A Study on Vocal Attack During Voice Therapy Exercises Using Photoglottography

*Aki Taguchi, †Naoki Hyodo, and *Osamu Shiromoto, *Mihara and †Matsuyama, Ehime, Japan

Summary: Objective. There are various exercises for voice therapy, but current evidence is insufficient to decide the most effective training technique for each type of dysphonia. This study focused on vocal attack as one of the causes of dysphonia. Hence, vocal attack during voice therapy exercises was investigated using photoglottogram (PGG).

Methods. Eighteen healthy adult subjects (10 males and 8 females) were included in this study. The first to fifth vocal waves during natural voice, hard and soft voice onset, and semi-occluded vocal tract exercises (SOVTE: humming, tubing, and lip trill) were assessed. We also calculated the 25% vocal fold open quotient (OQ) using a PGG and compared these parameters.

Results. The 25% OQ did not show any sex-related differences. In the first wave, the 25% OQ for hard attack was significantly lower than that for soft attack, tongue-out humming, and lip trill. In contrast, the 25% OQ value for soft attack was significantly higher than that for humming, 6 mm tube phonation, and lip trill. The 25% OQ values differed between SOVTE procedures; it was higher for the 13 mm tube phonation than for the 6 mm tube phonation. The 25% OQ at voice onset in the first to fifth waves differed depending on the SOVTE technique, indicating different voice onset patterns.

Conclusions. These results suggest that appropriate selection of SOVTE tailored for each patient may have benefit in further voice improvement. Future research should focus on conducting a similar study on patients with dysphonia.

Key Words: Vocal attack–Voice therapy exercise–PGG–SOVTE.

INTRODUCTION

Voice therapy consists of symptomatic voice therapy, which involves training according to voice symptoms, such as abnormalities in voice quality, pitch, and strength, and the process of comprehensively producing voice through breathing and vocalization techniques to enhance resonance training to improve speech abnormalities.[1] From these voice therapy methods, voice specialists and speech pathologists, after observing laryngeal findings, select and implement procedures based on clinical experience. There are various vocalization methods for patients; however, there are no manuals or flowcharts to decide the most effective training techniques for specific findings or vocalization conditions. It remains uncertain whether the choice of training procedure was the right one. Therefore, this study was conducted on vocalization patterns in speech therapy procedures, focusing on semi-occluded vocal tract exercise (SOVTE[2]). SOVTE is one of the most widely used voice therapies globally. It improves vocal dynamics and vocal tract resonance by increasing expiratory pressure and intraoral pressure by creating a blockage or constriction in the vocal tract, at a location other than the vocal folds.[2] SOVTE promotes an increase in vocal tract impedance, resulting in changes in the inertive reactance,[3–7] with favorable effects on voice production because of a reduction of phonation threshold pressure[6,8] and an increase in skewing of the glottal flow waveform.[5,6] The increasing vocal tract impedance can affect the glottal function through acoustic–aerodynamic interactions and mechano-acoustic interactions.[3,9,10] SOVTE includes techniques such as humming, tube phonation, lip trill, and water resistance. In this study, the SOVTE techniques of humming, tube phonation, and lip trill, in addition to hard and soft vocalizations, which are symptomatic voice therapies, were used.

This study focused on vocalization as a means of observing vocal attack. Vocal attack is the beginning of vocalization and represents the pattern of vocalization initiation.[11] Much of the state of vocal fold vibrations during vocalization is determined by the state of the larynx and expiratory flow at the moment of vocal attack. It can be regarded as a preparatory movement for stabilizing the vocal fold vibration toward the next stable vocalization. In 1969, Koike[12] investigated differences in laryngeal control during different vocal attacks using electromyography. In 1993, Ohtsu et al[13] conducted aerodynamic studies on voice disorders during vocal attack. In recent years, studies on vocal attacks using electroglottography (EGG)[14] and high-speed imaging[15,16] have been reported. However, these studies have been inconclusive regarding the treatment to be provided. In this study, a photoglottogram (PGG) was used to determine the open quotient (OQ) and compare normal vocalization and vocal attack during speech therapy. In previous studies on vocal attack, only the first wave of vocalization was observed. In contrast, this study investigated the first to fifth waves of vocalization to determine the trends of vocal fold vibration and glottal opening after a vocal attack.

Accepted for publication March 15, 2024.
From the *Faculty of Health and Welfare, Department of Health and Welfare, Prefectural University of Hiroshima, Mihara, Hiroshima, Japan; and the †Department of Rehabilitation, Matsuyama Rehabilitation Hospital, Matsuyama, Ehime, Japan.
Address correspondence and reprint requests to Aki Taguchi, Faculty of Health and Welfare, Department of Health and Welfare, Prefectural University of Hiroshima, Mihara, Hiroshima Japan. E-mail: akiaki@pu-hiroshima.ac.jp
Journal of Voice, Vol xx, No xx, pp. xxx–xxx
0892-1997

https://doi.org/10.1016/j.jvoice.2024.03.017

MATERIAL AND METHODS

Eighteen healthy adult participants (10 males and 8 females) were included in this study. The average age was 33.2 years old for males (standard deviation, 7.6 years) and 32.5 years old for females (standard deviation, 10 years). The research collaborators were staff members (speech therapist (ST), physical therapist, occupational therapist) at Matsuyama Rehabilitation Hospital who had no experience in voice therapy and consented to the study. The selection criteria were as follows: 1) no otolaryngological disease related to speech organs, 2) no indication of hoarseness, as assessed by an otolaryngologist or speech pathologist, 3) no experience of vocal training, and 4) no history of smoking. For research cooperation, oral and written explanations were provided to all the research collaborators, and written informed consent was obtained from the subjects. The study period was extended to avoid conducting the research during the period of COVID-19 infection spread.

An improved PGG (model OEEePGG01), jointly developed and commercialized by the Prefectural University of Hiroshima and Oisaka Denshiki Co., Ltd (Hiroshima, Japan), was used (Figure 1). The wavelength of the near-infrared light used for measurement in ePGG was in the 800 nm wavelength. This wavelength was selected as it avoids absorption by fat, water, and blood and maximizes light shielding by the vocal folds. The data measured using the PGG were recorded on a personal computer using a high-speed A/D BOX. The PGG, manufactured by Oisaka Electronics, has a floodlight mounted on its front neck (Figure 2a). In the study by Aoki et al,[17] an accurate glottal waveform could not be obtained because the female neck is narrower than the male neck and the floodlight could not be attached correctly. Therefore, they inserted the floodlight transnasally (Figure 2b). Similarly, our study too did not use a percutaneous floodlight; instead, a 6 Fr nasal tube (Figure 2c), which is thinner than the 10 Fr tube used by Aoki et al, was used. The transnasal light source emitted similar infrared light to the percutaneous source.

An otorhinolaryngologist inserted the floodlight from the nasal cavity into the upper part of the pharynx. Of the eight female subjects, the floodlight could not be inserted into the nasal cavity in one subject due to chronic rhinitis, and another subject was excluded from the study because PGG waveforms could not be extracted. For each waveform, the 25% value of the difference between the maximum and minimum values was calculated according to the EGG waveform analysis by Scherer et al.[18]

The PGG waveforms were recorded during normal natural vocalization /e/ and seven types of voice therapy exercises, which included speech training techniques of 1) hard attack and 2) soft attack and SOVTE techniques of 3) humming, 4) tongue out humming, 5) 6 mm tube phonation, 6) 13 mm tube phonation, and 7) lip trill. The /e/ vowel was used for the hard and soft attacks as well as the normal vocalization.

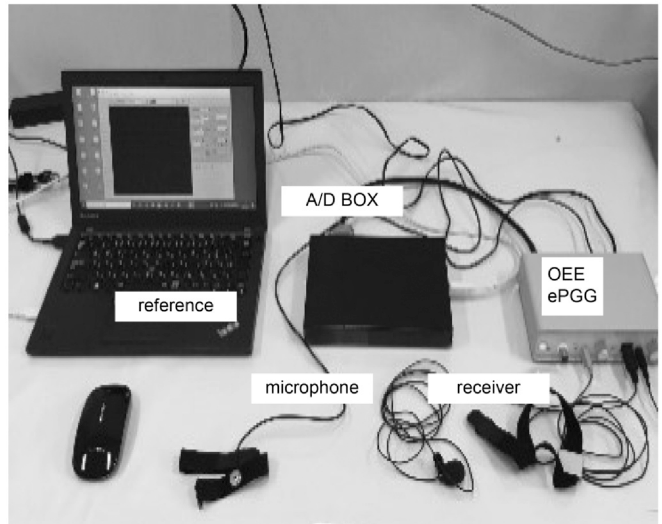

FIGURE 1. Improved photoglottogram used in this study.

This study extracted the PGG waveforms during natural voice and each voice training technique from the PGG waveforms displayed on a personal computer monitor screen (Figure 3), calculated the 25% OQ from the first to fifth waves, and compared the results. The waveform of the first PGG wave was defined as the first wave of voicing when the two researchers (otorhinolaryngologist and ST) reached an agreement (Figure 4). As a preliminary experiment, five waves were extracted at the time of stable utterance from the natural voice and each voice training technique, and compared with the average value of 25% OQ.

Before the experiment, the subjects were instructed to perform normal vocalization and the vocal training techniques. The experiments were conducted after ensuring that the subjects were able to perform each vocal training procedure according to the instructions. The ST gave verbal instructions, and the subjects practiced until they were able to pronounce the words and perform the techniques correctly. The practice time was 5–10 minutes. For the test, the subjects were instructed to produce natural voice and perform each vocal training technique five times for 3 seconds. It was ensured that a stable waveform was extracted. The volume and pitch of the voice were guided by the researchers to ensure that each subject could vocalize comfortably. To maintain consistency in the subject's voice pitch during practice, the subject's voice was recorded using a linear PCM recorder (TASCAM DR-05 VER3) before the examination, thus avoiding any potential influence from the projector inserted through the nasal cavity. Using the piano app "Piano Key Board Real," the subjects were presented with a sound of the same pitch as the subjects' recorded voice at the time of examination and were instructed to vocalize. For the tube phonation, a plastic straw with an inner diameter of 6 and 9 mm and a length of 210 mm was used.

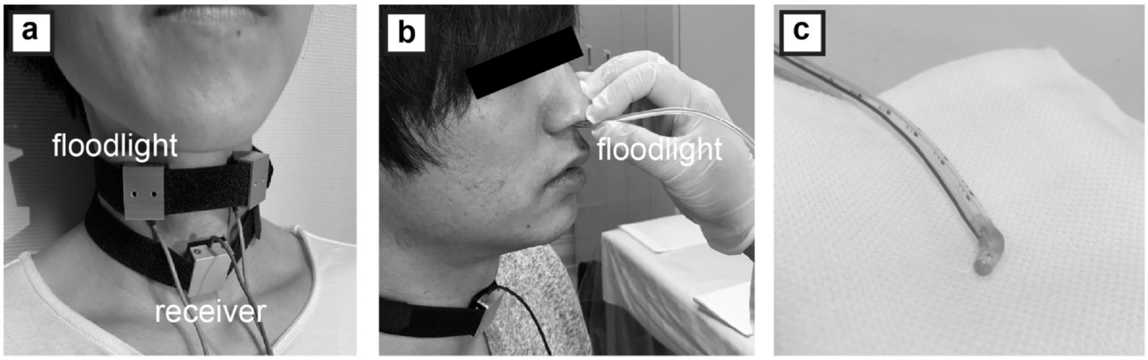

FIGURE 2. **a.** Transcutaneous floodlight. **b.** Transnasal floodlight. **c.** Small-diameter (6 Fr) floodlight.

The 25% OQ at stable vocalization and vocal attack were compared based on sex and the different voice therapy exercises. Moreover, the 25% OQ average value after the first wave of vocal attack by voice therapy methods (analysis of 25% OQ from the first to fifth waves, the third and fifth waves after vocal attack) was also determined.

Statistical analysis

Statistical analysis was performed using statistical processing software (*EZR* version 1.61). Before statistical processing, a normal distribution test was performed using Shapiro–Wilk test. In addition, the sample size was found to be appropriate as a result of power analysis. In all conducted tests, the significance probability (hereinafter referred to as P value) exceeds the significance level ($P > 0.05$). Additionally, normality assumptions remain unchallenged, suggesting a normal distribution within the dataset across all groups, thus confirming parametric data. The equality of variance was assessed using the F test for comparisons between two groups and the Bartlett test was used for comparisons involving three or more groups. In both instances, the calculated P values were greater than the significance level ($P > 0.05$), indicating no significant deviation from equal variances. Consequently, it was determined that the variances were equal.

Intra-examiner reliability: The same researcher analyzed the first wave of voicing for each voice training technique for all subjects twice, over a week apart. The intraclass correlation coefficient (ICC) was calculated for the results of this analysis. ICC (1, 1) was calculated using case 1 of the ICC. Intra-examiner reliability was 1.00 ($P < 0.01$) for natural voice, 1.00 ($P < 0.01$) for hard attack, 1.00 ($P < 0.01$) for soft attack, 1.00 ($P < 0.01$) for humming, and 1.00 ($P < 0.01$) for tongue out humming ($P < 0.01$), 6 mm tube 1.00 ($P < 0.01$), 13 mm tube 1.00 ($P < 0.01$), and lip trill 1.00 ($P < 0.01$), indicating high reliability.

Inter-examiner reliability: Another researcher examined the extraction of the first wave of vocalization for each vocal training technique. From the PGG waveform data of each vocal training procedure, data for all three subjects were randomly selected, and a total of 24 vocalization wave 1 data were analyzed. The ICC was calculated for the results of this analysis. ICC (2, 1) was calculated using case 2 of the ICC. Inter-examiner reliability was 1.00 ($P < 0.01$) for ground voice, 1.00 ($P < 0.01$) for hard voice, 0.923 ($P < 0.01$) for soft voice, 0.971 ($P < 0.01$) for humming, and 1.00 ($P < 0.01$), 6 mm tube 0.986 ($P < 0.01$), 13 mm tube 0.822 ($P < 0.01$), and lip trill 1.00 ($P < 0.01$), indicating high reliability.

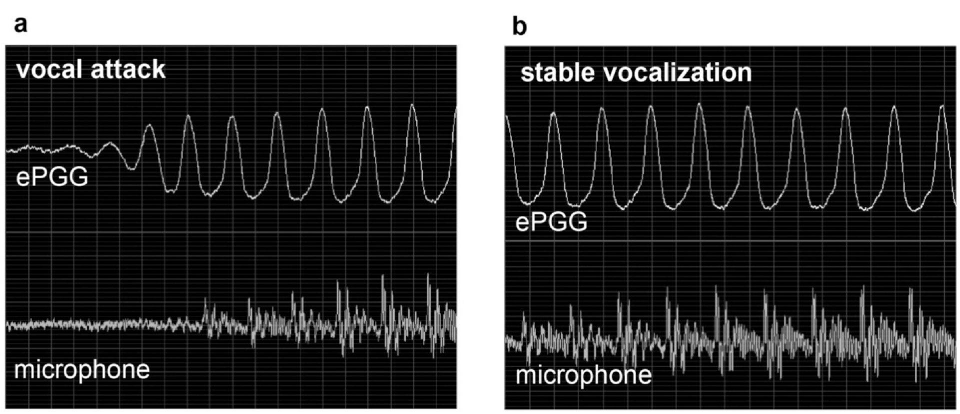

FIGURE 3. Photoglottogram waveform (vertical axis, 500 mV; horizontal axis, 10 ms). **a.** Vocal attack and **b.** stable vocalization.

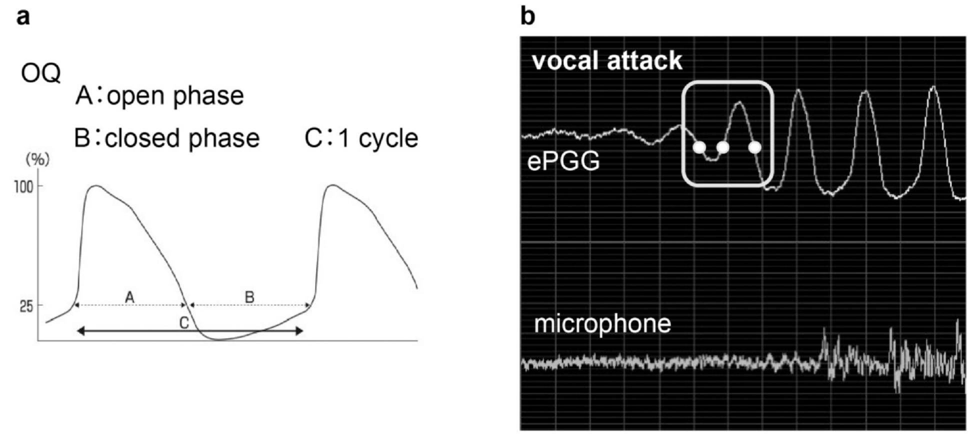

FIGURE 4. **a.** Glottal area waveform (cited from Reference[8]). **b.** Extraction of the first wave of vocal attack (white dots indicate the 25% open quotient) For the waveform of the first wave of the PGG, two researchers extracted points representing 25% of the difference between the maximum and minimum values of the waveform and defined the matching wave as the first wave of phonation.

Comparison of mean values between two groups: Data that showed normal distribution were subjected to an unpaired t test. A significance level of less than 0.05% was considered statistically significant.

Comparison of average values of three or more groups: Data that showed normality were subjected to one-way analysis of variance. Furthermore, the Tukey's method was used for multiple comparisons. In both cases, $P < 0.05$ was considered significant.

This study was conducted with prior approval (approval number: 21040501) from the Ethics Review Committee of Matsuyama Rehabilitation Hospital and in compliance with personal information protection guidelines.

RESULTS

The results were represented as a box and whisker plot (Figures 5–9). The five-number summary is the minimum, first quartile, median, third quartile, and maximum. In a box plot, a box was drawn from the first quartile to the third quartile. A horizontal line extended through the box at the median. The whiskers extended from each quartile to the minimum or maximum. Box is interquartile range. The numbers in the graph represent average values.

1) 25% OQ at stable vocalization

   According to sex

   In males, the 25% OQ for stable vocalization was the highest during tongue-out humming at 0.715 and the lowest during hard attack at 0.589 (Figure 5).

   Among females, the highest 25% OQ was 0.68 for tongue-out humming and 6 mm tube phonation and the lowest was 0.495 for lip trill. There were no significant differences between males and females for each voice therapy exercise.

   According to voice therapy exercises

   The mean value of the 25% OQ during stable vocalization was the highest at 0.702 for tongue-out humming and the lowest at 0.562 for hard vocalization, showing a significant difference (Figure 6). In addition, the 25% OQ value for lip trill greatly varied among subjects.

2) 25% OQ at vocal attack

   According to sex

   The average 25% OQ at vocal attack was the highest for both males and females during soft attack, at 0.705

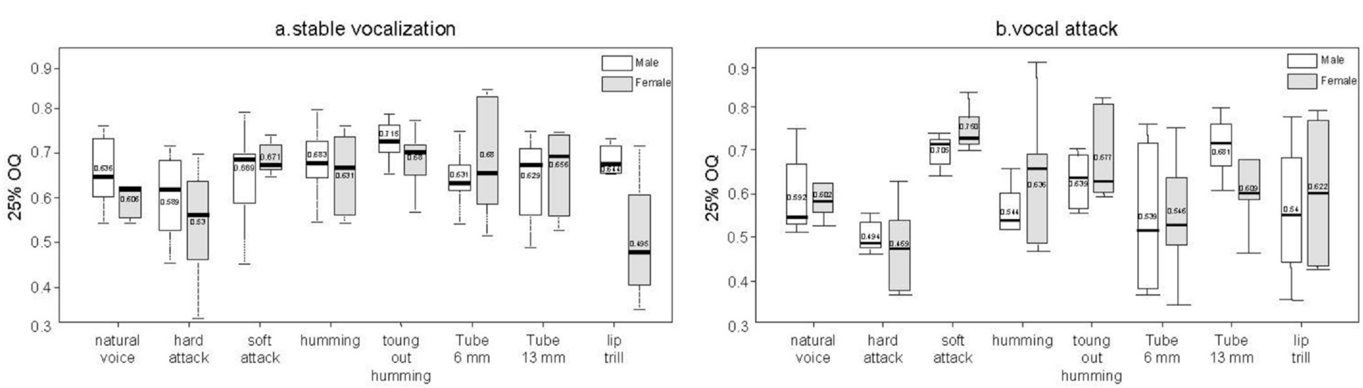

FIGURE 5. The average 25% OQ values by sex. **a.** Stable vocalization and **b.** vocal attack.

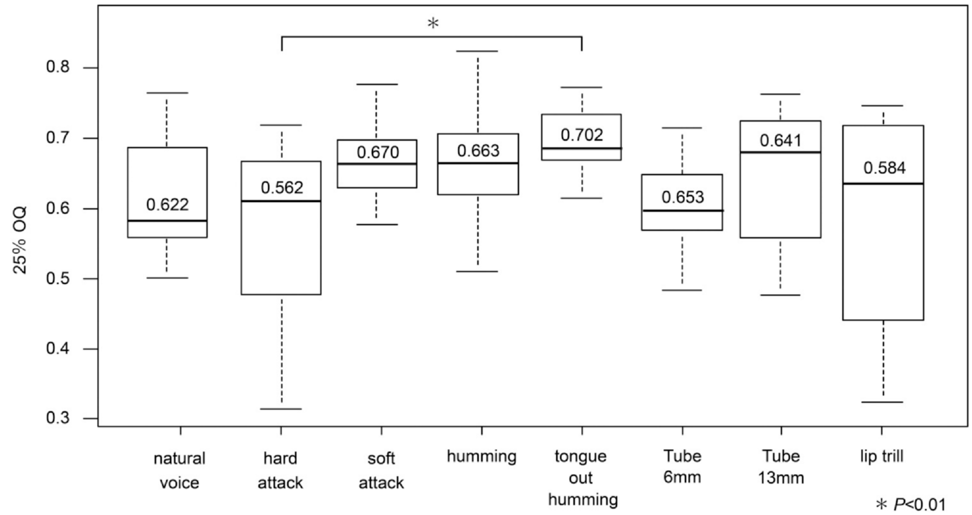

FIGURE 6. The average 25% OQ values during stable vocalization for voice therapy exercises. OQ, open quotient.

for males and 0.750 for females (Figure 5). The lowest 25% OQ for both males and females was during hard attack, at 0.494 for males and 0.470 for females. There was no significant difference in the 25% OQ at vocal attack for any procedure between men and women.

According to voice therapy exercises

Among the voice therapy exercises, the average value of the 25% OQ of phonation was the highest for soft attack (0.725) and lowest for hard attack (0.483).

The 25% OQ for hard attack was significantly lower than that for soft attack, tongue-out humming, and lip trill at the 0.01 level (Figure 7). In contrast, the 25% OQ value for soft attack was significantly higher than that for humming ($P < 0.05$), 6 mm tube phonation ($P < 0.01$), and lip trill ($P < 0.05$). The values in the first wave of the 25% OQ were lower for hard attack, humming, 6 mm tube phonation, and lip trill than those for natural voice. The variation in the 25% OQ values was conspicuous for lip trill.

3) 25% OQ average values after the first wave of vocal attack by voice therapy methods

25% OQ from the first to fifth waves

Figure 8 shows the average values of 25% OQ for the first to fifth waves of vocalization for the natural voice and each voice therapy methods. In the first wave of natural vocalization, the 25%OQ was 0.596, which was significantly lower than 0.622 during stable vocalization (Figure 8a). Second to fifth waves of the natural voice ranged from 0.652 to 0.690, with little variation in 25% OQ, which was significantly higher than the 25% OQ during stable vocalization. For hard attack, the 25% OQ gradually increased from 0.483 in

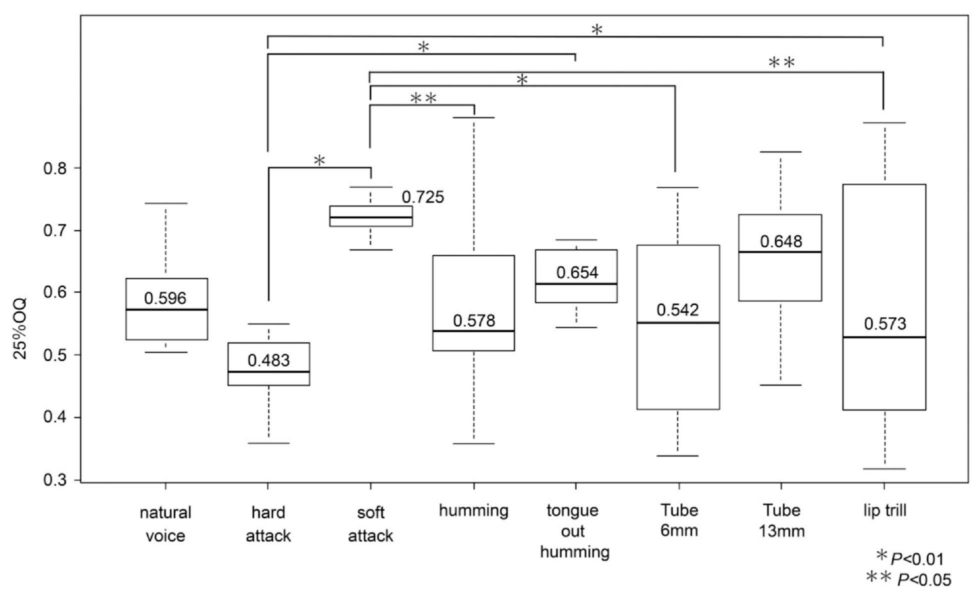

FIGURE 7. The average 25% OQ values during vocal attack for voice therapy exercises. OQ, open quotient.

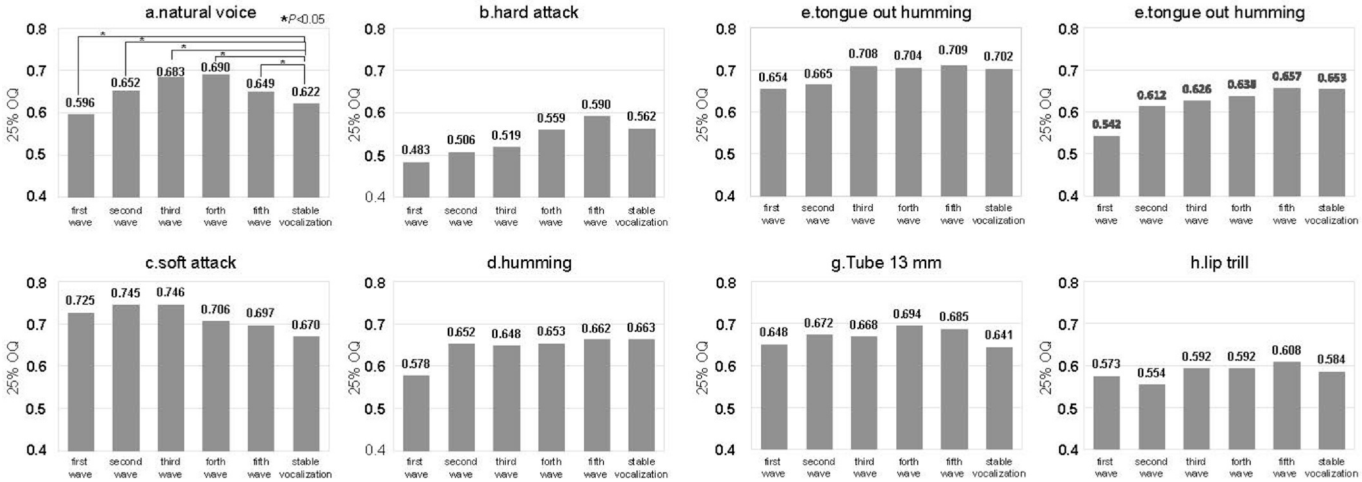

FIGURE 8. The average 25% OQ values from first to fifth waves of vocalization for voice therapy exercises. **a.** Natural voice, **b.** hard attack, **c.** soft attack, **d.** humming, **e.** tongue out humming, **f.** tube 6 mm, **g.** tube 13 mm, and **h.** lip trill.OQ, open quotient.

the first wave of vocalization to 0.590 in the fifth wave, but was slightly lower at 0.562 during stable vocalizations (Figure 8b). In addition, the 25% OQ for the first to fifth wave of hard attack and stable vocalization showed consistently lower values compared with the natural voice and other methods. The 25% OQ value for soft attack decreased gradually from 0.725 in the first wave to 0.697 in the fifth wave (Figure 8c). For humming, the 25% OQ was low at 0.578 only in the first wave, and the 25% OQ increased to 0.652 in the second wave, and remained unchanged until the fifth wave (Figure 8d). There was little difference in the 25% OQ for waves 2 to 5 of humming compared with stable vocalization. The 25% OQ from wave 1 to wave 5 during tongue-out humming was high at 0.654 to 0.709, and during stable vocalization it was 0.702, almost the same value as wave 5 (Figure 8e). The 25% OQ for 6 mm tube phonation was low at 0.542 only in the first wave of vocalization, and after the second wave, there was little change from 0.612 to 0.657 (Figure 8f). After the second wave of 6 mm tube phonation, at 25% OQ, there was little difference from stable vocalization. The 25th percentile OQ for phonation with a 13 mm tube ranged somewhat higher, between 0.648 and 0.685, across the first to fifth waves, showing minimal variation (Figure 8g). Specifically, the 25th percentile OQ of the initial wave of phonation using a 13 mm tube closely mirrored stable phonation, nearly identical at 0.641. Lip trill exhibited a slight variation in the 25th percentile OQ, ranging from 0.573 to 0.608 across the first to fifth waves of vocalization, with minimal deviation from stable vocalization (Figure 8h).

The third and fifth waves after vocal attack

The average 25% OQ values in the third and fifth waves were extracted. In the third wave, the highest 25% OQ was 0.746 for soft attack, followed by 0.708 for tongue-out humming (Figure 9a). In contrast, the 25% OQ value was lower for hard attack, at 0.519. A significant difference was observed between soft attack, which had the highest 25% OQ values, and 13 mm tube phonation, lip trill, and hard attack. In the third wave, the 25% OQ values were lower for natural voice and were the same as in the first wave for hard attack, humming, 6 mm tube phonation, and lip trill. The highest 25% OQ in the fifth wave was 0.710 for tongue-out humming; in contrast, the lowest 25% OQ was for hard attack, at 0.590, followed by lip trill, at 0.608 (Figure 9b). A significant difference was observed between the 25% OQ values for hard attack and tongue-out humming. In the fifth wave, hard attack and lip trill had lower 25% OQ values than that for natural voice.

Additionally, the transition from the first to fifth waves was analyzed. In the first wave, the 25% OQ was significantly lower for natural vocalization (0.596) than that for stable vocalization (0.622). The 25% OQ for natural voice tended to decrease gradually from the third to fifth waves but was significantly higher than that for stable voice. The 25% OQ value for soft attack gradually decreased from 0.725 in the first wave to 0.697 in the fifth wave. For hard attack, the 25% OQ gradually increased from 0.483 in the first wave to 0.590 in the fifth wave, but it decreased slightly to 0.562 during stable vocalization. For humming, the 25% OQ was as low as 0.578 in the first wave, was high in the second wave, and was almost the same as that during stable utterance from the second to fifth waves. The 25% OQ from the first to fifth waves for tongue-out humming was as high as 0.654–0.709, whereas for stable utterance, it was 0.702, almost the same as that for the fifth wave. The 25% OQ for the 6 mm tube phonation was as low as 0.542 in the first wave of vocalization and slightly increased to 0.612–0.657 in the second and

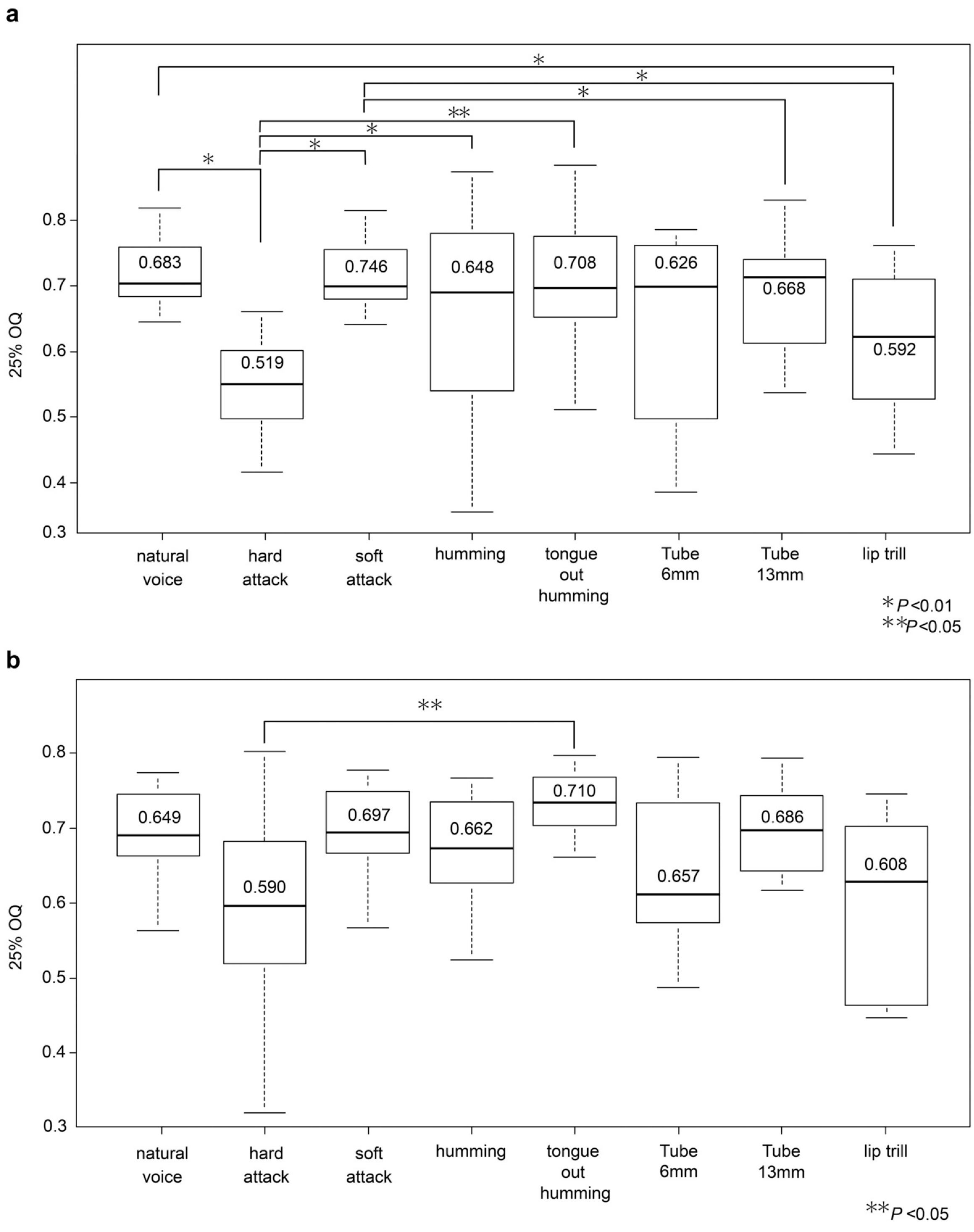

FIGURE 9. The average 25% OQ values after the first wave of vocalization for voice therapy exercises. **a.** Third wave and **b.** fifth wave. OQ, open quotient.

subsequent waves; moreover, no difference was observed between the stable vocalizations. The 25% OQ for the 13 mm tube phonation was slightly higher, ranging from 0.648 to 0.685 from the first to fifth waves. The 25% OQ of lip trill showed a small variation of 0.573–0.608 from the first to fifth waves, and the difference from that during stable utterance was also small. The transition from the first to fifth waves of vocalization was different for each procedure.

DISCUSSION

Vocal attack during voice therapy exercises was investigated using the PGG. The PGG uses a photoelectric transducer to record temporal changes in the glottic area associated with the opening and closing movements of the glottis during vocal fold vibration as temporal changes in the amount of light passing through the glottis. Sonesson entered the illumination light from the subglottis and recorded the electrical signals using a photoelectric transducer inserted orally into the supraglottis.[19] Sawashima[20] devised a method of inserting a laryngeal fiberscope transnasally, illuminating it from above the glottis, and recording the electrical signals with a photoelectric transducer placed below the glottis. With this method, natural articulation was possible and laryngeal images could be recorded simultaneously. However, little research has been conducted on PGG. Traditional visible light spectrum PGG has several inherent limitations. These include lack of uniformity in light density distribution across the vocal folds, light reflections from mucosal surfaces, changes in vertical laryngeal positions, changes in the position and angle of the light sensor, changes in the position and angle of the light beam, movements of articulator positions, and noise from the light source. Any of these factors could impair the validity and reliability of the signal and the accuracy of derived measurements.[21,22] This study used an improved PGG (OEEePGG01), jointly developed and commercialized by the Prefectural University of Hiroshima and Oisaka Electronics Co., Ltd. This device uses near-infrared light of a specific frequency as the light source and a photosensor that detects near-infrared light of that particular frequency to analyze vocal fold vibrations; hence, it is not affected by ambient light. The present PGG is one of the near-infrared photoglottographic (NIR PGG) systems, and its signals differ from those of conventional visible-light PGG showing stable glottal closure periods.[23] NIR PGG signals are thought to reflect changes in both glottal aperture and vocal-fold contact, partly because the vocal-fold tissues are semi-transparent to NIR light.[23] Vocalization is possible because a nasogastric tube is used as a floodlight. The floodlight can be placed much closer to the laryngopharynx than a percutaneous light, and more accurate data can be collected. It is easy to operate and does not require expensive equipment, such as high-speed cameras. These advantages indicate that PGG is suitable for analyzing the glottal release rate from vocal fold vibrations. Therefore, it was thought that this PGG has higher analysis accuracy than the conventional PGG, thus it was used. However, since this study did not compare visible light and NIR PGG, future studies are required to validate this methodology.

Analysis according to sex was performed because previous studies on EGG and PGG[24] included only a few male subjects, while female subjects were almost never included. This is probably because the female larynx is anatomically smaller, and the vocal fold is shorter than in males, making it difficult to measure vocal fold vibrations. Aoki et al[17] changed the floodlight from transcutaneous to transnasal and conducted their research with the improved PGG, which was used in this study. However, extraction of waveforms was not possible in five of the seven females. Therefore, the size of the nasal tube used for transnasal floodlight was modified from 10 to 6 Fr and succeeded in extracting vocal fold vibration waveforms in six of the eight female subjects. There were almost no sex differences in vocal attacks and voice therapy exercises. Our results showed that there were no sex differences in the vocal fold vibration patterns during vocal attacks. In addition, there were no gender differences in the results pertaining to voice therapy exercises in our study. Therefore, sex may not be a primary consideration when selecting voice therapy exercises. Winker and Sendlmeier[25] studied OQ measurements in EGG of young and elderly male and female speakers. Although they did not directly compare males and females, it is noteworthy that OQ values were slightly higher for females. Moreover, their results showed a strong positive relation between perceived age and OQ in male voices. In the future, examining OQ regarding age for males and females in PGG and comparing with EGG is desired.

Various methods for voice therapy exist, among which SOVTE has become one of the most commonly used techniques worldwide in recent years. Focusing on SOVTE, a comparative study of SOVTE, hard voicing, and soft voicing, which have been used for a long time to treat symptoms, was conducted. Lip trill, tube phonation, and humming were selected for their common usage in Japan among SOVTE, and their ease of acquisition. Water resistance was not used due to concerns about potential sound source detection issues with the microphone during vocalization. For future studies, a wider range of SOVTEs will be considered.

First, the 25% OQ values during stable vocalization were in the following order: hard attack < natural vocalization < SOVTE < soft attack < tongue-out humming. There were almost no differences in the 25% OQ values between the different SOVTE procedures. These results suggest that SOVTE does not cause glottal insufficiency as much as soft attack, and that SOVTE is a technique that can moderately promote glottal closure and change the shape of the vocal tract. In addition, because tongue-out humming showed a higher 25% OQ than soft attack, it was considered suitable for correcting overstrain.

The 25% OQ values in the first wave during vocal attack were in the order of hard attack < humming < 6 mm tube phonation < lip trill < natural voice < 13 mm tube phonation < tongue-out humming < soft attack. This result indicated that there was a difference in the 25% OQ values between SOVTE techniques during vocal attack. Compared with the 25% OQ values for natural voice, those of humming, 6 mm tube phonation, and lip trill were lower, whereas those of 13 mm tube phonation and tongue-out humming were higher. Therefore, SOVTE techniques with high 25% OQ, such as the 13 mm tube phonation and tongue-out humming, are recommended in patients with a hard attack on their vocal folds during vocalization. In addition, the use of humming, 6 mm tube phonation, and lip trill could be expected to further improve the voice in patients with glottic insufficiency or breathiness during vocalization. At vocal attack, the 13 mm tube phonation showed a higher 25% OQ value than the 6 mm tube phonation, suggesting that a larger tube diameter may cause greater vocal dynamics and vocal tract resonance by increasing expiratory pressure and intraoral pressure by creating a blockage or constriction in the vocal tract.

Based on previous reports[26,27] that the glottal closure rate and expiratory pressure differ depending on the tube diameter, it may be assumed that glottal closure also varies during vocalization depending on the tube diameter. There has been research on the flow and pressure relationships in different tubes used for SOVTE.[27,28] Additionally, several studies have been reported regarding OQ or contact quotient (CQ) during SOVTE using EGG.[29–32] However, no studies examining the relationship between tube diameter and PGG OQ or CQ have been conducted. Further research using other equipment, such as EGG and high-speed cameras, is necessary.

The 25% OQ values during the first to fifth waves of vocal attack and stable vocalization for each procedure were examined. The change in the 25% OQ value from the first to fifth waves differed depending on the procedure. In the natural voice, the 25% OQ value increased once after vocal attack and then decreased, until it settled into a stable voice. In contrast, the 25% OQ of hard attack gradually increased from a low value, whereas that of soft attack gradually decreased from a high value. Therefore, these two procedures are suitable for the correction of vocal attack. On comparing the 25% OQ of humming and tongue-out humming, it was found that the values were low only in the first wave of humming, and there was no change in the other values, whereas the values remained relatively high from the first wave onward for tongue-out humming. Presumably, tongue-out humming hinders oral closure because the tongue is sandwiched between the lips; moreover, the so-called semi-occlusion due to SOVTE is weak, resulting in glottal regurgitation insufficiency. The glottal regurgitation refers to the phenomenon in which exhaled air passes through the glottis and narrows the outlet called the lips, causing some of the exhaled air to flow backward from the lips toward the glottis. This observation suggests that tongue-out humming can remove hard vocalizations in cases where humming fails to remove hard attacks. In the 6 mm tube phonation, the 25% OQ value gradually increased from the first wave, while in the 13 mm tube phonation, the 25% OQ value remained high from the first wave. As reported by Winkler and Sendlmeier,[25] this observation seems to be related to changes in oral pressure and expiratory flow due to variations in the tube diameter. In contrast, variations in the 25% OQ for lip trill were observed during both stable vocalization and vocal attack. It is speculated that individual differences in the force applied during lip closure may affect glottal closure. In this study, the morphology of the vocal tract changed during the transition from phonation to stable phonation, depending on the SOVTE technique used.

Finally, the 25% OQ values in the first, third, and fifth waves during vocalization were compared among the different voice therapy techniques. In the first wave, except for the 13 mm tube phonation, the 25% OQ was lower than that for natural voice. In the third wave, the 25% OQ value was lower than that for natural voice for all SOVTE procedures. In the fifth wave, the 25% OQ value was higher than that for natural voice, except for lip trill. This suggests that vocalization during SOVTE involves glottis closure at a high level initially but gradually shifts to a moderate level of glottis closure.

This study focused on SOVTE and used PGG to study the vocal attack. Various vocalization patterns were observed among the SOVTE procedures. Individual differences were observed even within the same SOVTE procedure and there were variations depending on the procedure. These results are useful for selecting appropriate techniques for voice therapy. In the future, similar research on other voice therapy techniques, such as vocal function exercises, is planned. By comparing with this study, it is expected to become one of the indices for selecting voice therapy.

The limitation of this study is that PGG was used to study vocal attack, focusing on SOVTE. However, there are many other voice therapy procedures besides SOVTE. In the future, a research on vocalization is planned using PGG for other voice therapy procedures. Additionally, this study was conducted on healthy subjects. In the future, conducting this research on patients with dysphonia is desired.

CONCLUSION

Vocal attack during voice therapy exercises was investigated using the PGG. No sex differences were observed between the 25% OQ values of normal vocalization and voice therapy methods. Moreover, the 25% OQ value in the first wave of vocalization was low for hard vocalizations and high for soft vocalizations. Our study also showed that the 25% OQ values differed between SOVTE procedures; it was higher for the 13 mm tube phonation than for the 6 mm tube phonation. The changes in the 25% OQ values from the first to fifth waves of vocalization differed depending on the voice therapy method. These findings suggest that customizing SOVTE for individual patients could lead to additional improvements in voice quality.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

NA.

References

1. Verdolini K. *Guide to Vocology.* 1. New York: National Center of Voice and Speech; 1998:5–55.
2. Titze IR. Voice training and therapy with a semi-occluded vocal tract: rationale and scientific underpinnings. *J Speech Lang Hear Res.* 2006;49:448–459. https://doi.org/10.1044/1092-4388(2006/035).
3. Story B, Laukkanen A-M, Titze I. Acoustic impedance of an artificially lengthened and constricted vocal tract. *J Voice.* 2000;14:455–469.
4. Titze I. The physics of small-amplitude oscillation of the vocal folds. *J Acoust Soc Am.* 1988;83:1536–1552.
5. Rothenberg M. Acoustic interaction between the glottal source and the vocal tract. In: Stevens KN, Hirano M, eds. *Vocal Fold Physiology.* Tokyo, Japan: University of Tokyo Press; 1981:305–328.
6. Titze I, Story B. Acoustic interactions of the voice source with the lower vocal tract. *J Acoust Soc Am.* 1997;101:2234–2243.
7. Titze I, Laukkanen A. Can vocal economy in phonation be increased with an artificially lengthened vocal tract? A computer modeling study. *Logoped Phoniatr Vocol.* 2007;32:147–156.
8. Titze I. Phonation threshold pressure measurement with a semi-occluded vocal tract. *J Speech Lang Hear Res.* 2009;52:1062–1072.
9. Fant G, Lin Q. Glottal source-vocal tract acoustic interaction. *STL-QPSR.* 1987;1:13–27.
10. Guzman M, Laukkanen AM, Krupa P, et al. Vocal tract and glottal function during and after vocal exercising with resonance tube and straw. *J Voice.* 2013;27:e19–e34.
11. Asano K. An aerodynamic study on vocal onset. The differences among the vocal onsets from the air flow rate and subglottal pressure. *J Jûzen Soc.* 1971;80:617–631.
12. Koike Y. Experimental studies on vocal attack. *Practica Oto-Rhino-Laryngologica.* 1967;60:663–688. https://doi.org/10.5631/jibirin.60.663.
13. Ohtsu Y, Iwata S, Takeuchi K, et al. Aerodynamic studies of voice disorders in vocal attack. *Pract-Oto-Rhinolaryngol suppl.* 1993;65:171–178.
14. Watson BC, Baken RJ, Roark RM. Effect of voice onset type on vocal attack time. *J Voice.* 2016;30:11–14. https://doi.org/10.1016/j.jvoice.2014.12.004.
15. Yokonishi H, Imagawa H, Sakakibara K, et al. Relationship of various open quotients with acoustic property, phonation types, fundamental frequency, and intensity. *J Voice.* 2016;30:145–157. https://doi.org/10.1016/j.jvoice.2015.01.009.
16. Köster O, Marx B, Gemmar P, et al. Qualitative and quantitative analysis of voice onset by means of a multidimensional voice analysis system (MVAS) using high-speed imaging. *J Voice.* 1999;13:355–374. https://doi.org/10.1016/s0892-1997(99)80041-1.
17. Aoki R, Shiromoto O, Haji T, et al. Semi-occluded vocal tract ni yoru seitai-shindou heno eikyo -koden seimonzu niyoru kaiseki. *Jpn J Clin Res Dysarthria.* 2017;7:2–7.
18. Scherer R, Vail V, Rockwell B. Examination of the laryngeal adduction measure EGGW. *NCVS Status Prog Rep.* 1993;5:73–82.
19. Sonesson B. On the anatomy and vibratory pattern of the human vocal folds. *Acta Oto-Laryngol Suppl.* 1960;156:1–80.
20. Sawashima M. Use of the fiberscope in observing articulatory movements. *J Acoust Soc Jpn.* 1971;27:425–434.
21. Coleman RF, Wendahl RW. On the validity of laryngeal photosensor monitoring. *J Acoust Soc Am.* 1968;44:1733–1735.
22. Gerratt BR, Hanson DG, Berke GS, Precoda K. Photoglottography: a clinical synopsis. *J Voice.* 1991;5:98–105.
23. Chi Yujie, Honda K, Wei J. Near infrared photoglottography for measuring multiple glottal events. *JASA Express Lett.* 2022;2:1–6.
24. Yamanaka J. Usefulness of electroglottogram (EGG) and photoglottogram (PGG) for the analysis or vocal fold vibration-a high speed digital imaging study. *J Otolaryngol Jpn.* 2000;103:905–915.
25. Winkler R, Sendlmeier W. Open Quotient (EGG) measurements of young and elderly voices: results of a production and perception study. *ZAS Pap Linguist.* 2005;40:213–225. https://doi.org/10.21248/zaspil.40.2005.267.
26. Gaskill CS, Quinney DM. The effect of resonance tubes on glottal contact quotient with and without task instruction: a comparison of trained and untrained voices. *J Voice.* 2012;26:e79–e93. https://doi.org/10.1016/j.jvoice.2011.03.003.
27. Amarante Andrade PA, Wistbacka G, Larsson H, et al. The flow and pressure relationships in different tubes commonly used for semi-occluded vocal tract exercises. *J Voice.* 2016;30:36–41. https://doi.org/10.1016/j.jvoice.2015.02.004.
28. Smith SL, Titze IR. Characterization of flow-resistant tubes used for semi-occluded vocal tract voice training and therapy. *J Voice.* 2017;31:113.e1–113.e8.
29. Master S, Guzman M, Carlos de Miranda H, et al. Electroglottographic analysis of actresses and nonactresses' voices in different levels of intensity. *J Voice.* 2013;27:187–194. https://doi.org/10.1016/j.jvoice.2012.10.010.
30. Guzman M, Rubin A, Muñoz D, et al. Changes in glottal contact quotient during resonance tube phonation and phonation with vibrate. *J Voice.* 2013;27:305–311.
31. Andrade PA, Wood G, Ratcliffe P, et al. Electroglottographic study of seven semi-occluded exercises: LaxVox, straw, lip-trill, tongue-trill, humming, hand-over-mouth, and tongue-trill combined with hand-over-mouth. *J Voice.* 2014;28:589–595.
32. Calvache C, Guzman M, Bobadilla M, et al. Variation on vocal economy after different semioccluded vocal tract exercises in subjects with normal voice and dysphonia. *J Voice.* 2020;34:582–589. https://doi.org/10.1016/j.jvoice.2019.01.007.