

Resonance Characteristics of Hypopharyngeal Cavities

Kiyoshi Honda, Tatsuya Kitamura, Hironori Takemoto, Satoru Fujita & Parham Mokhtari

ATR Human Information Science Laboratories

{honda; kitamura; takemoto; fujita; parham}@atr.jp

Abstract

Hypopharyngeal cavities consisting of the laryngeal cavity and bilateral piriform fossa are located near the bottom of the vocal tract and influence vocal tract resonance in the higher frequencies. An MRI-based analysis on a male speaker revealed that the laryngeal cavity functions as a Helmholtz resonator to add a spectral peak at 3~3.5 kHz and that the piriform fossa causes troughs at 4~5 kHz. Based on the result, we propose an acoustic model of vowel production with hypopharyngeal cavity coupling and examine vowel-to-vowel variation of the hypopharyngeal resonance.

1. Introduction

The hypopharynx is a part of the vocal tract near the larynx that divides into three short tubes: the laryngeal cavity (vestibule) and the bilateral piriform fossa (sinuses), as shown in Fig. 1. Because these cavities are located near the closed end of the vocal tract, they exert significant acoustic effects on vowel spectra as reported by several researchers (Fant, 1960; Sundberg, 1987; Dang & Honda, 1996; Imagawa, et al., 2003). In this study, we aimed to evaluate acoustic characteristics of the hypopharyngeal cavities as possible factors deriving speaker characteristics.

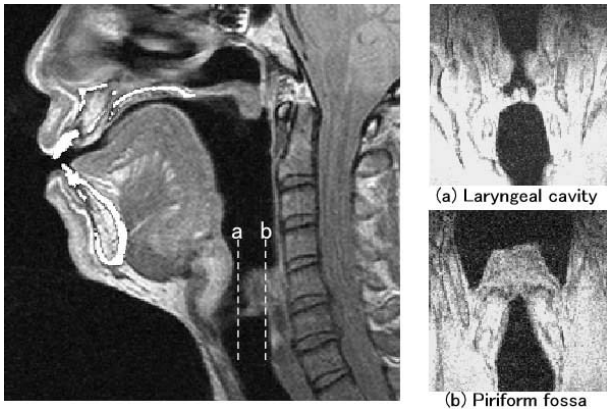


Fig. 1 Laryngeal cavity and piriform fossa in hypopharynx

We have also evaluated acoustic functions of the hypopharyngeal cavities in yielding individual characteristics of voices. Several experiments and simulations were performed based on three-dimensional MRI data obtained from male speakers producing the five Japanese vowels. 3D models of the whole vocal tract were made from these data to estimate the resonance characteristics using FEM and electric circuit simulations. Their mechanical replicas with a resin wall were used for acoustic experiments to directly measure

vowel spectra. The results obtained so far can be summarized as follows (Honda, et al., 2004; Kitamura, et al., 2004).

- Pharyngeal cavity volume influences the lower formants (F1, in particular).
- Laryngeal cavity functions as a Helmholtz resonator to enhance a spectral peak at a 3-3.5 kHz region.
- Piriform fossa provides short side-branches and causes spectral trough(s) at a 4-5 kHz region.

With these observations, we aim to develop an acoustic model of vocal characteristics. In this report, we analyze high-resolution MRI data to examine vowel-to-vowel variations in the resonance characteristics of hypopharyngeal cavities.

2. Method

The first author was a subject of MRI experiment to obtain 3D high-resolution images of the hypopharynx during production of the five Japanese vowels. MRI scans were performed using a custom larynx coil and phonation-synchronized scan method. The images of the lower pharynx were traced manually to reconstruct 3D geometry of the hypopharyngeal cavities. The data for the vowel /a/ and /i/ are shown in Fig. 2. The area functions of the hypopharynx were made to compute the frequency response of the cavities.

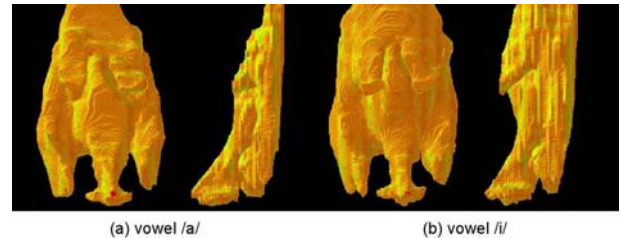


Fig. 2 Front and lateral view of pharynx in two vowels

3. Results

3.1. Laryngeal cavity

The laryngeal cavity forms a short tube above the vocal folds, which consists of the laryngeal ventricle and a narrow conduit (laryngeal tube) connecting the ventricle to the pharynx. This shape resembles a Helmholtz resonator with a long neck, and its resonance frequency is calculated to be about 3.5 kHz, which roughly agrees with the result from acoustic simulations based vocal tract area functions. In vowel spectra, this resonance is often associated with an independent peak in the spectrum, or combines with the fourth formant (F4), both resulting in an elevation of the spectral envelope in this frequency region. The peak frequencies for the vowel /a/ and /i/ are shown in Table 1. The laryngeal ventricle is wider in /i/ than in /a/, suggesting that the resonance frequency is lower

for /i/ and for /a/. However, the laryngeal tube is wider in /i/ than in /a/, which compensates for the ventricular volume difference, resulting in a stable resonance peak.

Table 1. Resonance of laryngeal cavity

	vowel /a/	vowel /i/
Peak frequency	3205 Hz	3255 Hz

3.2. Piriform fossa

The piriform fossa is each of the bilateral tubes behind the laryngeal vestibule, which form the entrance of the esophagus. If the vocal tract is thought of as a sound conduction route from the glottis to the lips, the piriform fossa can be considered a bilateral side branch of the vocal tract. Each of the cavities contributes an anti-resonance to vowel spectra and has a small influence on formant frequencies as well. The effect is also observed in acoustic experiments using mechanical models of the vocal tract with and without the piriform fossa. The length of the fossa is about 2 cm, but because of its funnel shape, the anti-resonance frequency is higher than that of the 1/4 wave-length resonance of an open tube of the same length. In adult male speakers, its effect is most pronounced at around 4-5 kHz. The trough frequencies for the vowel /a/ and /i/ are shown in Table 2. The bilateral cavities of the piriform fossa tend to shorter in /i/ than in /a/, resulting in higher anti-resonances in /i/ than in /a/.

Table 2. Anti-resonance of piriform fossa

	vowel /a/	vowel /i/
Right fossa	4176 Hz	4598 Hz
Left fossa	4645 Hz	4950 Hz

3.3. Vocal tract model with hypopharyngeal cavities

Taking into account the phenomena described above, we propose an acoustic model of vowel production with hypopharyngeal cavity coupling. Figure 3 shows a schematic drawing of the model with the data for vowel /e/. In this model, the vocal tract proper determines the lower formants, while the hypopharyngeal cavities add a peak at about 3-3.5 kHz and two troughs at about 4-5 kHz, both contributing to a sharp cutoff in the spectral envelope at about 4 kHz.

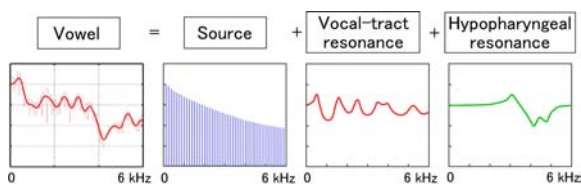


Fig. 3 An acoustic model of vowel production with hypopharyngeal cavity coupling.

4. Discussion

The hypopharyngeal resonance is characterized by a peak at 3.5 kHz and dips at 5 kHz, and it is accompanied by shifts of higher formants. Based on this observation, we propose an acoustic model of vowel production with hypopharyngeal

cavity coupling, as shown in Fig. 1 (Honda, et al, 2004). A vowel synthesis experiment has suggested that this model accounts for a causal mechanism of individual differences in vocal characteristics.

5. Acknowledgement

This research was supported in part by the National Institute of Information and Communications Technology.

6. References

- [1] Fant, G. (1960) *Acoustic Theory of Speech Production*. The Hague: Mouton.
- [2] Sundberg, J. (1987) *The Science of the Singing Voice*. Dekalb, Ill.: Northan Illinois Univ. Press.
- [3] Dang, J., & Honda, K. (1996) Acoustic characteristics of the piriform fossa in models and humans. *J. Acoust. Soc. Am.*, 101, 456-465.
- [4] Imagawa, H., Sakakibara, K., Tayama, N., & Niimi, S. (2003) The effect of the hypopharyngeal and supra-glottic shapes for the singing voice", *Proc. Stockholm Musical Acoustics Conf. 2003*, Vol.II, pp. 471- 474.
- [5] Honda, K., Takemoto, H., Kitamura, T., Fujita, S., & Takano, S. (2004) Exploring human speech production mechanisms by MRI. *IEICE Trans. Inf. & Syst.*, E87-D, 1050-1058.
- [6] Kitamura, T., Honda, K., & Takemoto, H. (2004) Individual variation of the hypopharyngeal cavities and its acoustic effects. *Acoust. Sci. & Tech.* (in print)