

A Fiberoptic Study of Laryngeal Gestures for Korean Intervocalic Stops

Dae-Won Kim, Hajime Hirose, and Seiji Niimi

Abstract

A fiberoptic investigation of laryngeal adjustments for Korean intervocalic stops was undertaken using /VCV/ nonsense words in carrier frames where the V was /i/. It was found that within a given tempo and across the place of articulation (i.e. /ph/ and /th/), neither the duration of oral closure nor the glottal width at release was systematically related to VOT, even in those cases where the peak glottal width precedes oral release, and that overall (i.e. across tempi and the place of articulation) the glottal width at release was somewhat positively related to VOT. It was inferred that, although a tense aspirated stop is differentiated from its unaspirated cognate by the size of glottal width at release and the degree of aspiration, the duration of aspiration is not simply a function of glottal width at release, as normally claimed in the existing literature, even in those instances where peak glottal width precedes oral release. It also was inferred that there seems to be a permissible range of variation in the relationship between VOT and glottal width at release, in relation to class, tempo, stress, etc. The effects of tempo on phonetic variables, the variation in interarticulator timing in tense aspirated stops, laryngeal adjustments for tense unaspirated stops, and the feature 'voice' in lax unaspirated stops were discussed.

1. Introduction

In a cineradiographic study of Korean word-initial stops in connected speech, Kim (1970) 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) and that "the instruction to close the glottis for all aspirated stops occurs at the time of oral release" (p.112). Other investigators' findings agreed with Kim's theory (e.g. for Korean isolated nonsense word-initial stops, see Kagaya, 1974; Hirose, Lee, and Ushijima, 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). Lofqvist (1980), however, claimed that the size of glottal width at release is functionally or positively related to degree of aspiration in voiceless stops, at least in those cases where peak glottal width precedes the release. This means that if peak glottal width follows the release, the size of glottal width at release would not be functionally or positively related to degree of aspiration in voiceless stops. This was supported by Dixit (1989). In a photo-electric glottographic study of Hindi bilabial stops at a normal speech rate, Dixit found that the degree of glottal width at oral release for the unvoiced aspirated plosives was much greater than for the voiced aspirated plosives, but the degree or duration of aspiration for the two types was about the same. For the voiced aspirated stops in Hindi, the peak glottal width occurred approxi-

mately 25 ms after release, and the size of the peak glottal width was much less than for the voiceless aspirated stops.

However, in a photo-electric glottographic study with intervocalic BrEng and Korean stops in connected speech, Kim (1991) found that VOT was not simply a function of the glottal area at release in an aspirated stop, even in those instances where the peak glottal area precedes the release. Kim also inferred that the speed of the glottal adduction following release for the following vowel was actively controlled to meet aerodynamic requirements in relation to the manner of articulation, stress and tempo.

In the present fiberoptic study, we wanted to duplicate Kim's experiment (1991) with the same Korean speaker (i.e. Kim, one of the authors) and the same speech items under the same conditions (eg. tempo). We wanted to know if Kim's claim holds under direct observations of the laryngeal adjustment in Korean stops. However, it was impossible to use the same speech items as in Kim (1991) due to the differences between a photo-electric glottograph and a fiberoptic device. The latter instrument could not provide a reasonable picture of the laryngeal gestures of a stop consonant because of the position of the tongue for the back vowel /a/. Thus, we substituted other speech items using the front high vowel /i/.

2. Method

2.1. Subject

A male Korean speaker served as the subject. He speaks a dialect close to that of Seoul, which is considered to be the standard dialect of Korea. The subject had no speaking defects.

2.2. Speech materials

Bisyllabic nonsense /VCV/ words, where the C was a stop and the V was the Korean front high vowel /i/, were constructed. These speech items were embedded in a carrier sentence in initial position, as shown in Table I.

2.3. Procedures

The speaker produced each sentence five times at a normal or moderately slow speech rate and five times at a fast speech rate, giving 90 utterances = 2 (tempi) x 5 (number of repetitions) x 3 (places of articulation) x 3 (manners of articulation).

Direct measurements were made from a fiberscope (PENTAX FNL-T013). The

output was synchronized with an audio microphone. For the printouts of the spectrograms, synchronized with the pulses of the video frames, a NEC computer was used. The fiberscope was connected to a U-matic video cassette recorder and the pictures of the laryngeal adjustment were taken on the U-matic video tape at a rate of 60 fields/s. The video images were transferred to a CRV disc (SONY, Laser videodisc Media LVM-3AAO) both for the slow motion display of the laryngeal gestures on a TV screen with fine controllers and for the development of desirable video-frames from a video printer (MITSUBISHI, SCT-CP 100).

Five tokens of each utterance (as indicated in Table I) were used to measure the time intervals between certain laryngeal and oral articulatory events indicated by vertical lines drawn through the spectrographic and audio signals (see Fig 1-5). The distance between the vocal processes of the arytenoid cartilages was measured frame by frame to obtain the time varying glottal width. The interval between frames was 16.6 ms.

Thus, the following measurements were made.

- (1) Time interval from implosion to peak glottal width
- (2) Time interval from peak glottal width to oral release
duration of oral closure phase
- (3) Voice onset time (VOT)
- (4) Glottal width (i.e. distance between vocal process) at
release
- (5) peak glottal width.

3. Results and discussion

3.1. The effects of tempo on the phonetic variables

As shown in Tables II and III, the fast speech was accompanied by a significant decrease in the phonetic variables, regardless of the place of articulation (for similar results for Swedish /t/ and /s/, see Lofqvist and McGarr, 1986; for a Danish /p/, see Andersen, 1981). This supports Kim's claim (1991) that fast speech is accompanied generally by a decrease of the phonetic variables of speech sounds, unless there exist some inhibitive factors such as the unstressed position-related VOT reduction rule in English (Kim, 1991).

3.2. Interarticulator timing in the tense aspirated stop

As seen in Figures 1 and 6, in the tense aspirated bilabial stop [ph], peak glottal width preceded oral release. This was the case in two out of five utterances in normal speech and three out of five cases in fast speech. In the rest of the utterances, the peak glottal width occurred at release. On the other hand, Figure 3 and 8 show that in the tense aspirated velar stop /kh/, peak glottal width followed the oral articulatory explosion. In normal speech, this was the case in three out of five utterances and in fast speech in three

out of four utterances (Figures 3 and 8). For the rest of the utterances, the peak glottal width for /kh/ occurred at oral release. On the other hand, in its alveolar cognate /th/ the peak glottal width occurred generally at oral release (see Figures 2 and 7).

Thus, some effects of the place of articulation on interarticulator timing in tense aspirated intervocalic stops could be seen. Kagaya (1974) mentions that in Korean aspirated stops the explosion occurs around the moment of maximum glottal aperture, but it appears from his figures that in aspirated bilabial stops, such as in /phe/ and /iphi/, the peak glottal width precedes the oral release (Figs 2 and 3), and that in their velar cognates, such as in /khe/, it follows the oral release (Fig. 2). As can be seen in Table II, in a tense aspirated stop the effects of the place of articulation on the duration of oral closure were greater than those on the interval from the implosion to peak glottal width, although for both parameters the further back the place of articulation, the shorter the duration (for similar results in English and Korean, see Kim, 1991). In normal speech, for example, the duration of oral closure for /kh/ was 34% shorter than that for /ph/, while the interval from the implosion to the peak glottal width for /kh/ was 18% shorter than for /ph/. In this case, thus, the different effects of place of articulation on each parameter may have had something to do with the variation in interarticulator timing, which may have resulted in a greater VOT for the tense aspirated velar stop /kh/ than for its bilabial and alveolar counterparts /ph/ and /th/.

However, considering Hutters's findings (1984) for Danish voiceless stops, the variation in interarticulator timing seems to be language specific. In Danish aspirated stops, the onset of explosion always results in the maximum aperture, regardless of the place of articulation, with a 10 ms lead-time on the average for /p/, for /k/ and /t/ the lead-times average 20 ms and 36 ms, respectively.

3.3. Laryngeal adjustments in tense unaspirated stops

As shown in Figure 4, for the Korean tense unaspirated bilabial stop [p'], the vocal processes began to open about 20 ms (i.e. at 16% of the duration of oral closure) after the onset of oral closure. The maximum glottal width occurred at 35% (i.e. 42 ms) of the duration of the segment after the onset of oral closure, and it continued up to the articulatory explosion, which in turn resulted in about 10 ms of VOT. For the subject used, this was generally consistent for the tense unaspirated stops, regardless of the place of articulation (for a similar result, see Kim, 1987). However, this contrasts with existing claims (cf. Kagaya, 1974; Hirose, Lee, and Ushijima, 1974). As quoted by Hong, Niimi and Hirose (1991), for example, Kagaya (1974) claimed that "[for a Korean tense unaspirated stop in the /CV/ nonsense word where the V was /i/ or /e/] complete contact of the vocal processes is achieved at a point from 80 to 100 ms before the articulatory explosion." (p.166) and it continued up to the onset of oral release. This was similar to a case with an intervocalic tense unaspirated stop in a /VCV/ nonsense word where the V was /i/ or /e/

(Kagaya, 1974, Fig 3). Thus, there are contrastive findings regarding the laryngeal adjustment for Korean tense unaspirated stops. This may be due to either interpersonal differences or the limitation of the devices used. However, considering Kagaya's report that "the [oral] articulatory explosion is found in just the same or one frame before voice onset" (p.166), the contrastive findings may be due mainly to the limitation of the devices used. One frame voice onset delay is about 20 ms in Kagaya (1974). This means that the VOT for the tense unaspirated stop ranged from 0 ms to 20 ms, giving an "average 15 ms" (Kagaya, 1974, p. 168) (for similar results for a Korean tense unaspirated stop associated with VOT, see Kim, 1965; Lisker and Abramson, 1964; Han & Weitzman, 1967; Hardcastle, 1973). This may imply that in a tense unaspirated stop the vocal processes are not completely closed immediately before the articulatory explosion and at the moment of oral release (cf. Kagaya, 1974; Hirose, Lee, and Ushijima, 1974), but are opened to some extent, which in turn may result in a voice onset delay for about 10 ms. The increased activity of the lateral cricoarytenoid muscle (i.e. one of the adductor muscles, Hardcastle, 1976) may have something to do in part with closing the opened glottis. In an EMG study (Hirose, Lee, and Ushijima, 1974), the lateral cricoarytenoid muscle showed a greater increase in activity immediately before the stop release in a Korean tense unaspirated stop, as compared with its lax counterpart. The increased activity of the vocalis muscle (i.e. one of the tensor muscles, Hirose and Gay, 1972; Hardcastle, 1976) and the lateral cricoarytenoid muscle would presumably result in an increase of the inner tension of the vocal fold, in the closure of the glottis, and in its constriction during or immediately after the release of oral closure. As a result, the tense unaspirated stop seems to have a higher voice pitch (e.g. Kagaya, 1974; Han & Weitzman, 1967) compared to its lax counterpart.

3.4. The feature 'voice' in lax unaspirated stops

As shown in Figure 5, on the other hand, in the Korean unaspirated bilabial stop [p] the glottal width gesture started at 57% of the duration of oral closure (i.e. 40 ms) after the onset of oral closure, and a clear glottal width -- which was slightly less than that for a tense unaspirated stop -- occurred about 10 ms before the onset of oral closure. It lasted about 10 ms after the oral release, which gave about 7 ms of VOT. For the subject used, this was a general finding for the lax unaspirated stops, except for the lax unaspirated velar stop /k/. For the lax velar stop, the vocal fold vibration occurred throughout the oral closure phase both in normal speech and in fast speech. There may be interpersonal and occasional differences in the laryngeal gestures for intervocalic lax unaspirated stops. Some speakers may have a vocal fold vibration throughout the oral closure phase, and others may have a certain devoiced period during the oral closure phase, as can be seen in Figure 5. In Kagaya (1974), one of two Korean speakers produced fully devoiced lax unaspirated stops in intervocalic position, but these were excluded from his discussion, probably because he thought the fully devoiced data were abnormal. This variation in voicing implies that the feature voice is linguistically insignificant for a Korean lax

unaspirated stop. For the English intervocalic stops /b/ and /d/, a similar result has been reported (Kim, 1989).

3.5. Relationship between VOT and glottal width at release

As seen in Figures 9 and 10, within a given tempo and across place of articulation (i.e. /ph/ and /th/), the correlation of VOT to the glottal width at release was insignificant: for normal speech, $r = 0.127$, $n = 10$, $p < 0.727$; for fast speech, $r = 0.069$, $n = 10$, $p < 0.851$. Thus, in this case, VOT was independently related to the glottal width at release, even in the cases where the peak glottal width precedes oral release. This would support Kim's claim (1991) that in a tense aspirated stop, VOT is not simply a function of glottal width at release, even in the instances where the peak glottal width precedes oral release although these data cannot be directly compared with Kim's data (1991), due mainly to differences in the speech items and experimental devices used. Figure 9 also shows that in normal speech, for the tense aspirated bilabial stop /ph/, VOT was significantly ($p < .008$) longer (mean 69 ms) than for its alveolar counterpart /th/ (mean 56 ms), whereas the glottal widths at release were more or less similar to each other (see Table II). This means that all else being equal in normal speech the speed of glottal adduction following the release for a following vowel was slower (about 19%) for the tense aspirated bilabial stop /ph/ than for its alveolar cognate /th/, which in turn resulted in a greater VOT for the tense aspirated bilabial stop /ph/ than for its alveolar counterpart /th/. On the basis of the findings, it was inferred that although the tense aspirated stop is differentiated from its unaspirated cognate by the size of the glottal width at release and the degree of aspiration, the size of the glottal width at release does not necessarily or automatically determine the duration or degree of aspiration for a tense aspirated stop, as is normally claimed in the literature (cf. Lofqvist, 1980; Dixit, 1989; Kagaya, 1974; Hirose, Lee, and Ushijima, 1974), even in those instances where the peak glottal width precedes oral release.

It also was recognized that across the place of articulation (/p/ and /t/), for example, in normal speech the permissible range of variation in VOT (62 ms + 9 ms) was much greater than in the size of the glottal width at release (51.6 + 4.6). On the other hand, in fast speech the variation in VOT and the glottal width at release ranged 46 ms + 4.9 ms and 40.7 + 4.4, respectively. In fast speech, the range of variation of VOT was decreased nearly by half, as compared to normal speech. However, the variation in the glottal width at release was nearly consistent, regardless of tempo, although there was a significant effect of tempo on the width of the glottis. On the basis of this finding, one may presume that each phonetic parameter may have its own permissible range of variation in relation to manner of articulation, stress, tempo, etc. and that the precision in hitting the VOT target point would be higher in fast speech than in normal speech, at least in those cases where the peak glottal width precedes the release or occurs at release (see Table II). If a speaker violates this rule associated with the precision in hitting the VOT, he or she may be misunderstood in communication or considered to be strange.

3.6. Overall relationship between VOT and glottal width at release

Within a given tempo and within a given place of articulation, either the size of the glottal width at release or the duration of the oral closure was found not to systematically relate to VOT (cf. Lofqvist, 1980; Andersen, 1981), even in those cases where the peak glottal width precedes the oral release. On the other hand, Figures 11 and 12 show that across tempi and within a given place of articulation, there was a positive relationship between VOT and glottal width at release: for /ph/, $r = 0.688$, $n = 10$, $p < 0.028$, and for /th/, $r = 0.575$, $n = 10$, $p < 0.082$. For the tense unaspirated bilabial stop /ph/, this finding was generally agreeable with Kim (1991, for an intervocalic /ph/, $r = 0.646$, $n = 10$, and $p < 0.05$). As seen in Table III, the effects of tempo on VOT and glottal width at release were significant, regardless of the place of articulation. Fast speech brought about a more or less similar amount of decrease both in VOT and in glottal width at oral release (see Table II). In this case, thus, the effects of tempo may be one of the main reasons for a positive relationship between VOT and glottal width at release.

For a tense aspirated alveolar stop, however, the effects of tempo on the glottal width at release was disagreeable with Kim (1991). In Kim (1991, Fig. 5), the effects of tempo on the glottal width at release was insignificant, whereas the effects of tempo on VOT was highly significant. The disagreement may be due to the differences between the two investigations in the experimental devices used, speech items, a combination of these, or due to occasional differences. Further investigation of these differences is called for.

Overall (across tempo and the place of articulation), however, there was a positive correlation of VOT to glottal width at release ($r = .665$, $n = 20$, $p < .001$). This agrees with the claims (Kim, 1970; Kagaya, 1974; Hirose, Lee and Ushijima, 1974; Lofqvist, 1980; Dixit, 1989; Kim 1991). However, considering the fact that within a given tempo and across the place of articulation VOT was independent of the glottal width at release, even in those cases where the peak glottal width preceded oral release, it is difficult to say that the duration of aspiration for a tense aspirated stop is simply a function of glottal width at release, even in those instances where the peak glottal width precedes oral release. As pointed out by Kim (1991), most investigators have concentrated on overall observations of the relationship between VOT and glottal width at release, particularly in normal speech. This may be one of the main reasons for the existing differences among reported findings.

Acknowledgement

This research was supported by the Japan-Korea Cultural Foundation. The authors would like to thank Hiroshi Imagawa, Atsushi Matsunaga (M.D.) and Masanobu Kumada (M.D.) for technical help.

References

- Andersen, P. (1981). The effect of increased speaking rate and (intended) loudness on the glottal behavior in stop consonant production as exemplified by Danish p, Ann. Rep. Inst. Phonetics. Univ. Copenhagen 15, 103-146.
- Butcher, A. (1977). Coarticulation in intervocalic voiceless plosives and fricatives in connected speech. Arbeitsberichte 8, 154-213 (Kiel).
- Han, M. S. & Weitzman, R. S. (1970). Acoustic features of /P, T, K/, /p, t, k/ and /ph, th, kh/, *Phonetica* 22, 112-128.
- Hardcastle, W. J. (1976). *Physiology of Speech Production: An Introduction for Speech Scientists*. London:Academic Press.
- Hardcastle, W. J. (1973). Some observations on the tense-lax distinction in initial stops in Korean, *Journal of Phonetics* 1, 263-272.
- Hirose, H. and T. Gay (1972). The activity of intrinsic laryngeal muscles in voicing control, *Phonetica* 25, 140-164.
- Hirose, H., C. Y. Lee, T. Ushijima (1974). Laryngeal control in Korean stop production, *Journal of Phonetics* 2, 145-152.
- Hong, K., S. Niimi, and H. Hirose (1991). Laryngeal adjustments for Korean stops, affricates and fricatives - an electromyographic study, *Ann. Bull, RILP* 25, faculty of medicine, Tokyo University, 17-31.
- Hutters, B. (1984). Vocal fold adjustment in Danish voiceless obstruent production, *Ann. Rep. Inst. Phonetics. Univ. Copenhagen* 18, 293-385.
- Kagaya, R. (1974). A fiberoptic and acoustic study of the Korean stops, affricates and fricatives. *Journal of Phonetics* 2, 161-180.
- Kim, C. W. (1970). A theory of aspiration, *Phonetica* 21, 107-116.
- Kim, C. W. (1965). On the autonomy of the tensivity feature in stop classification (with special reference to Korean stops), *Word* 21, 339-359.
- Kim, D. W. (1987). *Some Phonetic Aspects of Intervocalic Stop Consonants in British English and Korean*, Ph.D.thesis, University of Reading, England, U.K.
- Kim, D. W. (1989). The relationship between full voicing and the oral closure interval of a lax intervocalic stop, *Journal of Phonetics* 17, 239-243.

- Kim, D. W.(1991). The relationship between glottal area at release and voice onset time in stop production, *Linguistic research* 10, proceedings of the Kyung Hee international conference on linguistic studies, 85-103.
- Lisker, L. and Abramson, A. S. (1964). A cross-language study of voicing in initial stops: acoustic measurements. *Word* 20, 384-422.
- Lofqvist, A. (1980). Interarticulator programming in stop production, *Journal of Phonetics* 8, 475-490.
- Lofqvist, A. and N. S. McGarr (1986). Laryngeal dynamics in voiceless consonant production, Working Papers in Logopedics and Phoniatics 3, p. 49-65, DLP, University Hospital Lund, Sweden.
- Pertursson, M. (1976). Aspiration et activite' glottale. Examen Experimental a' partir de consonnes Islandaise. *Phonetica* 33, 169-198.

Table I. Korean speech items

-
- (a) ipi-ka ippi-ippi han-ta.
 - (b) ippi-ka iphi-iphi han-ta.
 - (c) iphi-ka ipi-ipi han-ta.
 - (d) iti-ka itti-itti han-ta.
 - (e) itti-ka ithi-ithi han-ta.
 - (f) ithi-ka iti-iti han-ta.
 - (g) iki-ka ikki-ikki han-ta.
 - (h) ikki-ka ikhi-ikhi han-ta.
 - (i) ikhi-ka iki-iki han-ta.
-

(where {p,t,k}, {pp,tt,kk} and {ph,th,kh} are lax unaspirated stops, tense unaspirated stops and tense aspirated stops, respectively).

Table II. Mean closure duration, aspiration, glottal width at release, peak glottal width in parentheses, and the intervals from implosion to peak glottal width and from peak glottal width to the oral release for the Korean aspirated tense stops (n = 5).

tempo and stop consonants		closure duration	VOT width	glottal to peak at release	Implosion glottal width	peak width to release
N /ph/	X	102.20	68.80	53.20*(54.0*)	96.88	5.32
	S	14.37	6.08	4.82 (5.66)	11.04	7.65
F /ph/	X	71.80	47.00	41.20 (41.60)	68.40	3.40
	S	4.08	3.49	5.40 (5.60)	2.70	3.21
N /th/	X	94.60	55.60	50.00 (50.00)	94.60	0.00
	S	5.77	7.73	4.24 (4.24)	5.77	0.00
F /th/	X	63.20	45.40	40.20 (40.20)	63.00	0.00
	S	5.54	3.91	3.63 (3.63)	5.54	0.00
N /kh/	X	69.60	77.40	52.20 (54.40)	79.56	-9.96
	S	10.04	9.26	3.49 (2.61)	10.51	9.09
F /kh/	X	42.00	56.80	26.00 (33.00)	56.53	-14.53
	S	7.83	10.89	9.38 (6.63)	12.75	10.44

(ms, N = normal speech, F = fast speech, *an arbitrary unit)

Table III. The effects of tempo on the phonetic variables

stops	duration of segment	VOT	glottal width at release	implosion to peak glottal width	peak glottal width
/ph/	p < .002	p < .000	p < .006	p < .001	p < .008
/th/	p < .000	p < .030	p < .004	p < .000	p < .004
/kh/	p < .002	p < .012	p < .001	p < .021	p < .000

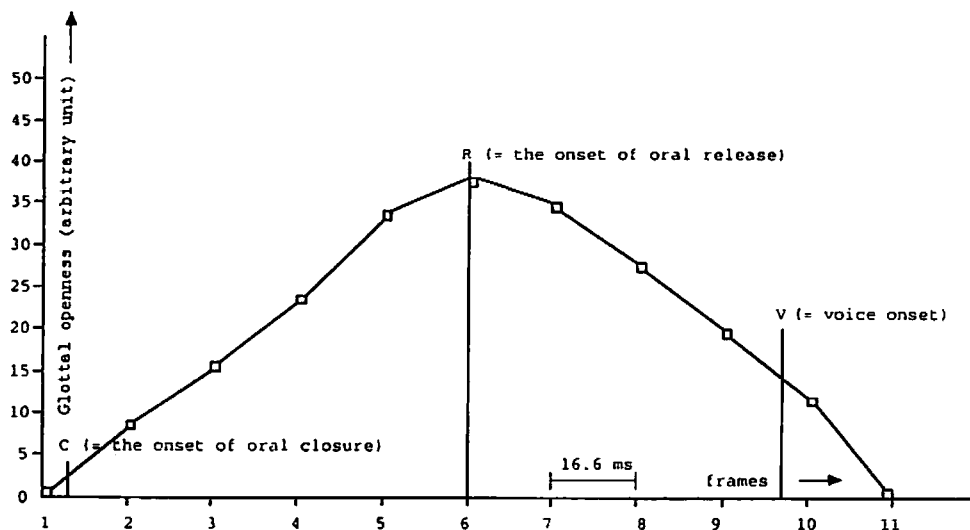


Figure 7. Laryngeal adjustment for a Korean intervocalic tense aspirated stop [th] in fast speech.

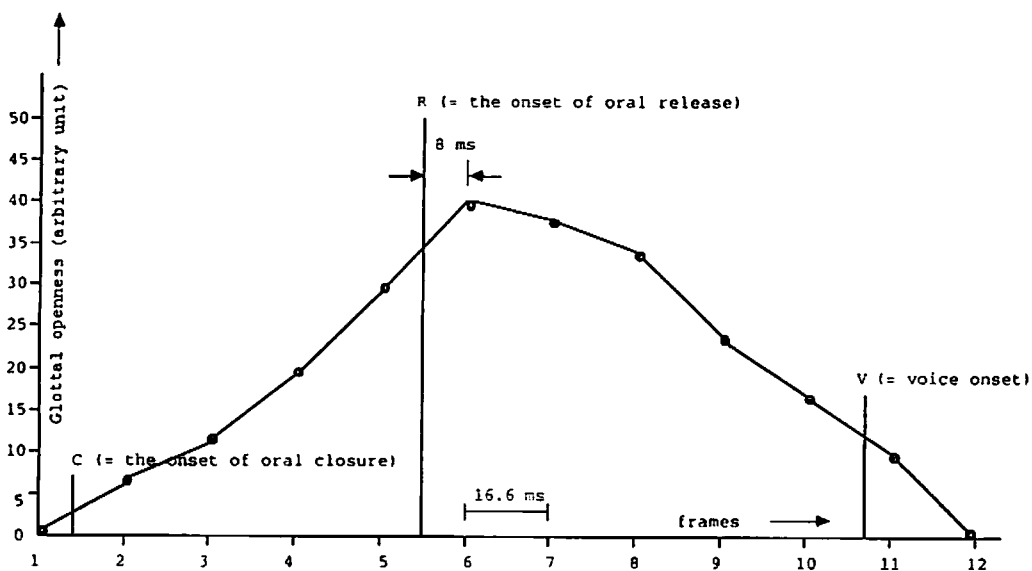


Figure 8. Laryngeal adjustment for a Korean intervocalic tense aspirated stop [kh] in fast speech.

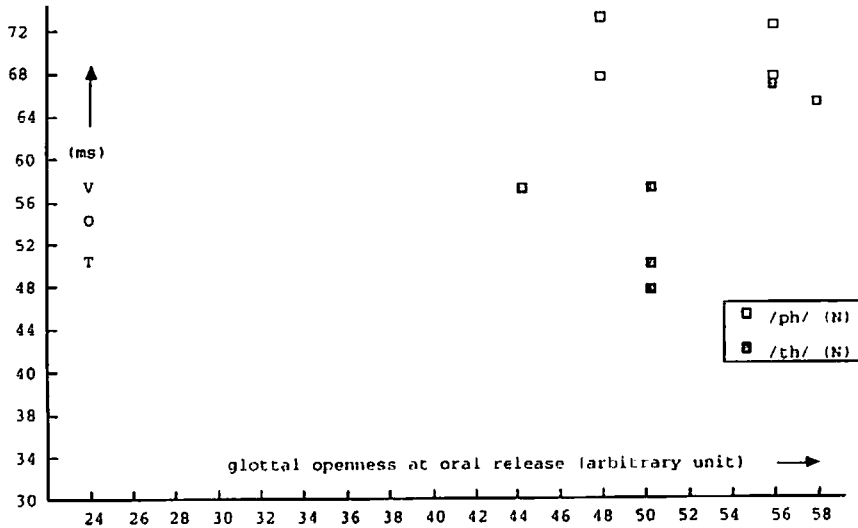


Figure 9. Scatter diagram of VOT against glottal openness at oral release during the Korean aspirated tense stops /ph/ and /th/ in nonsense /VCV/ words embedded in a carrier sentence (N = normal speech).

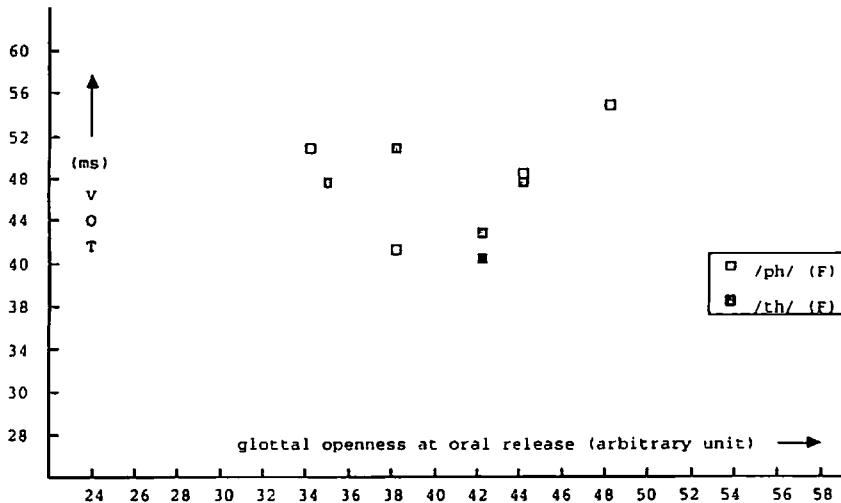


Figure 10. Scatter diagram of VOT against glottal openness at oral release during the Korean aspirated tense stops /ph/ and /th/ in nonsense /VCV/ words embedded in a carrier sentence (F = fast speech, = overlapped).

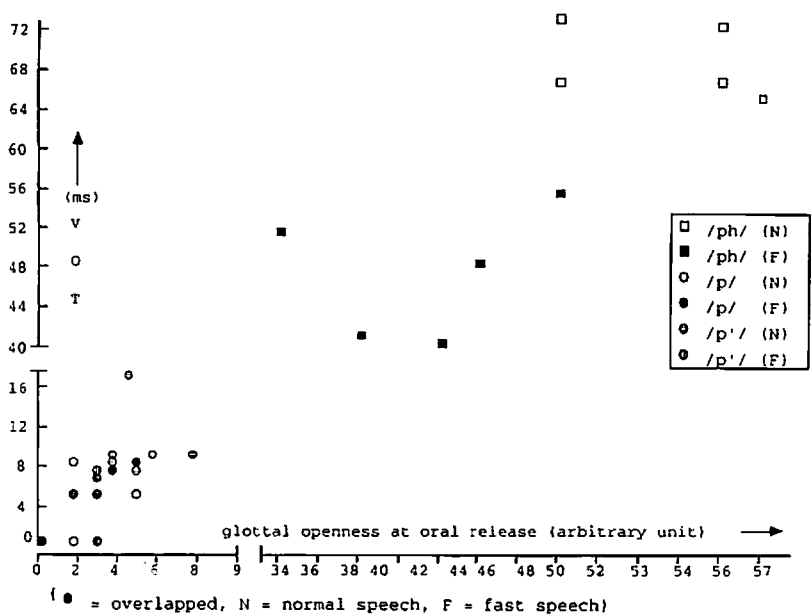


Figure 11. Scatter diagram of VOT against glottal openness at oral release during the Korean stops /p, p', ph/ in nonsense /VCV/ words embedded in a carrier sentence, where /p/ indicates an unaspirated lax stop, /p'/ an unaspirated tense stop, and /ph/ an aspirated tense stop.

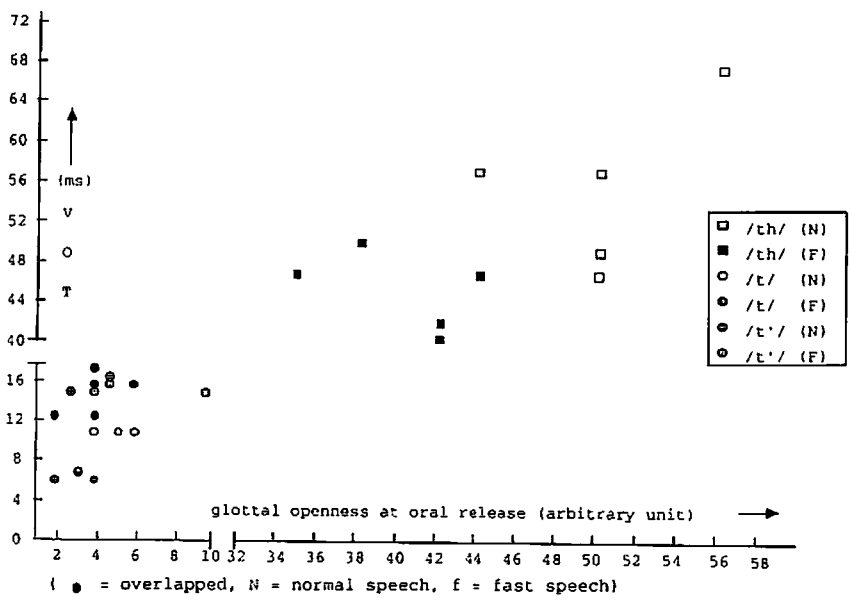


Figure 12. Scatter diagram of VOT against glottal openness at oral release during the Korean stops /t, t', th/ in nonsense /VCV/ words embedded in a carrier sentence, where /t/ indicates an unaspirated lax stop, /t'/ an aspirated tense stop, and /th/ an aspirated tense stop.