Whistling shares a common tongue with speech: Bioacoustics from real-time MRI of the human vocal tract

Michel Belyk[1,2,3], Benjamin G. Schultz[3,4], Joao Correia[3,5], Deryk S. Beal[2,6], Sonja A. Kotz[3,7]

[1]Department of Speech, Hearing and Phonetics, University College London, UK.

[2]Bloorview Research Institute, Holland Bloorview Kids Rehabilitation Hospital, Toronto, Canada

[3]Faculty of Psychology and Neuroscience, University of Maastricht, Maastricht, The Netherlands

[4]Institute of Logic, Language, & Computation, University of Amsterdam, Amsterdam, The Netherlands

[5]Basque Center on Cognition, Brain and Language, Donostia-San Sebastian, Spain.

[6]Department of Speech-Language Pathology, University of Toronto, Toronto, Canada

[7]Department of Neuropsychology, Max Planck Institute for Human and Cognitive Sciences, Leipzig, Germany

RUNNING TITLE: Bioacoustics of whistling

Keywords: whistle, speech, communication, tongue, magnetic resonance imaging, evolution

Correspondence to:

Michel Belyk, Ph.D.

University College London

Speech, Hearing and Phonetic Sciences

London, United Kingdom

e-mail: belykm@gmail.com

Abstract

Most human communication is carried by modulations of the voice. However, a wide range of cultures has developed alternate forms of communication that make use of a whistled sound source. For example, whistling is used as a highly salient signal for capturing attention, can have iconic cultural meanings such as the cat-call, enact a formal code as in boatswain's calls, or stand as a proxy for speech in whistled languages. We used real-time magnetic resonance imaging to examine the muscular control of whistling to describe a strong association between the shape of the tongue and the whistled frequency. This bioacoustic profile parallels the use of the tongue in vowel production. This is consistent with the role of whistled languages as proxies for spoken languages, in which one of the acoustical features of speech sounds are substituted with a frequency modulated whistle. Furthermore, previous evidence that non-human apes may be capable of learning to whistle from humans suggests that these animals may have similar sensorimotor abilities to those that are used to support speech in humans.

Introduction

Whistling produces a loud and pitched sound that approximates a sine-wave. These sounds travel well over large distances [1] and are easy to discern from other biological sounds by the rare occurrence of pure tone sine waves in nature. These features have made whistling a viable alternative sound source for human communication when signal fidelity may be more important than signal complexity [2,3].

Whistling may be a more robust channel in contexts where the voice is unreliable, such as communication over long distances or in poor weather. For example, naval vessels maintain a traditional code of boatswain's calls, in which arbitrary combinations of whistles correspond to simple commands [4,5]. Furthermore, a number of cultures have developed whistled proxies of spoken language [6]. In these languages, the whistled frequency stands in for one of the acoustical features that would normally be carried by the voice [6–8]. Whistled languages encode less information from which to identify the intended speech sounds than voiced speech, but are more robust to long distance communication. The narrow frequency band of the whistle gives it more power per unit of spectral bandwidth, increasing its signal-to-noise ratio and the effective range of communication [1–3].

Whistles are physical phenomena that occur when airflow interacts with objects to produce a positive feedback loop. For example, the hole-tone whistle is produced when a jet of air passes through two constrictions [9,10]. The space between the constrictions forms a resonator that selectively amplifies particular frequencies. Pressure fluctuations at the surface of the jet

propagate backwards and are amplified according to the characteristics of the resonator. 11,12]. Ring-shaped vortexes are formed in the downstream jet at the frequency of the resonant cavity. The perceptual property of this periodic waveform is the pitch of the whistle.

Many whistled codes used by humans are produced with the aid of the hands or an instrument, but the most basic form of whistling is the bilabial whistle. Though common knowledge suggests that whistling is primarily determined by the action of the lips, the tongue has an active role. Shadle [11] hypothesised that the lips form a constriction through which a jet of air is forced and that a resonant cavity behind the lips and bounded by the tongue determines the frequency that is whistled.

The tongue is a muscular organ that is divided into extrinsic and intrinsic muscle groups [13,14]. The extrinsic lingual muscles originate in osseous structures, such as the mandible and hyoid bone, and insert in the body of the tongue with the primary function of changing the tongue's position. The intrinsic lingual muscles make up the body of the tongue itself and serve to reconfigure the shape of the tongue to produce the dexterous movements required by both swallowing and speech. The human tongue in particular receives dense and complex innervation, which may support fine motor-control [15–17]. The changing shape of the tongue is used during speech to create narrow constrictions in the oral cavity that divide the vocal tract into a series of resonant cavities [18–20]. Together these cavities selectively amplify a combination of frequency bands that encode vowel sounds of speech [20–22].

Two previous studies provide support for the role of the tongue in whistling. Kaburagi et al. [23] used magnetic resonance imaging to gather still images of one individual. Qualitatively, it appeared from these images that the configuration of the tongue varied by the frequency being whistled with gross similarity to vowel production. Azola et al. [24] gathered x-ray cineradiographic images from two individuals providing evidence that the space between the tip of the tongue to the incisors forms a resonant cavity as with speech.

We used real-time anatomical MRI to collect videographic data with high contrast between the tongue and the surrounding vocal tract. We modeled the cross-sectional shape of the tongue to quantify tongue movements in their entirety and applied functional principle components analysis to explore variation in the tongue shapes that whistlers employed. This approach provided a quantitative, generalizable, and holistic description of the tongue's role in whistling.

Methods

Participants

Six participants (three male, including authors MB and BS) with no speech-motor or auditory deficits were recruited from Maastricht University. Participants had varied cultural backgrounds including German, Dutch, Canadian, Australian, and American and ages ranging from 20 to 33.

Procedure

Each participant performed three sound production tasks while undergoing real-time magnetic resonance imaging (rtMRI). In separate runs, each participant was instructed to 1) whistle a continuous siren spanning the range

of frequencies that they could reliably produce, 2) whistle a chromatic scale of discrete notes over the same range, and 3) produce a whistle with conventionalizsed meaning (a “cat call” was familiar to all participants despite diverse cultural backgrounds). Participants were instructed to produce sound as part of a breath phrase of approximately eight seconds and to breathe normally. Participant’s heads were constrained by foam pillows in the MRI.

Real-time magnetic resonance imaging

Real-time MRI collects a series of anatomical images from a mid-sagittal slice of the head and neck (Figure 1). Images were collected on a Siemens 3T MAGNETOM Prisma Fit at the Maastricht Brain Imaging Centre with the LiveView pulse sequence [25]. T1-weighted images were collected with an acquisition time of 60ms (sampling rate 16.67 Hz) over a single mid-sagittal slice with thickness = 8 mm, in-plane resolution = 2 mm by 2 mm, field-of-view = 256 by 256 mm, repetition time = 2.58 ms, echo time = 1.64 ms, and flip angle = 8°. K-space was sampled over 125 radial spokes. Scan durations were controlled manually and ranged from 88 to 98 seconds per run. Two runs (call, discrete scale) from one participant were discarded due to scanner malfunction or poor signal-to-noise ratio in imaging data. A third run (discrete scale) from a separate participant was discarded due to poor audio recording quality.

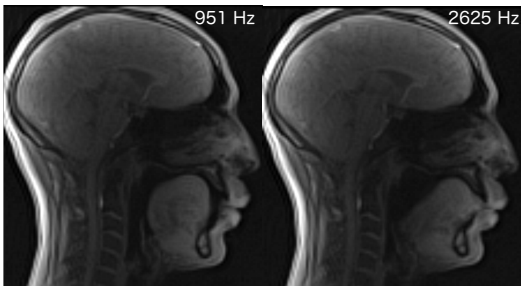

Figure 1: Still images of the highest and lowest frequency whistled by one participant whistling a continuous siren.

Acoustical measurement

Audio recordings were collected continuously throughout the scanning session using an MRI compatible microphone attached to the side of the head coil (sampling frequency 44100 Hz, quantization 32-bit). Recordings were anti-alias filtered by applying a low-pass filter at the sampling frequency prior to digitization. Audio and rtMRI data were synchronised by aligning the onset of acoustical artefacts associated with MRI acquisition with the first image volume.

Acoustical MRI artefacts where then removed using the noise reduction algorithm in Audacity [26]. The spectral profile of the MRI acoustical artefact was estimated from a period after the onset of the MRI related noise in each scanning run and before the onset of whistling and this noise component was down-weighted across the recording (noise reduction = 48 dB, sensitivity = 1.5, frequency smoothing = 3 bands). Two iterations of this procedure sufficiently filtered the acoustical waveform. Recorded waveforms were visually inspected in Praat (v6.0.36) [27] by an experienced acoustical analyst (MB) to identify remaining artefacts and omit corresponding time points from further analysis.

Whistling frequency measurements were extracted semi-automatically with an in-house Praat script. The script extracted the fundamental frequency by autocorrelation and calculated the mean frequency within a window equal to half the rtMRI sampling rate centered at each image acquisition.

Tongue shape measurement

The edge of the tongue was detected in each frame automatically using a custom MATLAB script [28,29]. A trace was then computed from tongue-edge maps using the tongue root as a reliably identifiable point of origin. This produced a continuous function of Y (anterior-posterior) and Z (ventral-dorsal) coordinates that capture the shape of the tongue. The coordinate values were centered to create an image space with the origin at the center of mass of the tongue in each frame for analytical purposes. Figures are plotted with origins at the tongue root to facilitate visualization.

Functional data analysis

Spatially smooth representations of the tongue contour were created by modeling each tongue trace with a B-spline with a basis set of cubic polynomials placed at every second sample along the trace using the *fda* package implemented in R [30,31]. Smoothing parameters were chosen by generalised cross validation. The length of each trace was normalised to the mean to remove the confounding influence of the cross-sectional size of the tongue and to ensure that tongue splines were modeled with a consistent number of knots.

Variation in tongue shape was explored using functional principal components analysis (fPCA) [32]. Functional PCA explores patterns of variation in the shapes of functions around a mean shape [33]. Much like discrete PCA, fPCA seeks principal components that maximize variation between observations [34–36]. The principal components of discrete PCA are eigenvectors that map each component back onto a set of discrete variables. The principal components of functional PCA are eigenfunctions that map each component back onto variations in shape (Supplementary file 1). FPCA was conducted simultaneously on functions of Y and Z coordinates to produce a two-dimensional description of tongue shape variation. This approach has the benefit of partitioning anterior-posterior (Y) and dorsal-ventral (Z) functions and yields separate Y and Z principal component subscores. These subscores can be interpreted as the degree to which the shape of tongue is deformed along the Y or Z plane in each MRI frame.

A separate examination of the functional principal components for each participant and each whistling task confirmed that the components were highly consistent across participants and tasks. The data were therefore combined and fitted to a linear mixed-effects model with the dependent variable of whistled frequency, with Y and Z subscores as regressors. The model accounted for random factors of Participant and Condition with random slopes for the effects Y and Z subscores at each level of the random factors [37]. This approaches the maximal random effects structure [38], though random effects of Condition were not nested within Participant due to a failure of more complex models to converge [Hz ~ 1+Y*Z + (1+Y*Z | Whistler) + (1+Y*Z | Condition)]. Significance was assessed by *F*-tests with degrees of freedom

214 determined by Satterthwaite's approximation for degrees of freedom, at an
215 alpha level of 0.05.

216 **Results**

217 The first two functional principal components (fPC) accounted for 62% and
218 17% of the total variance in tongue contour. These functional components
219 describe a dimension from 1) low-forward to high-back tongue position and 2)
220 high-forward to low-back tongue position (Figures 2 and 3). As each fPC
221 describes a distinct dimension of tongue shapes, we report separate models
222 for each principal component. Across both fPCs Y subscores were clearly
223 associated with whistled frequency. Supplementary materials 2-4 contain
224 representative MR videography clips synced with audio. This relationship is
225 summarised by the ß parameters reported below. These indicate the increase
226 in whistled frequency per increase of 1 unit of fPC score (equivalent to
227 adjacent tongue shapes in Figures 2 and 3).

228 Y subscores on the first functional principal component predicted the
229 frequency that was being whistled ($F(1, 8.5) = 6.5$, $p = 0.03$, ß = 399.1, *CI* =
230 [92.9, 705.3]). Z subscores were poor predictors of whistled frequency ($F(1,$
231 $5.5) = 2.1$, $p = 0.21$, ß = -200,1, *CI* = [-473.5, 75.3]), and no interaction was
232 apparent between subscores ($F(1, 8.4) = 0.01$, $p = 0.94$, ß = 4.5, *CI* = [-112.1,
233 121.0]). An anterior-ventral tongue position was associated with high
234 frequency whistling ($R^2 = 0.61$; Figure 2).

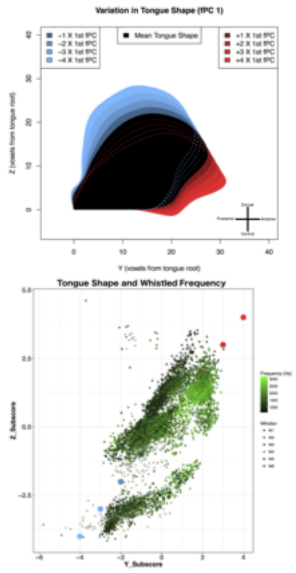

Figure 2: Top) Mean shape of the tongue (black) framed by shapes marking the first functional principle component (red to blue). Successive shades of red mark tongue shapes with fPC1 scores of +1 to + 4. Successive shades of blue mark tongue shapes with fPC1 scores of -1 to -4. Dashed lines continue each shaded area where they would otherwise be obscured. Bottom) Scatterplot showing Y and Z subcomponent scores of fPC1 for each frame. Colour hue indicates the frequency being whistled at each frame. Symbols indicate the whistler that contributed each point. Large background circles are fictive data points plotted for the purpose of facilitating the interpretation of fPC scores only. Each fictive point indicates the fPC1 score associated with the tongue shape of the same colour in the top panel. The origin corresponds to the mean tongue shape.

Y subscores on the second functional principal component also predicted the frequency that was being whistled ($F(1, 8.8) = 5.8$, $p = 0.04$, ß = 297.0, CI = [56.2, 537.8]). Z subscores were poor predictors of whistled frequency ($F(1, 5.6) = 1.9$, $p = 0.22$, ß = 222.7, CI = [-96.5, 541.8]) and no interaction was apparent between subscores ($F(1, 8.6) = 0.04$, $p = 0.84$, ß = 13.8, CI = [-115.2, 142.8]). An anterior-dorsal tongue position was associated with high frequency whistling ($R^2 = 0.60$; Figure 3).

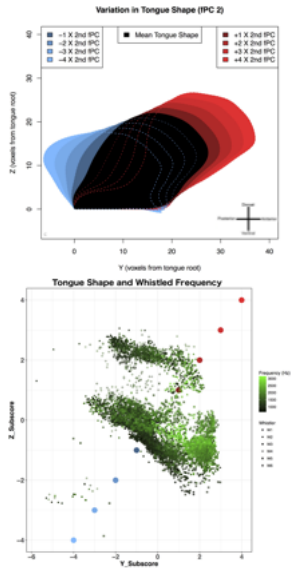

Figure 3: Top) Mean shape of the tongue framed by shapes describing the second functional principle component. Bottom) Scatterplot showing Y and Z subcomponent scores of fPC2 for each frame.

Discussion

We used rtMRI to demonstrate that the shape of the tongue is strongly associated with the frequency of bilabial whistling in humans, such that forward configuration of the tongue produced the highest frequencies regardless of tongue height. This mechanism was consistent across contexts, including simple but highly artificial siren sounds, music-like discrete chromatic scales, and complex calls with culturally imposed meaning. This is consistent with Shadle's hypothesis that the tongue shapes a resonant cavity behind the lips to determine the whistled frequency [11]. Tongue configurations that reduce the size of the resonant cavity between the incisors and the tongue amplify pressure fluctuations with shorter wavelengths (i.e., higher frequencies). Our approach modeled the entire cross sectional surface of the tongue. Though the posterior tongue may be ancillary to whistled sound production, its movement is none the less a part of whistled sound production.

A shared bioacoustical mechanism with speech

The same mechanism determines the frequency of whistling and the identities of spoken vowels. Vowel sounds are produced by shaping resonant cavities within the vocal tract [19,20,32]. These resonant cavities selectively amplify certain frequency bands of the voice, called formants, which together encode the identity of spoken vowels [18,39]. For, example a low-back tongue position produces a high first formant (F1) and low second formant (F2), as in the sound /ɑ/ (odd). A high-forward tongue position produces a low F1 and high F2, as in the sound /i/ (even). The most anterior of these resonant cavities, which determines the second formant in the context of speech, is a strong driver of whistled frequency. We observed whistled frequencies ranging from 600 Hz to 3100 Hz, which spans the values of the second formant that encode vowel sounds [40–42].

We observed two functional principal components of tongue shape: One capturing variation from low-back to high-forward tongue configurations and a second capturing variation from high-back to low-forward tongue configurations. Forward configurations of the tongue were associated with high frequencies across components, suggesting that multiple tongue configurations may produce similar bioacoustical effects. These alternative modes of production may provide an avenue to study motoric degrees of freedom in this muscular system [43].

The shared bioacoustics of whistling and vowel production may inform the study of whistled languages. Twelve whistled languages have been documented, though anecdotal reports suggest that they may be more abundant [2,6]. The most well studied of these is Silbo Gomero of the Canary

Islands, in which whistling is used as a sound source in place of the voice [44]. Silbadors produce loud hand-assisted whistles to communicate over long distances over mountainous terrain. They describe producing Silbao as whistling while moving ones tongue as though to pronounce words in spoken Spanish [44]. The effect is to approximate spoken Spanish with the whistled frequency standing in for the second formant (F2) of Spanish vowels [6,7]. Similar whistled-proxies have been described of French [45], Turkish [46], and Greek [47], among other languages [2,6]. Though the simpler acoustical structure of whistling encodes less information than the voice, even amateur whistlers are highly precise [48]. The common bioacoustical mechanisms of speaking and whistling may explain the emergence of whistled proxies across diverse languages and cultural groups.

A bioacoustical clue to the evolution of speech

Whistling may provide a novel avenue to understand the evolution of speech motor abilities through the comparative study of human and non-human apes. Though whistling has not been observed in non-human apes in the wild, at least one species (*Pongo* spp.) can learn to whistle in captivity [49,50]. In most instances it has not been possible to determine whether these animals spontaneously imitated their caretakers or were explicitly trained. In one case, this behaviour was observed to transfer between cohabitating animals, demonstrating the potential for cultural transmission [49]. This behaviour has provided evidence that orangutans have voluntary control over the upper lip, lower-lip, and respiratory muscles, which are readily accessible to external observation. Our study, along with that of Azola et al. [24], demonstrates the strong involvement of the tongue in human bilabial whistling. Whistling in non-

325 human apes may provide a useful animal model for the study of sensorimotor
326 capacities that support speech. Dynamic imaging in non-human apes is
327 needed to confirm that the tongue is similarly involved when these species
328 whistle, in line with broad similarities in vocal tract anatomy [51].

329 **Conclusions**

330 The tongue is a strong determinant of the frequency of oral whistling, with
331 forward tongue configurations associated with higher frequencies. This lingual
332 component of whistling corresponds with the bioacoustical mechanism that
333 produces the second formant in vowel production. This finding is consistent
334 with the link between whistled languages and the spoken languages for which
335 they act as proxy. Comparative research with non-human apes that have
336 learned to whistle may provide further insights into the evolution of the lingual-
337 motor skills that support speech.

338 **Ethics**
339 This study was approved by the Ethical Review Committee for Psychology and Neuroscience
340 at Maastricht University.
341
342 **Data Accessibility**
343 Data and code are available from the Dryad repository (doi.org/10.5061/dryad.kb56cd1).
344
345 **Author Contributions**
346 M.B. conceived of the study, designed the study, analysed the data, and drafted the
347 manuscript; B.G.S and J.C. helped design the study and provided analytical support, D.B.
348 contributed to drafting the manuscript. S.A.K. contributed to the design of the study. All
349 authors contributed comments and critical revisions.

350 **Competing Interests**
351 The authors report no competing interests.
352
353 **Funding**
354 This work was funded by grants from the Natural Sciences Research Council of Canada
355 [PDF502954-2017] and the Kimmel Family Opportunity Fund, and the Biotechnical and
356 Biological Sciences Research Council of the UK [BB/M009742/1].
357
358 **Acknowledgements**
359 We thank Nicolas Iuorio for engineering support.

360 **Figure Legends**

361
362

References

1. Wiley RH, Richards DG. 1978 Physical constraints on acoustic communication in the atmosphere: Implications for the evolution of animal vocalizations. *Behav. Ecol. Sociobiol.* **3**, 69–94. (doi:10.1007/BF00300047)

2. Meyer J. 2004 Bioacoustics of human whistled languages: An alternative approach to the cognitive processes of language. *An. Acad. Bras. Cienc.* **76**, 405–412. (doi:/S0001-37652004000200033)

3. Meyer J. 2005 Whistled speech: A natural phonetic description of languages adapted to human perception and to the acoustical environment. In *9th European Conference on Speech Communication and Technology*, pp. 49–52.

4. Manwaring GE. 1922 The boatswains whistle. *Mar. Mirror* **8**, 98–101. (doi:10.1080/00253359.1922.10655133)

5. Lieutenant A. 1911 The boatswain's call. *Mar. Mirror* **1**, 9–15. (doi:10.1080/00253359.1911.10654462)

6. Rialland A. 2005 Phonological and phonetic aspects of whistled languages. *Phonology* **22**, 237–271. (doi:10.1017/S0952675705000552)

7. Meyer J. 2008 Typology and acoustic strategies of whistled languages: Phonetic comparison and perceptual cues of whistled vowels. *J. Int. Phon. Assoc.* **38**, 69–94. (doi:10.1017/S0025100308003277)

8. Shadle CH. 1981 The acoustics of whistling. *J. Acoust. Soc. Am.* **70**, S12. (doi:10.1119/1.2341241)

9. Chanaud RC. 1970 Aerodynamic whistles. *Sci. Am.* **222**, 40–47.

10. Wilson TA, Beavers GS, DeCoster MA, Holger DK, Regenfuss MD. 1971 Experiments on the fluid mechanics of whistling. *J. Acoust. Soc. Am.* **50**, 366–372. (doi:10.1121/1.1912641)

11. Shadle CH. 1983 Experiments on the acoustics of whistling. *Phys. Teach.* **21**, 148–154. (doi:10.1119/1.2341241)

12. Henrywood RH, Agarwal A. 2013 The aeroacoustics of a steam kettle. *Phys. Fluids* **25**, 107101. (doi:10.1063/1.4821782)

13. Sonntag CF. 1925 The comparative anatomy of the tongues of the Mammalia. –XII. Summary, classification and phylogeny. *J. Zool.* **95**, 701–761.

14. Sanders I, Mu L. 2013 A three-dimensional atlas of human tongue muscles. *Anat. Rec.* **296**, 1102–1114. (doi:10.1002/ar.22711)

15. Mu L, Sanders I. 2010 Human tongue neuroanatomy: Nerve supply and motor endplates. *Clin. Anat.* **23**, 777–791. (doi:10.1002/ca.21011)

16. Weddell G, Harpman JA, Lambley DG, Young L. 1940 The innervation of the musculature of the tongue. *J. Anat.* **74**, 255–267.

17. Stål P, Marklund S, Thornell LE, De Paul R, Eriksson P-O. 2003 Fibre composition of human intrinsic tongue muscles. *Cells Tissues Organs* **173**, 147–161. (doi:10.1159/000069470)

18. Vorperian HK, Kent RD. 2007 Vowel acoustic space development in Children: A synthesis of acoustic and anatomic data. *J. Speech Lang. Hear. Res.* **50**, 1510. (doi:10.1044/1092-4388(2007/104))

19. Stone M, Davis EP, Douglas AS, Aiver MN, Gullapalli R, Levine WS, Lundberg AJ. 2001 Modeling tongue surface contours from cine-MRI images. *J. Speech Lang. Hear. Res.* **44**, 1026. (doi:10.1044/1092-4388(2001/081))

20. Story BH, Titze IR, Hoffman EA. 1996 Vocal tract area functions from magnetic resonance imaging. *J. Acoust. Soc. Am.* **100**, 537–554. (doi:10.1121/1.415960)

21. Titze IR. 2008 Nonlinear source–filter coupling in phonation: Theory. *J. Acoust. Soc. Am.* **123**, 2733. (doi:10.1121/1.2832337)

22. Joliveau E, Smith J, Wolfe J. 2004 Tuning of vocal tractresonance by sopranos. *Nature* **427**, 116.

23. Kaburagi T, Shimizu T, Uezu Y. 2018 A morphological and acoustic analysis of the vocal tract during the act of whistling. *Acoust. Sci. Technol.* **39**, 198–206. (doi:10.1250/ast.39.198)

24. Azola A, Palmer J, Mulheren R, Hofer R, Fischmeister F, Fitch WT. 2018 The physiology of oral whistling: A combined radiographic and MRI analysis. *J. Appl. Physiol.* **124**, 34–39. (doi:10.1152/japplphysiol.00902.2016)

25. Zhang S, Block KT, Frahm J. 2010 Magnetic resonance imaging in real time: Advances using radial FLASH. *J. Magn. Reson. Imaging* **31**, 101–109. (doi:10.1002/jmri.21987)

26. Audacity Team. 2018 Audacity(R): Free audio editor and recorder. (v2.1.3). http://audacityteam.org

27. Boersma P, Weenink D. 2016 Praat: Doing phonetics by computer. (v6.0.36). http://www.fon.hum.uva.nl/praat/

28. Mathworks. 2017 MATLAB. (R2017a). https://www.mathworks.com/

29. Correia JM. 2017 Online tongue segmentation using rtMRI. http://github.com/joao-mendonca-correia/rtMRI_tongueTracking

30. Ramsay JO, Wickham H, Graves S, Hooker G. 2017 fda: Functional Data Analysis.

441 31. R Core Team. 2019 R: A language and environment for statistical
442 computing. (v3.6.0) http://cran.r-project.org/.

443 32. Slud E, Stone M, Smith PJ, Goldstein M. 2002 Principal components
444 representation of the two-dimensional coronal tongue surface.
445 *Phonetica* **59**, 108–133. (doi:10.1159/000066066)

446 33. Ramsay JO, Hooker G, Spencer S. 2009 *Functional Data Analysis With*
447 *R and MATLAB*. New York: Springer Science+Busness Media Inc.

448 34. Locantore N *et al.* 1999 Robust principal component analysis for
449 functional data. *Test* **8**, 1–73. (doi:10.1007/BF02595862)

450 35. Ramsay JO. 2000 Functional components of variation in handwriting. *J.*
451 *Am. Stat. Assoc.* **95**, 9–15. (doi:10.2307/2669518)

452 36. Levitin DJ, Nuzzo RL, Vines BW, Ramsay JO. 2007 Introduction to
453 functional data analysis. *Can. Psychol. Can.* **48**, 135–155.
454 (doi:10.1037/cp2007014)

455 37. Bates D, Maechler M, Bolker B, Walker S. 2015 Fitting linear mixed-
456 effects models using lme4. *J. Stat. Softw.* **67**, 1–48.
457 (doi:10.18637/jss.v067.i01)

458 38. Barr DJ, Levy R, Scheepers C, Tily HJ. 2013 Random effects structure
459 for confirmatory hypothesis testing: Keep it maximal. *J. Mem. Lang.* **68**,
460 255–278. (doi:10.1016/j.jml.2012.11.001)

461 39. Rakerd B, Verbrugge RR. 1985 Linguistic and acoustic correlates of the
462 perceptual structure found in an individual differences scaling study of
463 vowels. *J. Acoust. Soc. Am.* **77**, 296–301. (doi:10.1121/1.392393)

464 40. Hillenbrand J, Getty LA, Clark MJ, Wheeler K. 2005 Acoustic
465 characteristics of American English vowels. *J. Acoust. Soc. Am.* **97**,
466 3099–3111. (doi:10.1121/1.411872)

467 41. Adank P, Hout R Van, Smits R. 2004 An acoustic description of the
468 vowels of Northern and Southern Standard Dutch. *J. Acoust. Soc. Am.*
469 **116**, 1729–1738. (doi:10.1121/1.1779271)

470 42. Heald SLM, Nusbaum HC. 2015 Variability in vowel production within
471 and between days. *PLoS One* **10**, 1–14.
472 (doi:10.1371/journal.pone.0136791)

473 43. Bernstein NA. 1967 *The co-ordination and regulation of movements*.
474 Oxford: Pergamon Press.

475 44. Classe A. 1957 The unusal whistle language of the Canary Islanders:
476 'Bring me two horses', he whistled. *UNESCO Cour.* **11**, 30–32.

477 45. Busnel RG. 1964 Document sur une langue sifflée Pyrénénne.

478 46. Güntürkün O, Güntürkün M, Hahn C. 2015 Whistled Turkish alters
479 language asymmetries. *Curr. Biol.* **25**, R706–R708.

480 (doi:10.1016/j.cub.2015.06.067)

481 47. Charalambakis C. 1994 A case of whistled speech from Greece. In
482 *Themes in Greek linguistics: Papers from the 1st International*
483 *Conference on Greek Linguistics* (eds I Philippaki-Warburton, K
484 Nicolaidis, M Sifianous), pp. 389–396. Amsterdam & Philidelphia:
485 Benjamins.

486 48. Belyk M, Johnson JF, Kotz SA. 2018 Poor neuro-motor tuning of the
487 human larynx: A comparison of sung and whistled pitch imitation. *R.*
488 *Soc. Open Sci.* **5**, 171544. (doi:10.1098/rsos.171544)

489 49. Lameira AR, Hardus ME, Kowalsky B, de Vries H, Spruijt BM, Sterck
490 EHM, Shumaker RW, Wich SA. 2013 Orangutan (Pongo spp.) whistling
491 and implications for the emergence of an open-ended call repertoire: A
492 replication and extension. *J. Acoust. Soc. Am.* **134**, 2326–2335.
493 (doi:10.1121/1.4817929)

494 50. Wich SA, Swartz KB, Hardus ME, Lameira AR, Stromberg E, Shumaker
495 RW. 2009 A case of spontaneous acquisition of a human sound by an
496 orangutan. *Primates* **50**, 56–64. (doi:10.1007/s10329-008-0117-y)

497 51. Fitch WT, de Boer B, Mathur N, Ghazanfar AA. 2016 Monkey vocal
498 tracts are speech-ready. *Sci. Adv.* **2**, e1600723.
499 (doi:10.1126/sciadv.1600723)

500