Ann. Bull. RILP, No. 10, 37-46 (1976)

A PRELIMINARY REPORT ON THE ELECTROMYOGRAPHIC STUDY OF THE PRODUCTION OF THE JAPANESE SEMIVOWEL /j / *

Kuniko Kakita, Hajime Hirose, Tatsujiro Ushijima
and Masayuki Sawashima

Abstract

This report gives certain preliminary results of an investigation of the articulatory properties of the Japanese semivowel /j/ by use of electromyography (EMG). Specifically, the EMG activity of one lingual muscle, the genioglossus muscle (GG), during the production of /j/ in various phonetic contexts was analyzed and compared with corresponding EMG data on the vowel /i/, since these two speech sounds are very much alike and are usually described as having a common phonetic property, viz., "high-front". It was found that while the production of /i/ involves the contraction of the GG as a whole, /j/ is produced with a characteristic decrease in the activity of its posterior fibres. The EMG results were then compared with x-ray data on the tongue configuration for /j/. A characteristic backward displacement of the tongue root for /j/ in comparison with /i/ was interpreted to be in correlation with the decrease in the EMG activity of the posterior GG fibres.

0. Introduction

Some tentative results of the analysis of the EMG data on Japanese vowels have been presented elsewhere (Miyawaki et al., 1975). To recapitulate those results very briefly, five pairs of electrodes were placed in different portions of the GG, i.e., from "relatively anterior" to "relatively posterior" fibres, of one subject, a native speaker of Japanese, and the EMG activity was recorded simultaneously from these five locations during the production of various $/V_1V_2/$ sequences. It was found that all GG locations show high activity for /i/, whereas they show very low activity for /a/. The production of /e/ was found to involve primarily the anterior fibres, while the production of /u/ involves the posterior fibres. The activity for /o/ was low for all GG locations except the posterior fibres, which showed a fair amount of activity.

---

* This is a slightly revised version of a paper which appeared in The Tsuda Review, No. 20, 1975.

In this report we will examine the EMG activity during the production of /j/ in various contexts. Our specific interest here will be to compare the EMG activity for a "semivowel" /j/ ([j]) with that for a "vowel" /i/ ([i]), especially because these speech sounds are closely related phonetically.

The general description of the two semivowels [j] and [w] has been given by a number of phoneticians (Jones, 1918; Hattori, 1951; Heffner, 1964; Pike, 1964 and Abercrombie, 1967). They unanimously point out 1) that in the production of a semivowel, the primary articulators assume, or approximate, a position similar to that of an analogous vowel, viz. [i] in the case of [j] and [u] in the case of [w], and 2) that a semivowel is characterized by its 'gliding' property and consequently by its non-syllabicity. Thus, for example, Jones writes: "In pronouncing the most usual English j the speech-organs start at or near the position for the English 'short' i... and immediately leave this for some other sound.... "(Jones, 1918; p.209).

However, research efforts directed specifically at disclosing the phonetic characteristics of semivowels, either articulatory or acoustic, seem rather scarce. With regard to Japanese, Miyawaki et al. have tentatively shown, by means of Dynamic Palatography, that the palato-lingual contact pertaining to [j] observed during the production of selected [VjV] sequences is characteristically different from that of [i] (Miyawaki et al., 1973, 1974). Hattori has pointed out that the actual tongue height for [j] varies with phonetic context (Hattori, 1951). Thus, for example, [j] of [kaja] ('mosquito-net') is produced with a tongue position describable as [eᴛ], in a narrow phonetic transcription, whereas that of [heja] ('room') would be transcribed as [ɛ⊥]. That the articulation of [j] is largely dependent on the context has also been shown recently by Hiki and Imaizumi (1974). By using an analysis-by-synthesis method, they attempted to derive from the actual EMG data the analog time-varying patterns of neuromotor commands, in terms of the optimum target value and its beginning for each phoneme. They found that the variability of the target value due to contextual difference was considerable both for [i] and for [j]. An acoustic study of the semivowels from the viewpoint of speech recognition has also been carried out recently by Fujisaki and his colleagues (Fujisaki et al., 1975). To state briefly, an abstract command for each phoneme, in terms of the value of the formant target, its duration, and the rate of transition between consecutive phonemes, was extracted as parameters from time-varying patterns of measured formant frequencies. The results of perceptual experiments employing these parameters as control variables indicated that, of the three parameters, the command duration was "most effective in discriminating semivowels [j] and [w] from vowels [i] and [u]" (Fujisaki et al., 1975; p. 1).

We have, then, a partial knowledge of various phonetic properties of semivowels. The present study hopefully offers additional information to the existing knowledge as surveyed above, revealing some of the physiological correlates of the Japanese semivowel [j], specifically in comparison with those of the cognate vowel [i].

1. Materials and methods

The EMG activity of a given subject during the production of various $/V_1jV_2/$ sequences was recorded simultaneously from the above-mentioned five GG locations (GG1 through GG5, corresponding to anterior through posterior fibres). The $/V_1jV_2/$ sequences were randomized together with the $/V_1V_2/$ sequences selected for the vowel analysis referred to above, and were read by the subject during a single recording session. A carrier sentence /kono ___ onokosimasu/ ('We keep this ___.') was used to embed both $/V_1jV_2/$ and $/V_1V_2/$) sequences. No accent kernel was placed on the $/V_1jV_2/$(and $/V_1V_2/$) sequences. The number of utterance samples for each utterance type was 12.

The list of $/V_1jV_2/$ sequences which served as speech material is given in Table 1. The total number of possible $/V_1jV_2/$ combinations in Japanese is 15, but certain combinations were of necessity omitted from

V_1 \ V_2	a	o	u
i	(ija)	(ijo)	(iju)
e	x	x	eju
a	aja	x	aju
o	oja	ojo	oju
u	uja	ujo	uju

Table 1: The list of $/V_1jV_2/$ sequences that served as speech material. An 'x' indicates a lack of data for that particular sequence. (See text for explanation of bracketed sequences.)

the list due to an experimental restriction on the total amount of speech material. The bracketed sequences /ija/, /ijo/, and /iju/ are not included in the analysis for a reason which will be given below.

The data was obtained by use of an EMG data collection and processing system at Haskins Laboratories. A detailed technical description of this system is given in Kewley-Port (1973, 1974). The EMG signals were recorded simultaneously with sound signals by use of bipolar hooked-wire electrodes. The use and insertion of wire electrodes are discussed in Hirose (1971). The raw EMG signals thus obtained were full-wave rectified, integrated, and then averaged over different samples of each utterance type with reference to a line-up point, viz., implosion of the first /n/ in the carrier sentence.

Figure 1 shows an example of the averaged EMG curves. The utterance is /konoojoonokosimasu/ ('We keep this ojo.'). A curve in each row displays, as a function of time, the averaged EMG activity recorded from one of the five electrode locations. An arrow indicates the peak in the EMG curve which corresponds to the articulation of /j/ in the given context /...o_o.../. It may also be seen that every GG location except GG5 shows

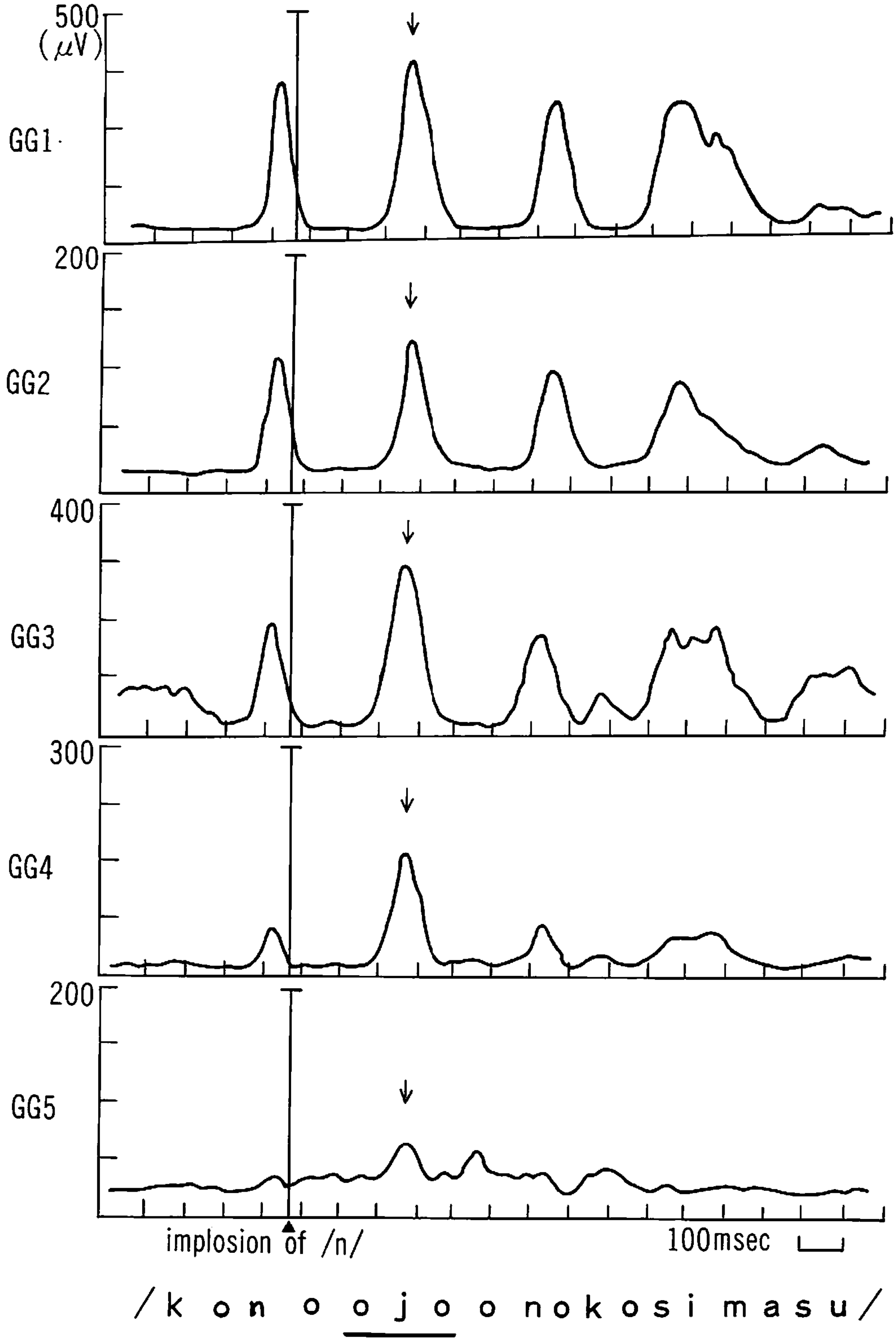


Figure 1. Averaged EMG curve for /konoojoonokosimasu/. An arrow indicates the peak in the EMG curve which corresponds to /j/.

activity for /n/ ([n]) and /si/ ([ʃi]) portions of the utterance.

For the analysis, the measurements of the EMG activity values were made from the averaged EMG curves as shown in Fig. 1 above. More precisely, for each $/V_1jV_2/$ sequence the peak values corresponding to /j/ were identified for each electrode location and the values were measured off the curve. This was feasible in most cases; however, when it was difficult to identify a 'peak' in any one of the GG channels, the location of the peak in each of the other channels was consulted for the same utterance type. In $/V_1jV_2/$ sequences where $/V_1/$ was /i/, it was impossible to identify a peak corresponding uniquely to /j/ since all GG locations showed peaks corresponding to combined activity for both /i/ and /j/. These sequences were therefore omitted from the present analysis.

3. Results

Table 2 gives the EMG activity values corresponding to /j/ in various

electrode location / utterance type	GG1	GG2	GG3	GG4	GG5
/eju/	70	54	64	60	55
/aja/	83	66	67	63	22
/aju/	90	70	92	89	45
/oja/	96	79	79	60	19
/ojo/	100	95	85	70	29
/oju/	96	100	89	67	39
/uja/	100	79	58	51	22
/ujo/	100	95	77	44	22
/uju/	100	95	64	44	36
mean	93	81	75	61	32
standard deviation	(9.7)	(15.0)	(11.5)	(13.3)	(12.1)

Table 2. EMG activity values (peak values) corresponding to /j/ in various contexts.

contexts. The values measured from the averaged EMG curves have been converted into percentage values, taking the highest activity value obtained for each electrode location to be 100%.

As in the case of /i/, all GG locations show activity for the production of /j/. There is, however, a tendency for a more posterior location to show a lesser amount of activity compared with /i/. Since both /i/ and//j/ are described phonetically as having a 'high-front' tongue configuration, it is of interest to compare them more closely at the EMG level. Figure 2 compares the EMG values for /j/ in the sequences /oja/, /ojo/, and /oju/ with those of /i/ in the sequences /ia/, /io/ and /iu/: these /V_1V_2/ sequences are preceded by a vowel /o/ of the carrier sentence /kono ___ onokosimasu/, thus providing phonetic contexts reasonably comparable with those of /j/, i.e., /o_a/, /o_o/, and /o_u/. Data points for each electrode location indicate the amount of EMG activity corresponding to /j/ when the activity values for /i/ in the corresponding contexts are taken to be 100%. It is clearly shown that the more posterior the GG location, the smaller the amount of EMG activity for /j/ when compared with /i/.

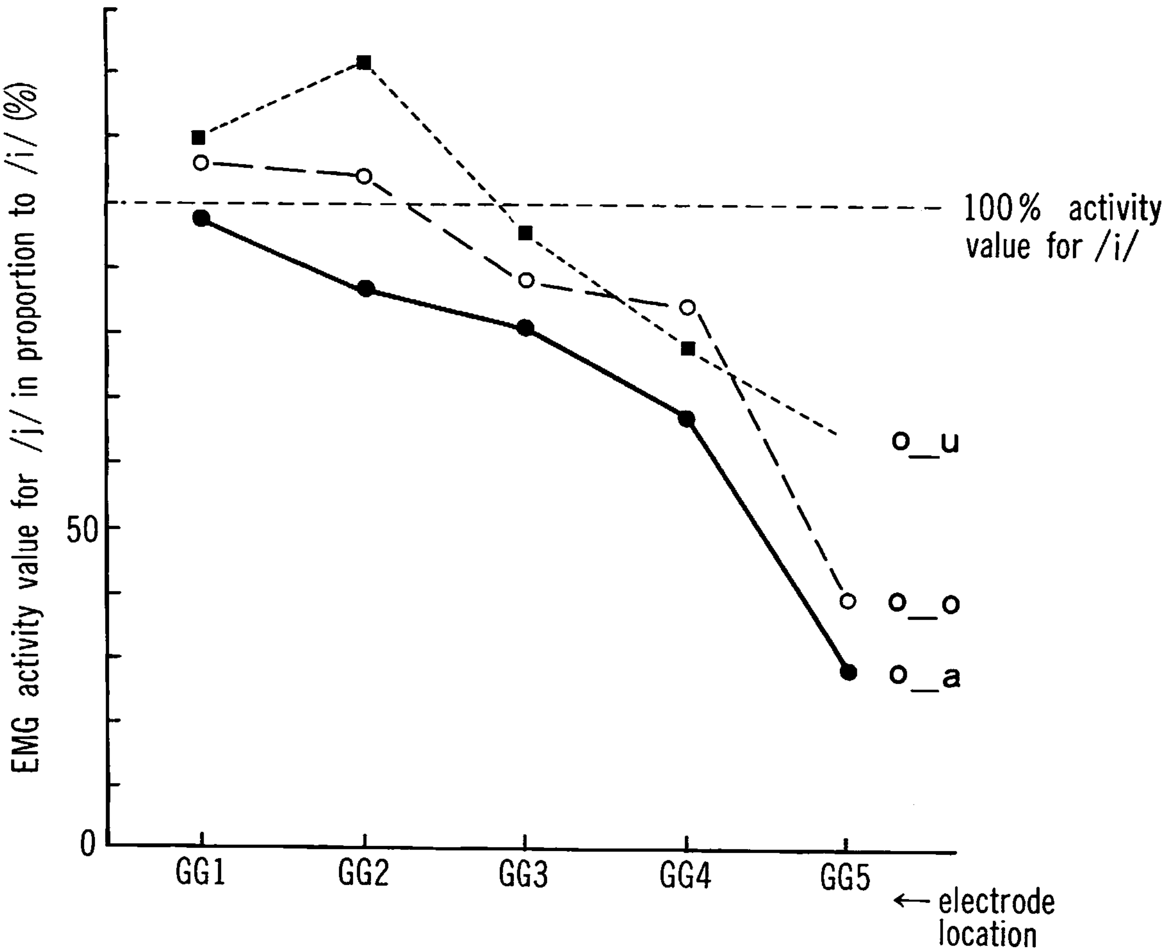


Figure 2. EMG activity values for /j/ in proportion to activity values for /i/ (set at 100% for each electrode location) in matched contexts.

3. Comparison with x-ray data [1)]

Now, if it is the posterior fibres of the GG that show a reduced amount of activity in the production of /j/ compared with that of /i/, we would expect the tongue root to be pulled less anteriorly, or displaced backward towards the pharyngeal wall, during the production of /j/ than it would be in the case of /i/. This is due to the fact that the posterior fibres of the GG are the sole agent for pulling the tongue root forward. X-ray data (though of a different subject SK) does in fact show that the root of the tongue is indeed pulled less forward during the articulation of /j/ compared with that of /i/ (Figure 3).

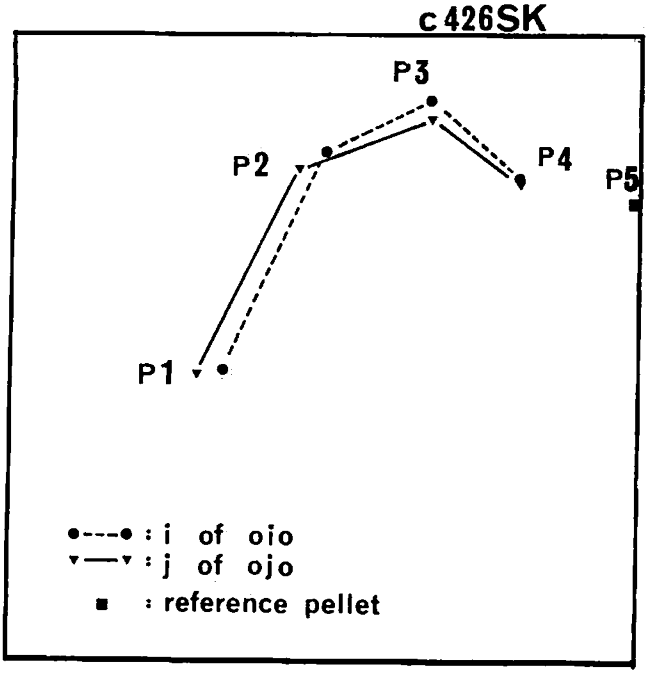
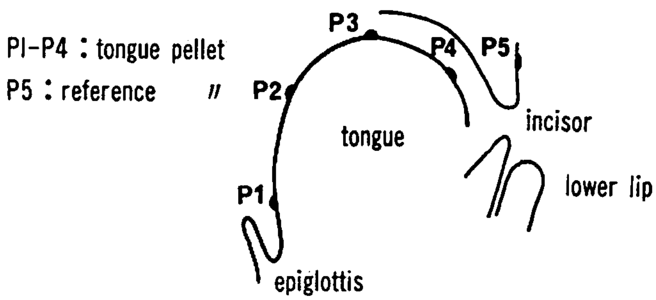


Figure 3. Pellet positions for /i/ and /j/ in the context /o _ o/.



1) The x-ray data presented here is taken from a larger body of data obtained by means of a computer-controlled x-ray microbeam system developed at the Research Institute of Logopedics and Phoniatrics, Faculty of Medicine, University of Tokyo (Kiritani et al., 1975; Kiritani and Fujimura, 1975). The entire body of data is now being analyzed on a full scale and the results will be reported on elsewhere.

Figure 3 shows the schematic tongue contours for /i/ and /j/ in the context /o _ o/, produced once each without any carrier sentence, in terms of the positions of pellets (P1 - P4) placed along the midsagittal line on the dorsum of the tongue. 2) These pellet positions pertain to the points in time when the pellets are maximally displaced from the position of the preceding vowel /o/. 3) The two tongue contours are superposed with reference to a pellet placed on the upper incisor ("REFERENCE PELLET,P5" in the figure).

Note the position of P1. As expected, in the production of /j/, it is characteristically displaced more posteriorly towards the pharyngeal wall than during the production of /i/. This is in agreement with the aforementioned result of the EMG analysis that the posterior fibres of the GG are not contracting to an extent that they would for /i/. It may be argued that some other muscles, e. g., the hyoglossus muscle or the pharyngeal constrictor, are contracting to pull the tongue root backward: if so, however, it would be all the more probably that the posterior fibres of the GG would decrease their activity in order to cooperate in the backward displacement of the tongue root.

With regard to the GG, then, the semivowel /j/ is distinct from the vowel /i/ in that the contraction of the muscle fibres, or more precisely, the degree of contraction specified for different portions of the GG, is characteristically different for /j/ and /i/.

The fact that the fibres of the GG do not show a spatially uniform contraction is plausible when we consider the unique morphology of the GG, its fan-shaped radiation. Normally, a muscle contraction takes place between two locations, resulting simply in a shortening of the distance between them; however, in the case of GG, the contraction takes place between one origin and 'many insertions', the result of which is rather complex.

It is interesting to note in passing, however, that the EMG peaks for the five locations all occur simultaneously. It seems that, although the specification of the degree of contraction varies according to the various portions of this muscle, the temporal specification is nevertheless identical (i. e., simultaneous) for all muscle areas. It therefore seems reasonable to speak of the GG as "one muscle" temporally, even though it is "more than one muscle" spatially.

In general, we may say that two types of specification are required for a motor command to contract the GG: 1) a spatial specification, i. e., specifying which portion of the GG is to contract to what degree, and 2) a temporal specification, i. e., specifying when to contract.

---

2) The same kind of comparison was made of /i/ and /j/ in the context /a _ a/. The results reported on in the text also pertain to this comparison pair.

3) P1 is placed about 1.5 cm above the valecula and P4, about 2 cm from the tip of the tongue at rest position; P2 and P3 are so placed that the distance between the four pellets would be approximately the same. Since the pellets are placed along the midsagittal line, they may be considered to reflect the state of the GG directly, especially since we know that the intrinsic muscles do not show any signifiant activity during the production of vowels and semivowels. (The GG fibres reach the tongue dorsum medially and paramedially, but not laterally).

4. Further problems.

The present study was concerned exclusively with comparing /i/ and /j/ at a specific temporal point in the course of articulatory movement. That is to say, the EMG activity corresponding to /i/ and /j/ was measured at the moment in time when they reached maximum value, and these measurements were then compared with the tongue contours observed at the specific moment when the most i-/j-like tongue configurations were attained. In other words, the study does not take into account the temporal factor intrinsic to any kind of articulatory gesture.

As is pointed out by many phoneticians, and also suggested by the study of Fujisaki et al. (1975), the temporal element is essentially non-negligible in the analysis of semivowels. For instance, the time course of the EMG activity curves should be compared with that of the pellet movements of the x-ray observation. It would also be necessary to look into the temporal dimension of the acoustic signals, e.g., the time-varying formant frequencies which reflect the time-varying shape of the vocal tract.

An analysis along these lines is now under way.

5. Summary

In this paper we have reported on some of the results of the EMG study of the production of the semivowel /j/. It was found a) that /j/, as compared with the vowel /i/, is produced with a characteristically reduced amount of activity in the posterior fibres of the GG, and b) that the more posterior the GG location, the greater is the degree of reduction.

The EMG results were also compared with x-ray data showing the tongue configuration for /j/. A characteristic backward displacement of the tongue root for /j/ in comparison with /i/ was interpreted to be in correlation with the decrease in the EMG activity of the posterior GG fibres.

Acknowledgments

The electromyographic experiments were conducted at Haskins Laboratories, U.S.A., partly supported by Grant No. DE-01774 from the National Institute of Dental Research, National Institutes of Health. I would like to express my gratitude to all members of the laboratory for their kind assistance in the data collection and processing.

References

Abercrombie, D. (1967), Elements of General Phonetics, Univ. Press, Edinburgh.

Basmajian, J. V. (1962), Muscles Alive: Their Functions Revealed by Electromyography, Williams and Wilkins, Baltimore.

*Fujisaki, H. et al. (1975), 'Recognition of Semivowels in Spoken Words' Handout for the Meeting on Speech Science (Onseikenkyuukai shiryoo) S-75-01.

*Hattori, S. (1951), Onseigaku [Phonetics], Iwanami-shoten, Tokyo.

Heffner, R-M. S. (1964), General Phonetics, The University of Wisconsin Press, Madison.

*Hiki, S. and S. Imaizumi (1974), 'Rules of Generating Motor Commands of Tongue Muscles in Connected Vowels and Semivowels' Handout for the Meeting on Speech Science (Onseikenkyuukai shiryoo) R-73-48.

Hirose, H. (1971), 'Electromyography of the articulatory muscles: current instrumentation and technique' Status Report on Speech Research (Haskins Laboratories) SR 25/26 (pp. 77-86).

Jones, D. (1918), An Outline of English Phonetics, (9th edition 1960: Maruzen Reprint, Tokyo.)

Kewley-Port, D. (1973), 'Computer processing of EMG signals at Haskins Laboratories' Status Report on Speech Research (Haskins Laboratories) SR-33, (pp. 173-184).

Kewley-Port, D. (1974), 'An experimental evaluation of the EMG data processing system: time constant choice for digital integration' Status Report on Speech Research (Haskins Laboratories) SR-37/38 (pp. 65-78).

Kiritani, S. et al. (1975), 'Tongue Pellet Tracking and Other Radiographic Observations by a Computer Controlled X-ray Microbeam System' Annual Bulletin (Research Institute of Logopedics and Phoniatrics, Univ. of Tokyo) No. 9 (pp. 1-13).

Kiritani, S. and O. Fujimura (1975), 'Computer-controlled X-ray study of Japanese consonants and vowels--a preliminary report' Journal of the Acoustical Society of America 58, Suppl. No. 1 (p. S-12).

*Miyawaki, K. et al. (1974), 'Observation of Some Selected Articulations in Japanese by Use of Dynamic Palatography' Handout for the Meeting on Speech Science (Onseikenkyuukai shiryoo) S-73-47.

Miyawaki, K. et al. (1974). 'Palatographic Observation of VCV Articulations in Japanese' Annual Bulletin (Research Institute of Logopedics and Phoniatrics, Univ. of Tokyo) No. 8 (pp. 51-57).

Miyawaki, K (1974), 'A study on the musculature of the human tongue: observations on transparent preparations of serial sections' Annual Bulletin (Research Institute of Logopedics and Phoniatrics, Univ. of Tokyo). No. 8 (pp. 23-50).

Miyawaki, K. et al. (1975), 'A preliminary report on the electromyographic study of the activity of lingual muscles' Annual Bulletin (Research Institute of Logopedics and Phoniatrics, Univ. of Tokyo) No. 9 (pp. 91-105).

Pike, K. L. (1964), Phonetics, The Univ. of Michigan Press, Ann Arbor.

* Written in Japanese.