Effects of Volume, Pitch, and Phonation Type on Oscillation Initiation and Termination Phases Investigated With High-speed Videoendoscopy

*Melda Kunduk, †Takeshi Ikuma, ‡David C. Blouin, and †Andrew J. McWhorter, *‡*Baton Rouge and* †*New Orleans, Louisiana*

Summary: Objectives. This study aimed to investigate the effects of varying volume, pitch, and phonation types on the initiation and termination phases of vocal fold oscillation using high-speed digital videoendoscopy. Specifically, it addressed the effects of the variation of volume, pitch, and phonation type (normal, pressed, and breathy) on the transient duration of the vibrating glottal length (length transient duration, T_{len}), the transient duration of the glottal area waveform (area transient duration, T_{area}), the time offset between the beginning (or the end) of the full-length vibration and the full-amplitude vibration, T_{Δ}, and the variation of the fundamental frequency during the vocal fold oscillation initiation and termination segments (pitch instability, %PI).

Methods. A female subject with no voice problem produced voices with varying pitch and loudness, including comfortable pitch and comfortable loudness, normal pitch loud, high pitch and comfortable loudness, and high pitch and loud. Breathy and pressed phonations were also recorded. Each of the six phonation types was recorded six times, which resulted in 72 transient segments (each recording included both initiation and termination phases). Mixed model statistical analyses were employed to the five objective high-speed digital videoendoscopy parameters.

Results. Preliminary findings demonstrated significant findings for voice type effects for the length and area transient durations for the oscillation initiation segment but not for the oscillation termination segment.

Conclusions. This study demonstrates that voice types appear to influence vibration initiation patterns more than the vibration termination patterns.

Key Words: High-speed videoendoscopy–Vocal fold oscillation initiation segment–Vocal fold oscillation termination segment–Oscillation initiation–Oscillation termination.

INTRODUCTION

The vibratory patterns during the initiation and termination of vocal fold vibration are thought to be sensitive to changes created by voice disorders and when examined may expose the limitations of the phonatory system. These phases of vocal fold vibration may serve as areas of assessment in the evaluation of vocal function, its pathology, and its treatment outcomes. The goal of this study was to determine the effects of the variations of volume, pitch, and phonation type (normal, pressed, and breathy) on the vocal fold vibration characteristics during these transient phases.

Previous studies have investigated and supported the value of the initiation and termination phases. Ludlow and Connor[1] reported on a reaction time task (using acoustic and laryngograph signals) and found that the time between the onset of laryngeal movement and the onset of phonation was significantly increased in patients with spastic dysphonia compared with normal controls. Phonation threshold pressure (PTP) is an aerodynamic parameter that estimates the minimum lung pressure needed to initiate and sustain vocal fold oscillation.[2,3] Studies have demonstrated that PTP is increased by dehydration[4–6] and following a vocally fatiguing task.[7] PTP is thought to relate to vocal fold tissue properties and glottal configuration, and it has been considered as an indicator of ease of phonation and vocal fold health.[2,3] Another aerodynamic parameter, the phonation threshold flow (PTF), is defined as the minimal glottal airflow to initiate and sustain phonation as the airflow alternate to PTP.[8–10] PTF is also considered to be sensitive to changes in the biomechanical properties of the larynx and its function. Zhuang and colleagues used PTF to differentiate between normal and pathologic voices.[10] Direct observation of changes in vocal fold vibratory patterns during the initiation and termination phases of phonation holds promise in the assessment and treatment of voice disorders. Disease processes can potentially change tissue elasticity, mass, length, and the pre-phonatory gap, which may alter the vibratory characteristics in the initiation and termination phases.

Several studies using various approaches—acoustic[11–17] and aerodynamic measures,[4–7] physical modeling,[18] excised larynges,[19–21] electromyography,[22,23] electroglottography,[24–28] videostroboscopy,[29] and high-speed videoendoscopy (HSV)—have investigated vocal fold vibration at initiation and termination phases.[30–40] Unique among these methodologies, HSV allows direct observation of these very short and transient vocal fold vibrations. It can assess the vocal fold vibration within a cycle, including variation in the oscillation patterns along the vocal folds and between them individually. This ability to directly evaluate subtle variations in vibratory behavior is potentially useful in the determination of voice-disordered vocal function. Moore[37] and other investigators have used HSV to examine vocal fold

Accepted for publication August 23, 2016.

From the *The Department of Communication Sciences and Disorders, Louisiana State University, Baton Rouge, Louisiana; †Department of Otolaryngology-Head and Neck Surgery, Louisiana State University Health Sciences Center, New Orleans, Louisiana; and the ‡Department of Experimental Statistics, Louisiana State University, Baton Rouge, Louisiana.

Address correspondence and reprint requests to Melda Kunduk, The Department of Communication Sciences and Disorders, Louisiana State University, 68 Hatcher Hall, Field House Drive, Baton Rouge, LA 70803. E-mail: mkunduk@gmail.com

Journal of Voice, Vol. 31, No. 3, pp. 313–322
0892-1997

http://dx.doi.org/10.1016/j.jvoice.2016.08.016

vibratory patterns during the initiation phase[30–40] and the termination phase.[28,32,35]

Duration of the transient phase is one of the primary target measurements in the study of vocal fold vibration.[30–40] The transient phases are proposed to be the time periods from vibration onset until the vocal folds enter steady-state vibration and from the time they release from steady-state vibration until the disappearance of vibratory behavior. Previous HSV studies vary in their methodology in determining the beginning and the end of the steady-state oscillation as well as the timing of the vibration onset and offset. Accurately measuring the transient durations is challenging because the steady-state vibration is inherently not truly steady. Thus, the definition of the boundaries between the transient state and the steady-state is subjective (even for objective measures). In previous reports on the initial vocal fold vibration, steady-state oscillation is based mainly on the amplitude growth of vocal fold vibration either in the glottal width or in the glottal area. Using this methodology, Eysholdt and colleagues analyzed vocal fold vibration in normal *vs* pathology and found that the vocal fold with a lesion starts oscillating 100 ms later than the healthy vocal fold.[30] Kunduk et al[35] investigated age-related changes and reported that an older subject had longer duration of initiation and termination phases than a younger subject.

Other studies have noted change in the glottal length variation during the initiation period. Moore[38] suggested that initial oscillation starts where the folds are least resistant, eventually extending to the full vocal fold length. Werner-Kukuk and von Leden[39] studied vocal vibratory patterns with HSV on imitated soft, hard, and breathy vocal initiations in subjects with normal voice. A rapid gain of amplitude of vibration was observed with a hard vocal attack compared with a soft attack, but the soft attack achieved full-length vibration first. Age has also been seen as a factor, where older women achieved full-length vibration later than young women during the initiation phase.[34]

Inspired by the reported variations in the glottal length during the initiation, we previously introduced an automatic detection algorithm for the timing of attainment and loss of full-length vibration.[32] We found that vibration length is less susceptible to variations in oscillatory characteristics during the steady-state phase than the glottal area root mean square (RMS) waveform. The observed measurement consistency is desirable to provide a steadier transient duration measurement than those based on the amplitude waveforms (ie, less measurement noise because of the expected variation in the steady state).

Change in the fundamental frequency (F_0) has also been investigated as another transient measurement of interest. Acoustic studies have investigated the short-term changes in the F_0 immediately before and after voiceless consonants (ie, offset and onset of voicing) using the relative fundamental frequency (RFF) measure.[15–17] The RFFs were measured from the periods of the first observable cycles around the voiceless consonant relative to the 10th cycle furthest from the consonant. Stepp and her colleagues[16] reported significant lowering of onset and offset RFF in individuals with vocal hyperfunction than in controls. This result was interpreted to be caused by an increase in overall laryngeal muscle tension that decreased the expected phonetically mediated vocal fold tension.[16] F_0 variation can also be assessed with HSV, although the variation behavior may differ from in-speech assessment because the voiceless consonant is not present for sustained vowel phonation in the HSV examination. Measurement of F_0 variation still may be helpful in assessing the effect of the change in tension and stiffness induced by voice disorders. The frequency variation may represent a vocal effort level used to initiate vocal fold vibration or pliability at the termination of vibration.

This study investigated the effects of the variations in volume, pitch, and phonation type (normal, pressed, and breathy) on the HSV objective parameters during transient phases. Breathy and pressed (or hard) phonation types were chosen because they are among the most frequently identified phonation types in subjects with voice disorders, with varying etiologies in clinic settings. The parameters include the transient duration of the vibrating glottal length (length transient duration, T_{len}) and the variation of the F_0 during the vocal fold oscillation initiation and termination segments (pitch instability, %PI). The %PI is measured by the discrepancy of the number of glottis cycles measured against the expected number during the transient phase. In addition, two other outcome parameters were examined in the study: the transient duration of the glottal area waveform (area transient duration, T_{area}) and the time offset between the beginnings (or the ends) of full-length vibration and the full-amplitude vibration, T_{Δ}.

METHODS

Subjects and HSV data collection

The HSV data were recorded at 8000 frames per second using a Color High-Speed Video System, Model 9710 (KayPENTAX, Montvale, NJ) while a female subject (M.K.) with healthy voice produced voices with varying pitch and loudness, including comfortable pitch and comfortable loudness (NPNL), normal pitch loud (NPL), high pitch and comfortable loudness (HPNL), and high pitch and loud (HPL). In addition, the same subject imitated breathy and pressed phonation voices using comfortable volume and pitch. The subject, who is a speech language pathologist with an extensive training and experience in assessment and treatment of vocal function in a clinic setting, is capable of imitating breathy and pressed voice precisely and consistently. Block randomization was used for the HSV data collection to address the fatigue effect, and each phonation type was collected six times (a total of 72 samples, as one each of initiation and termination samples were captured in each recording).

One aim in data collection was to obtain samples that are consistent in both pitch and volume within each voice type. During the HSV data collection, the F_0 and volume of the voice was monitored using KayPENTAX *Computerized Speech Lab* (Model 4150) with a headset microphone. The subject could see the volume and pitch measurements to ensure that the volume and pitch were on target. The target pitch and loudness was determined before the HSV data collection. Target frequencies and volumes were set from pre-trial measurements. A typical comfortable pitch of this speaker was 217 Hz, and the sound pressure