Ann.Bull.RILP(1986)No.20, 55-67

HOW VOICELESS SOUND SEQUENCES ARE ORGANIZED IN TERMS OF GLOTTAL OPENING GESTURES

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

I. ABSTRACT

The purpose of the present paper is to summarize the results of our serial investigations on the way how voiceless sound sequences, such as voiceless obstruent clusters and devoiced vowels in a voiceless consonant environment, are organized in terms of their glottal opening and closing gestures, using native speakers of American English, Dutch, Japanese, Icelandic, and Swedish. The techniques used in this experiment were simultaneous recordings of photo-electric glottography, fiberoptic filming and laryngeal electromyography. Our conclusion is 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 phonetic nature of the voiceless segments involved: A voiceless obstruent specified with aspiration noise tends to require an independent opening gesture, while an unaspirated stop appearing in a voiceless environment can be produced within the opening gestures attributed to an adjacent aspirated stop or fricative. As for a voiceless fricative, a continuously wide open glottis is required. The velocity of the initial opening movement varies depending on the properties of the initial voiceless segment: When the initial unvoiced segment in the cluster is a fricative, the speed of the abduction is significantly faster than when the initial voiceless sound is an aspirated or unaspirated stop, regardless of the nature of the following voiceless segments. As for the devoicing phenomenon in Japanese, the data show that the glottal opening gesture for a voiceless sequence containing a devoiced vowel is characterized by a monomodal pattern, unless the vowel occurs between a voiceless fricative and a geminate one, such as /siQs/, where a bimodal pattern may occur. The peak glottal opening attained during a devoiced vowel segment is larger when a fricative either precedes or follows than when the vowel is surrounded by single or geminate stops.

II. INTRODUCTION

In recent years, there has been a significant increase in our understanding of the role of the larynx in accomplishing voicing control. In particular, the approximation and separation

* University of Tsukuba
** Lund University
*** University of Antwerp

of the bilateral vocal folds have been demonstrated to be crucial conditions together with aerodynamic factors--such as the transglottal pressure drop--in determining the initiation, maintenance and cessation of voicing (e.g., Yoshioka, 1984). Since in running speech we produce voiceless sounds as well as voiced ones, it is conceivable that we adjust and accommodate these glottal conditions in an appropriate fashion for succesive speech sound productions.

These notions are, however, usually investigated using simple speech materials, such as regularly alternating sound sequences in terms of the voicing distinction. In other words, there have been surprisingly few systematic observations of laryngeal behaviour during continuously voiceless sound sequences. Our preliminary data implied that such sequential unvoiced sounds are not produced with a static, open glottis, but with continuously changing opening and closing gestures in a regular fashion. Thus, we decided to conduct a series of investigations to explore the temporal organization of the glottal movements during the production of such sequences of unvoiced sounds.

We also noticed that each language permits only a limited number of voiceless sequences. For example, English pronunciation does not include word-initial /pf/ except in loan words, while some other European languages allow such combinations. As for Japanese, its phonology does not allow voiceless, "pure" obstruent clusters other than geminates. Syllable-final obstruents also rarely occur. On the other hand, in conversational speech of the Tokyo dialect, a high front vowel, such as /i/ and /u/, surrounded by voiceless obstruents is typically produced without any vocal pulses including the vowel segment.

Therefore, the purpose of this cross-language study is twofold. One is to explore the general rule governing the organization of the glottal opening and closing gestures during voiceless sound sequences across languages. The other is to reveal why such phonological rules with regard to voiceless sound combinations exist for each language, and how these language-specific unvoiced strings are organized in terms of glottal movements.

III. METHOD AND PROCEDURE

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

The EMG data were obtained using bipolar hooked-wire electrodes. The electrodes, consisting of a pair of platinum-tungsten alloy wires (50 μm in diameter with isonel coating), were inserted perorally into the posterior cricoarytenoid muscle (PCA)--the most involved glottal abductor--under indirect laryngoscopy with the aid of a specially designed curved probe. 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 ms 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 ensemble-averaged tokens with a 35 ms time constant.

The glottal opening and closing movements were recorded by the transillumination technique. A flexible laryngeal fiberscope (Olympus VF-O type, 4.5 mm in outer diameter) inserted through a nostril provided illumination of the larynx. The amount of light passing through the glottis was sensed by a photo-transistor (Philips BPX81) placed on the neck just below the lower edge of the cricoid cartilage. The electric output was stored on another channel of the FM tape. These signals were sampled at 200 Hz and processed with the digital system. The glottal view through the fiberscope served as the monitor for the lighting angle and detecting fogging of the lens.

A direction-sensitive microphone was used for recording the audio signal in a direct mode on another channel of the FM tape. The audio signal was used to determine the line-up points for further processing. In the averaging process, the rectified audio signal was integrated over 15 ms.

Table I Linguistic material in English. All the carrier sentences in Set I were combined with all the words in Set II. The abbreviated phonetic transcriptions indicate the voiceless sounds with which the experiment is concerned. (V=Vowel)

Set I		Set II	
I may	/-V#/	aid	/#V-/
I make	/-k#/	cave	/#k-/
My ace	/-s#/	sale	/#s-/
I mask	/-sk#/	scale	/#sk-/
He makes	/-ks#/		
He Maskes	/-sks#/		

Native speakers of American English, Dutch, Japanese, Icelandic and Swedish served as the subjects. Among the possible voiceless sequences, only meaningful combinations were selected as the linguistic material for each language. For example, the combination of /s/ and /k/ is optimum in forming the greatest possible number of meaningful contexts in English. Therefore, as shown in Table I, sentences containing the phonemes /s/ and /k/ in many combinations were constructed as test utterances. For the Dutch material, the various voiceless obstruents and obstruent clusters in various positions--with a word boundary preceding, following, or intervening within the clusters--are listed in Table II. As for the devoiced vowel in Japanese, we chose test words that contained the devoiceable vowel /i/ between the voiceless obstruents /s/ and /k/. The production of the first word in Table III, /kikee/, may be transcribed as having an unvoiced string [ki̥k], i.e., [k] plus the devoiced vowel [i], followed by a slightly aspirated [k].

Table II Linguistic material in Dutch. All the carrier sentences in Set I were combined with all the words in Set II. The abbreviated phonetic transcriptions indicate the voiceless sounds with which the experiment is concerned. (V=Vowel)

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

Table III Linguistic material in Japanese. All the test words were uttered in the frame sentence "sorewa___desu". The phonetic transcriptions indicate the voiceless sounds with which the experiment is concerned. (Q= geminate phoneme)

奇形	/kikee/	[ki̥k]
吉慶	/kiQkee/	[ki̥kk]
規制	/kisee/	[ki̥s]
吉世	/kiQsee/	[ki̥ss]
詩形	/sikee/	[si̥k]
失敬	/siQkee/	[si̥kk]
姿勢	/sisee/	[si̥s]
失政	/siQsee/	[si̥ss]

IV. RESULTS

Single voiceless obstruents are usually produced with an open glottis. More specifically, for the production of word-initial voiceless obstruents in any languages so far examined, the glottal articulation is characterized by a "ballistic" opening and closing gesture. For example, Figure 1 gives results for the word-initial fricative [s] and the stop [k] in American

English. The top row (GW) represents the glottograms, representing the opening and closing patterns obtained by the transillumination technique. This figure also contains the corresponding averaged activity patterns of the abductor muscle, the posterior cricoarytenoid muscle (PCA), together with audio envelope curves. Here, the glottal movement data and the EMG signals clearly demonstrate that the single voiceless fricative [s] and the stop [k] are produced with a single-peaked monomodal pattern of the glottal opening. A closer comparison of these patterns, however, reveals some differences between the two groups. The peak value of the glottal opening for [s] is significantly larger than that for [k], although the peak is attained much earlier for the fricative than for the stop. Consequently, the velocity of the glottal abduction turns out to be comparably faster for the fricative. These notions have been proved to be valid across other languages.

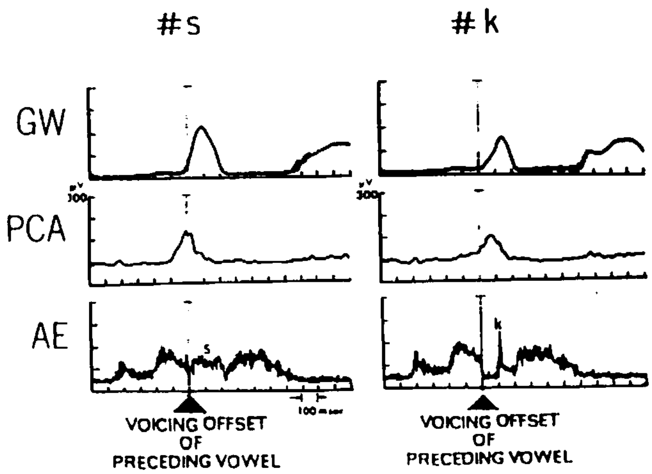

Figure 1 English data. Representative glottograms, averaged electromyograms and audio envelopes for word-initial fricative /s/ and stop /k/.

Figure 2 gives the glottographic patterns for the first eight tokens of the English material containing a voiceless fricative [s] in various contexts including a geminate combination. It should be mentioned that the geminate combination /s#s/ was produced with prolonged continuous frication noise. The glottographic curves among these three types are all characterized by one single opening peak, regardless of the word boundary position. Nevertheless, the detailed patterns of these curves differ in several details. The maximum opening for word-final /s/ is essentially smaller than that for word-initial /s/ or geminate /s#s/, and the frication period is shorter in the single, final position. As for the sharpness of the peak, the geminate /s#s/ is blunt compared to the others. Therefore, the opening and closing velocities are not significantly different in each case. In other words, the glottal opening for the geminate sequence reaches its maximum as quick as that for the single consonants, and the glottis is kept as open as possible during the prolonged frication period.

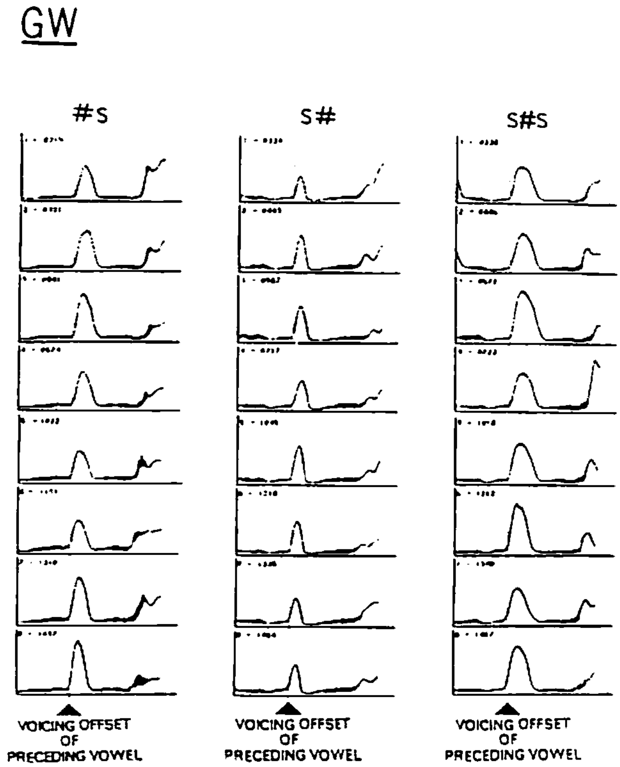

Figure 2 English data. Glottograms for eight productions of three utterance types containing the fricative /s/ in various contexts.

Figure 3 shows the production of the geminate combination /k#k/ in contrast with the corresponding single stops. The acoustic signals revealed that the geminate sound was uttered with a longer duration of the closure period, followed by a degree of aspiration comparable to that for the single, word-initial aspirated stop [k^h]. The pattern of the glottal opening for the geminate may be characterized by a single peak similar to that for the word-initial aspirated cognate. In addition, the stop burst, indicated by arrows in the graphs, shows that, at least in this American English subject, the glottal opening is always at its maximum in the middle of aspiration period for the single stop, while it peaks around the burst for the geminate. From the Swedish and Dutch data, however, we noticed that the burst for single, word-initial stops constantly occurs just after the peak glottal opening is attained. Considering that the degree of aspiration in American English appears to be comparably large, the discrepancy in the time from burst to peak glottal opening for aspirated stops may be attributed to language differences. Rather, it should be mentioned here that the timing of this glottal opening and closing movements in relation to the supraglottal articulatory gesture is particularly tightly controlled for aspirated stop productions, regardless of language difference. In other words, the velocity of the abduction itself may be varied, slower for the geminate than for the single segment. In contrast to these findings, the pattern for the word-final stop /k/ is quite different. The pointed triangles at the left correspond to the implosion of the final stop, and those

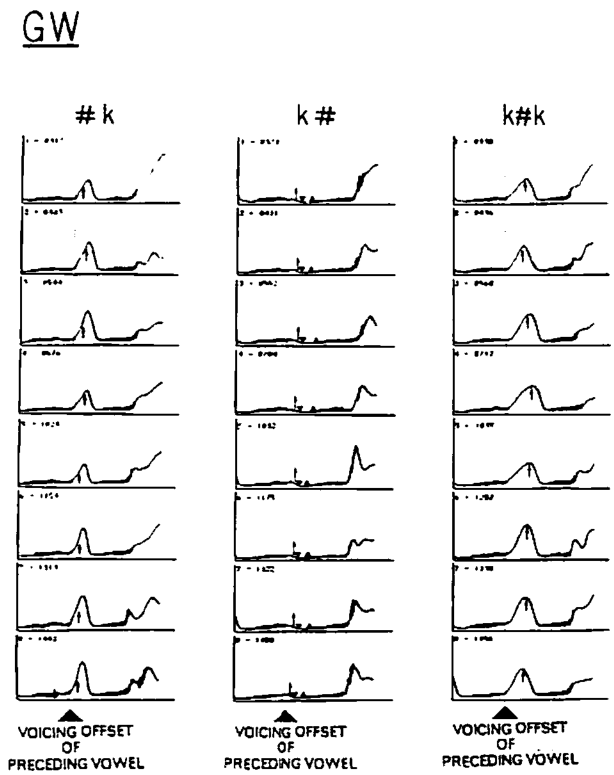

Figure 3 English data. Glottograms for eight productions of three utterance types containing the stop /k/ in various contexts.

at the right indicate the release for the glottal attack of the following word-initial vowel. The data clearly demonstrate that the word-final stop was actually produced with a nearly closed glottis, restricting the airflow from the lungs. The Dutch data also show the same result for this so-called "glottalization" gesture.

Figure 4 shows the glottographic patterns for eight tokens of three utterance types where the place of the word boundary in the voiceless cluster /sk/ sequence varies. An overall survey reveals that, although there are some variations within each utterance type, the glottograms for the word-initial /#sk/ and word-final /sk#/ sequences show one single opening gesture. On the other hand, those for the /s#k/ combination regularly demonstrates two separate opening gestures during the same phoneme sequence. In order to illustrate the corresponding activity patterns of the abductor muscle (PCA), Figure 5 shows the averaged electromyographic curves for this muscle for the same three types, together with representative glottograms and the corresponding audio envelopes. These curves further confirm that

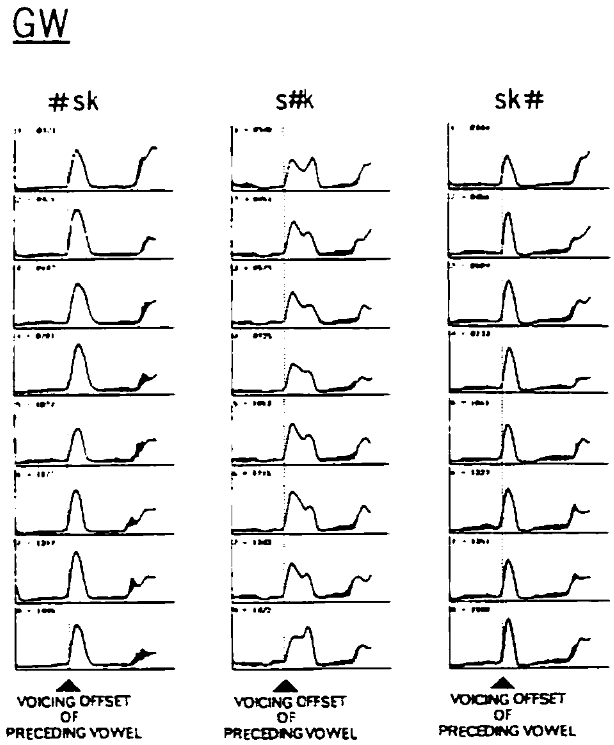

Figure 4 English data. Glottograms for eight productions of three utterance types containing the /sk/ sequence in various contexts.

the difference in the number of peak openings among the productions of the /sk/ combination is inherently linked to an underlying difference in the activity pattern of the abductor muscle, with some time lead. A closer comparison of the acoustic signals with the glottograms reveals that the local maximum openings of the glottis are always reached during the fricative segment [s] and the aspirated stop [k^h] if present. No specific opening gesture may be detectable for the unaspirated stop [k]. Rather, the glottal articulation during the unaspirated stop seems to be a continuation of the closing phase of the glottal gesture for the preceding [s] segment. This result that, in a cluster of fricative plus unaspirated stop, only one glottal gesture is found, has been further confirmed for complicated consonant clusters even in other languages, such as Swedish and Icelandic, as shown in Figures 6 and 7, respectively.

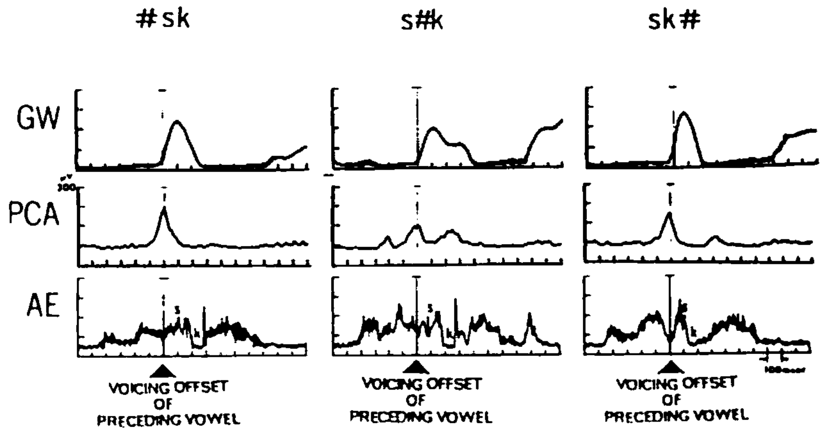

Figure 5 English data. Representative glottograms, averaged electromyograms and audio envelopes for three utterance types containing the /sk/ sequence in various contexts.

As for sequences of voiceless fricative clusters, the glottal opening gesture shows a rather simple pattern, with a single blunt peak, or sometimes with a plateau around the peak. Figure 8 contains various types of fricative clusters, some of which include unaspirated stops. Interestingly, these Dutch data show that, unlike the case of the geminates in American English, the muscle activity patterns show a separate opening gesture for each segmental fricative, although at the movement level the glottal opening gesture is rather monomodal. In other words, the adjustments for each voiceless fricative originate separately and become fused at last. Of another interest is that the velocity of the initial opening movement varies depending on the properties of the initial voiceless segment. When the first unvoiced segment in the cluster is a fricative, the speed of the abduction is significantly faster than when the initial voiceless sound is an aspirated or unaspirated stop, regardless of the nature of the following voiceless segments. This fact has been also disclosed in other languages.

Although Japanese voiceless clusters containing a devoiced vowel are not "pure" consonant clusters, the glottal movement and its adjustment resemble those for fricative clusters in European languages. Figure 9 shows various utterance types containing the devoiceable vowel /i/ in a voiceless sequence composed of two different obstruents, /s/ and /k/. It is obvious here that the glottograms are characterized by a monomodal pattern, unless the vowel occurs between a voiceless fricative and a geminated one, such as /siQs/, where a bimodal pattern may occur. The peak

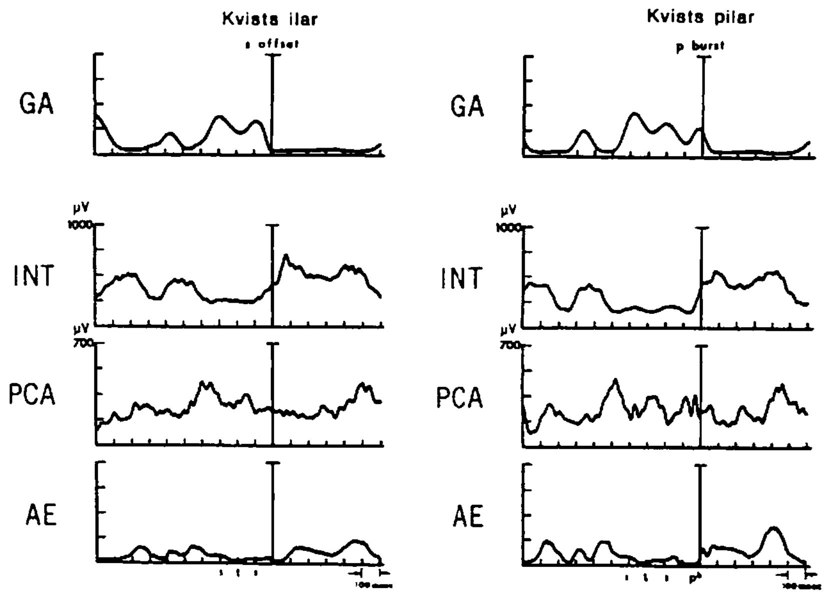

Figure 6 Swedish data. Glottograms, averaged electromyograms and audio envelopes for various voiceless obstruent clusters.

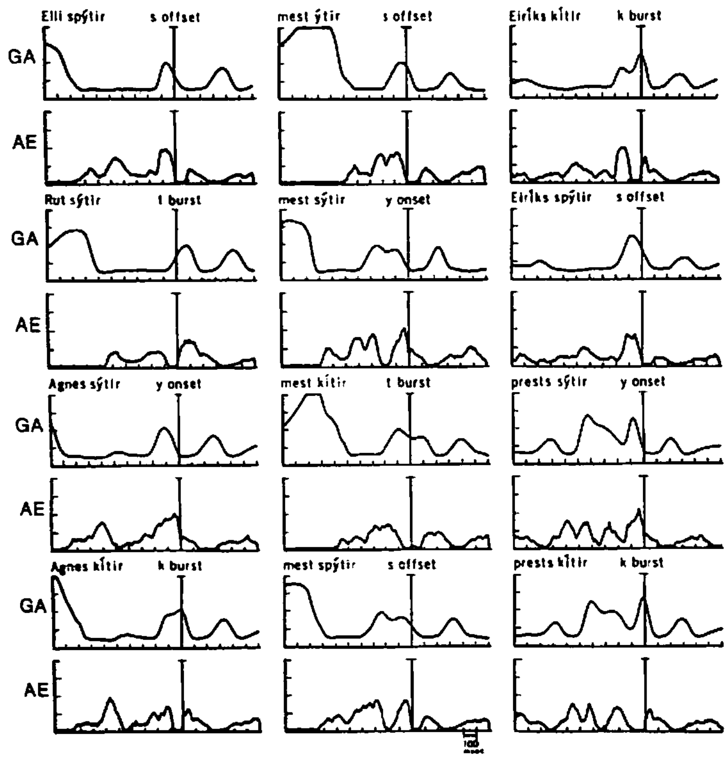

Figure 7 Icelandic data. Glottograms, and audio envelopes for various voiceless obstruent clusters.

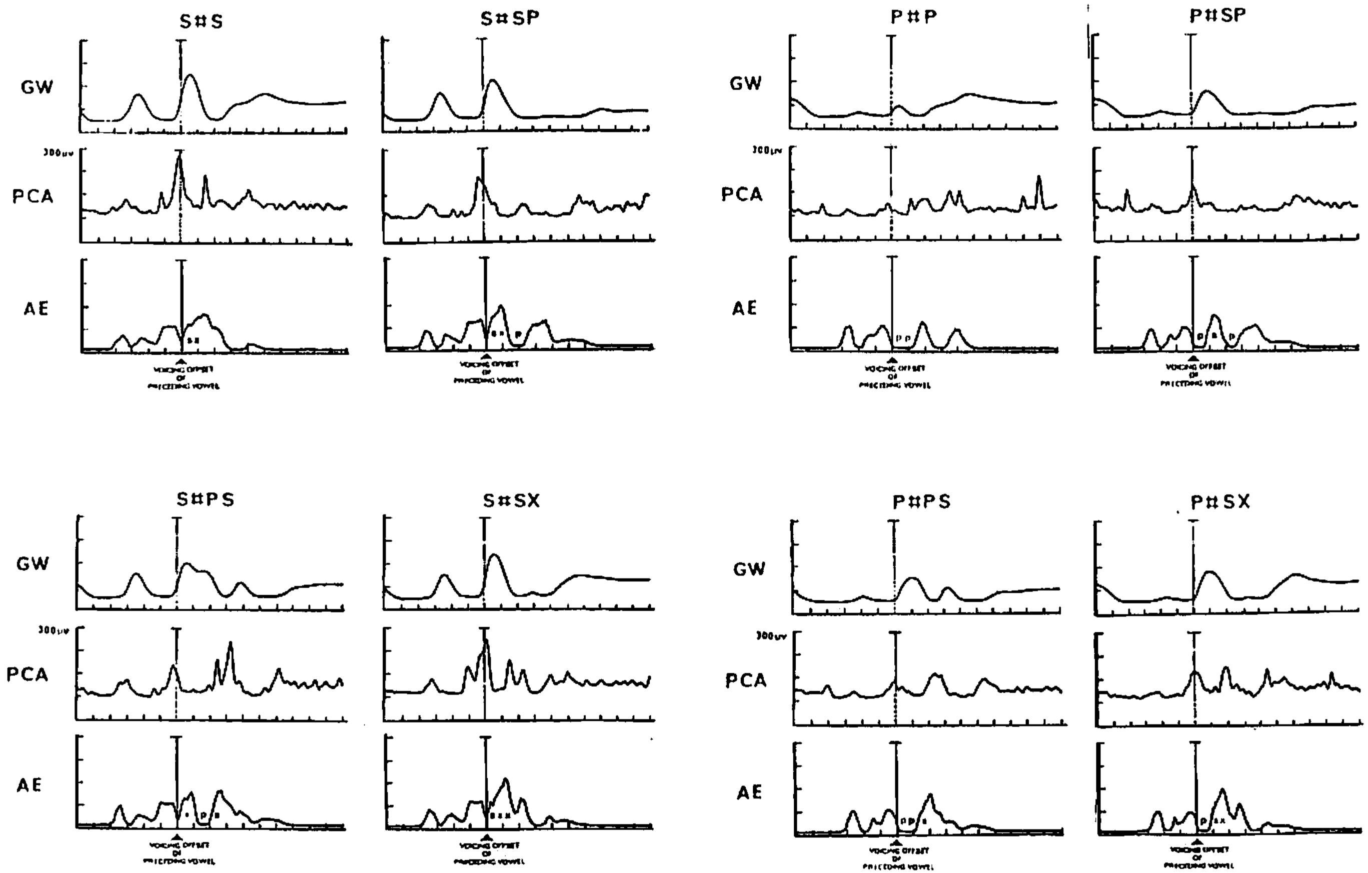

Figure 8 Dutch data. Glottograms, electromyograms and audio envelopes for various voiceless obstruent clusters.

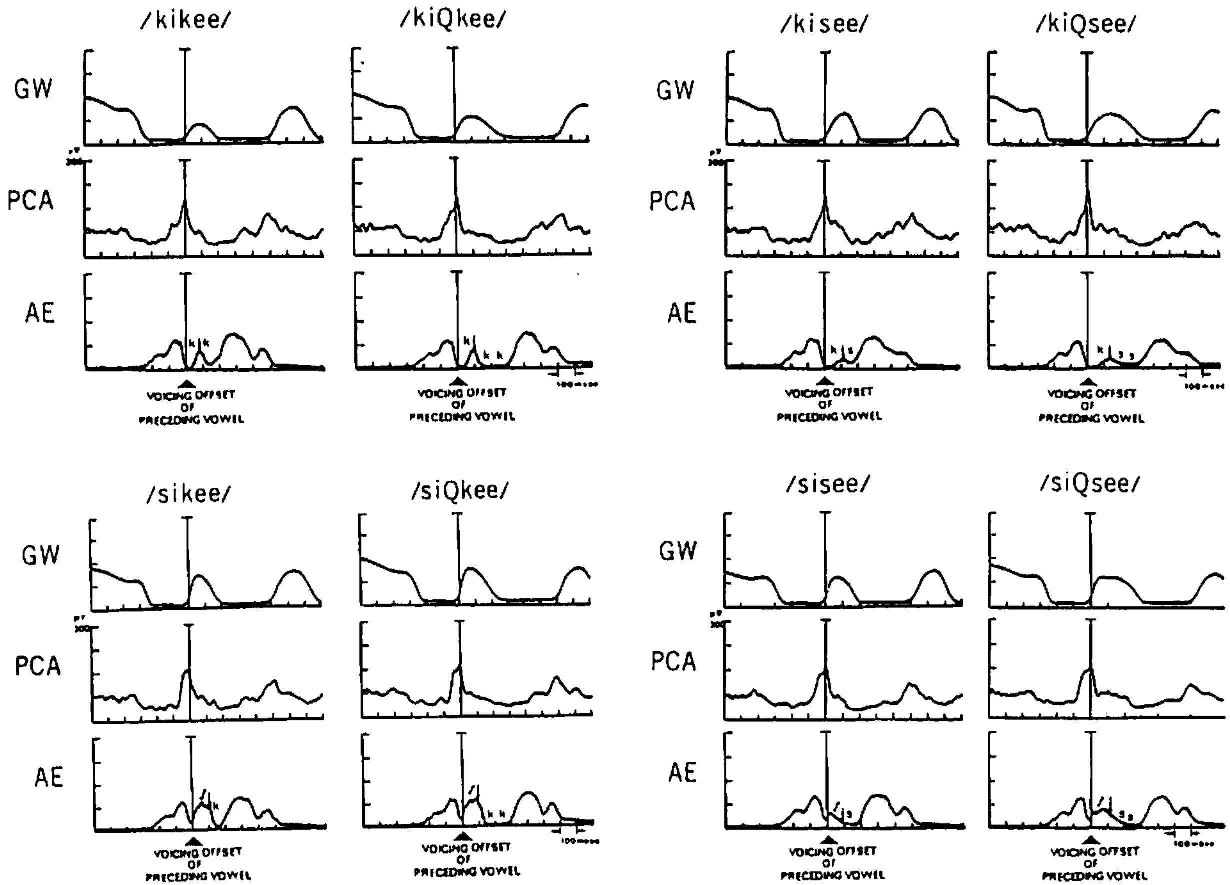

Figure 9 Japanese data. Glottograms, electromyograms and audio envelopes for various utterance types containing the devoiced vowel /i/.

glottal opening, attained during the devoiced vowel segment, is larger when a fricative either precedes or follows than when the vowel is surrounded on both sides by single or geminated stops. The velocity of the initial opening phase is also different, depending upon the nature of the initial obstruent. When a voiceless fricative precedes the devoiced vowel, the opening movement is faster than when voiceless stop precedes the vowel. Therefore, we may conclude that the devoicing phenomenon in Japanese vowels is a substitute for consonant clusters at least from phonetic viewpoint.

VI. ACKNOWLEGMENTS

The collection of the original data was supported by NINCDS Grants NS-13617 and NS-13870, and BRS Grant RR-05596 to Haskins Laboratories, when the senior author was a visiting researcher.

REFERENCES

Löfqvist, A. and H. Yoshioka: Laryngeal Activity in Swedish Obstruent Clusters. Journal of the Acoustical Society of America, 68, 792-801, 1980.

Löfqvist, A. and H. Yoshioka: Interarticulator Programming in Obstruent Production. Phonetica, 38, 21-34, 1981.

Löfqvist, A. and H. Yoshioka: Laryngeal Activity in Icelandic Obstruent Production. Nordic Journal of Linguistics, 4, 1-18, 1981.

Löfqvist, A. and H. Yoshioka: Intersegmental Timing: Laryngeal-oral Coordination in Voiceless Consonant Production. Speech Communication, 3, 279-289, 1984.

Yoshioka, H.: Laryngeal Adjustments in the Production of the Fricative Consonants and Devoiced Vowels in Japanese. Phonetica, 38, 236-251, 1981.

Yoshioka, H., A. Lofqvist and H. Hirose: Laryngeal Adjustments in the Production of Consonant Clusters and Geminates in American English. Journal of the Acoustical Society of America, 70, 1615-1623, 1981.

Yoshioka, H., A. Lofqvist and H. Hirose: Laryngeal Adjustments in Japanese Voiceless Sound Production. Journal of Phonetics, 10, 1-10, 1982.

Yoshioka, H., A. Lofqvist and R. Collier: Laryngeal Adjustments in Dutch Voiceless Obstruent Production. Ann. Bull. RILP, 16, 27-35, 1982.

Yoshioka, H.: Glottal Area Variation and Supraglottal Pressure Change in Voicing Control. Ann. Bull. RILP, 18, 45-49, 1984.