys.* **73**, 971–995. (doi:10.3758/s13414-010-0073-7)
19. Kupers R, Ptito M. 2014 Compensatory plasticity and cross-modal reorganization following early visual deprivation. *Neurosci. Biobehav. Rev.* **41**, 36–52. (doi:10.1016/j.neubiorev.2013.08.001)
20. Pisanski K, Cartei V, McGettigan C, Raine J, Reby D. 2016 Voice modulation: a window into the origins of human vocal control? *Trends Cogn. Sci.* (doi:10.1016/j.tics.2016.01.002)
21. Johnson K, Mullennix JW. 1997 *Talker variability in speech processing*. Morgan Kaufmann Publishers Inc.
22. Irino T, Patterson RD. 2002 Segregating information about the size and shape of the vocal tract using a time-domain auditory model: the stabilised wavelet-Mellin transform. *Speech Commun.* **36**, 181–203. (doi:10.1016/S0167-6393(00)00085-6)
23. Smith DRR, Patterson RD, Turner R, Kawahara H, Irino T. 2005 The processing and perception of size information in speech sounds. *J. Acoust. Soc. Am.* **117**, 305–318. (doi:10.1121/1.1828637)
24. Peña M, Mehler J, Nespor M. 2011 The role of audiovisual processing in early conceptual development. *Psychol. Sci.* **22**, 1419–1421. (doi:10.1177/0956797611421791)