Estimating tongue stiffness during phonation using ultrasound passive shear wave elastography

Chenhao Chiu[1], Wei-Cheng Hsiao[2], Huang-Yu Shih[2], Bao-Yu Hsieh[2,3], Yining Weng[1]
[1]National Taiwan University (Taiwan), [2]Chang Gung University (Taiwan), [3]Chang Gung Memorial Hospital at Linkou (Taiwan)
chenhaochiu@ntu.edu.tw, v25293626@gmail.com, lamboipe888@gmail.com, byhsieh@mail.cgu.edu.tw, wynjdws@gmail.com

Recent biomechanical modeling studies have proposed that sound variants or alternations may be associated with physiological preferences, such as muscle-induced stress and strain of the tongue surface [1,2]. But the account of physiological preference was largely based on simulation results, while *in vivo* measurements on the articulatory apparatus (e.g., the tongue) remain sparse. A recently developed technique of ultrafast imaging captures up to thousands of ultrasound images per second, making it possible to analyze the mechanical waves that travel through the scanned tissues [3,4]. By calculating the amplitude and velocity of this mechanical wave, the stiffness of the scanned tissues can then be estimated (a method known as ultrasound elastography). Miura et al. [5] employed ultrasound elastography with ultrafast imaging to estimate the stiffness (hardness, in their term) and pressure of the tongue when external forces were applied. However, it is yet to be determined whether the stiffness of the tongue may be different during phonation, or vary during productions of different sounds.

We used ultrasound passive elastography with ultrafast imaging to capture the state of the tongue during vowel articulation. One native speaker of Mandarin participated in data collection. The speaker produced multiple repetitions of Mandarin [a], [i], and [u] vowels, all in high level tone. Each vowel vocalization lasted for 4 ~ 5 seconds. During each vocalization, a roughly 100 ms long ultrasound video was captured at a frame rate of 3k to 9k fps. Ultrasound ultrafast imaging allowed us to track the propagation of the mechanical waves produced by the intrinsic vibrations of the vocal folds through the tongue tissues. We then measured the wave amplitude, propagation velocity, and the associated stiffness of the tongue, as well as spectral properties of the mechanical waves such as vibration frequency. Additionally, acoustic recordings were carried out in parallel with ultrasound data collection. The acoustic fundamental frequencies (F0) of vowel vocalizations were calculated for cross comparison.

Our results show that the frequencies of the mechanical waves measured at the tongue match with those measured from acoustic recordings (Fig. 1), affirming that the vibrations of the vocal folds propagate mechanical waves all the way through the tongue. This suggests that the vocal folds can serve as an intrinsic source of vibration, suitable for shear wave estimation and quantification of tongue stiffness using ultrafast imaging. The propagating velocities of the mechanical waves also differ for each vowel (Fig. 2, bottom), indicative of vowel-dependent local tongue stiffness. Taken together, the findings of the current study demonstrate the viability of ultrasound passive shear wave elastography for examining speech production. This technique could be a promising tool for probing into the detailed physiological operations of the speech apparatus during articulation, and potentially further account for patterns of sound combinations or changes.

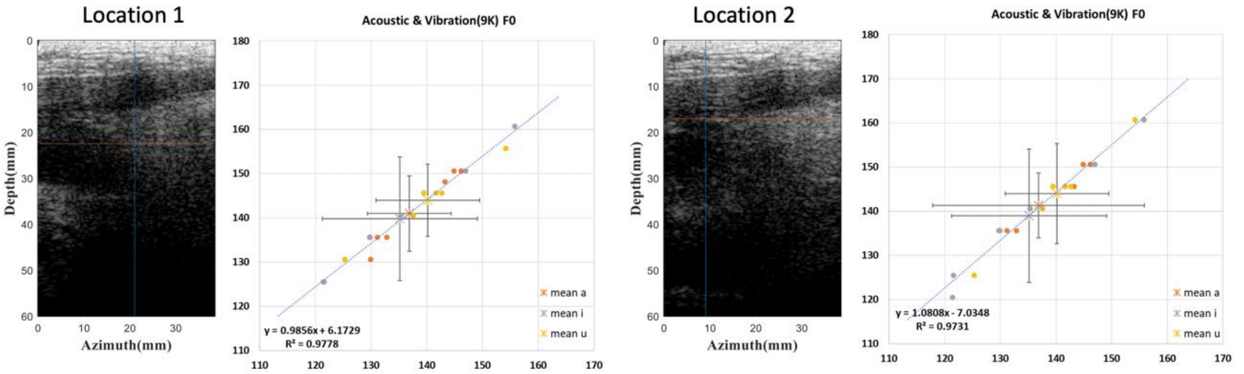

Fig.1 Correlations between estimated ultrasound vibration frequency (y axis) and recorded acoustic F0 frequency (x axis, both in Hz) at two different locations of the midsagittal tongue. The B-mode images indicate where the ultrasound vibration frequency was obtained.

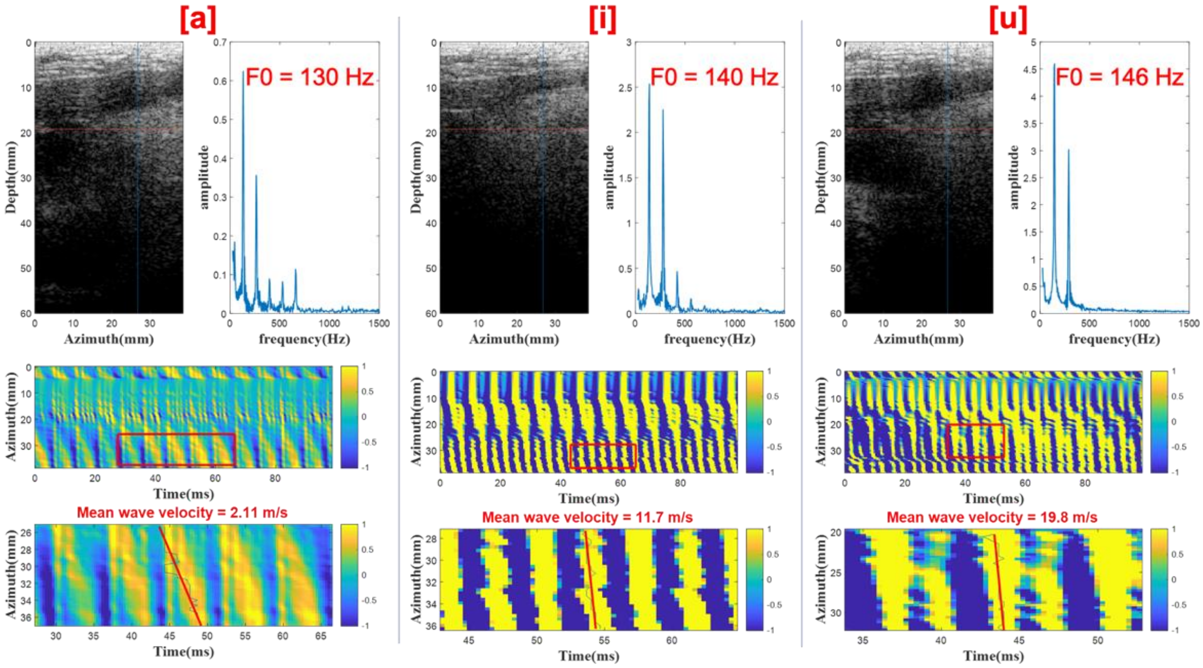

Fig.2 Mechanical vibrations of tongue tissues at the intersection of the blue and orange lines in the B-mode images (top left) were submitted to spectral analysis (top right). Shear wave velocities were calculated for each vowel and visualized in M-mode images (bottom).

References

[1] Stavness, I., Gick, B., Derrick, D., & Fels, S. (2012). Biomechanical modeling of English/r/variants. *The Journal of the Acoustical Society of America*, *131*(5), EL355-EL360.

[2] Chiu, C., & Lu, Y. A. (2021). Articulatory evidence for the syllable-final nasal merging in Taiwan Mandarin. *Language and Speech*, *64*(4), 771-789.

[3] Jing, B., Tang, S., Wu, L., Wang, S., & Wan, M. (2016). Visualizing the vibration of laryngeal tissue during phonation using ultrafast plane wave ultrasonography. *Ultrasound in Medicine & Biology*, *42*(12), 2812-2825.

[4] Jing, B., Ge, Z., Wu, L., Wang, S., & Wan, M. (2018). Visualizing the mechanical wave of vocal fold tissue during phonation using electroglottogram-triggered ultrasonography. *The Journal of the Acoustical Society of America*, *143*(5), EL425-EL429.

[5] Miura, K., Ohkubo, M., Sugiyama, T., Tsuiki, M., & Ishida, R. (2021). Determination of the Relationships Between intra-and Extraoral Tongue Hardness, Thickness, and Pressure Using Ultrasonic Elastography. *Dysphagia*, *36*(4), 623-634.