Glottal Area and Voice Onset Time

Dae-Won Kim

Abstract

There is general agreement that voice onset time (VOT) is functionally related with the glottal opening at the moment of the oral release of a stop. However, systematic investigations of tempo and the place of articulation as affecting the glottal opening and VOT have relatively neglected. Various instrumental techniques were used to verify the claim with BrEng and Korean speakers, under controlled experimental conditions, tempo being one of them. It was found that voiceless aspiration (i.e. VOT) is not simply a function of the glottal area at the moment of the oral release of a stop as it is normally defined in the existing literature. Within a given place of articulation and across tempo, VOT was generally insignificantly related to the glottal area. It is inferrd that the glottal adduction onset time for the following vowel is actively controlled by the speaker to meet aerodynamic requirements in relation to class (i.e. aspirated and unaspirated) and tempo. Some possible underlying physiological mechanisms for various phonetic aspects of intervocalic stops, associated with the glottal area and VOT, were discussed.

I. Introduction

There are some investigations showing the functional relationship between the glottal opening at the stop release and voice onset time. C-Kim (1970), for example, claimed that "Aspiration is defined as a function of the glottal opening at the time of release of the oral closure of a stop"(p. 115). This has been in general agreement with other findings (e.g. for Korean isolated nonsense word-initial and intervocalic stops, see Kagaya, 1974; for Icelandic word-initial stops, Petursson, 1976; for German intervocalic stops in connected speech, see Butcher, 1977; for Danish word-initial stops, see Hutters, 1984). In a study, for instance, with a photo-electric glottograph combined with fiberoptic technique, and partly EMG, Hutters (1984, p.352) reported that for Danish word-initial stops /p^h, b, t^h, d, k^h, g/ as in pile-bile, tie-die, kilde-gilde, etc. embedded in the carrier phrase [di vi si: __], "the duration of aspiration as it is defined by C-kim (1970)- and as it is normally defined." However, the existing data associated with the relationship between GRA and VOT are limited to the overall observation at the normal speech rate. In other words, the investigation of the effect of tempo and the place of articulation both on the glottal opening at the explosion (GRA) and on the voice onset time (VOT) has been relatively neglected, and an investigation to verify the existing claim, using controlled speech materials, is required. In this study, I will confine my attentions mainly to the tense aspirated stops, since the unaspirated stops(lax and tense) are characterized by having the adduction of the vocal folds during the oral closure phrase(eg. Kagaya, 1974).

2. Method

2.1. Subjects

Two male subjects were used in the study. One of the two subjects was British and the other was Korean. The British subjects has English parents, but he had his school education at an American school in Africa. After school he had lived in England for most of his life. He said that he tried deliberately to avoid an American accent. He was adjudged to have a near RP accent. The Korean subject (the author) speaks a near Seoul accent, which is considered to be the standard Korean. None of the subjects had any speaking defects.

2.2. Speech materials

In order to obtain natural data, an attemp was made to construct bisyllabic /'CVCVC/ English words in pairs of natural sentences where the C was a stop. The English test items were structured such that were nearly similar syntactically and prosodically as follows:

British English speech materials

(a) /a: ðeɪ ˈmɒbɪŋ ðə ˚stɔ:z/

Are they mobbing the stores?

(b) /a: ðeɪ ˈmɒpɪŋ ðə ˚flu/

Are they mopping the floor?

(c) /a: ju ˈhi:dɪŋ maɪ əd˚vaɪs/

Are you heeding my advice?

(d) /a: ju ˈhi:tɪŋ maɪ ˚ba:θ/

Are you heating my bath?

(e) /a: ju ˈbægɪŋ maɪ ˚si:t/

Are you bagging my seat?

(f) /aː ju ˈbægɪŋ mɪ ˈʌp/
Are you backing me up?

Korean speech materials
(a) /apa-nɨn tarɨl pomjə ap'a ap'a han-ta /
Apa is saying 'dad' 'dad' looking at the moon.
(b) /pam-i ap^ha ap^ha hamjə on-ta /
The night comes saying pain, pain.
(c) /aka-nɨn ap^hita-rɨl at'a at'a han-ta /
The child says at'a at'a for pain.
(d) / at^ha-ka musɨn t'ɨs-i-ja /
What do you mean by 'at^ha'?
(e) / we tataiksən-ɨl ataiksən ira-he /
Why do you say 'ataiksən' for tataikson?
(f) /dʒəː pupu-nɨn mak^hao-eso manna-s-ta /
We met that couple at Makao.
(g) /ak'a-pon saram-i ak^hino-ja /
The man who you saw just now is Akyno.

(where /p,t,k/ are lax unaspirated stops, /p^h,t^h,k^h/ tense aspirated stops, and /p',t',k'/ tense unaspirated stops.)

In Korean, however, it was almost impossible to design such sentence pairs differing only in the location of stop consonants in question, maintaining /VCV/ structures where the vowel was /a/ and the consonant was an oral stop. As seen in the test items, four of the nine words with the type stop consonants were in sentence-initial position, and the rest of them were positioned within sentences. The number of syllables of the words containing the type stop consonants was not always consistent, such as /ataiks ən/, /at'a/, /at^{h}a-ka/, etc.

There are some observations showing that vowels are affected to a more significant extent than consonants by various factors, such as position (Oller, 1973), stress placement (Fry, 1955), pause (Klatt, 1976), tempo (Gay, Ushijima, Hirose, and Cooper, 1974), etc. For example, in a spectrographic study with seven American speakers using nonsense words in /CVC/, /CVCC/, etc sequences embedded in the frame: "Say _____.", Oller (1973) claimed that "final-syllable vowel increments were approximately 100 ms and final consonant increments were about 20 ms" (p.1235). Klatt (1976) also stated that "the syllable before the pause increases by about 60 to 200 ms, with omst of the durational increment restricted to the vowel and any postvocalic sonorant or fricative consonants" (p.1211). Thus, it is conceivaboe that the position in the sentence and the number of following syllables affect the duration of a predceding vowel in Korean as in the case with English. In justifying the constructed Korean speech items, therefore, my attention was confined to the effect of the position and the number of syllables on the duration of the preceding vowels. In an informal pilot experiment, however, with an electropalatograph and laryngograph using /ataiksən/,/ataata/,/at'a-ka/,/at'a at'a/, etc. in sentence pairs. The results of the experiment showed that the effect of the number of syllables of the duration of the preceding vowel followed by the alveolar stops was not only insignificant but also inconsistent. Also the results demonstrated that the position effect of the duration of the preceding vowel followed by the bilabial stops was insignificant, although overall sentence-initial stops gave a 4% decrease of the preceding vowel than sentence-medial stops.Thus,on the basis of the informal experiment, it was adjudged that the constructed Korean speech materials are generally reasonable.

2.3. Procedures

The speakers produced the sentences with a rise-fall intonation. Each sentence was produced five items at a normal or moderately slow speech rate and five times at a fast speech rate, giving 2(tempi) x 5 (the number of repitition) x 3(the place of articulation) x 3(the manner of articulation) = 90 utterances in Korean.

Simultaneous measurements were made from an electro-palatograph and a lip sensor (both manufactured by the Phonetics Labratory, University of Reading) and a photo-electric glottograph was used for sensing the relative glottal area. Their outputs were synchronized with an electro-laryngograph (manufactured by the Phonetics Labratory, University College London), accelerometer (Knowles electronics BU-1771) and audio microphone. The electro-palatograph (EPG) was interfaced to a Commodore micro-computer (CBM 3032). For print-outsof EPG data, the Commodore microcomputer was used. For more detailed discussion of the instrumentation and the characteristics of the devices, see Kim (1987, p.163-168). Mingograms were made of signals of the devices and of the EPG synchronization pulse. Duration measurements were made out for intervals corresponding to the voice onset time (VOT) from laryngograms synchronized with electropalato-

grams and lip sensor signals. For measuring the relative glottal area at the onset of the release of oral stop closure, a base line indicating a complet closure of the of the vocal folds (i.e. glottal stop) was constructed on the basis on the photo-electro-glottographic signal on the mingograms. At the onset of the release of oral closure, the amplitude of the photo-electrical trace was measured to obtain the relative glottal area (GRA) in mm. For alveolar stops, the onset of the release of oral closure was defined as the begining of a gap in the full lingual contact pattern in the electropalatogram.
An abrupt fall of signal from the lip sensor was considered to be clear indication of the onset of the release of the bilabial stops.

For velar stops, difficulties were experienced in constructing reference lines for the onset/offset of oral closure. The electro-palatographic data did not give a clear-cut indication of either the onset of the oral colsure or the release of the oral closure. One possible reason for this may be that contact for the stops was made further back than the most posterior row of electrodes They may have approximated uvular stops as the last row of electrodes was placed at the junction between the hard and the soft palates. Thus, the measurement of the velar stops was excluded. The experiment was undertaken in the phonetics labratory, the University of Reading. For the statistical analysis, T-test and Pearson's correlation on Statworks were used on the Macintosh Plus computer.

3. Results and discussion

Within a given place of articulation for the BrEng tense aspirated intervocalic stops and across tempo, the correlation of GRA to VOT was observed to be insignificant at the 5% level. This was generally similar to the case with the Korean intervocalic stops, except for the tense aspirated bilabial intervocalic stops(df=B, r=0.646,p<0.05) (see Figs 3 and 5). Both in BrEng and in Korean,therefore, within a given place of articulation of an intervocalic stop and across tempi, VOT was generally independent of GRA, except for the single case associated with the Korean tensed aspirated bilabial intervocalic stops.

The effect of tempo both on GRA and on VOT. As seen in Fig. 4, in fast speech the BrEng tense aspirated bilabial intervocalic stops were manifested with a highly significant decrease (mean 62%) in GRA, compared with normal speech: the mean ratio was 1 (F) : 2.6 (N). This was similar to the case with the Korean tense aspirated bilabial intervocalic stops (mean 37%): the mean ratio was 1 (F) : 1.6 (N) (see Fig. 3), although the Korean tense aspirated bilabial stops showed a less tempo effect on GRA than the BrEng aspirated bilabial stops. This difference may be due to the differences in language, surrounding vowels, subject, etc. Both in BrEng and in Korean, on the other hand, for the tense aspirated bilabial stop the effect of tempo on VOT was insignificant, although the Korean tense aspirated bilabial intervocalic stops tended to have considerably shorter(13%) VOT in fast speech than in normal speech. Contrasted to this, both in BrEng and in Korean, the tense aspirated alveolar intervocalic stops were produced with a highly significantly less VOT in fast speech than in normal speech (see Figs 5 and 6). Overall, the tempo effect on VOT was 33% (BrEng) and 32% (Korean). For GRA, however, this was insignificant. Thus, both in BrEng and Korean the effects of tempo on VoT and GRA were conditioned by the place of articulation. In other words, for the tense aspirated bilabial stops the tempo effect on GRA was highly significant, whereas for their alveolar counterparts it was insignificant. For VOT, however, the reverse was observed to be true. All else being equal, this may be one of the main reasons for the fact that for the tense aspirated intervocalic stops, within a given articulatory place and across tempi, VOT was generally insignificantly correalted to GRA.

The possible underlying physiological mechanisms for the alveolar stop-related insignificant tempo effect on GRA. All else being equal, the BrEng alveolar stop-related insignificant tempo effect on GRA may be due mainly to a combination of (1) the physiological connections between the tongue and the larynx (Hardcastle, 1976; Fink, Basek, and Epanchin, 1956), (2) the decreased thickness of the vocal folds caused by the surrounding high pitch vowel /i:/ (Hollien, 1962; Van den Berg and Tan, 1959), and (3) the nature of the photo-electric glottograph which is designed to signal even the trans-illuminated light of the closed glottis (for detailed discussion of these, see Kim, 1987, section 5.11). However, considering the case associated with the Korean tense aspirated alveolar intervocalic stops where the surrounding vowels were /a/, the alveolar stop-related negligible tempo effect on GRA may be due mainly to the physiological connections between the tongue and the larynx. Considering the alveolar stop-related insignificant tempo effect on GRA and the fact that during the tense aspirated alveolar intervocalic stops there was a highly significant difference in VOT between the two tempos, one can presume that during the tense aspirated alveolar intervocalic stops the onset of the glottal adduction for the following vowels was actively controlled by the speaker to meet aerodynamic requirements in relation to tempo. In other words, both in BrEng and in Korean the speed of the onset of the glpttal adduction was conditioned by tempo: at the fast tempo it was greater than at the normal tempo.

Possible reasons for the BrEng tense aspirated bilabial stop-related negligible eggect of tempo on VOT. There are some phoneticians claiming that the amount of aspiration given to the English /p/ in the unstressed syllable will always tend to be less than that in the stressed syllable (e.g. Gimson, 1962; Jones, 1972). This agrees with Kim's (1987) findings. In a laryngographic study with eight BrEng speakers using isolated nonsense VCV words, Kim(1987) found that voiceless stops in the unstressed position generally gave a significantly shorter VOT than in the stressed position. Considering this, one may presume that in fast speech for the phonetic realization of the /p/ in the unstressed position, the English speaker may have three tasks to carry out simultaneously; (1) the fast speech-motivated reduction of VOT, and (3) the production of a significantly greater VOT for the tense aspirated stop than for its lax unaspirated cognate. However, considering the fact that there was practically no difference in VOT between the two tempi, it seems that at the fast tempo the task (1) was cancelled for the tasks (2) and (3). In this study, at the normal tempo the amount of VOT given to the BrEng unstressed /p/ was observed to be an average 32 ms (SD=0.580). This is much less than the findings associated with VOT for the word-initial stressed /p/ in the existing literature. In a spectrographic study (Weismer, 1979), for example, the data obtained from nine AmEng speakers using a set of CVC words embedded in the frame "Say _____ instead," where C was one of the three voiceless stops /p,t,k/, showed that in normal speech the amount of VOT given to the word-initial /p/ in the stressed position was an average 52.8 ms (SD = 14.59). Thus, although they are not directly comparable due to differences in dialect, speech item, speaker, the experimental devices used, etc., the value for VOT associated with the present study was much less (39%) than that for the stressed position-related /p/ in the existing literature. Considering this, it seems that at the normal tempo the unstressed position-related reduction of VOT (i.e. the second task)was manifested as usually claimed in the existing literature. If this is the case, it seems reasonable to speculate that in the fast speech the speaker may have to avoid the tempo-motivated reduction of VOT to save the distinction between /p/ and /b/. If this assumption is true, this may be one of the possible reasons for the fact that in the fast speech for the BrEng aspirated bilabial stops the difference in VOT between the two tempi was insignificant. At the level of physiology, however there was a fast tempo motivated significant reduction of GRA(see Fig. 4).

In the Korean tense aspirated bilabial stops, on the other hand, the fast tempo-motivated reduction of VOT was manifested significantly. Thus, the unstressed position-related reduction of VOT may be language specific. In other words, in the Korean phonology the unstressed position-related VOT reduction rule may not operate. This assumption would be supported by Kim's (1987) findings that for the Korean tense aspirated intervocalic stops in the isolated /VCV/ words the stress placement effect on VOT was generally insignificant, regardless of the place of articulation. It seems, however, universal that in order to save time the fast speech is expected to be manifested generally with a decrease of the phonetic variables of speech sounds,unless there are some inhibitive factors such as the unstressed position-related VOT reduction rule,the physiological connection between the tongue and the larynx,etc.

It also is worthwhile to note that the same amount of GRA given to the BrEng bilabial intervocalic stops was observed to yield a clearly different VOT according to the phonological information, such as class (i.e. tense and lax, traditionally voiced and voiceless). Fig. 4 shows that the shortest GRA (i.e. 3 mm) for the tense aspirated bilabial intervocalic stop yeilded either 24 ms or 38 ms of VOT. For its lax unaspirated cognates, however, the same size of GRA (i.e. 3 mm) gave either 8 ms or 20 ms of VOT. Thus, considering the results, it seems reasonable to presume that during the tense aspirated intervocalic stops the glottal adduction onset time for the following vowel was conditioned both by tempo and by class (i.e. tense/lax in English and aspirated/unaspirated in Korean.

This means that the onset of the glottal adduction for the following vowel is actively controlled to meet aerodynamic requirements in relation to class and tempo. In other words, the voiceless aspiration time (i.e. VOT) is not simply a function of the glottal area at the moment of the oral release of a stop as it is normally defined in the existing literature (cf. C-Kim, 1970; Hutter, 1984; Kagaya, 1974; Petursson 1976). This implies that VOT and GRA are autonomously controlled receiving motor commands from the high level neural-system in the brain in relation to tempo and class.

In general, other investigators did not take into consideration tempo and the place of articulation as affecting GRA and VOT, and they usually made overall observations of the relationship between GRA and VOT in normal speech only. Probably, this is the main reason why my results are different from others in the existing literature. My data also show that in normal speech, overall there is a significant relationship between GRA and VOT, particularly in Korean. Overall (i.e. across tempo, class and the place of articulation), for the Korean intervocalic stops GRA was significantly correlated to VOT: the greater the GRA, the longer the VOT; the less GRA, the shorter the VOT ($r = 0.808$, $df = 58$, $p < 0.001$).

Figs 1 and 2 show that for the BrEng intervocalic stops the overall relationship between GRA and VOT was considerably less than that for the Korean intervocalic stops. One of the possible reasons for this would be the alveolar stop-related greater GRA (see Figs 4 and 6). All else being equal, the alveolar stop-related greater GRA may be due mainly to a combination of (1) the decreased thickness of the vocal folds caused by the surrounding high pitch vowel /i:/ (Hollien, 1962; Van den Berg and Tan, 1959) and (2) the nature of the photo-eletric glottogragh which is designed to signal even trans-illuminated light of the closed glottis (for detailed discussion of these, see Kim, 1987, section 5.11).

This research was supported in part by grant from the Organized Research Fund of the Ministry of Education of the Korean Government, and forms part of the author's Ph. D. thesis (1987), at the University of Reading. The author thanks Dr. Hardcastle for much advice, assistance and critical comment.

References

Berg, J. Van den and Tan, T.S. (1956). "Results of experiments with human larynxes." Pract. otorhino-laryng 21, 425-450.

Butcher, A. (1977) "Coarticulation in intervocalic voiceless plosives and fricatives in connected speech." Arbeitsberrichte 8, p.154-213 (Kiel).

Fink, V.E., Basek, M. and Epanchin, V. (1956). "The mechnism of opening of the human larynx." Laryngoscope Vol. 66, p.410-425.

Fry, D.B. (1955). "Duration and intensity as physiological correlates of linguistic stress." Journal

of Acoustic Society of America 27, 765-768.
Gay, T., Ushijima, T., Hirose, H. and Copper, F.S.(1974). "Effect of speaking rate on labial consonant-vowel articulation." Journal of Phonetics 2, 47-63.

Gimson, A.C. (1962). An Introduction to the Pronunciation of English , Edward Arnold.

Hardcastle, W.J. (1976). Physiology of Speech Production, London; Academic Press.

Harris, M.S. and Umeda, N. (1974). "Effect of speaking mode on temporal factors in speech; vowel duration." Journal of Acoustic Society of America 56, 1016-1018.

Hollien, H. (1962). "Vocal fold thickness and fundamental frequency of phonation." Journal of Speech and Hearing Research Vol. 5, No. 3, 237-243.

Hutters, B. (1984). "Vocal fold adjustments in Danish voiceless obstruent production." Annual Report of the Institute of Phonetics 18. 163-184 (Copenhagen University).

Jones, D. (1972). An Outline of English Phonetics, 9th ed. Cambridge; Heffer.

Kagaya, R. (1974). "A fiberscopic and acoustic study of the Korean stops, affricates and fricatives." Journal of Phonetics 2, p. 161-180.

Kim, Chin-Wu (1970), "A Theory of Aspiration." Phonetica 21, pp.107-116.

Kim, D.-W. (1987). Some phonetic aspects of intervocalic oral stop consonants in British English and Korean, Ph. D. thesis, University of Reading (Available from the Pusan National University, Pusan 609 Korea, or the University of Reading, England).

Klatt, D.H. (1976). "Linguistic uses of segmental duration in English; Acoustic and perceptual evidence." Journal of Acoustic Society of America 59, 1208-1221.

Oller, D.K. (1973). "The effect of position in utterance on segment duration in English." Journal of Acoustic ociety of America 54, 1235-1247.

Petursson, M.(1976). "Aspiration et activite glottale, Examen experimental apartir de consonnes Islandaises." Phonetica 33, 169-198.

Weismer, B. (1979). "Sensitivity of voice onset time (VOT) measure to certain segmental features in speech." Journal of Phonetics 7, 197-204.