The Movements of Vocal Folds during Voice Onset Time of Korean Stops

Ki Hwan Hong* · Hyun Ki Kim · Yoon Soo Yang* · Bum Kyu Kim* · Sang Heon Lee***

ABSTRACT

Voice onset time (VOT) is defined as the time interval from the oral release of a stop consonant to the onset of glottal pulsing in the following vowel. VOT is a temporal characteristic of stop consonants that reflects the complex timing of glottal articulation relative to supraglottal articulation. There have been many reports on efforts to clarify the acoustical and physiological properties that differentiate the three types of Korean stops, including acoustic, fiberscopic, aerodynamic and electromyographic studies. In the acoustic and fiberscopic studies for stop consonants, the voice onset time and glottal width during the production of stops has been known as the longest and largest in the heavily aspirated type followed by the slightly aspirated type and unaspirated types. The thyroarytenoid and posterior cricoarytenoid muscles were physiologically inter-correlated for differentiating these types of stops. However, a review of the English literature shows that the fine movement of the mucosal edges of the vocal folds during the production of stops has not been well documented. In recent years, a new method for high-speed recording of laryngeal dynamics by use of a digital recording system allows us to observe with fine time resolution. The movements of the vocal fold edges were documented during the period of stop production using a fiberscopic system of high speed digital images. By observing the glottal width and the visual vibratory movements of the vocal folds before voice onset, the heavily aspirated stop was characterized as being more prominent and dynamic than the slightly aspirated and unaspirated stops.

Keywords: Voice Onset Time, High Speed Digital Images, Vocal Fold Movements

1. Introduction

VOT is a temporal character of stop consonants that reflects the types of stop such as the aspirated or non-aspirated types. The stop consonants in Korean are classified into three types according to the manner of articulation as unaspirated (UA), slightly aspirated (SA) and heavily aspirated (HA) stops. Both the UA and the HA types are always voiceless in any environment. The SA type is voiced in word-medial position when both the preceding and the following sounds are voiced, but voiceless in other

* Dept. of Otolaryngology-Head and Neck Surgery, Chonbuk National Univ.
** Institute for Speech Sciences, Chonbuk National Univ.

environments. Generally, the length in time from articulatory explosion to voice onset stands for the voice onset time (VOT) in the initial phase of utterance. The VOT of the UA type shows within 20 msec, and about 40-50 msec in the SA and 50-70 msec in the HA[1] [2].

There have been many reports on efforts to clarify the acoustical and physiological properties that differentiate these manner categories, including acoustic, fiberscopic, aerodynamic and electromyographic studies. Umeda et al.[3] studied the Korean stops, affricates and fricatives with sound spectrography. The fundamental frequency at voice onset after both the UA and HA consonants was higher than that for the SA consonants, and the voice onset times were longest in the HA followed by the SA and UA. Han et al.[1] reported in their speech synthesis and perception studies that the SA and UA stops differed primarily in terms of a gradual versus a relatively rapid intensity build-up of the following vowel after the stop release. Lee et al.[4] measured both the intraoral and subglottal air pressure simultaneously during the production of the three kinds of stops. They found that the subglottal pressure was higher for the HA stop than for the other two stops. They also compared the dynamic pattern of the subglottal pressure slope for the three categories and found that the HA stop showed the most rapid increase in subglottal pressure in the time period immediately before the stop release. They concluded that the HA stop was the most dynamic in this respect.

Kagaya[5] reported fiberscopic and acoustic studies of the Korean stops. He mentioned that the UA type may be characterized by a completely adducted state of the vocal folds, stiffened vocal folds and the abrupt decreasing of the stiffness near the voice onset, increasing subglottal pressure and/or lowering of the glottis before explosion, while the HA type may be characterized by an extensively abducted state of the vocal folds and a heightened subglottal pressure. On the other hand, none of these positive gestures are observed for the SA type.

Hong et al.[6] conducted an electromyographic study using thyroarytenoid and posterior cricoarytenoid (PCA) muscles during stop production. He reported that the most marked and earliest activation of the PCA muscle associated with a steep reactivation of the thyroarytenoid muscle before the voice onset were most characteristic for the production of the HA consonants. For the production of the UA consonants, little or no activation of the PCA muscle associated with the earliest and most marked reactivation of the thyroarytenoid muscle was characteristic. For the SA consonants, he reported a more moderated activation of the PCA muscle than for the UA consonant, and the least and the latest reactivation of the thyroarytenoid muscle among the three types of Korean consonants.

In the present paper, the results of observation of the vibratory movements of the vocal fold edges before voice onset will be discussed with physiological characterizations in terms of laryngeal gestures according to the different types of stop consonants. The movements of vocal fold edges were evaluated using high speed digital images. The EGG signals and acoustic waveforms were also evaluated and compared with the vibratory movements of vocal fold edges during stop production.

2. Methods

A native Korean speaker of the Chonbuk dialect was the subject in this experiment and did not have any laryngeal lesions. The test words were prepared so as to place the consonants in different phonological environments, and were all meaningful words. These words were uttered in the frame sentence / *i k əs i* <u>*CVCV*</u> *i d a* / (This is /CVCV/). The bilabial stops /p, p^h, p'/ for consonant /C/ were used, where /V/ was always /i/. The /p/ stands for SA, /p^h/ for HA and /p'/ for UA. The subject was asked to phonate with comfortable pitch and loudness and the recording was done in a soundproof room.

The high speed digital imaging system which was developed at the Department of Speech Physiology, Tokyo University, Japan, was used in this study[7]. The fiberoptic nasopharyngoscope was inserted through the nostil. In this high speech digital imaging system, frame rate and data length (maximum recording time) are in a trade-off relation. We chose a frame rate of 4,500 frames per second. The duration of the recording was automatically set at 0.7 second. The speech and electroglottographic (EGG) signals were recorded simultaneously with video-images. They were formatted with MS-DOS binary files and the A/D conversion was performed by 12 bit and 2 channels mode. The A/D conversion rate was synchronized to the external clock pulse and the clock rate was 18.0068 Hz (4 clock pulses per frame). The speech and EGG file contained an extra 1021 frames of data in addition to the initial 3072 frames of data. The file format of the image data file was MS-DOS binary file and one image file contained 64 frames image data. The recorded images were displayed on the computer in slow motion and the edges of both vocal folds were automatically detected. After extraction of the threshold between the vocal fold edges, the brightness curves of sequential frames were drawn based on the threshold (Figure 1). The time course of the movement and vibration of the vocal folds during the production of stops was displayed.

The parameters evaluated in this study were the glottal widths at the moment of stop explosion and the movements of the vocal fold edges during voice onset time (after stop explosion to vowel production) according to the three types of stop. The relationships of EGG signals and acoustic waveforms were also discussed with the vibratory movements of the vocal fold edges.

3. Results

At the moment of the explosion of stops the SA type was characterized by a widely abducted state of the vocal processes of vocal folds and more widely abducted in the HA stop as in Figure 2. However, the vocal processes of both vocal folds were nearly adducted in the UA stop. After stop explosion the movement of vocal processes was different according to the types of stop. The vocal processes were approximated shortly in the UA, but gradually in the SA and HA stops, by observing the video recorded images.

Figure 3, 4 and 5 show the speech and EGG signals and the brightness curves of sequential frames by extraction of the threshold between the vocal fold edges during the production of stops. In the SA type (Figure 3) the vocal fold edges moved inward, gradually and smoothly, after stop explosion and before vowel production. After an appropriate inward movement of vocal processes, both vocal fold edges vibrated symmetrically by moving laterally and then medially after maximal lateralization. During this period the vocal processes moved still inward, but not completely approximated. After this incomplete vibratory curve of vocal fold edges, the vocal processes closed completely and the vocal fold edges vibrated periodically for vowel production. The EGG signal was started at the end of a first incomplete vibratory curve. However, the acoustic signal was not distinct during this period, but showed only slight traces.

In the HA stop (Figure 4) the vocal fold edges moved inward gradually, but not smoothly comparing to the SA after stop explosion. With appropriate adduction of vocal processes the vocal fold edges vibrated also symmetrically with small amplitude (a first symmetrical vibratory curve). After this first curve, the vocal fold edges moved inward again looking as nearly approximated, but not completely closed as shown by the video images. This approximation of vocal fold edges resulted from the large vibratory amplitude of mucosal waves with the incomplete closure of vocal processes by the careful observation of video images. After this approximation, the vocal fold edges moved laterally again and then moved medially after maximal lateralization. During this period the vocal processes were seen still moving gradually inward. After this second vibratory curve of the vocal fold edges, the vocal processes closed completely and the vocal fold edges vibrated periodically for vowel production. The EGG signal was not produced at a first symmetrical vibratory curve, and started at the end of the second symmetrical vibratory curve of the vocal fold edges. The acoustic signal was not produced distinctly at the period of the first vibratory curve of vocal fold edges, but very small amplitude was documented. The acoustic signals were started distinctly from the second vibratory curve of vocal fold edges.

In the UA stop (Figure 5), the vocal processes kept nearly closed after stop explosion.

With a nearly adducted status of vocal processes, both vocal fold edges vibrated symmetrically, moving laterally and medially. During this period, the vocal processes looked nearly approximated when observing the video images, but not completely closed. After this vibratory curve of the vocal fold edges, the vocal processes closed completely and the vocal fold edges vibrated periodically for vowel production. The EGG signal was started at the end of this first symmetrical vibratory curve. The acoustic waveform was not produced during this vibratory curve.

4. Discussion

It has been well known that the temporal change in glottal width occurs at the explosion of voiceless stop consonants in the initial phase of utterance. And then narrowing of the glottis begins immediately after the glottal width has reached its local maximum. In this study, the glottal widths showed largest in the HA consonant and followed by the SA type, but there was a small gap during the production of the UA consonant. According to the literature review, Hong et al.[2] reported same results using ordinary nasopharyngeal fiberscope. Kagaya[5] also reported that the glottis closes gradually before articulatory explosion and the glottis shows opened still at the explosion of stop with some fluctuations of glottal width with an ordinary fiberscopic examination. These differences of glottal width and condition result from the different activities of the laryngeal muscles. Hong et al.[6] reported that during the HA stop production the marked and earliest activation of the posterior cricoarytenoid (PCA) muscle before activation of the thyroarytenoid muscle was most characteristic, but there was no activation of the PCA muscle and marked activation of the thyroarytenoid muscle in the UA stop. He also found a more moderated activation of the PCA muscle and the least reactivation of the thyroarytenoid muscle in the SA stop. Hirose et al.[8] [9] also reported on the laryngeal control in stop production that the HA stop appeared to be characterized by a marked suppression of all the adductor muscles immediately before the stop release, and the SA stop by a less predominant suppression of the adductor activities before release. In the UA stop, they reported that the thyroarytenoid muscle and the lateral cricoaryetenoid muscles showed a marked increase in activity before the stop release, which presumably resulted in an increase in inner tension of the vocal folds as well as in constriction of the glottis during or immediately after the articulatory release.

However, the fine movements of vocal fold edges during production of stops has not been well documented in the English literature because of too short of a period of stop production and small frames of laryngeal images with the ordinary fiberscope. On the ordinary fiberscopic data of laryngeal movements during stop production, Kagaya[5] reported that the vocal folds always start to vibrate as soon as they are set in the adducted position in the HA and SA consonants. In the UA, in contrast, adducted vocal folds without any vibration are observed during the period between the articulatory release and the voice onset. He suggested that this difference must be explained by some other independent physical quality beside the glottal width, something pertaining to the physical condition of the vocal folds themselves, such as stiffness[10]. They suggested that a considerable amount of airflow through the slightly open membranous portion (the spindle-shaped glottis) must exist in the HA and SA stops, because turbulent noise is being produced at the articulatory stricture during this period between explosion and voice onset.

In this study using high speed digital images, the HA stop was characterized by the existence of two distinct vibratory curves of the vocal fold edges and one distinct vibratory curve in the SA before the complete closure of vocal processes for vowel production. In the UA stop, a small and incomplete vibratory curve was also noted before complete closure of vocal processes. These differences could be explained by aerodynamic characteristics of Korean stops. The different values of the airflow in /CV/ context are dynamically related to the activity of laryngeal adjustment during the production of stops. Hong et al.[11] reported that the glottal airflow during production of stops was highest in the HA stops, then followed by the SA stops, and in the UA stops it was lowest. Dart[12] also reported a low airflow after UA stops release and a higher airflow after SA stop release. The airflow was determined by the glottal width between the vocal folds at the release of stops. The volume of air used during a articulatory event depends mainly on the subglottal pressure, vocal tract impedance, and the events duration. In the Korean stops the subglottal pressure was determined by indirectly measuring the intraoral pressure during pronunciation of /p/, /p^h/ and /p'/, respectively. This intraoral pressure is a key of articulatory variable. The peak air pressure of the UA is higher than the SA. But its pressure of UA was not clearly distinguished between the HA and the UA stops[11].

The difference in glottal width and vibratory curves of the vocal fold edges between the HA and the SA types could also be compared in the acoustic analyses. On the visual observation of sound-spectrograms for the two types, it is found that the apparent intensity of aspiration noise is higher for the HA type than for the SA type, while the opening degree of the glottis is larger for the HA type. If the subglottal pressure for the HA type is considerably higher than that for the SA type, the intensity of aspiration noise for the HA type should be strong when compared with that for the SA type. The comparatively low subglottal pressure for the SA type together with the somewhat abducted position of the vocal folds at the moment of articulatory release causes the delay of voice onset in the /CV/ utterance. In the UA stop, the vocal fold vibration may be initiated by an increase of transglottal pressure difference caused by the articulatory release.

As a result, by observing the vibratory movements of vocal fold edges during stop production, the HA stops can be said to be characterized by more active articulatory dynamics than the SA and UA stops because the intraoral pressure and oral airflow for the HA is higher than for the SA and UA stops. The high subglottal pressure and airflow result in more dynamic vibration of the vocal fold edges. Even though with small airflow, narrow glottal width and stiff vocal folds during production of the UA stop, the high subglottal pressure results in weak vibratory movement of the vocal folds edges before the complete closure of vocal processes. Finally, on the glottal width and the visual vibratory movements of vocal folds before voice onset, the HA stop was characterized as being more prominent than the SA and UA stops.

References

[1] Han, M. S. & Weitzman R. S. 1970. "Acoustic features of Korean /P, T, K/, /p, t, k/ and /p^h,t^h,k^h/." *Phonetica,* 22, 112-28.

[2] Hong, K. H., D. S. Chon, Y. J. Kim & K. Y. Jung. 1992. "Laryngeal adjustments for Korean stops-acoustic, electromyographic and fiberoptic analysis." *Korean J Otolaryngol Head Neck Surg,* 35, 770-82.

[3] H, Umeda & Umeda N. 1985. "Acoustical features of Korean 'forced' consonants." Gengo Kenkyu. *J Linguistics Soc Japan,* 48, 23-33.

[4] Lee, C. Y. & Smith T. S. 1972. "Oral and direct subglottal pressure in Korean stops." *J Acoust Soc Am,* 51: S102.

[5] R, Kagaya. 1974. "A fiberscopic and acoustic study of the Korean stops, affricates and fricatives." *J Phonetics,* 2, 161-80.

[6] Hong, K. H., S. Niimi & H. Hirose. 1991. "Laryngeal adjustments for the Korean stops, affricates and fricatives-an electromyographic study." *Ann Bull RILP,* 25, 17-31. University of Tokyo.

[7] Kiritani, S., H. Imagawa & H. Hirose. 1990. "High-speed digital image recording for the observation of the vocal fold vibration." *Proc ICSLP*, Kobe, 61-4.

[8] Hirose, H., C. Y. Lee & T. Ushijima. 1974. "Laryngeal control in Korean stop production." *J Phonetics,* 2, 145-52.

[9] Hirose H., H. S. Park, H. Yoshioka, M. Sawashima & H. Umeda. 1981. "An electromyographic study of laryngeal adjustment for the Korean stops." *Ann Bull RILP,* 15, 31-43. University of Tokyo.

[10] Halle, M. & K. N. Stevens. 1971. "A note on laryngeal features." *Quarterly progress report,* Research Laboratory, Electronics, M.I.T., No 101.

[11] Hong, K. H., Y. H. Kim & H. K. Kim. 1994. "Aerodynamic characteristics of Korean stops using Aerophone II." *Langauge Studies,* 20, 35-45. Chonbuk National University Press,

[12] Dart, S. N. 1987. "An aerodynamic study of Korean stop consonants." *J Acoust Soc*

Am, 81, 138-47.

Received: Jan. 25, 2002.
Accepted: Feb. 24, 2002.

▲ Ki Hwan Hong, M.D.
Dept. of Otolaryngology-Head and Neck Surgery
Chonbuk National University, Medical School
Chonju, Chonbuk, 560-182, Korea
Tel: +82-63-250-1990 Fax: +82-63-250-1986
E-mail: khhong@chonbuk.ac.kr

▲ Hyun Ki Kim, M.D.
Institute for Speech Sciences, Chonbuk National University
Chonju, Chonbuk, 560-182, Korea

▲ Yoon Soo Yang, M.D.
Dept. of Otolaryngology-Head and Neck Surgery
Chonbuk National University, Medical School
Chonju, Chonbuk, 560-182, Korea

▲ Bum Kyu Kim, M.D.
Dept. of Otolaryngology-Head and Neck Surgery
Chonbuk National University, Medical School
Chonju, Chonbuk, 560-182, Korea

▲ Sang Heon Lee, M.D.
Dept. of Otolaryngology-Head and Neck Surgery
Chonbuk National University, Medical School
Chonju, Chonbuk, 560-182, Korea

LEGEND FOR FIGURES

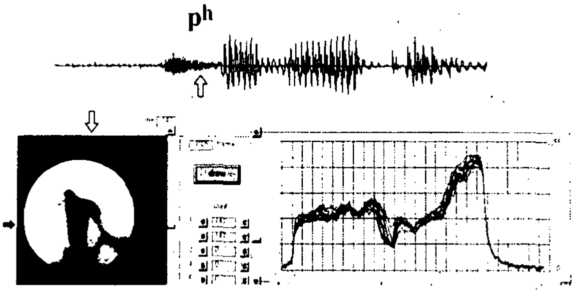

Fig. 1 Glottal width measurement between vocal fold edges. The white arrows stand for the point of the laryngeal image during stop production. The cursor (lower left, black arrow) indicates the brightening threshold between the vocal fold edges. The brightness curves (lower right) of sequential frames are drawing.

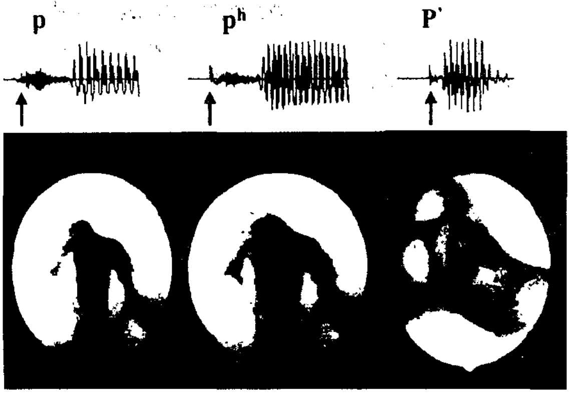

Fig. 2 The glottal widths at stop explosion for three types of stop. The arrows stand for the moments of explosion acoustically. The glottal width is largest in HA and followed by SA and UA stops.

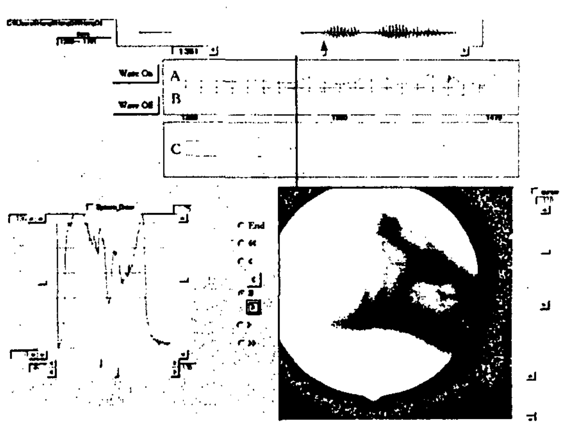

Fig. 3 Movement pattern of vocal fold edges during slightly aspirated stop production. A stands for acoustic waveform, B for EGG signal, and C for glottal width

measure. The vocal fold edges are moving inward, but not completely, with incomplete closure of vocal process during stop production (just before occurring the EGG signals). Vertical line stands for the first complete closure of the vocal fold edges for vowel production.

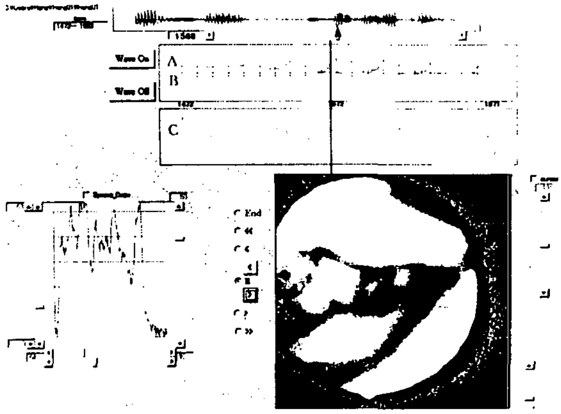

Fig. 4 Movement pattern of Vibration of vocal folds during strong aspirated stop production. The vocal fold edges are moving medially with a near approximation of vocal process during stop production (just before occurring the EGG signals). After occurrence of EGG signals(B), the vocal folds are vibrating completely for vowel production.

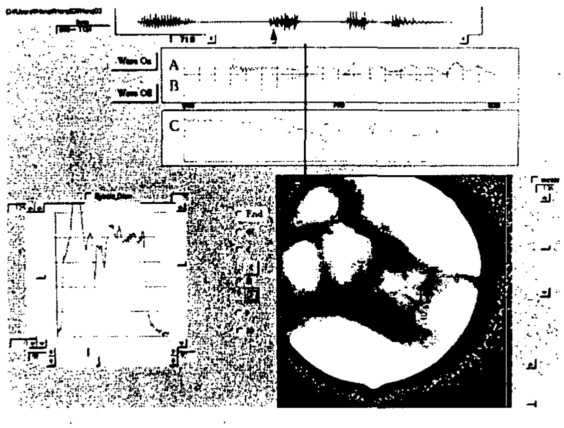

Fig. 5 Vibration of vocal folds during unaspirated stop production. The vocal fold edges are moving inward slightly with a nearly approximation of vocal process during stop production (just before occurring the EGG signals). After the occurrence of EGG signals, the vocal folds are vibrating completely for vowel production.