Relationship of Various Open Quotients With Acoustic Property, Phonation Types, Fundamental Frequency, and Intensity

*Hisayuki Yokonishi, *Hiroshi Imagawa, †Ken-Ichi Sakakibara, *Akihito Yamauchi, *Takaharu Nito, *Tatsuya Yamasoba, and ‡Niro Tayama, *Bunkyo-Ku, Tokyo, Japan, †Ishikari-Gun, Hokkaido, Japan, ‡Shinjuku-Ku, Tokyo, Japan

Summary: Introduction. In the present study, we examined the relationship between various open quotients (O_qs) and phonation types, fundamental frequency (F_0), and intensity by multivariate linear regression analysis (MVA) to determine which O_q best reflects vocal fold vibratory characteristics.

Methods. Using high-speed digital imaging (HSDI), a sustained vowel /e/ at different phonation types, F_0s, and intensities was recorded from six vocally healthy male volunteers: the types of phonation included modal, falsetto, modal breathy, and modal pressed phonations; and each phonation was performed at different F_0s and intensities. Electroglottography (EGG) and sound signals were simultaneously recorded with HSDI. From the obtained data, 10 conventional O_qs (four O_qs from the glottal area function, four kymographic O_qs, and two EGG-derived O_qs) and two newly introduced O_qs ($\overline{O_q^{\mathrm{edge}}}^{+}$ and $\overline{O_q^{\mathrm{edge}}}$) were evaluated. And, relationships between various O_qs and phonation types, F_0, and intensity were evaluated by MVA.

Results. Among the various O_qs, $\overline{O_q^{\mathrm{edge}}}^{+}$ and $\overline{O_q^{\mathrm{edge}}}$ revealed the strongest correlations with an acoustic property and could best describe changes in phonation types: $\overline{O_q^{\mathrm{edge}}}^{+}$ was found to be better than $\overline{O_q^{\mathrm{edge}}}$. O_q^{MLK}, the average of five O_qs from five-line multiline kymography was a very good alternative to $\overline{O_q^{\mathrm{edge}}}$. EGG-derived O_qs were able to differentiate between modal phonation and falsetto phonation, but it was necessary to consider the change of F_0 simultaneously. MVA showed the changes in O_q values between modal and other phonation types, the degree of involvement of intensity, and no relationship between F_0 and O_qs.

Conclusions. Among O_qs evaluated in this study, $\overline{O_q^{\mathrm{edge}}}^{+}$ and $\overline{O_q^{\mathrm{edge}}}$ were considered to best reflect the vocal fold vibratory characteristics.

Key Words: Open quotient–Voice–Normal–High-speed digital imaging–Kymography–Kymogram–Electroglottgraphy–Modal–Pressed–Breathy–Falsetto–Multivariate linear regression analysis.

INTRODUCTION

Voice quality is primarily determined by vibratory motions of the vocal fold. Open quotient (O_q) is one of the most important vibratory parameters, which is closely associated with vocal acoustics.

O_q has a close relationship with vocal qualities such as "breathy" and "pressed" phonations.[1,2] Furthermore, the O_q of falsetto phonation is smaller than that of modal phonation.[3–5] In terms of the vocal spectrum, O_q is closely associated with $H1^* - H2^*$, the difference in amplitude between the first two harmonics of an acoustic signal spectrum after formant-based correction.[6,7]

Various studies have been performed to assess the relationship between O_q and fundamental frequency (F_0). Earlier studies revealed no or only a weak positive correlation between O_q and F_0 in male speakers[8–12] and a positive correlation in female speakers.[12,13] Later, Henrich et al[4] investigated the interrelationship among O_q, F_0, and intensity at the same phonation type in consideration of the impact of laryngeal mechanism: in modal phonation, O_q showed no correlation with F_0 and a negative correlation with intensity, and in falsetto phonation, O_q showed a negative correlation with F_0 and no correlation with intensity.

Another study applied multiple regression analysis to the vibratory data obtained from 10 excised canine larynges model to analyze the relationship between O_q and various vibratory characteristics and revealed direct relationships between O_q and vocal fold tension, glottal width, and F_0.[14]

The choice of O_q, according to the study design, is still a moot point, however. Various methods can be used to derive O_qs, depending on the instrument used to measure the O_q. Photoglottography (PGG) and Electroglottography (EGG) are the most common methods used to indirectly measure the O_q. O_q by EGG is usually obtained by tracking the maximum positive peak in the first derivative of the EGG, which approximates the instant of the glottal opening, and its maximum negative peak, which approximates the instant of the glottal closing.[8,15,16] O_q from PGG is obtained by tracking the maximum positive peak in the third derivative of the PGG wave, which often approximates the instant of the glottal opening, and its maximum negative peak, which often approximates the instant of the glottal closing.[3,9,13,17,18] High-speed digital

Accepted for publication January 30, 2015.

From the *The Department of Otolaryngology, The University of Tokyo Hospital, Bunkyo-Ku, Tokyo 113-8655, Japan; †The Department of Communication Disorders, The Health Sciences University of Hokkaido, Ishikari-Gun, Hokkaido 061-0293, Japan; and the ‡The Department of Otolaryngology and Tracheo-Esophagology, The National Center for Global Health and Medicine, Shinjuku-Ku, Tokyo 162-8655, Japan.

Address correspondence and reprint requests to Hisayuki Yokonishi, The Tokyo Metropolitan Police Hospital, Nakano 4-22-1, Nakano-Ku, Tokyo 164-8541, Japan. E-mail: yokonishi-tky@umin.ac.jp

Journal of Voice, Vol. ■, No. ■, pp. 1-13
0892-1997/$36.00

http://dx.doi.org/10.1016/j.jvoice.2015.01.009

imaging (HSDI) are used for direct measurement of the O_q. O_qs can also be derived from the glottal area function or kymography.[8,10] Furthermore, *OT-50* is a videostroboscopic parameter related to O_q, which calculates the time duration between the midpoints of the glottal opening and closing phases, using the glottal area function.[19] There are several advantages and disadvantages of calculating O_qs. First, O_q derived from the glottal area function is not effective in the assessment of cases with a steady posterior glottal gap, which is often observed in vocally healthy female subjects, because O_q derived from the glottal area function becomes 1, despite the presence of normative vocal fold vibrations. This is also true in cases of incomplete glottal closure (eg, a female falsetto phonation or a patient with unilateral vocal fold paralysis). Second, O_q obtained from threshold or a differentiation technique such as *OT-50* tends to be lower than O_qs derived by other methods. A systematic comparison of these O_qs in response to the change in phonation type, F_0, and intensity has not yet been performed.

Therefore, the purpose of the present study was to further investigate the relationship between O_q and an acoustic property, phonation types, F_0, and intensity by multiple regression analysis using an HSDI device under various conditions of phonation types, F_0, and intensity and to determine which O_q best reflects the vocal fold vibratory characteristics by comparing the various O_qs that were simultaneously measured.

MATERIALS AND METHODS

Subject and instrumental setup

Data were collected from six vocally healthy male volunteers (22, 25, 31, 33, 34, and 43 years old) who were not professional but accustomed to change voice quality because of chorus experience. For these subjects, a sustained vowel /e/ at different phonation types, F_0s, and intensities was recorded. The types of phonation included modal phonations at seven different frequencies (G2 [98 Hz], C3 [131 Hz], E3 [165 Hz], G3 [196 Hz], C4 [262 Hz], E4 [330 Hz], and G4 [392 Hz]), falsetto phonations at five different frequencies (C4 [262 Hz], E4 [330 Hz], G4 [392 Hz], C5 [523 Hz], and E5 [659 Hz]), modal breathy phonations at four different frequencies (G2 [98 Hz], C3 [131 Hz], E3 [165 Hz], and G3 [196 Hz]), and modal pressed phonations at two different frequencies (E3 [165 Hz] and G3 [196 Hz]). Modal phonation was induced by instructing the examinees to phonate as they usually spoke. Falsetto phonation was induced by instructing the examinees to phonate in falsetto. Modal breathy phonation was induced by instructing the examinees to phonate with a sufficient amount of air. Modal pressed phonation was induced by instructing the examinees to phonate with strong glottal closure. Each phonation was performed at three different intensities (weak and strong). The vowel /e/ was chosen to obtain optimal exposure during the endoscopic examination.

A high-speed digital camera (FASTCAM-1024 PCI; Photron, Tokyo, Japan) was used in this study. The rigid endoscope (#4450.501; Richard Wolf, Knittlingen, Germany) was connected to this camera via an attachment lens ($f = 35$ mm; Nagashima Medical Instruments, Tokyo, Japan). The recording was performed at a frame rate of 4500 fps with an image resolution of 400 × 512 pixels, 8-bit grayscale, and memory size of 12 GB, which allowed a sampling duration of 5.57 seconds. EGG and sound signals were simultaneously recorded with HSDI. EGG signals were recorded using a 1-channel electroglottograph (Laryngograph, Greater London, United Kingdom). Sound signals were recorded using a dynamic microphone (SM58; Shure Inc., Chicago, United States), which was fixed 30 cm anterior to the mouth of the examinees. Those data were modified by a microphone amplifier (FP11; Shure Inc.) and sampled at 25 kHz as the 16-bit data by an analog-to-digital converter (PCI-360116; Interface, Hiroshima, Japan). Newly HSDI-derived O_qs

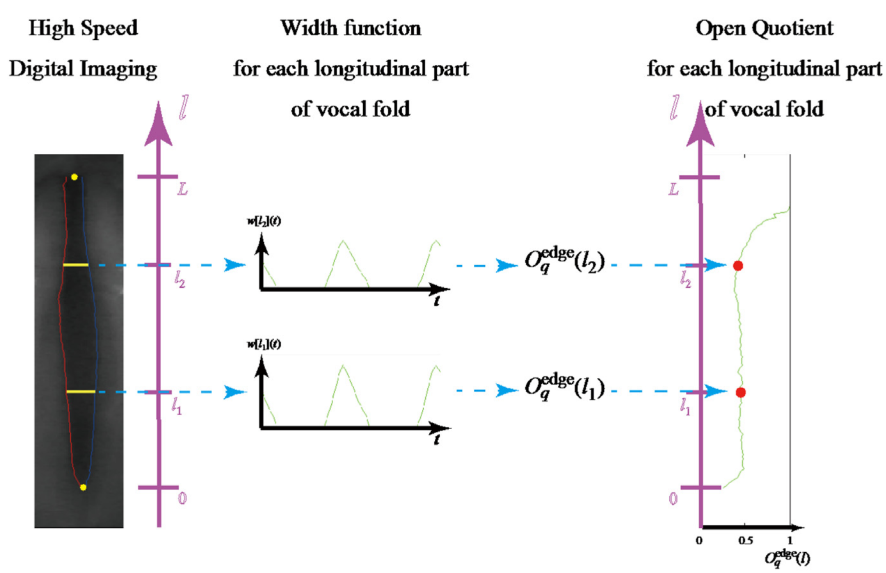

FIGURE 1. Procedure used to calculate $\overline{O_q^{\text{edge}}}$ from high-speed digital imaging. Using the program implemented in MATLAB, the coordinates of the free edge were extracted in pixels from high-speed digital imaging, and each O_q^{edge} was calculated from the edge width-time function on each line.