

Female vocal tracts are optimized for speech

Tiena Danner¹ and Steven Moran^{*1,2}

^{*}Corresponding Author: steven.moran@unine.ch

¹Institute of Biology, University of Neuchâtel, Switzerland

²Linguistic Research Infrastructure, University of Zurich, Switzerland

Sexual dimorphism is pervasive in the animal kingdom, where it may lead to adaptations that are sexually selected for, but not necessarily fortuitous. Consider for example barn swallows. Males have long spines at their tail feathers, which impairs flying. However, this feature simultaneously signals strength and “good genes” because despite this impairment, they are successful at providing for their offspring. Female barn swallows do not have long spines, so their tail feathers are optimized “just” for flying. Here we make a similar argument regarding the human larynx, which is typically lower in men and results in an overall longer vocal tract and lower f_0 and formants. Although this feature may signal size or attractiveness (Fitch, 2000; Feinberg, Jones, Little, Burt, & Perrett, 2005), we show empirically that the lower larynx deteriorates the optimal articulation of vowel sounds. Our findings therefore support the hypothesis that the male vocal tract is an evolutionary trade-off between transmitting speech effectively *and* other factors, whereas female vocal tracts are optimized for speech.

1. Introduction

Sex differences in human vocal tract anatomy affect voice characteristics and speech acoustics (Simpson, 2002, 2009). The most pronounced anatomical difference is the lower position of the larynx in men, which results in a longer overall vocal tract, and consequently, lower and more densely spaced formant frequencies. However, it remains unclear how sexual dimorphism in laryngeal position interacts with the overall shape of the vocal tract and whether this entails systematic differences in articulatory strategies between sexes. To investigate these questions, we used geometric morphometrics (GM) to quantitatively analyze vocal tract shapes and their covariation with speaker-specific articulation and resulting acoustics.

GM is a quantitative framework originally developed in physical anthropology to analyze and compare shapes of biological objects independent of their position, orientation, and scale (Bookstein, 2018), e.g., to study the ontological shape differences in the human and Neanderthal cranium and mandible (Ponce de León & Zollikofer, 2001). Unlike traditional linear measurements that make it difficult to separate size from shape, GM considers the spatial configuration of multiple anatomical landmarks simultaneously, allowing shape differences to be

captured more comprehensively and analyzed statistically. This makes GM ideally suited for investigations of vocal tract morphology and speech articulation (Zollikofer, Haenni, Suter, & Ponce de León, 2011; Dediu, Moisik, Baetsen, Bosman, & Waters-Rist, 2021; Gully, 2021; Danner, 2025).

2. Materials & Methods

Here we co-analyzed vocal tract shapes, speaker-specific articulatory strategies, and acoustic data at synchronous time stamps to determine the morpho-phonetic differences between men and women. Our sample includes resting state vocal tract MRI images of 73 individuals (38f, 35m) and 32 (16f, 16m) of them producing the five target American English vowels [i, e, a, o, u] in carrier words in the publicly available USC MRI database (Lim et al., 2021).

To assess sex-related static and articulatory differences while controlling for overall size of the vocal tract, we conducted permutation-based multivariate analyses of covariance (MANCOVA) on vocal tract landmark configurations, i.e., shape regression (Mitteroecker & Gunz, 2009; Collyer & Adams, 2018; Adams, Collyer, Kaliontzopoulou, & Baken, 2022), with sex as a fixed factor and vocal tract size as a covariate. This allowed us to quantify the proportion of shape variation attributable to sex after accounting for overall vocal tract size differences in the sample. To examine potential sex differences in articulatory coordination, defined as the covariation between articulatory subsystems (e.g., tongue and palate) we applied partial least squares (PLS) analysis to tongue and palate landmark sets to test whether the strength or pattern of tongue–palate covariation differed between sexes (Mitteroecker & Gunz, 2009; Mitteroecker, Gunz, Neubauer, & Müller, 2012). Finally, to evaluate sex differences in how articulatory differences between vowels are reflected in their acoustic realizations, we computed pairwise Euclidean distances among vowels in articulatory space and the corresponding distances in acoustic space, and regressed these measures to quantify the strength of the relationship between articulation and acoustics in female and male speakers.

3. Results

Our results show no major *shape* differences between the “average” male and female vocal tract morphology (Fig. 1, A). Despite this overall similarity, shape regression indicates that both size of the vocal tract and sex explain small but significant proportions of total vocal tract shape variation (size: $r^2 = 0.03$, $p = 0.013$; sex: $r^2 = 0.05$, $p = 0.001$). Male vocal tracts are on average longer (and larger), primarily due to a lower larynx position and larger body size. Interestingly, we observe that within both sexes, there is substantial morphological integration between facial prognathism and larynx position, suggesting a developmental mechanism that ensures an “optimal” vocal tract length, i.e., evolutionary stabilization for vocal communication.

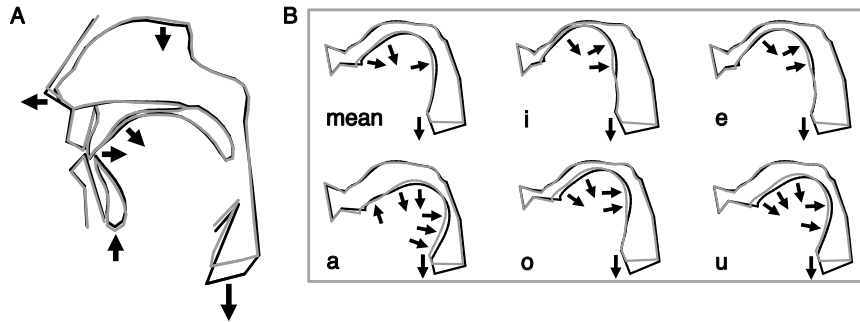


Figure 1. Sexual dimorphism in static vocal tract shape (A) and in the production of the five target vowels (B). Average female shapes in gray, average male shapes in black. Black arrows indicate major shape changes between sexes (female to male).

We find statistically significant and systematic articulatory differences in the production of vowels between sexes (Fig. 1B). Women and men differ in their articulatory configurations and in the coordination of their active and passive articulators. Shape regression of articulatory landmark configurations, while pooling data across vowels and controlling for centroid size of the vocal tract, shows that vowel identity explains the largest proportion of articulatory shape variation ($r^2 = 0.4$, $p = 0.001$), followed by sex ($r^2 = 0.06$, $p = 0.001$) and centroid size ($r^2 = 0.03$, $p = 0.001$). These results indicate robust sex-related differences in vowel articulation beyond overall vocal tract size effects. Moreover, sex differences in articulatory configuration vary across vowels: evaluating the same model separately for each vowel reveals stronger sex effects for back vowels ([a]: $r^2 = 0.13$; [o]: $r^2 = 0.11$; [u]: $r^2 = 0.15$) than for front vowels ([i]: $r^2 = 0.07$; [e]: $r^2 = 0.09$), with the effect for [i] not reaching statistical significance. In addition to these differences in articulator positioning, women exhibit stronger articulatory “integration” than men, reflected in tighter covariation between tongue shape and palate shape (women: slope = 0.67, $r^2 = 0.47$; men: slope = 0.43, $r^2 = 0.30$), indicating more tightly coordinated tongue–palate configurations. Here, articulatory integration is defined as the tendency of distinct anatomical regions to covary in a structured manner due to functional, developmental, or genetic linkage.

Likewise, the relationship between articulatory movements and their acoustic outcomes – i.e., how changes in tongue and palate configurations translate into changes in vowel sounds – is stronger in women. Linear regression analyses of Euclidean distances in articulatory versus acoustic space reveal higher slopes and greater explained variance for females (slope = 45.4, $r^2 = 0.34$, $p < 0.001$) than for males (slope = 30.0, $r^2 = 0.21$, $p < 0.001$). These results indicate that the relationship between articulatory and acoustic distances is stronger in women, with

higher regression slopes and greater explained variance (slope = 45.4, intercept = 12.1, $R^2 = 0.34$, $p < 0.001$) than in men (slope = 30.0, $R^2 = 0.21$, $p < 0.001$).

Thus, our results confirm earlier findings showing that men tend to move their articulators over a larger range to produce the same (or even less distinct) vowel quality contrasts (Weirich & Simpson, 2014). Women achieve a larger acoustic vowel space with smaller articulatory movements and men must reach more extreme articulatory positions, e.g., for low and back vowels [u, o, a] because of their lower larynx position.

4. Discussion

We hypothesize that the lower larynx position causes the observed articulatory differences and deteriorates the optimal articulation of vowel sounds by men. Although there appears to be a morphological balancing mechanism that ensures optimal vocal tract length within both sexes, male vocal tracts have to compensate more for their lowered larynx. This suggests that the transmission of speech sounds for men is at odds with other factors, including signaling size and attractiveness (Fitch, 2000; Feinberg et al., 2005), but also puts them at higher risk of pharyngeal collapse and for developing obstructive sleep apnea (Malhotra et al., 2002; Yamashiro & Kryger, 2012). Female vocal tracts on the other hand seem to be evolutionarily optimal for the transmission of speech, which has been previously hypothesized based on the results of articulatory computer models (De Boer, 2010). Here we provide empirical evidence to support this hypothesis, and we conclude that male vocal tracts are indeed an evolutionary trade-off between transmitting speech effectively *and* size signaling, whereas women's vocal tracts are optimized for speech.¹

Acknowledgments

TD and SM were funded by the Swiss National Science Foundation (PCEFP1_186841, EVOPHON). We also thank Valeriia Vyshnevetska and two anonymous reviewers for their comments and feedback.

References

- Adams, D., Collyer, M., Kaliontzopoulou, A., & Baken, E. (2022). *Geomorph: Software for geometric morphometric analyses. R package version 4.0.4*.
- Bannan, N., Dunbar, R., & Bamford, J. (2024). The evolution of gender dimorphism in the human voice: The role of octave equivalence. *Current Anthropology*, 65(3), 503–527.

¹An anonymous reviewer points out alternative hypotheses, e.g., differences in larynx position due to proposed evolutionary origins of singing for defensive signaling (Podlipniak, 2025) and to retain octave equivalence for social bonding (Bannan, Dunbar, & Bamford, 2024).

- Bookstein, F. L. (2018). *A course in morphometrics for biologists: Geometry and statistics for studies of organismal form*. Cambridge University Press.
- Collyer, M. L., & Adams, D. C. (2018). RRPP: An R package for fitting linear models to high-dimensional data using residual randomization. *Methods in Ecology and Evolution*, 9, 1772–1779.
- Danner, T. (2025). *Morpho-phonetics: A framework for studying speech and its evolution*. PhD thesis, University of Neuchâtel.
- De Boer, B. (2010). Investigating the acoustic effect of the descended larynx with articulatory models. *Journal of Phonetics*, 38(4), 679–686.
- Dediu, D., Moisik, S. R., Baetsen, W., Bosman, A. M., & Waters-Rist, A. L. (2021). The vocal tract as a time machine: Inferences about past speech and language from the anatomy of the speech organs. *Philosophical Transactions of the Royal Society B*, 376(1824), 20200192.
- Feinberg, D. R., Jones, B. C., Little, A. C., Burt, D. M., & Perrett, D. I. (2005). Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices. *Animal Behaviour*, 69(3), 561–568.
- Fitch, W. T. (2000). The evolution of speech: a comparative review. *Trends in Cognitive Sciences*, 4(7), 258–267.
- Gully, A. J. (2021). Quantifying vocal tract shape variation and its acoustic impact: A geometric morphometric approach. In *Proc. Interspeech 2021* (pp. 3999–4003).
- Lim, Y., Toutios, A., Bliesener, Y., Tian, Y., Lingala, S. G., Vaz, C., Sorensen, T., Oh, M., Harper, S., Chen, W., et al.. (2021). A multispeaker dataset of raw and reconstructed speech production real-time MRI video and 3D volumetric images. *Scientific Data*, 8(1), 187.
- Malhotra, A., Huang, Y., Fogel, R. B., Pillar, G., Edwards, J. K., Kikinis, R., Loring, S. H., & White, D. P. (2002). The male predisposition to pharyngeal collapse: importance of airway length. *American Journal of Respiratory and Critical Care Medicine*, 166(10), 1388–1395.
- Mitteroecker, P., & Gunz, P. (2009). Advances in geometric morphometrics. *Evolutionary Biology*, 36(2), 235–247.
- Mitteroecker, P., Gunz, P., Neubauer, S., & Müller, G. (2012). How to explore morphological integration in human evolution and development? *Evolutionary Biology*, 39(4), 536–553.
- Podlipniak, P. (2025). The role of musical pitch in long-distance defensive signaling. *Annals of the New York Academy of Sciences*, 1550(1), 206–226.
- Ponce de León, M. S., & Zollikofer, C. P. E. (2001). Neanderthal cranial ontogeny and its implications for late hominid diversity. *Nature*, 412(6846), 534–538.
- Simpson, A. P. (2002). Gender-specific articulatory–acoustic relations in vowel sequences. *Journal of Phonetics*, 30(3), 417–435.
- Simpson, A. P. (2009). Phonetic differences between male and female speech. *Language and Linguistics Compass*, 3(2), 621–640.

- Weirich, M., & Simpson, A. P. (2014). Articulatory vowel spaces of male and female speakers. In *Proceedings of the 10th international seminar on speech production* (pp. 449–452).
- Yamashiro, Y., & Kryger, M. (2012). Is laryngeal descent associated with increased risk for obstructive sleep apnea? *Chest*, *141*(6), 1407–1413.
- Zollikofer, C. P. E., Haenni, S., Suter, S. K., & Ponce de León, M. S. (2011). MRI-based morphometric analysis of the human vocal tract during speech formation and implications for fossil hominin vocal abilities. In *American Journal of Physical Anthropology* (Vol. 144, pp. 318–318).