Morphological Variation in the Adult Vocal Tract: A Study Using rtMRI

Adam Lammert, Michael Proctor and Shrikanth Narayanan
University of Southern California

Abstract

No two vocal tracts are exactly identical. Human vocal tracts display considerable morphological variation in terms of size, proportions and shape. Previous studies have focused on the dramatic variations present during development [1], [2], [3]. However, substantial differences also exist between fully developed individuals in terms of size, shape and proportions. We describe and quantify broad differences in morphology, explain certain systematic aspects in the variation (e.g., sexual dimorphism) and review some relevant characteristics of vocal tract structures.

Quantifying morphological variation is motivated by both practical and theoretical considerations. It has long been known that morphological differences affect acoustic output by altering the resonant properties of the vocal tract. Growing evidence also suggests that morphology affects production behavior. For example, palate shape explains certain differences in the production of sibilant fricatives [4], [5]. We ultimately hope to predict individual production patterns from observations of morphology. In addition, the utility of articulatory data (e.g., pooling data across subjects) is limited by poor understanding of morphological variation. Several studies have used general warping functions to impose normalization of various vocal tract landmarks [6]. We believe that a firm understanding of the space of variations will facilitate a normalization scheme which is more meaningful.

We collected realtime magnetic resonance imaging (rtMRI) data of 30 adult subjects (11 female, 19 male) from diverse linguistic (15 English, 8 German, 5 Mandarin, 1 Hindi) and racial (24 caucasian, 6 asian) backgrounds. Subjects spoke while lying supine in the scanner. The use of rtMRI allows subjects use their articulators naturally (i.e., there is no posing of the articulators). Images were reconstructed at a resolution of 68×68 pixels. Further details of our acquisition protocol can be found in [7]. For each subject, we identified images representing absolute rest position during breathing, with clenched teeth. We used Canny edge detection [8] with manual linking and correction to trace the hard

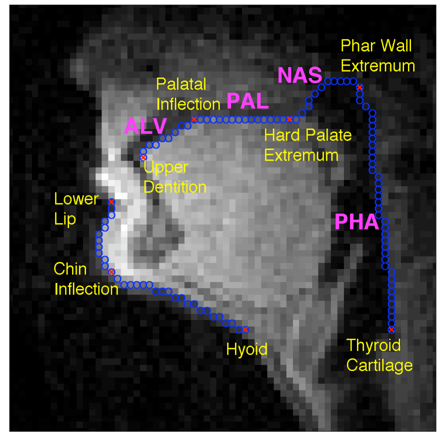

Fig. 1: *Hard structure traces and anatomical landmarks for one midsagittal vocal tract image.*

structures of the vocal tract, including (1) along the chin, and (2) along the passive articulators inside the vocal tract. The latter trace followed the upper dentition, the maxilla and palatine bones, along the posterior surface of the vomer, the pharyngeal tonsil and down the pharyngeal wall to the thyroid cartilage. Several anatomical landmarks were identified for the purposes of analysis (see Fig. 1).

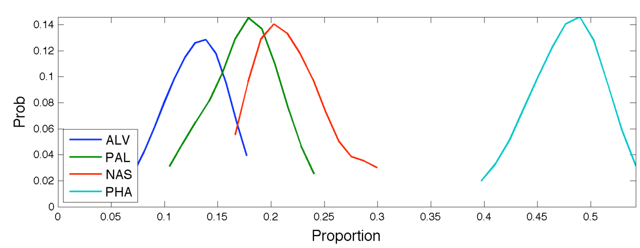

Fig. 2: *Distribution of the vocal tract proportions, corresponding to the segments shown in Figure 1.*

We calculated the length between each landmark as a proportion of the entire trace length. The distributions of these proportions for the interior trace are shown in Fig. 2. The distributions are highly unimodal, with the pharyngeal wall segment making up the nearly half of the entire length. Interestingly, pharyngeal wall length is proportionally longer in males than in females to a statistically significant degree (Student's t-test, $p = 0.045$). There was also a significant difference between chin proportions by

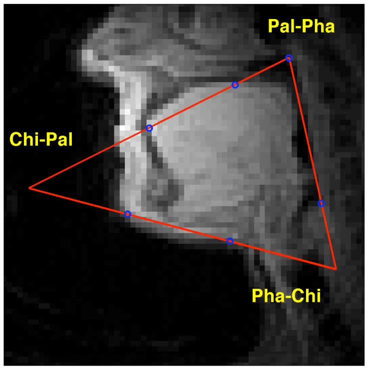

Fig. 3: *The triangle formed by one subject's anatomical landmarks, with characteristic angles labeled.*

sex, such that the anterior chin segment was proportionally longer in men (Student's t-test, $p = 0.025$).

We also explored the major characteristic angles by treating the vocal tract as a triangle whose sides are the palate, chin and pharyngeal wall (see Fig. 3). The angles of the triangle form the distributions in Fig. 4. A bimodal distribution can be seen in the Chin-Palate angle, with modes around 34 and 43 degrees. During teeth clenching, the chin and palate move as a unit and the angle between them is stable. The other two angles change as a function of nodding motion and spinal flexion, so defining a reference position for comparison is difficult. Our subjects were instructed to lay comfortably in the scanner, which we considered a natural definition of reference position.

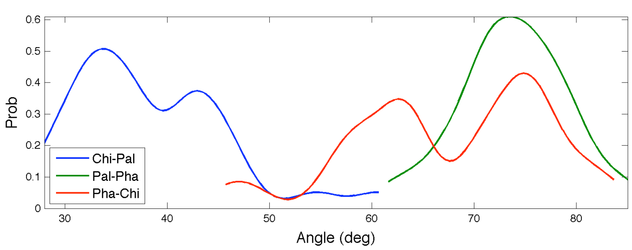

Fig. 4: *Distribution of characteristic angles, corresponding to the angles shown in Figure 3.*

Variations in shape for the palate and the rear pharyngeal wall were clustered into four categories, using the K-means algorithm. The results are displayed in Fig. 5. The clustered palates show considerable variety, from large circles (A) and small circles (D), to sharply angular (B) and low and flat (C). In contrast, the pharyngeal wall shapes reflect straightforward differences in curvature from nearly linear to quite concave. It should be noted that these differences in concavity are unrelated to spinal flexion or head orientation. The rear pharyngeal wall

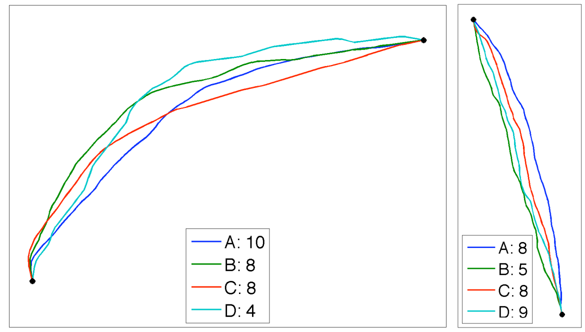

Fig. 5: *Cluster representations of the palatal (left) and pharyngeal (right) shapes. The legend shows the label and number of constituents of each cluster.*

is a solid mass of connective tissue which does not readily distend [9]. Indeed, there is essentially no correlation between the concavity of the pharyngeal wall (i.e., area of deviation from linear) and the Palate-Pharynx angle ($r = 0.01$, $p = 0.95$).

ACKNOWLEDGEMENTS

Work supported by NIH Grant DC007124. Thanks to Dr. Michael Proctor for linguistic expertise.

REFERENCES

[1] W. Fitch and J. Giedd, "Morphology and development of the human vocal tract: A study using magnetic resonance imaging," *Journal of the Acoustical Society of America*, vol. 106, no. 3, pp. 1511–1522, 1999.

[2] L. Boe, G. Captier, J. Granat, M. Deshayes, J. Heim, P. Birkholz, P. Badin, N. Kielwasser, and T. Sawallis, "Skull and vocal tract growth from fetus to 2 years," in *Eighth International Seminar on Speech Production*, 2008.

[3] H. Vorperian, S. Wang, M. Chung, E. Schimek, R. Durtschi, R. Kent, A. Ziegert, and L. Gentry, "Anatomic development of the oral and pharyngeal portions of the vocal tract: An imaging study," *Journal of the Acoustical Society of America*, vol. 125, no. 3, pp. 1666–1678, 2009.

[4] M. McCutcheon, A. Hasegawa, and S. Fletcher, "Effects of palatal morphology on /s, z/ articulation," *Journal of the Acoustical Society of America*, vol. 67, no. 1, pp. 94–94, 1980.

[5] S. Fuchs, P. Perrier, C. Geng, and C. Mooshammer, "What role does the palate play in speech motor control? insights from tongue kinematics for german alveolar obstruents," in *Towards a better understanding of production processes*, J. Harrington and M. Tabain, Eds. Psychology Press, 2006.

[6] J. Wei and J. Dang, "Morphological normalization of vocal tract shape," in *ICASSP 2010*, 2010.

[7] S. Narayanan, K. Nayak, S. Lee, A. Sethy, and D. Byrd, "An approach to real-time magnetic resonance imaging for speech production," *Journal of the Acoustical Society of America*, vol. 115, no. 4, pp. 1771–1776, 2004.

[8] J. Canny, "Computational approach to edge detection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, 1986.

[9] L. Penning, "Radioanatomy of upper airways in flexion and retroflexion of the neck," *Neuroradiology*, vol. 30, pp. 17–21, 1988.