Ann. Bull. RILP (1982) No. 16, 27-35

LARYNGEAL ADJUSTMENTS IN DUTCH VOICELESS OBSTRUENT PRODUCTION

Hirohide Yoshioka, Anders Löfqvist, and René Collier***

Introduction

We have recently reported how voiceless sound sequences, such as voiceless obstruent clusters, are organized in terms of their glottal opening and closing gestures, using native speakers of Swedish (Löfqvist and Yoshioka, 1980), American English (Yoshioka, Löfqvist and Hirose, 1981), Icelandic (Löfqvist and Yoshioka, 1981), and Japanese (Yoshioka, Löfqvist and Hirose 1982). The conclusion of these studies was that, in the production of sequential unvoiced sounds, the glottal opening gesture is characterized by a one, two or more-than-two peaked pattern in a regular fashion according to the nature of the voiceless segments. A voiceless obstruent specified with aspiration or frication noise tends to require a separate opening gesture, while an unaspirated stop in a voiceless environment can be produced within the opening gesture attributed to an adjacent aspirated stop or fricative. For example, the /sk#sk/ sequence in English was produced in most cases with two separate opening gestures. In contrast, an /sks#k/ string was in general accompanied by three opening gestures (Yoshioka, Löfqvist and Hirose, 1981).

Furthermore, the velocity of the initial opening movement was shown to vary depending on the properties of the initial voiceless segment. When the first unvoiced segment in the cluster was a fricative, the speed of the abduction was significantly faster than when the initial voiceless sound was an aspirated or unaspirated stop, regardless of the nature of the following voiceless segments. This also means that the difference in velocity during the initial opening phase held true despite the fact that, for most cases beginning with a voiceless unaspirated stop, the peak glottal opening occurred during the following fricative segment.

In order to examine the validity of these notions across different languages, the current experiment was carried out in cooperation with a native speaker of Dutch, using the same combined techniques of photo-electric glottography, fiberoptic filming and laryngeal electromyography. In addition, we paid particular attention to voiceless obstruent clusters containing two consecutive fricatives, since the other languages we published research on so far do not contain enough such combinations.

Method and Procedure

The techniques used in the present experiment were simultaneous recordings of photo-electric glottography, fiberoptic filming and laryngeal electromyography (EMG), in parallel with audio signal.

---

* Lund University, Sweden

** University of Antwerp, Belgium

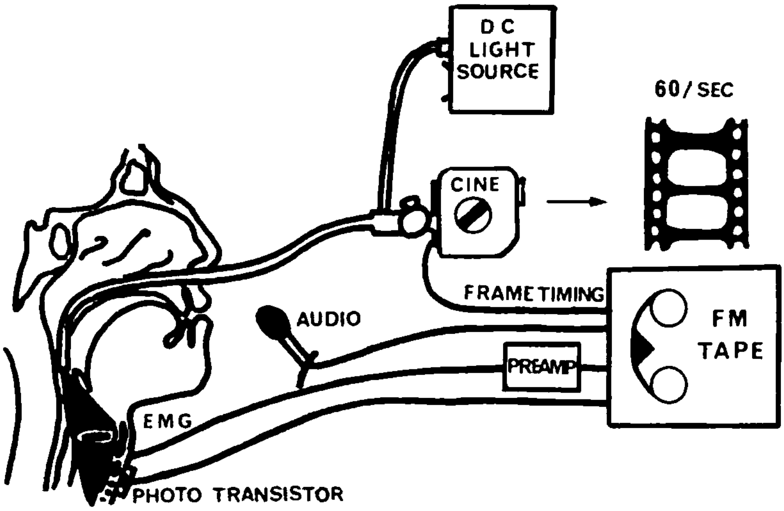


Fig. 1 *Diagram of experimental procedure.*

The EMG data were obtained using bipolar hooked-wire electrodes (Basmajian and Stecko, 1962; Hirano and Ohala, 1969). The electrodes, consisting of a pair of platinum-tungsten alloy wires (50 microns in diameter with isonel coating), were inserted perorally into the posterior cricoarytenoid muscle (PCA) under indirect laryngoscopy with the aid of a specially designed curved probe (Hirose, Gay, Strome and Sawashima, 1971). Before insertion, topical anaesthetic was applied to the mucous menbrane of the hypopharynx using a small amount of 4% Lidocaine spray (Xylocaine). For verification of the electrode position, the subject was instructed to perform several non-speech and speech maneuvers that are well understood in terms of the manner of the PCA involvement, including inspiration and expiration during breathing, voluntary swallowing, pitch changes including register shifts, glottal attack, and voiced-voiceless sound contrasts. The EMG signal was monitored on an oscilloscope during the verification gestures and also during the entire recording session.

The interference voltages of the EMG signals, after high-pass filtering at 80 Hz, were recorded on a multi-channel FM recorder together with the audio signal. After full-wave rectification and integration over a 5 msec time window, the action potentials were fed into a computer at a sampling rate of 200 Hz for further processing to obtain the muscle activity patterns for the ensemble-averaged tokens with a 35 msec time constant (Kewley-Port, 1977). The figures to be presented in this paper represent activity patterns aligned with reference to the voicing offset of the vowel preceding the voiceless sequence.

For the movement data, the glottal view through a flexible laryngeal fiberscope (Olympus VF-O type, 4.5 mm in outer diameter) was photographed with a cine camera at a rate of 60 frames/sec. A synchronization signal was registered on the FM recorder to identify each frame. Then, frame by frame analyses were made with the aid of a mini-computer to calculate the distance between the vocal processes, which is considered one of the indicators of glottal width (Sawashima and Hirose, 1968; Sawashima, 1976).

A cold DC light source (Olympus CLS), providing illumination of the upper glottal area, also served as the light source for the photo-electric glottography. The amount of light passing through the glottis was sensed by a photo-transistor (Philips BPX 81) placed on the neck just below the lower edge of the cricoid cartilage. The electrical output was recorded on another channel of the FM tape. These signals were sampled at 200 Hz and processed with a digital system.

The linguistic material consisted of Dutch voiceless obstruents and obstruent clusters in various positions, with a word boundary preceding, following, or intervening within the cluster. Table 1 shows sentences containing various combinations of voiceless fricatives and voiceless stops. The words in Set I (most of them "verbs") were combined with those in Set II ("nouns"). The abbreviated phonetic transcriptions, except V (correspondings to a voiced vowel), indicate the type of clusters containing the various voiceless sound sequences with which the experiment is concerned.

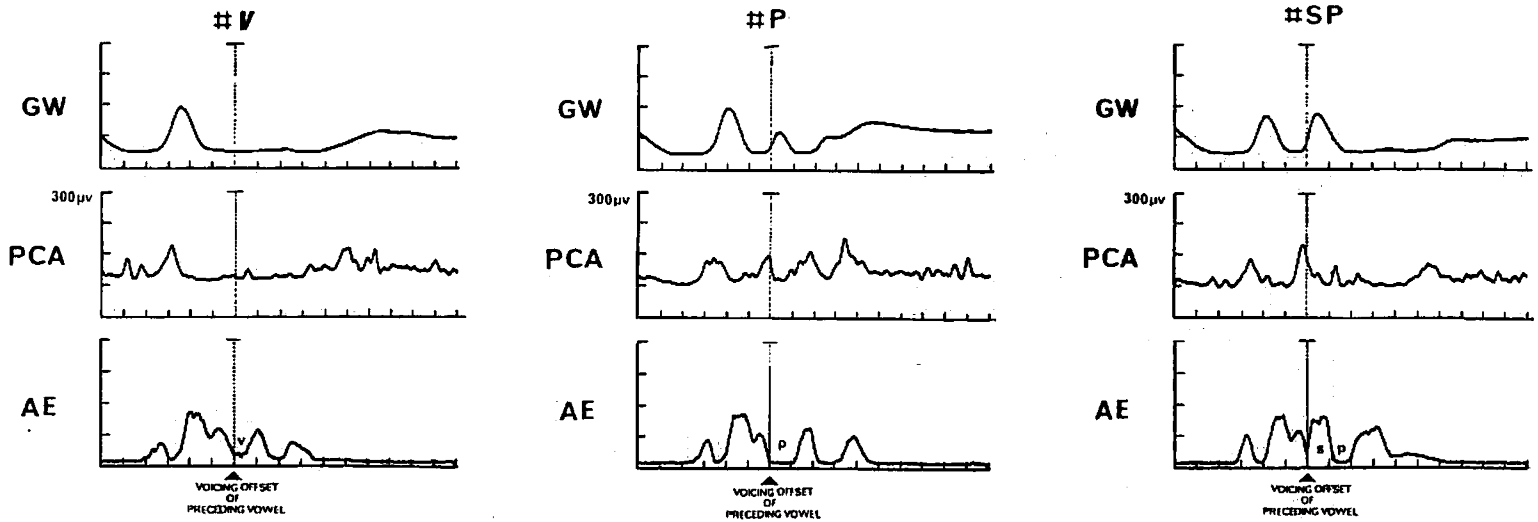
Table 1 *Linguistic material. All the words in Set I following "Ik" were combined with those in Set II. The abbreviated phonemic transcriptions indicate the voiceless sounds with which the experiment is concerned.*

	Set I		*Set II*	
Ik	zie	/-V#/	Iman	/#V-/
	riep	/-p#/	Piet	/#p-/
	vries	/-s#/	Sieb	/#s-/
			Chiro	/#x-/
			Spiro	/#sp-/
			Psycho	/#ps-/
			Schipper	/#sx-/

A native male speaker of Dutch (actually the third author), read all of the items more than 12 times each, in random order. Simultaneous recordings of the photo-electric output and fiberoptic filming together with the EMG signals were made during the first two repetitions of each utterance type, followed by more than 10 additional recordings of only the photo-electric and EMG signals. Therefore, more than 12 repetitions of each utterance type were needed for averaging. During the latter part of the session, the glottal image was constantly monitored through a fiberoptic view finder. Such careful monitoring is mandatory for reliable interpretations of large amounts of photo-electric recordings, as we have discussed elsewhere (Löfqvist and Yoshioka, 1980; Yoshioka, Löfqvist and Hirose, 1981). Although no particular instructions were given to the subject about vocal intensity or speaking rate, a gross survey of his audio waveforms and acoustic envelopes revealed that the intra-session variability for each utterance type was comparably small.

Results

Fig. 2 illustrates the results for the various types of word-initial voiceless obstruent sequences following a word-final voiced vowel. Also included is the pattern



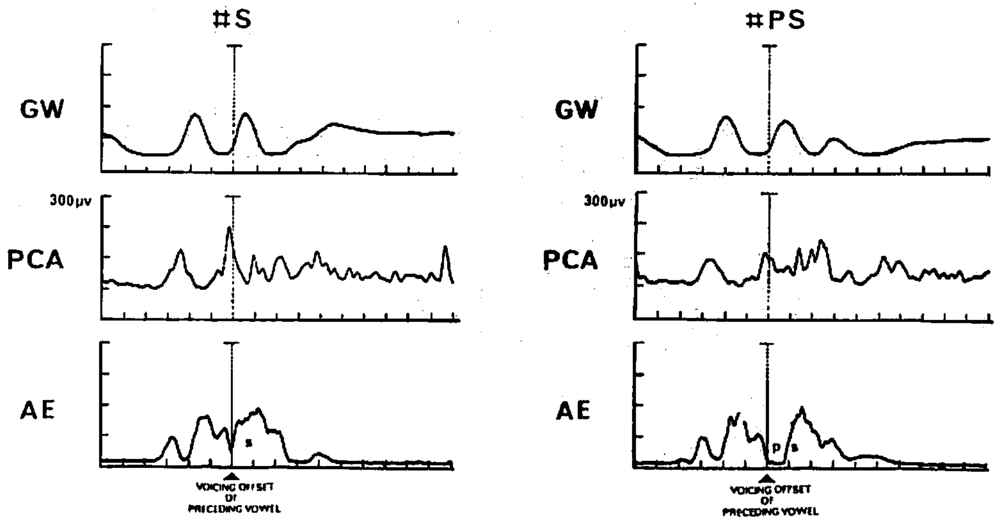


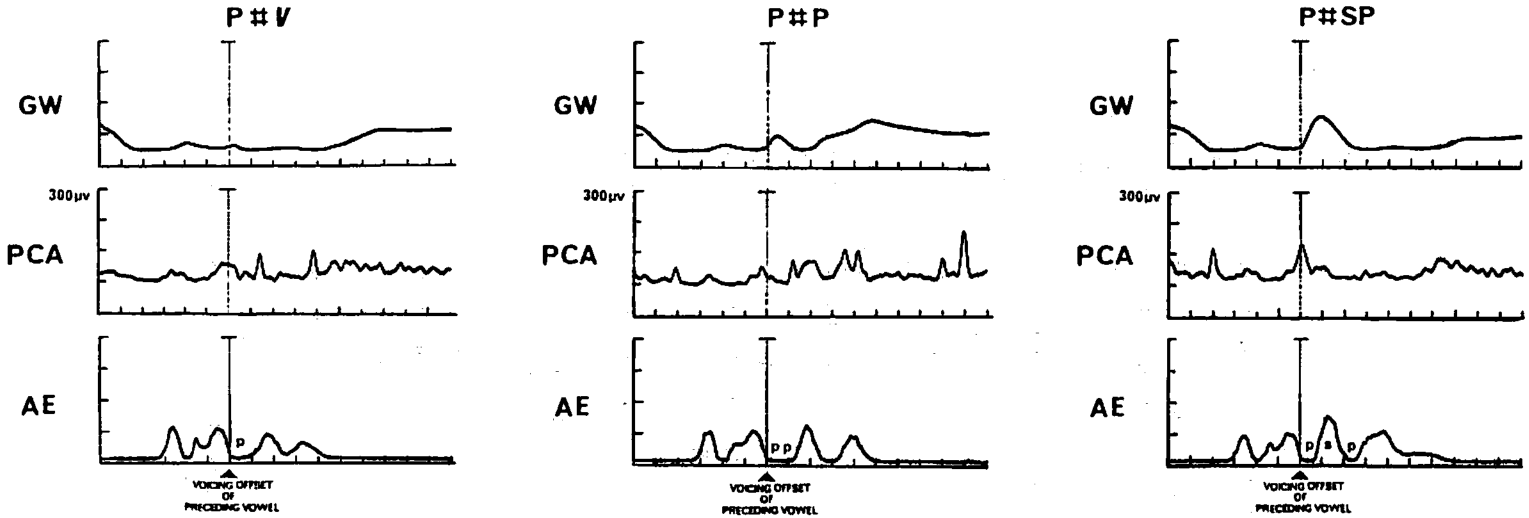
Fig. 2 *Averaged glottograms (GW), averaged electromyograms (PCA), and averaged audio envelopes (AE) for the utterance types containing various word-initial voiceless sequencies preceded by a word-final voiced vowel. (V = vowel).*

for the words beginning without any voiceless sound. Since glottal opening patterns obtained by photo-electric glottography have been shown to be practically identical to those obtained by plotting the distance between the vocal processes from fiber-optic cine-films, we will focus on our photo-electric glottograms as an index of the glottal width change during the pertinent voiceless sequence production. In Fig. 2, the top row (GW) represents the averaged glottograms. As for the laryngeal EMG, the figure contains the corresponding averaged activity patterns of the abductor muscle – the posterior cricoarytenoid muscle (PCA) – that has been demonstrated to substantially control glottal aperture (Hirose, 1976; Yoshioka, 1981). These signals were aligned with respect to the voicing offset of the preceding word-final vowel which is shown by the vertical dotted lines in the audio envelope curves at the bottom.

It is obvious here that a single glottal opening gesture always occurs during voiceless sound sequences, regardless of the type and number of the voiceless obstruents with which the experiment is concerned.[1] The size and the peak timing of the opening, however, differ among these combinations. For the production of the word-initial stop /p/, the glottal movement is rather slow and small in comparison with that for the word-initial fricative /s/. As for voiceless sequences composed of two different obstruents, the data for word-initial /sp/ and /ps/ show that the glottal opening gesture is characterized by a mono-modal pattern even for these two-phone combinations.[2] The size of the opening for these is comparable to that for the single fricative /s/. The timing of the peak for these two-phone combinations is different. The peak timing for /sp/, which is attained during the frication noise for /s/, is rather earlier than that for /ps/, in which the peak timing approximately coincides with the timing of the oral release for the stop /p/. Consequently, the velocity of the glottal abduction appears to be a little faster for the word-initial /sp/ cluster than for the word-initial /ps/ sequence.

Fig. 3 compares the production of various word-initial voiceless obstruent combinations following a word-final voiceless stop /p/. Again, the activity pattern for an utterance with a word-initial voiced vowel following a word-final /p/ is added as a reference. Here, several points are worth mentioning. First, the word-final stop /p/ is generally produced with a negligibly small opening gesture of the glottis, particularly when it is followed by a word-initial vowel. This fact appears to be related to the glottalization of a stop sound in this position. As for the production of the geminate combination /p#p/, the audio envelope curve reveals that it was uttered with a slightly longer duration of the closure period than that for a single word-initial aspirated stop /p/ (shown in Fig. 2). Nevertheless, the curve of the abduction activity pattern, as well as that of the glottal opening movement for the geminate, was characterized by one single peak similar to, or even smaller than, that for the word-initial single stop /p/, although a word boundary intervenes within the geminate and the duration of the open phase is longer.

1. The opening before the line-up corresponds to the voiceless cluster /k#s/ in the carrier sentence "Ik zie ____."
2. The extra opening found in the GW curve well after the line-up for the utterance containing a /#ps/ sequence corresponds to the voiceless fricative [X] in the word "Psycho."



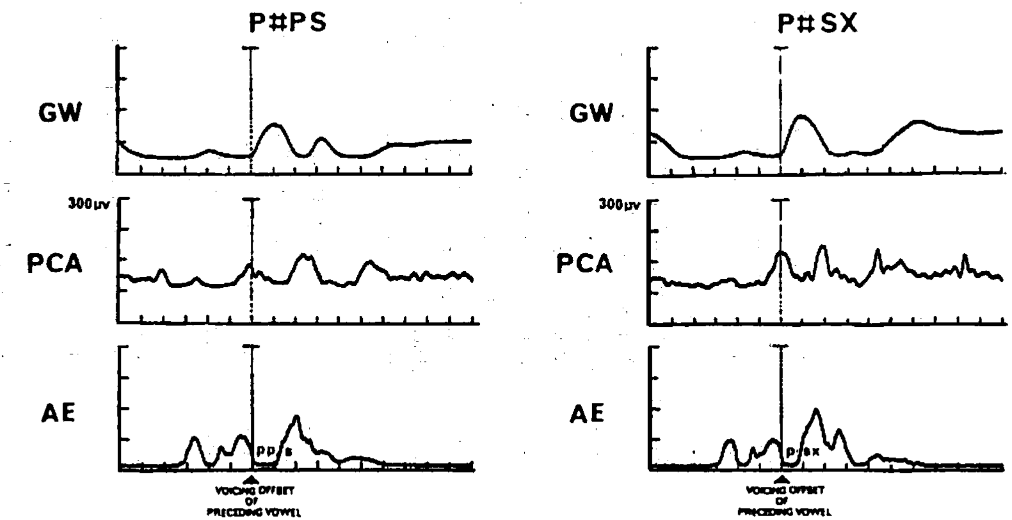
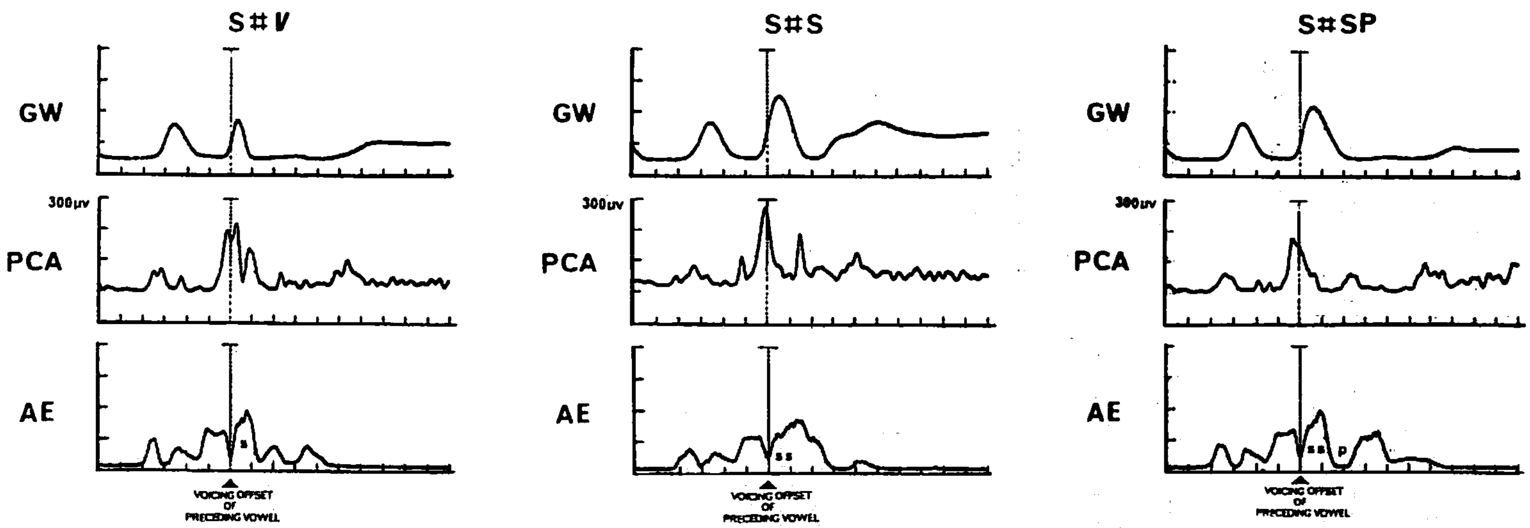


Fig. 3 *Averaged glottograms (GW), averaged electromyograms (PCA), and averaged audio envelopes (AE) for the utterance types containing various word-initial voiceless sequences preceded by a word-final voiceless stop /p/. (V = vowel).*



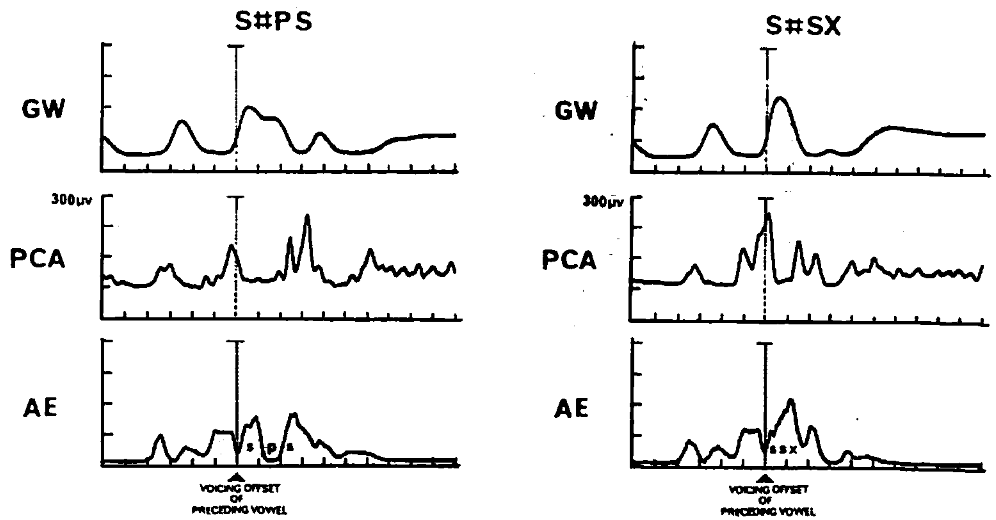


Fig. 4 *Averaged glottograms (GW), averaged electromyograms (PCA), and averaged audio envelopes (AE) for the utterance types containing various word-initial voiceless sequences preceded by a word-final voiceless fricative /s/. (V = vowel).*

Fig. 3 includes various types of three-phone combination beginning with a word-final /p/. The averaged abductor activity curves as well as the glottographic patterns among these three types are all characterized by a single peaked opening gesture (see again footnote 2). In addition, the glottal opening movement during the voiceless sequences always reaches its maximum not during the stop period but during the frication noise production. In other words, glottal opening gestures specific to unaspirated stops, such as word-final /p#/, can not be identified. As for the amplitude of the peak glottal opening, the /sx/ sequence, a combination of two consecutive voiceless fricatives, was the largest. These findings imply that a voiceless fricative tends to require a wide opening gesture, while an unaspirated stop in a voiceless environment can be produced within the opening gesture attributed to the adjacent fricative production.

Fig. 4 contains various voiceless obstruent clusters following a word-final voiceless fricative /s/. It can be observed that the maximum opening for word-final /s/ followed by a vowel is smaller than that for the geminate /s#s/ and the word-initial single /s/ (shown in Fig. 2), although the abductor shows considerably larger activities even for the word-final /s/. The frication period is shortest in this isolated word-final position. Consequently, the velocity of the opening as well as the closing phase do not seem significantly different in either position.

Fig. 4 includes various types of three-phone voiceless clusters beginning with the word-final fricative /s/. It can be seen that the glottal opening for /s#ps/ is characterized by two separate opening gestures, while those for /s#sp/ and /s#sx/ show mono-modal patterns. Again, the local maximum openings are always reached during the fricative segments, regardless of the type. Therefore, it is conceivable that the word-initial stop /p/ in the cluster /ps/ is produced rather during the closing movement of the glottis. In other words, we may conclude that a voiceless fricative requires a wide glottal opening gesture, while an unaspirated stop embedded in a voiceless environment can be produced within the opening gesture attributable to an adjacent fricative.

Acknowledgments

This work was supported in part by Grant in Aid for Scientific Research, Ministry of Education, Japan (No. 57771036) to Hirohide Yoshioka, and by NINCDS Grants NS-13617 and NS-13870, and BRS Grant RR-05596 to Haskins Laboratories, New Haven CT U.S.A., where the original data were collected.

References

Basmajian, J. and G. Stecko (1962): A new bipolar indwelling electrode for electromyography. Journal of Applied Physiology 17, 849.

Hirano, M. and J. Ohala (1969); Use of hooked-wire electrodes for electromyography of the intrinsic laryngeal muscles. Journal of Speech and Hearing Research 12, 362-373.

Hirose, H. (1976); Posterior cricoarytenoid as a speech muscle. Annals of Otology, Rhinology and Laryngology 88, 334-343.

Hirose, H., T. Gay, M. Strome and M. Sawashima (1971); Electrode insertion technique for laryngeal electromyography. Journal of the Acoustical Society of America 50, 1449-1450.

Kewley-Port, D. (1977); EMG signal processing for speech research. Haskins Laboratories Status Report on Speech Research SR-50, 123-146.

Löfqvist, A. and H. Yoshioka (1980); Laryngeal activity in Swedish obstruent clusters. Journal of the Acoustical Society of America 68, 792-801.

Löfqvist, A. and H. Yoshioka (1981); Laryngeal activity in Icelandic obstruent production. Nordic Journal of Linguistics 4, 1-18.

Sawashima, M. (1976); Current instrumentation and technique for observing speech organs. Technocrat 9-4, 19-26.

Sawashima, M. and H. Hirose (1968); A new laryngoscopic technique by use of fiberoptics. Journal of the Acoustical Society of America 43, 168-169.

Yoshioka, H. (1981); Laryngeal adjustments in the production of the fricative consonants and devoiced vowels in Japanese. Phonetica 38, 236-251.

Yoshioka, H., A. Löqvist and H. Hirose (1981); Laryngeal adjustments in the production of consonant clusters and geminates in American English. Journal of the Acoustical Society of America 70, 1615-1623.

Yoshioka, H., A. Löfqvist and H. Hirose (1982); Laryngeal adjustments in Japanese voiceless sound production. Journal of Phonetics 10, 1-10.