ACOUSTIC CONTRIBUTIONS OF THE HYPOPHARYNGEAL CAVITIES IN GENERATING SINGER'S FORMANT: A CASE STUDY

Hironori Takemoto[1], Itsuki Shishime[1], Natsuki Toda[1], Jun Takahashi[2], Yukiko Nota[3]

[1] Chiba Institute of Technology, Chiba, Japan
[2] Osaka University of Arts, Osaka, Japan
[3] National Institute of Japanese Language and Linguistics, Tokyo, Japan

***Keywords*: Singer's Formants, Laryngeal Cavity, Piriform Fossae, FDTD**

Abstract:

Objectives / Introduction: The singer's formant, a spectral emphasis around 3 kHz commonly present in vowels produced by male opera singers [1], has been theorized to result from a cluster of third, fourth, and fifth formants (F3, F4, and F5) [1]. The hypopharyngeal cavities, encompassing the laryngeal cavity and bilateral piriform fossae, generates this cluster [1]. However, the specific mechanisms remain ambiguous. This case study investigates the acoustic contributions of the laryngeal cavity and piriform fossae to the generation of the singer's formant. We employed MRI (Magnetic Resonance Imaging) to capture the three-dimensional vocal tract shape of a tenor during singing and subsequently analyzed transfer functions when modifying the hypopharyngeal cavities.

Methods: In an MRI scanner (Siemens, MAGNETOM Prisma fit 3T), a professional tenor vocalized the vowel /a/ while supine for 15 s at a pitch corresponding to F3 (174.6 Hz). Scanning commenced approximately 2 s after initiating singing. The singing voice was recorded using an optical microphone (Optoacoustics Optimic 1140). The vowel spectrum was extracted from recordings obtained before scanning due to loud noise during scanning. Volumetric data of the vocal tract were reconstructed from the images, with a pixel resolution and slice thickness of 1 × 1 mm and 2.5 mm, respectively. Dental images were superimposed onto the volume data [2], and the vocal tract shape from the glottis to the lips was extracted. The vocal tract wall thickness was increased outward by 3 mm to maintain the vocal tract shape, resulting in a vocal tract model [3]. The model was discretized to enhance spatial resolution to 0.5 mm. The transfer function from the glottis to the lips in the original model was calculated using the finite-difference time-domain method [3]. Subsequently, the laryngeal cavity was excluded from the original model, and the transfer function from the pharynx to the lips was computed. The bilateral piriform fossae were omitted from the original model, and the transfer function from the glottis to the lips was calculated. Henceforth, these transfer functions are denoted as TF-org, TF-lc, and TF-pf (Fig. 1).

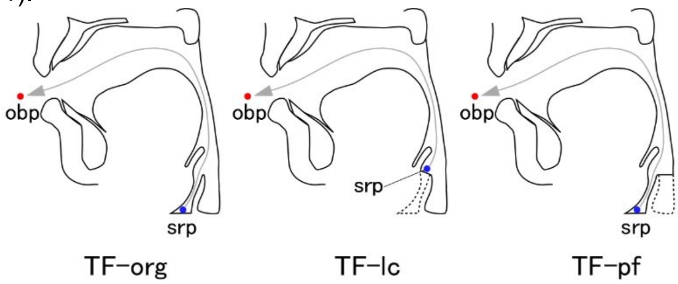

Fig. 1 Calculation of the three transfer functions (srp: source point, obp: observation point)

Results: Fig. 2 displays the TF-org alongside the spectrum of the sung vowel. TF-org aligned closely with the spectrum below 3.5 kHz. The singer's formant, representing a formant cluster of F3, F4, and F5, was evident on the spectrum in the frequency range from 2.3 kHz to 3.1 kHz. Correspondingly, TF-org exhibited clusters of the third, fourth, and fifth resonances (fR3, fR4, and fR5, respectively) within the same frequency range. Hence, modifications to the original vocal tract model offered insights into the mechanism underlying the generation of the singer's formant.

Fig. 3 presents the transfer functions: TF-org, TF-lc, and TF-pf. A comparison between TF-org and TF-lc revealed the role of the laryngeal cavity in generating the singer's formant. Removing the laryngeal cavity induced significant changes in the resonance peak cluster: fR4 disappeared, fR3 underwent a frequency increase of 50 Hz and a level decrease of 23 dB, and fR5 decreased in frequency and level by 100 Hz and 29 dB, respectively. However, this removal resulted in only minor changes in the other peaks and dips. These findings suggest that the laryngeal cavity generates fR4—potentially the principal component of the singer's formant—with limited impact on other resonances and anti-resonances, as corroborated by a previous study [4]. This phenomenon arises from the upper part of the laryngeal cavity constricting and the lower part of the pharyngeal cavity expanding during opera singing, rendering the laryngeal cavity acoustically independent from the remainder of the vocal tract.

A comparison between TF-org and TF-pf showed the contribution of piriform fossae to generating the singer's formant. The elimination of the piriform fossae removed a deep spectral dip at 3.5 kHz and impacted the resonance peak cluster as follows: fR3 increased in frequency by 50 Hz and in level by 4 dB, fR4 increased only in level by 1 dB, and fR5 increased in frequency by 250 Hz and in level by 4 dB. These findings indicate that the vocal tract without piriform fossae generated fR3 and fR5, while the piriform fossae decreased their frequencies. The degree of frequency decrease occurred at fR5, which was near the dip caused by the piriform fossae. The piriform fossae acted as side branches of the vocal tract, generating pole-zero pairs on the transfer function [4]. Consequently, the deep spectral dip at 3.5 kHz on TF-org corresponded to a zero. fR5 plausibly corresponds to the pole, as indicated by Vampola *et al.* [5].

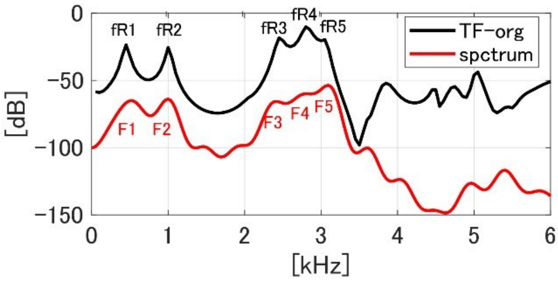

Fig. 2 Original transfer function and vowel spectrum

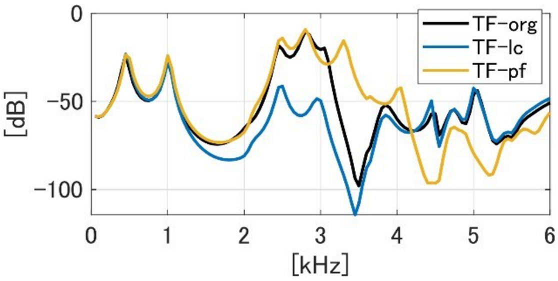

Fig. 3 Three transfer functions

Conclusions: The three-dimensional vocal tract shape of the tenor while singing the vowel /a/ was measured using MRI. The calculated transfer function from the vocal tract shape aligned well with the vowel spectrum below 3.5 kHz, and a resonance peak cluster corresponding to the singer's formant was observed. Transfer functions were computed and analyzed with modification to the hypopharyngeal cavities to explore the acoustic role of the hypopharyngeal cavities in generating the singer's formant. The results indicated that the laryngeal cavity generated fR4, the primary component of the resonance peak cluster, significantly increasing the level of the cluster. However, the piriform fossae generated a deep spectral dip directly above the cluster and decreased the frequency of fR5. Consequently, the clusters became tighter. fR5 may represent a pole of the pole-zero pair generated by the interaction between the piriform fossae and vocal tracts other than the hypopharyngeal cavities.

The mechanism for generating a singer's formant can be hypothesized as follows based on the foregoing discussion: The upper part of the laryngeal cavity narrowed during opera singing, acoustically isolating it from other vocal tract parts. The first resonance of the laryngeal cavity appeared as fR4 on the transfer function. Constricting the upper part of the laryngeal cavity (laryngeal vestibule) and expanding the lower part (laryngeal ventricle) decreased the frequency of fR4 to approximately fR3. The frequency of fR5 can be decreased to a certain degree (e.g., 3.5 kHz) by the deformation of the vocal tract other than the piriform fossae [6]. The piriform fossae elongated owing to lowering the larynx during singing, decreasing dip frequencies from 4 kHz to 3.5 kHz. Hence, the frequency of fR5 can be further decreased to be close to the clusters of fR3 and fR4. Consequently, the resonance peak clusters fR3, fR4, and fR5 were generated around 2.8 kHz.

Acknowledgements:

This work was supported by JSPS KAKENHI, Grant Number 23K11172, and funding from the Center for Language Resource Development, National Institute for Japanese Language and Linguistics.

References:

[1] Sundberg, J. (1974). Articulatory interpretation of the "singing formant." *The Journal of the Acoustical Society of America*, *55*(4), 838–844.

[2] Takemoto, H., Kitamura, T., Nishimoto, H., & Honda, K. (2004). A method of tooth superimposition on MRI data for accurate measurement of vocal tract shape and dimensions. *Acoustical Science and Technology*, *25*(6), 468– 474.

[3] Takemoto, H., Mokhtari, P., & Kitamura, T. (2010). Acoustic analysis of the vocal tract during vowel production by finite-difference time-domain method. *The Journal of the Acoustical Society of America*, *128*(6), 3724–3738.

[4] Takemoto, H., Adachi, S., Kitamura, T., Mokhtari, P., & Honda, K. (2006). Acoustic roles of the laryngeal cavity in vocal tract resonance. *The Journal of the Acoustical Society of America*, *120*(4), 2228–2238.

[5] Vampola, T., Horáček, J., & Švec, J. G. (2015). Modeling the influence of piriform sinuses and valleculae on the vocal tract resonances and antiresonances. *Acta Acustica United with Acustica*, *101*(3), 594–602.

[6] Story, B. H. (2006). Technique for "tuning" vocal tract area functions based on acoustic sensitivity functions. *The Journal of the Acoustical Society of America*, *119*(2), 715–718.