FIBERSCOPIC OBSERVATION OF VELAR MOVEMENTS DURING SPEECH

T. Ushijima and M. Sawashima

Introduction

A new technique of the direct visualization of velar movements through the nasal cavity, by use of a fiberscope, was presented in the last issue of our progress report.[1)]

This fiberscopic observation provides us with useful knowledge about the dynamic characteristics of velar movements during speech. In this paper we report several findings so far obtained for Japanese speech samples.

Experimental Procedure

1. Speech Samples

Nonsense syllables and meaningful words were prepared as shown in Table 1. Meaningless syllables, which were pronounced in isolation, consisted of a series of Japanese five vowels, repetitions of /CV/ syllables, and a repetition of /teN/, where /N/ is the syllable final nasal in Japanese. The meaningful words contained nasal sounds in various phonological environments in combination with the vowel [e] and non-nasal consonants. These words were pronounced with a frame sentence of "... desu" (it is ...). No accent kernel was attached to every test word.

2. Instrumentation and Procedures

The fiberscope used is the one we call a "wide angle model," specifications of which have been reported elsewhere.[1)] This wide angle model was inserted through the nose of subject, positioned in the inferior nasal meatus, and clamped at the nostril. The distance from the tip of the scope to the nostril was approximately 6 cm. In this position the nasal surface of the velum was clearly observed except when it was lowered to an extreme position as in the case of the rest position.

The velar view was photographed by a 16-mm cinecamera at a rate of

Table 1. List of test words

Nonsense syllables	Meaningful words
/aiueoaiueoa/	/see'ee/
/tetetete/	/see'eN/
/sesesese/	/deenee/
/dededede/	/seN'ee/
/zezezeze/	/seNsee/
/nenenene/	/zeNsee/
/teNteNteNteN/	/neNsee/
	/seeneN/
	/seN'eN/
	/teNseN/
	/deNseN/
	/heNseN/
	/seNteN/

50 frames per sec, with simultaneous recording of the speech signal and synchronization time marks. Sound spectrograms of the speech signal with the time marks were made to obtain correspondence between film frames and speech sounds.

Two male Japanese, both speakers of the Tokyo dialect, served as subjects. Each subject practiced reading the prepared speech samples: the nonsense syllables twice and the meaningful sentences three times. All of the repeated samples were photographed to analyse the frames.

3. Analysis of the Frames

The photographed film frames were examined, frame by frame, by means of a motion picture analyzer. Fig. 1 shows a schematic view of the velum with superimposed traces of its three positions; 1) at the rest position, 2) at a moderately elevated position, and 3) at a highly elevated position. In the view field a convex outline delimits the nasal surface of the velum against

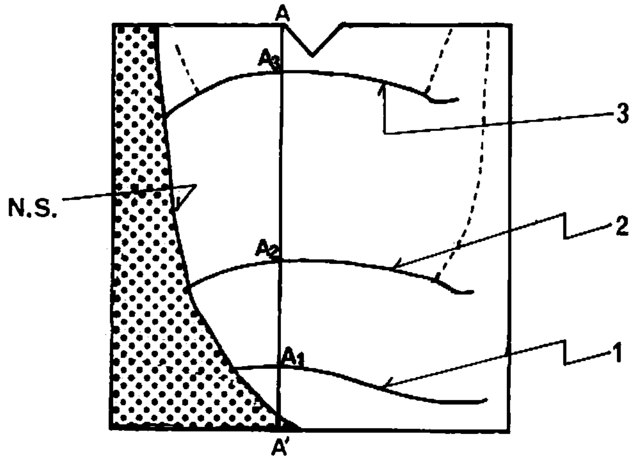

Fig. 1. Schematic drawing of the outline of the upper surface of the velum:
1: at rest
2: at a moderately elevated position
3: at a highly elevated position.
Velar height is defined as the displacement along line AA' from A1. N. S. : the nasal septum.

the posterior wall of the epipharynx.

The contour of the posterior edge of the nasal septum, which runs down to the nasal floor, served as a reference to detect a shift in the position of the tip of the scope. Generally the scope stayed fixed in respect to the maxilla during utterances. In a very few occasions, however, we found some shift of the view field which took place between different series of utterance. The head movement would result in a slight vertical and/or horizontal shift as well as a subtle twist of the view field. Correction of the measurement of the velar height was then made with reference to the contour of the nasal septum.

In order to estimate the velar height, we used a vertical line AA' on the frame as shown in Fig. 1. This line was obtained by connecting the highest point of the velum in each frame. Velar height was then defined as the displacement along line AA' in millimeters from A1; the point of intersection across the outline 1 at rest position of the velum.

Results and Remarks

1. Velar Height and Velopharyngeal Distance

The opening of the velopharyngeal passage results in an acoustic coupling of the nasal cavity to the vocal tract. In order to find an observable measure that relates to the degree of this coupling, the relationship between the velopharyngeal distance and the velar height, as observed in the technique above, has been examined. At the end of the experiment on subject 1, we filmed velar movements for the utterance of /teNteNteNteN/, and then the fiberscope was pushed further to the posterior pharyngeal wall with its tip bent downward. There, we could get a good view of the velopharyngeal passage. The subject was then instructed to try to repeat the same type of utterance at the same rate and effort level. Fig. 2 compares two series of successive frames of the two views, from the beginning of the utterance up to the second /t/. Even though the two series pertain to two separate utterances, the time courses as examined by sound spectrograms turned out to be reasonably comparable. When the velum is elevated more than halfway (see series (a), Fig. 2), the velopharyngeal contact is achieved with small lateral openings of the nasal passway (see series (b)). Further elevation of the velum is necessary for the complete closure of the passage.

The velar height and the velopharyngeal distance measured on the film frames in these two comparable series are plotted as functions of time in Fig. 3. The ordinate gives an arbitrary scale for the two measures. The curve with filled circles shows the change of velar height. The ascending slope of the curve corresponds to the velar elevation. The curve with crosses represents the velopharyngeal distance. The descending curve shows the closing process of the velopharyngeal port. The minimal velar height to achieve the velopharyngeal contact (velopharyngeal distance zero) was approximately 3.7 on this arbitrary scale for subject 1.

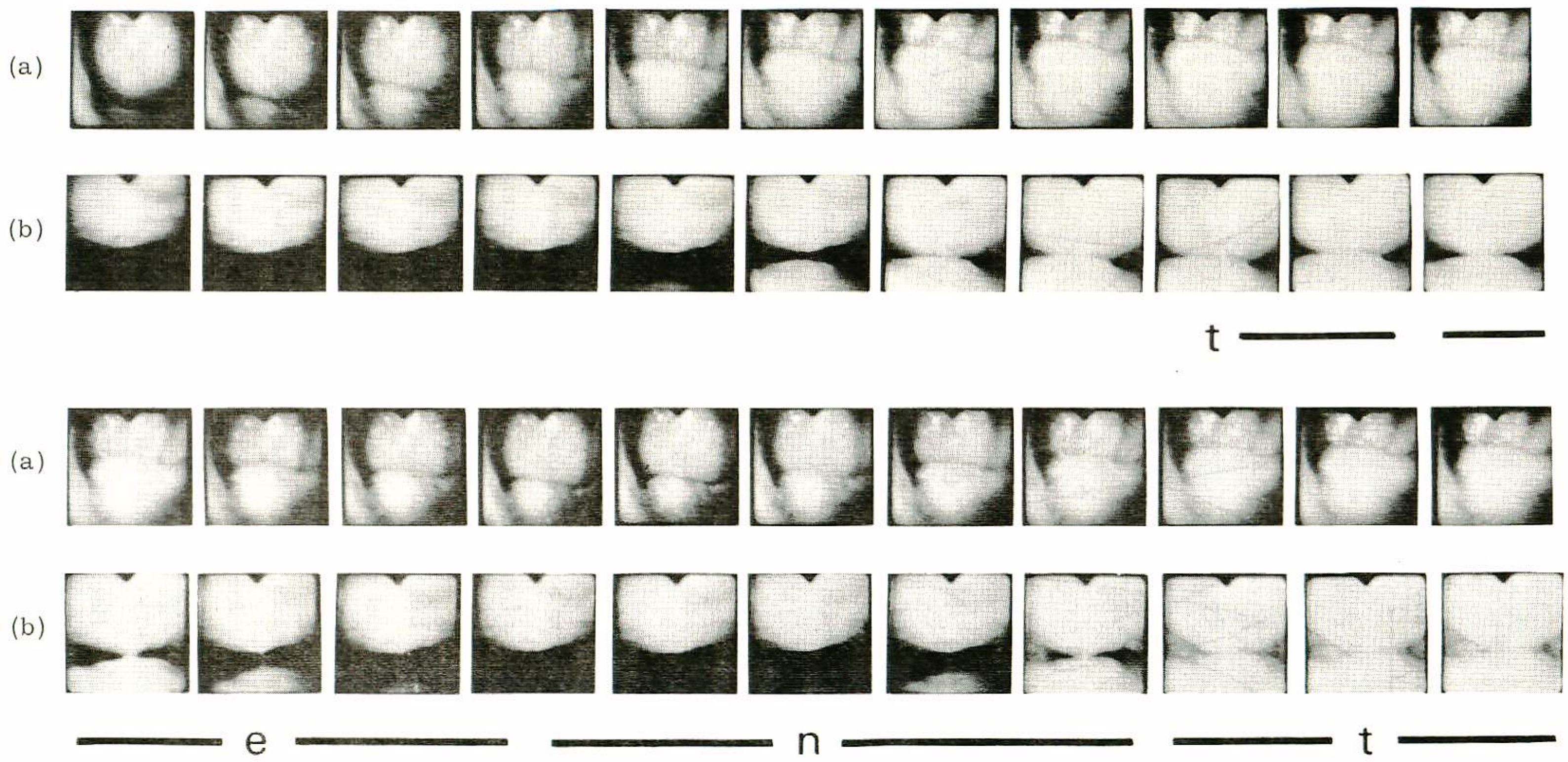

Fig. 2. Successive frames of the velar view (a), and of the view of the velopharyngeal passage (b) for the first four phonemes in the utterance /teNteNteNteN/, on subject 1.

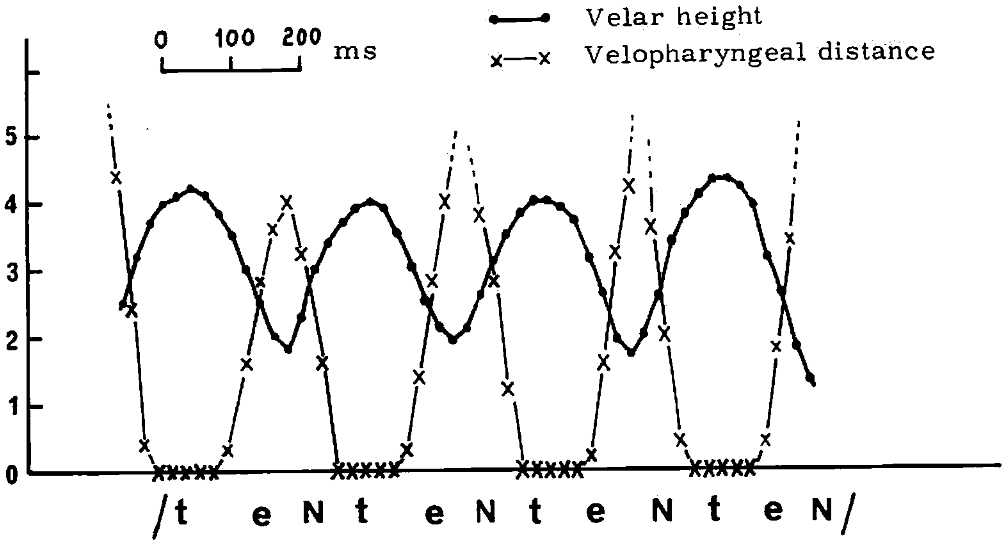

Fig. 3. Velar height and velopharyngeal distance (subject 1). The ordinate is given an arbitrary linear scale.

2. Nonsense Syllables

i. /aiueoaiueoa/

Velar height for articulations of five Japanese vowels was examined at the steady state of each vowel. The vowel sequence /aiueoaiueoa/ was pronounced: a) in a continuous utterance without voice secession and b) with pauses between adjacent vowels. The relation between the observed velar height measure and the velar opening as discussed above was used in interpreting their data when necessary.

Among the vowels, the velar level showed the highest value for /i/ and /u/, the lowest for /a/, and intermediate values for /e/ and /o/. Fig. 4 exemplifies a velar height curve for subject 1. Short vertical bars across this curve indicate the voice onset or offset as estimated in the spectrogram. Because of little scatter in successive data points, a smooth curve was obtained simply by connecting the plotted points one after another.

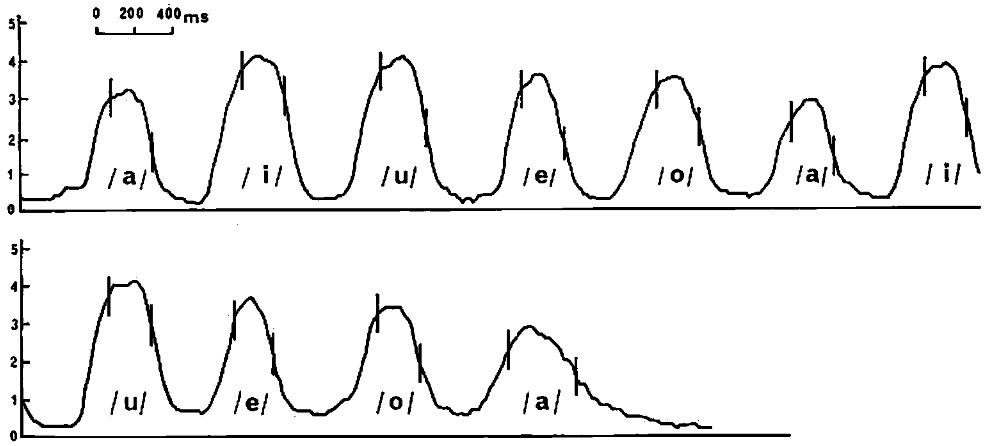

Fig. 4. Velum height of the vowel sequence /a, i, u, e, o, a, i, u, e, o, a/ on subject 1.

It should be pointed out that the accurate timing of the beginning of velar elevation from the rest position cannot be determined on the film frame, because the posterior part of the velum was not visible for the rest position.

From Fig. 4 the following points are noted:

a) Velar elevation starts in advance of voice onset.

b) Velar elevation reaches its peak during the individual vowel.

c) The velum is more elevated at voice onset than at voice offset.

d) At voice offset the velopharyngeal port is presumably open even in the case of /i/ or /u/.

e) Moreover, peak values for /e/, /o/ and /a/ seem to indicate some opening of the velopharyngeal port even for the highest position of the velum.

Concerning the velar elevation for vowels, many authors have reported results obtained cineradiographically[2, 3)] or electromyographically.[2, 7, 8)] The velar elevation has been estimated by the activities of the palatal levator muscle. Iwashita[7] did not prove the difference of the EMG activities between the five Japanese vowels. Most of the other investigators, however, noted that the extent of the velar elevation seemed to vary according to the tongue position. Lubker et al.[8] noticed that, with position and context held constant, levator activity for the close vowel /i/ was greater than for the open vowel /a/. From Fritzell the means of the activities for high vowel /i/ and /u/ were exactly the same, markedly higher than the means for the low vowels /æ/ and /a/, and the means for those two were approximately the same.

The results obtained here are consistent with Fritzell[2] and Umeno.[3] The former stated that there was a significant difference in the velar height between high and low vowels but not between front and back vowels.

ii. /CVCVCVCV/ (C=t, s, d, z. V=e)

In Fig. 5 (a), one example from subject 1 is shown as a velum height of /tetetete/. Short vertical bars across the curve of velum height indicate the articulatory closure or release of [t] as estimated on the spectrograph.

Fig. 5 (a) shows some difference in velum height between [t] and [e]. However, this consonant-vowel difference was not consistent for other samples either for subject 1 or 2. The highest velar elevation tends to occur for utterance initial consonants, as seen in Fig. 5 (a). A general tendency that articulatory gestures are intensified in the utterance initial position has been observed also for tongue movements in palatographic studies.[9] During the utterance final vowel segment, there was observed a significant lowering of the velum.

Our data obtained in this study did not reveal any consistent difference in velum height, either between stop and fricative or between voiced and voiceless consonants.

Differences in the degree of velar elevation between different consonants have been reported and discussed by many investigators. For example, Umeno[3] found in his cineradiographic studies the greatest elevation for voiced stop and fricative consonants, and velum height was lower for voiceless stops and still lower for voiceless fricatives.

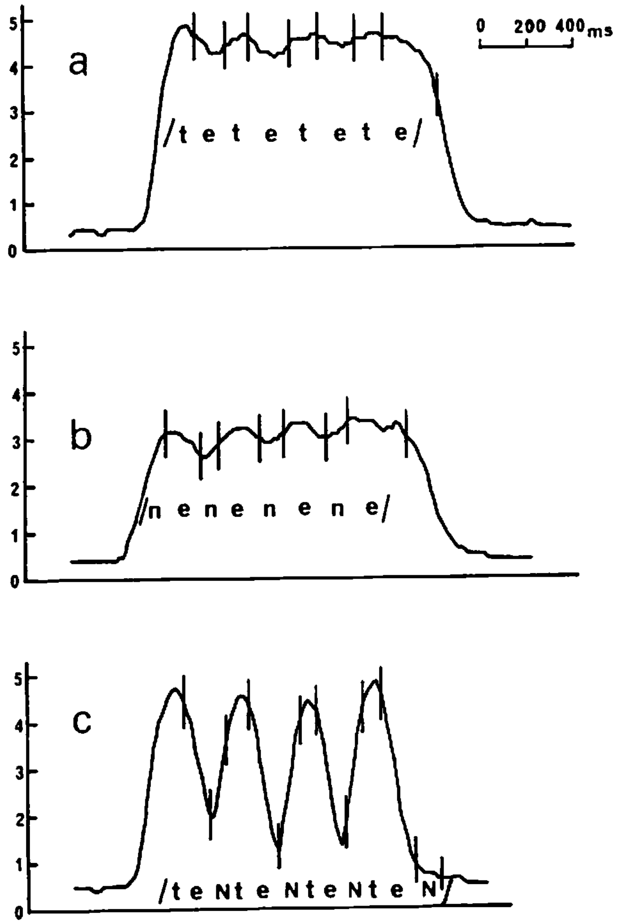

Fig. 5. Velum height of three nonsense syllables:
(a) /tetetete/
(b) /nenenene/
(c) /teNteNteNteN/
on subject 1.

Iwashita[7] ranked the EMG activities of the palatal levator muscle for voiceless stops, voiced stops, and fricatives in that order from high to low. Lubker[8] inferred a tighter velopharyngeal closure for voiceless consonants than for voiced counterparts. Subtelny,[5] however, reported no significant difference between voiced and voiceless consonants. Berti[6] claimed a possibility of greater elevation of the velum for voiced consonants to increase the pharyngeal cavity size as compared to the voiceless counterparts.

Our results at this point agree with Subtelny's, although further study with more data is necessary for a firm conclusion.

iii. /CVCVCVCV/ (C=n V=e)

Velum height of /nenenene/ is shown in Fig. 5 (b). Dips in the curve clearly indicate lowering of the velum for /n/. Elevation of the velum is seen in correspondence to the occurrence of each vowel segment. It should be noted, however, that the velar height for [e] between [n] is considerably lower than for either the same vowel placed between [t] (Fig. 5 (a)) or in a sequence of vowel utterances (Fig. 4).

iv. /CVNCVNCVNCVN/ (C=t V=e N-syllable final nasal)

Velum height of /teNteNteNteN/ in Fig. 5 (c) shows a regular oscillatory movement of the velum in correspondence to the rhythmic repetition of a syllable [teN]. In each cycle of the velar movement, there is a peak of elevation for [t] followed by a marked descent for [N]. The peak values for subject 1 indicate, as mentioned before, the presence of the complete velopharyngeal closure for all [t]'s.

The descent of the velum for [N] is far greater than that for the syllable initial [n] for /nenenene/, although the syllable final nasal is often said to be phonetically [n] when it is followed by [t], [d], or [n] in Japanese. There is no apparent difference in the duration of the nasal segments between the two series, since the mean of the three measured durations is 90 msec for [n] and 80 msec for [N]. The contextual effect, presence or absence of [t] in the syllable, may lead us to expect less lowering of the velum for [tentententen] than for [nenenene], which is contrary to the result actually obtained.

Our preliminary data, therefore, seem to imply an inherent difference between the syllable initial and syllable final nasal consonants, with a greater nasalization for the latter in Japanese. In fact, in Japanese the syllable final nasal /N/ has some special features, which syllable initial ones, /n/ and /m/, do not possess. The phoneme /N/ has a time segment of utterance corresponding to one mora and the specification of the place of lingual or labial articulation for the segment seems entirely dependent on that of the following phoneme.

Velar position varies in time to a great extent during vowel segments as well as during nasal murmurs. This characteristic variation is apparently due to an anticipatory coarticulation of the velar gestures. Lowering of the velum for /N/ starts before the release of /t/, and elevation for /t/ already at the beginning of the preceding nasal segment.

Another finding to be mentioned in this series of utterances is that each peak showed a symmetric pattern of ascending and discending curves. This implies that the speed of velar elevation is almost equal to that of velar lowering.

3. Meaningful Sentences

The selected examples of the curves of velum height shown in Fig. 6 clearly reveal a descent of the velum for the nasal sounds and an elevation for /t/, /d/, /s/, and /z/. Measurement of the velar height showed no significant difference between voiceless and voiced sounds, nor between stop and fricative sounds.

Comparison of /seN'ee/ samples (Fig. 6 (d)) and /deenee/ samples (Fig. 6 (c)) reveals again a greater degree of the velar lowering for the syllable final /N/ than for syllable initial /n/.

Concerning the sequence of /----NseN----/ in /teNseN/ (Fig. 6 (h)), /deNseN/, or /heNseN/, there is a peak for /s/ between the dips for the two /N/'s. The peak values indicate a complete velopharyngeal closure for /s/ on subject 1. A similar result was obtained for /t/ in the same context, as shown in /seNteN/ (Fig. 6 (i)).

These results are consistent with those obtained for nonsense syllables.

4. Coarticulatory Effect in the Velar Movements

For the phoneme sequences, such as /seN/, /teN/, /seen/, /deen/, /Need/ /Ns/ and /Nd/, velar height varies to a great extent during the vowel segment or during the nasal segment. This suggests that the coarticulatory (smoothing) effect of the velar movements for the transition from high velum (stop and fricative) sounds to low velum (nasal), or the reverse, takes place

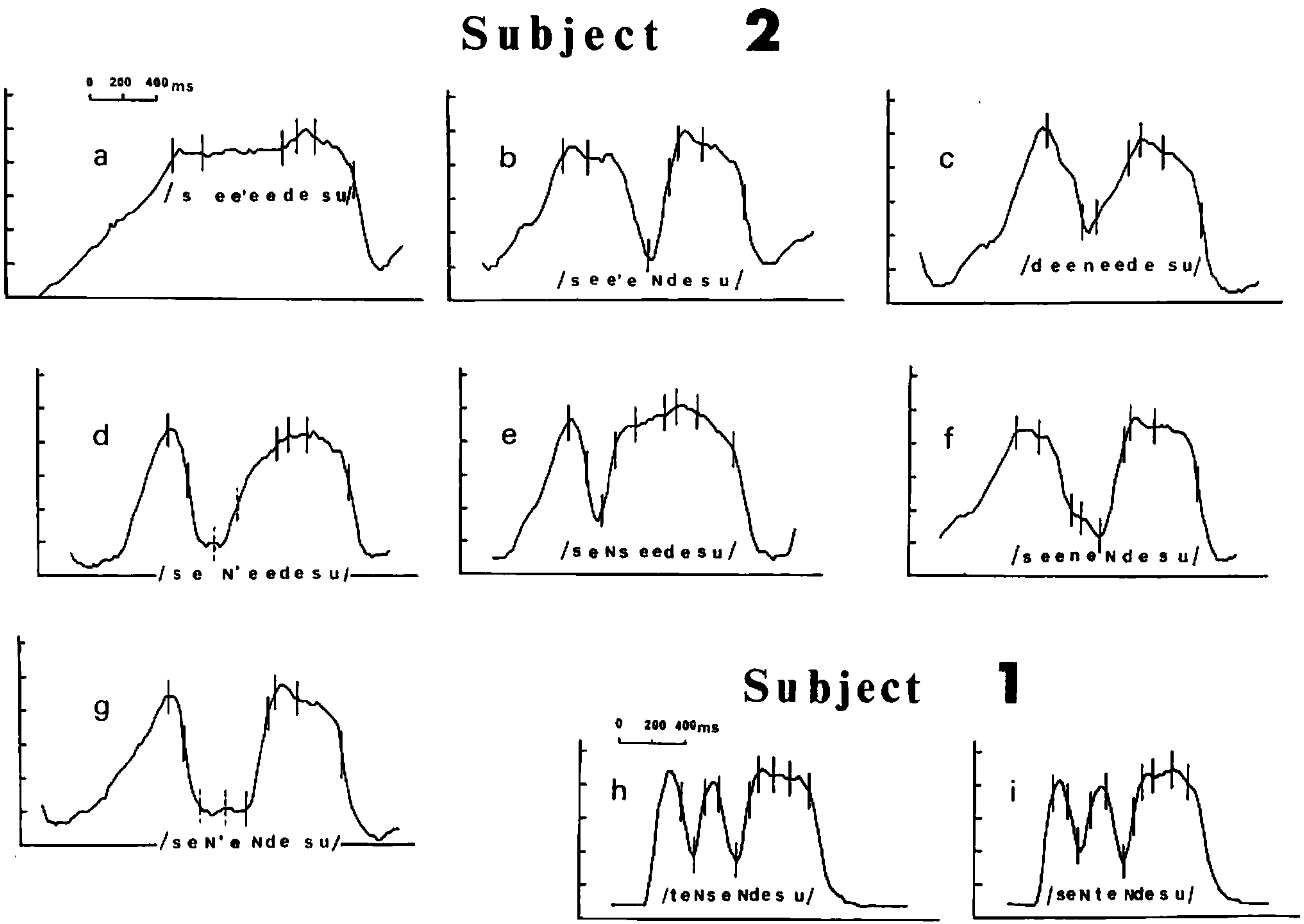

Fig. 6. Selected examples of the velum height (a): /see'ee/, (b): /see'eN/, (c): /deenee/, (d): /seN'ee/, (e): /seNsee/, (f): /seeneN/, and (g): /seN'eN/ on subject 2. (h): /teNseN/ and (i): /seNteN/ on subject 1.

during interposed vowel as well as during the nasal sound segment. A similar smoothing effect was also seen for the nonsense /teNteNteNteN/ series.

Apart from this transitional smoothing effect, the velar height for the vowel itself is very inconsistent. The velum keeps a high position when the vowel is placed between fricative or stop consonants, while it stays as low as the nasals when the same vowel is in between nasal sounds, as shown in /seN'eN/ (Fig. 6 (g)). Our results confirm Moll and Daniloff's suggestions of "unspecified" velar position for the vowel in English.

In /seeneN/ (Fig. 6 (f)) the velar lowering for /n/ starts already at a time point near the /s/-/e/ boundary. In other words, the velar opening coarticulates over as many as two vowel phonemes or morae preceding the nasal. This articulatory behavior can be well explained by Moll's and Daniloff's model.

On the other hand, in all of /see'eN/ samples the velar lowering is not initiated near the beginning of the first vowel in the sequence. Actually the velum seems to keep its high level for a certain time period after the onset of the first vowel. This does not apply to Moll and Daniloff's model. In the case of /see'eN/, there are three vowels (or morae) interposed between high velum and low velum sounds. Furthermore, there is a syllable boundary in the vowel sequences.

Anticipatory (coarticulatory) velar lowering seems to be influenced by either the number of the interposed vowel phonemes (or morae) or the presence of syllable boundary, or both. In the present study, however, the prepared speech samples are insufficient for examination of these factors in comparison with the ceteris paribus. Further study with more systematically controlled speech samples is necessary.

Coarticulatory velar lowering towards the nasal is seen in some contexts, even during /s/ which may be taken as a high velum consonant. In all of /seN----/ sequences, such as /seN'ee/, /seNsee/ and /seN'eN/ (Fig. 6 (d) (e) (g)), the velum begins to lower consistently before the onset of the following vowel. This is not seen in the /seeneN/ samples. The reasons for such differences seem also related to the factors influencing the coarticulatory velar movement above-mentioned.

The velar lowering effect is observed, to some extent, also for /s/'s in

/--Ns--/ series, such as in /seNsee/, /zeNsee/ and /neNsee/. In these samples the velar position is slightly but consistently lower at the beginning of its frication for [s] than its highest value during [s].

References

1) M. Sawashima and T. Ushijima, "Use of the Fiberscope in Speech Research," Annual Bulletin (Research Institute of Logopedics and Phoniatrics, University of Tokyo) No. 5, 25-34 (1971).

2) B. Fritzell, "The Velopharyngeal Muscles in Speech, An Electromyographic and Cineradiographic Study," Acta Oto-laryng. Suppl. 250 (1969).

3) M. Umeno, "A Cineradiographic Study of the Movements of the Soft Palate and Tongue Root during Speech, "Pract. Otol. Kyoto Vol. 55, No. 8, 305-336 (1962) (in Japanese).

4) K. Moll and R. Daniloff, "Investigation of the Timing of Velar Movements during Speech," J. Acoust. Soc. Am. Vol. 50, 678-684 (1971).

5) J. Subtelny, G. Kho, and R. McCormack, "Multidimensional Analysis of Bilabial Stop and Nasal Consonants-Cineradiographic and Pressure-Flow Analysis," Cleft Palate Journal Vol. 6, 263-289 (1969).

6) F. Berti, "The Velopharyngeal Mechanism: An Electromyographic Study, A Preliminary Report," Status Report on Speech Research (Haskins Laboratories) SR-25/26, 117-129 (1971).

7) A. Iwashita, "Electromyographic Study on the Articulation Mechanism of Japanese Speech Sounds," Pract. Otol. Kyoto Vol. 58, No. 10, 712-723 (1965) (in Japanese).

8) J. Lubker, B. Fritzell, and J. Lindqvist, "A Velopharyngeal Function: An Electromyographic Study," Speech Transmission Laboratory Quarterly Progress and Status Report No. 4, 9-20 (1970).

9) I. Fujii, "Phoneme Identification with Dynamic Palatography," Annual Bulletin (Research Institute of Logopedics and Phoniatrics, University of Tokyo) No. 4, 67-73 (1970).