# The Effects of Humming on the Prephonatory Vocal Fold Motions Under High-Speed Digital Imaging in Nondysphonic Speakers

***Toshihiko Iwahashi, *Makoto Ogawa, *,†Kiyohito Hosokawa, *Chieri Kato, and *Hidenori Inohara,** **Suita, †Osaka, Japan*

**Summary: Objectives.** This study aimed to investigate whether humming affects the adductive motion of the vocal folds and transient glottal closure in the prephonatory adjustment phase of vocal onset using high-speed digital imaging (HSDI) and a motion analysis software program.

**Methods.** Twenty normal healthy adults without any vocal abnormalities were enrolled. While a transnasal flexible fiberscope connected to a high-speed camera was inserted, each participant was asked to perform three phonatory tasks—natural /e:/ phonation, loud /e:/ phonation, and humming /m:/ phonation—and laryngeal HSDI movies (4000 frame/s) were recorded. On each HSDI movie, the duration of the prephonatory glottal closure was measured. In addition, using motion analysis, the changes in the angle between the bilateral vocal folds during vocal fold adduction and the average angular velocity in the ranges of 100%–80%, 80%–20%, and 20%–0% from all of the angular changes were analyzed.

**Results.** The angular changes showed sigmoid and polynomial-like curves during the natural/humming and loud phonation, respectively, and the 80%–20% and 20%–0% average velocities were the highest during the natural/humming and loud phonation, respectively. The humming phonation decreased all of the average regional velocities, eliminated the transient prephonatory glottal closures observed during the natural and loud phonation, and induced a greater value for the minimal angle than the natural phonation.

**Conclusions.** The present study demonstrates that humming encourages easy vocal initiation by decelerating the vocal fold adductive motion throughout the prephonatory adjustment phase and alleviating transient prephonatory laryngeal closure, leading to gradual and smooth vocal fold positioning.

**Key Words:** Humming–Loud phonation–Vocal onset–High-speed digital imaging–Motion analysis.

## INTRODUCTION

Humming is a well-known vocal training technique for producing a resonant voice.[1–5] Previous textbooks on voice therapy[1–3] have emphasized the importance of feeling resonance in the nose, cheeks, or lips during humming in order to effectively induce a hum. Yiu et al[4,5] also described that speakers should produce the sound /m:/ followed by gliding the pitch to the most comfortable and natural level during humming as if sincerely agreeing with someone. Indeed, voice therapy sessions using humming have been reported to improve the perceptual vocal quality in patients with vocal nodules or laryngitis[4] and those diagnosed to have muscle tension dysphonia (MTD) with supraglottic compression,[6] and to decrease the computed perturbation parameters of acoustic and electroglottographic (EGG) signals in MTD patients.[7] To elucidate the mechanisms underlying the vocal improvement by humming, our research group has assessed the laryngeal changes that occur during sustained humming phonation using objective methodologies.[8–10] The results of these studies revealed multifarious effects of humming in both dysphonic patients and nondysphonic speakers. The effects were as follows: (1) an immediate decrease in the degree of supraglottic compression in both MTD patients and nondysphonic speakers;[8] (2) an immediate decrease in the perturbation parameters of EGG signals, reflecting the irregularity of the contact between the bilateral vocal folds in the vibratory cycles in both MTD patients and nondysphonic speakers;[9] (3) an immediate induction of a slight increase in the contact quotient (CQ), ie, the temporal ratio of the vocal fold closed phase to each period during phonation, in MTD patients;[9] and (4) an immediate decrease in the perturbation parameters of the glottal area waveforms derived from the laryngeal movies recorded by high-speed digital imaging (HSDI) in a group of nondysphonic speakers and dysphonic patients with benign vocal mass lesions combined.[10] These results suggest that humming modulates the spatial interrelationship between the laryngeal structures, leading to the optimization of the phonatory dynamics in the larynx.

In addition, a textbook by Colton and Casper[2] mentions encouragement of easy vocal initiation as one of the rationales for the use of humming. Werner-Kukuk and von Leden[11] define vocal initiation/attack/onset as "a function of vocal cord placement at the start of phonation which in turn depends on the precise coordination between the subglottic pressure and the resistance at the level of larynx." Recently, Wittenberg et al[12] divided the process of vocal initiation into four main events: adduction of the vocal folds, prephonatory glottis closure, onset of the oscillation, and steady-state oscillation. In contrast, Watson et al[13] stated that "the initiation of phonation includes two somewhat distinct phase: (1) the prephonatory adjustment phase that is associated with setting the appropriate tension, gross adduction,

Accepted for publication September 8, 2016.

Conflict of interest: The authors declare no conflicts of interest in association with the study.

Disclosure: The authors alone are responsible for the content and writing of the paper.

From the *Department of Otorhinolaryngology—Head and Neck Surgery, Osaka University Graduate School of Medicine, Osaka, Japan; and the †Department of Otorhinolaryngology, Osaka Police Hospital, Osaka, Japan.

Address correspondence and reprint requests to Makoto Ogawa, Department of Otorhinolaryngology—Head and Neck Surgery, Osaka University Graduate School of Medicine, 2-2 Yamadaoka, Suita, Osaka 565-0871, Japan. E-mail: mogawa@ent.med.osaka-u.ac.jp

Journal of Voice, Vol. 31, No. 3, pp. 291–299

0892-1997



http://dx.doi.org/10.1016/j.jvoice.2016.09.008

and aerodynamic forces, and (2) the attack phase that is associated with the onset of the vocal fold oscillation and sound generation." Thus far, variable patterns of vocal onset have been characterized using various experimental methodologies, such as high-speed cinematography,[11,14] pneumotachography,[15,16] electromyography (EMG) of the intrinsic laryngeal muscles,[17,18] EGG,[13,19] image analyses of laryngeal video movies,[20] and computed motion analyses of HSDI movies of the larynx.[21–23] In particular, Werner-Kukuk and von Leden [11]expounded on the three main types of physiological vocal attacks, namely the breathy, the hard, and the soft attacks, as follows: (1) in the breathy attack, the vocal folds assume "a paramedian position prior to sound production"; (2) in the hard attack, "the glottis is closed tightly before the onset of phonation"; and (3) the soft attack is "a smooth onset of phonation without any aspirate sound, and is produced through gradual adduction of the vocal cords towards the median line, with the gentle closure of the glottis." Especially regarding the prephonatory adductive vocal fold movement in the three patterns of phonatory onset, the velocity of the vocal fold adduction in the prephonatory adjustment phase has been reported to be faster and slower in the hard and breathy attacks, respectively.[11,20] Colton and Casper[2] also describe that rapid and complete adduction of the vocal folds prior to the initiation of phonation characterizes a hard glottal attack. In recent decades, with advancements of digital technology, computed image analyses of the normal laryngeal videos now enable the determination of the kinematics of the vocal fold adductive motion in the prephonatory adjustment phase, but in a discontinuous manner.[20,24,25] Later, the use of HSDI has allowed for analyzing the continuous changes in parameters derived from the laryngeal images.[26,27] In particular, we recently applied the combination of HSDI and a motion analysis software program to analyze the vocal fold adductive motion during throat clearing (TC) and normal phonation, and confirmed that the kinematics of the angular changes between the vocal folds were different: a sigmoid and a polynomial-like curve during normal phonation and TC, respectively.[27]

Regarding the patterns of prephonatory transient glottal closure at the three patterns of vocal onset, Wittenberg et al[21] reported the digital kymographic characteristics as follows: (1) in the normal phonation, the prephonatory vocal fold adduction movement is followed by a period of prephonatory glottal closure, and during this period the ventricular folds show a discrete adduction movement; however, the view at the vocal folds is never obscured; (2) in the hard attack, a more prolonged period of prephonatory standstill, ie, a transient laryngeal closure where the adductive movement of the ventricular folds that temporarily cover the vocal folds can be observed; (3) in the breathy attack, the vocal folds do not end up in a vocal fold contact, and the glottal closure remains incomplete during the phase of prephonatory buildup of the muscular tension with a shorter prephonatory standstill.

In the present study, by applying our previous methodology using the combination of HSDI and a motion analysis software program,[27] we aimed to verify (1) whether or not humming affects the velocity of the adductive motion of the vocal folds; (2) whether or not humming changes the minimal vocal fold angle, approximate to the terminal phonatory position of the vocal folds; (3) whether or not humming modulates the duration of the prephonatory transient glottal closure in the vocal onset of nondysphonic speakers; and (4) whether or not loud phonation like shouting, a representative of inappropriate phonation in the view of vocal hygiene, changes the above-mentioned measures.

## METHODS

### Participants

The protocol of this study complied with the Declaration of Helsinki, and an institutional review board approval was obtained from the Osaka University Graduate School of Medicine (No. 15592). Twenty healthy nondysphonic participants without any musical or theatrical training (medical staff members and postgraduate students) who showed neither functional nor organic laryngeal abnormalities (13 men and 7 women; median age: 32.5 years; range: 23–60 years) were included. Before the recording of the laryngeal HSDI videos, each patient underwent a routine ENT examination to confirm the absence of laryngeal abnormalities.

### Recording of the laryngeal HSDI videos and EGG signals

In a common ENT consultation room without soundproofing, the laryngeal HSDI movies and EGG signals during the phonatory tasks were recorded synchronously by a single laryngologist (TI). After placing the EGG electrodes (Model EG2-PCX2; Glottal Enterprises Inc, Syracuse, NY) on the neck skin surface and topically administering 1% lidocaine and 0.02% adrenalin into the nasal cavity, each participant in a seated position underwent the insertion of a flexible rhino-laryngo fiberscope (ENT-P4; Olympus, Tokyo, Japan) connected to a monochrome high-speed camera system (Phantom Miro eX4; Vision Research Inc, Wayne, NJ) and a 300W xenon light source (CLV-S40 Pro; Olympus). The tip of the fiberscope was located at a level just below the uvula. Subsequently, under the observation of the larynx, the laryngologist asked the participant to perform three tasks, namely (1) stable natural phonation of "/e:/" at a habitual pitch and loudness for more than 3 seconds and (2) loud phonation of the vowel "/e:/" like a shout for more than 3 seconds, followed by (3) humming phonation to close his or her lips and hum "/m:/" for more than 3 seconds in a relaxed manner while feeling resonance in the nose or lips, without changing the pitch as possible. In cases where an individual failed to produce a hum appropriately or sufficiently loud voice, the performance of the tasks could be repeated up to three times. Consequently, all of the participants were judged to be able to perform all of the tasks satisfactorily. While the participant was performing the tasks, the HSDIs and EGG signals were recorded synchronously. HSDIs of $256 \times 256$ pixels in size were captured at 4000 frames/s at 16-bit resolution. The digitalization and synchronous data acquisition were achieved using a data acquisition device (NI-USB 6341; National Instruments, Tokyo, Japan) set at a sampling frequency of 48,000 Hz with 16-bit quantization. The recorded data were transferred to a *Windows PC* (ThinkPad E540; Lenovo, Tokyo, Japan), converted to the cinefile format and then stored. The data recordings and transfer were completed for all of the participants within 20 minutes, because the transfer of data of each task requires a few minutes.