DOI: 10.61782/fa.2023.0071

forum acusticum 2023

ROLES OF THE GLOTTAL CONSTRICTION ON PRODUCTION OF [h] IN HIGH FREQUENCY ENERGY

Tsukasa Yoshinaga[1*] **Kikuo Maekawa**[2] **Akiyoshi Iida**[1]
[1] Department of Mechanical Engineering, Toyohashi University of Technology, Japan
[2] National Institute for Japanese Language and Linguistics, Japan

ABSTRACT

The glottal fricative [h] is known to show no clear place of articulation in the vocal tract, and the sound source is argued to be located either in the oral cavity or at the glottis. In this study, to investigate the roles of the glottal constriction on the production of [h], we conducted aeroacoustic simulations of [h] in the contexts of /ha/ and /ho/ in Japanese with two glottal configurations: adducted vocal folds and open glottis at rest. The vocal tract geometry was extracted from magnetic resonance imaging (MRI) scans of a male Japanese speaker sustaining [h]. The airflow characteristics in the vocal tracts were predicted by solving the compressible Navier-Stokes equations. The results showed that the turbulent airflow was generated at both the oral and glottal constrictions in the contexts of /ha/, whereas the turbulent airflow was observed only near the glottis in the context of /ho/. These results indicate that oral constriction produces sound sources like pharyngeal fricatives in /ha/, while glottal constriction is needed to produce [h] in /ho/.

Keywords: *speech production, glottal fricative, numerical simulation, turbulence*

1. INTRODUCTION

The fricative sounds are known to be produced by turbulent airflows at the constricted vocal tract, and the sound source location of the glottal fricative [h] has been argued for many years. After the major revision of the IPA chart in 1989, the manner of articulation of the glottal consonants was discussed by many researchers. Ladefoged [1] and Kloster-Jensen [2] commented that [h] is produced by the turbulent airflow passing through the oral cavities and it should be removed from the 'glottal' fricative. In contrast, Laufer [3] argued that the glottal constriction was observed in Hebrew, Arabic, and Finnish, and [h] should be listed as the glottal fricative. This discussion is still far from settled and the IPA chart is expected to be examined in near future [4].

The acoustic characteristics of [h] have been theoretically investigated by Stevens [5]. Based on the assumptions of vocal tract and glottal configurations, he predicted the acoustic characteristics of [h] and showed that the sound sources near the glottis become prominent, although some sound sources in the vocal tract may affect when the vocal tract constriction is around 0.3 cm^2 for /hi/ or /hu/. However, it is still unclear whether the main sound source is located near the glottis or the vocal tract because no measurement data, which can be typically obtained from structural imaging, was provided for the analysis.

Therefore, in this study, we investigate the sound source location of [h] by conducting a numerical flow simulation on the vocal tracts extracted from magnetic resonance imaging (MRI). The vocal tracts of [h] in contexts of /ha/ and /ho/ were extracted, and the sound source location was explored under the two glottal configurations in the rest and adducted positions.

2. METHOD

2.1 Vocal tract geometry

The three-dimensional vocal tract geometry of [h] was extracted from a static MRI measurement. The measurement was conducted as a side product for the construction of the real-time MRI articulatory movement

*__Corresponding author__: yoshinaga@me.tut.ac.jp

database [6]. The subject is an adult male (age: 50) who speaks standard Japanese, and he sustained [h] for 15 s in two phonemic contexts: /ha/ and /ho/. These contexts were chosen because other vowel contexts in Japanese are pronounced with different places of articulation like /hi/ as [çi] and /hu/ as [ɸu]. The subject was told to hold his vocal tract shape, even if he could no longer hold his breath. The resolution of images was 256 × 256 with 28 slices, and the voxel spacing became 1 × 1 × 3 mm^3.

The vocal tract geometry was extracted based on luminosity values using the software itk-SNAP [7]. The extracted vocal tracts are shown in Fig. 1 along with the mid-sagittal images of the MRI. The geometry of the incisors was neglected since no close contact between the upper and lower incisors was observed. The small opening of the velopharyngeal valve was cut to assume the velum closure during the production of [h].

For the glottal configuration, two vocal fold states were applied for the extracted vocal tracts: adducted vocal folds and open vocal folds at rest position. The adducted vocal folds were made with a simple triangular slit with a glottal width of 3 mm and a length of 15 mm as described in [5]. The thickness of the slit was set to 3 mm. As a result, the glottal constriction with a cross-sectional area of 22.5 mm^2 was set at the inlet of the vocal tract. The rest glottal shape was extracted from another static MRI at rest position, and the vocal tract just below the epiglottis was replaced with the glottis in rest.

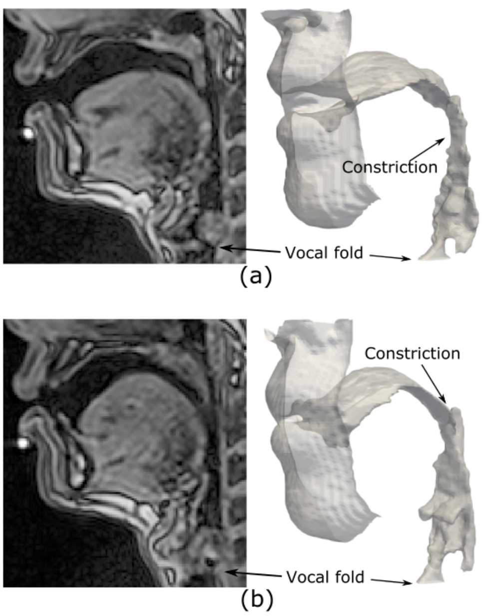

Figure 1. Vocal tract geometry of [h]. The [h] in the context of /ha/ is shown in (a), whereas the [h] in the context of /ho/ is shown in (b).

2.2 Numerical simulation

The turbulent airflow in the vocal tracts is simulated by solving the three-dimensional Navier-Stokes equations; these are partial differential equations which describe the motion of viscous fluid substances, including airflow. The equations were calculated by the high-accuracy finite difference method. The spatial derivatives were solved by the sixth-order-accuracy compact scheme while the time integrals were calculated by the third-order-accuracy Runge-Kutta method, which is a numerical tool used to approximate and estimate a function's value using its derivative. To prevent the numerical instability due to the turbulence, tenth-order-accuracy spatial filter was applied as an implicit turbulence model of the large eddy simulation. In addition, to express the complex geometry in the structured grids, the volume penalization method [8] was employed for the governing equations. The details are reported in [9].

The structured computational grids were constructed for each vocal tract geometry. A total number is approximately 146 million for each case. The minimum grid size near the glottis and the vocal tract constriction was set to 0.1 mm, and a time step was set to 1.0×10^{-7} s considering the Courant-Friedrichs-Lewy (CFL) number to be below 0.4. An acoustic chamber with a volume of 76.6 cm^3 was put below the vocal folds, and a uniform velocity was set at the inlet of the chamber to maintain the volume flow rate of 417 cm^3/s. In the wall region, the velocity was set to 0 to realize the nonslip wall. At the outlet of the computational domain, the non-reflecting boundary was set to prevent acoustic reflection from the outlet. After 1.6×10^5 preliminary calculation steps, 8×10^4 samples were collected for the calculation of root mean square (RMS) values.

3. RESULTS AND DISCUSSION

The instantaneous velocity fields on the mid-sagittal plane of the vocal tract of [h] are shown in Fig. 2. In the context of /ha/, the turbulent airflows were formed around the glottal outlet, pharynx, and soft palate when the vocal folds were adducted. At this time, the minimum cross-sectional areas of the glottis and vocal tract were 22.5 and 28.1 mm^2, respectively. In contrast, by opening the glottis, the turbulent jet disappeared at the glottal outlet, while the fluctuating flow was still observed near the pharynx and soft palate. The area of the glottal opening at rest was approximately 71.9 mm^2.

In the vocal tract of [h] in /ho/, the turbulent jets with large velocities (> 10 m/s) were observed both near the pharynx and the glottis, when the vocal folds were adducted, in the same way as in /ha/. The minimum cross-sectional area of the vocal tract was approximately 40.6 mm^2. Meanwhile, with the glottal opening in /ho/, flow velocities were less than 10 m/s all over the vocal tracts, and velocity fluctuation was not observed in the vocal tract.

The RMS values of the velocity fluctuations on the mid-sagittal plane of the vocal tract [h] are plotted in Fig. 3. The velocity fluctuations in vocal tracts can be considered as one of the aeroacoustic sound sources for broadband noise, and the RMS values of the velocity fluctuations are correlated with amplitudes of the sound source in Lighthill's analogy [10]. Hence, by assuming that the higher RMS values in Fig. 3 are the prominent sound sources in the vocal tract, the main sound source locations were estimated to be near the glottal outlets, pharynx, and soft palate in /ha/ with adducted vocal folds and only in the area from upper part of pharynx to soft palate in /ha/ with the open glottis. In contrast, the strong sound source appeared only near the glottal outlet when the vocal folds were adducted in /ho/, and no clear sound source was observed in the vocal tract of /ho/ with the open glottis. These results indicate that the aeroacoustic sound source can be generated in the vocal tract without the glottal constriction in the context of /ha/, whereas the glottal constriction is needed to produce the sound source for the pronunciation of [h] in /ho/. The preliminary mechanical experiments with vocal tract replicas [11] suggested that the sound source locations were variable depending on the

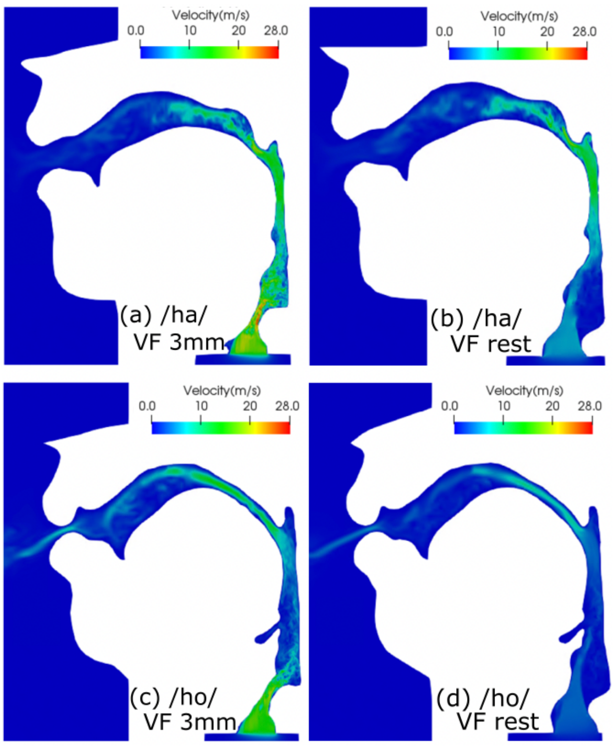

Figure 2. Instantaneous velocities on the mid-sagittal plane in the vocal tracts. The [h] in the contexts of /ha/ for adducted and rest vocal folds is shown in (a) and (b), whereas the [h] in the contexts of /ho/ for adducted and rest vocal folds is shown in (c) and (d), respectively.

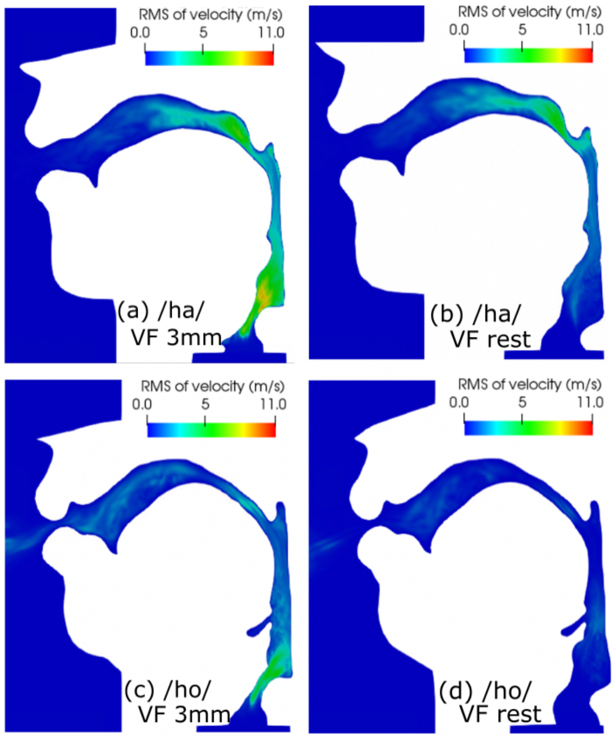

Figure 3. Root mean square (RMS) values of velocity fluctuations on the mid-sagittal plane in the vocal tracts. The [h] in the contexts of /ha/ for adducted and rest vocal folds is shown in (a) and (b), whereas the [h] in the contexts of /ho/ for adducted and rest vocal folds is shown in (c) and (d), respectively.

contexts among /ha/, /he/, and /ho/ in Japanese, and these results are consistent with the current findings.

It should be noted that, when the vocal folds were adducted, the velocity fluctuations near the pharynx were stronger than those with open glottis in the context of /ho/, although the vocal tract shapes above the epiglottis were exactly the same. Further analysis is needed to elucidate how the jet flow at the glottis affects the sound source formation near the soft palate and how each sound source affects the acoustic characteristics of [h].

From the current results, we can conclude that the sound source locations of [h] are highly dependent on the phonemic context, and the Japanese [h] may need to be replaced by the other phonetic symbols like pharyngeal fricative [ħ] or velar fricative [x] if the sound source is located far from the glottis. This information has a potential to be highly useful when we construct the phonetic theories for fricative consonants. However, these results are limited to Japanese, and from a point of view of phonetics, such as in view of the inter-speaker variability of articulation, the number of subjects should be increased to further validate these phenomena.

4. CONCLUSIONS

In this study, the numerical flow simulations were conducted to investigate the sound source locations of fricative [h] in two contexts /ha/ and /ho/. The turbulent airflow and its sound sources were generated near the glottis, pharynx, and soft palate when the vocal folds were adducted in /ha/, while the sound source also appeared near the pharynx and soft palate with the open glottis in /ha/. This indicates [h] can be pronounced without the glottal constriction in the context of /ha/. In contrast, the prominent sound source was generated only near the glottis in /ho/ when the vocal folds are adducted, and the glottal constriction was necessary to produce the sound in the context of /ho/. These results suggest that reconsideration of the phoneme [h] is needed based on the precise data of constriction and its sound source locations for each context in many languages. For future work, it is necessary to increase the number of subjects and consider the inter-speaker variability in many languages.

5. ACKNOWLEDGMENTS

This work was partly supported by Japan Society for the Promotion of Science (JSPS) KAKENHI (Grant Number: JP19K21641 and JP20H01265), and MEXT as "Program for Promoting Researches on the Supercomputer Fugaku" (Project ID: hp230032).

6. REFERENCES

[1] Ladefoged, P.: "Some proposals concerning glottal consonants," *J. Int. Phonetic Assoc.*, vol. 20, no. 2, pp. 24-25, 1990.

[2] Kloster-Jensen, M.: "Letters to the Editor," *J. Int. Phonetic Assoc.*, vol. 21, no. 1, pp. 42-42, 1991.

[3] Laufer, A.: "The 'glottal fricatives.'" *J. Int. Phonetic Assoc.*, vol. 21, no. 2, pp. 91-93, 1991.

[4] Garellek, M.: "Theoretical achievements of phonetics in the 21st century: Phonetics of voice quality," *J. Phonetics*, vol. 94, 101155, 2022.

[5] Stevens, K. N.: *Acoustic phonetics.* (MIT press, Cambridge), 2000.

[6] Maekawa, K.: "Production of the utterance-final moraic nasal in Japanese: A real-time MRI study," *J. Int. Phonetic Assoc.* vol. 53, no. 1, pp. 189-212, 2021.

[7] Yushkevich, P. A., Piven, J., Hazlett, H. C., Smith, R. G., Ho, S., Gee, J. C., Gerig, G.: "User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability," *Neuroimage*, 31(3), 1116-1128, 2006.

[8] Liu, Q., and Vasilyev, O.V.: "A Brinkman penalization method for compressible flows in complex geometries," *J. Comput. Phys*. vol. 227, no. 2, pp. 946–966, 2007.

[9] Yoshinaga, T., Maekawa, K., Iida, A.: "Aeroacoustic differences between the Japanese fricatives [ɕ] and [ç]," *J. Acoust. Soc. Am.*, vol. 149, no. 4, pp. 2426-2436, 2021.

[10] Yoshinaga, T., Nozaki, K., Wada, S.: "Experimental and numerical investigation of the sound generation mechanisms of sibilant fricatives using a simplified vocal tract model," *Phys. Fluids,* vol. 30, no. 3, 035104, 2018.

[11] Yoshinaga, T., Maekawa, K., Iida, A.: "On the sound source locations of 'glottal fricative' [h]," *Proc. 20th International Congress of Phonetic Sciences (ICPhS 2023)*, (accepted).