PITCH DEPENDENCY OF VOCAL TRACT TRANSFER FUNCTIONS

*Nobuhiro Miki**,
*Pierre Badin***, *Kenji Takemura**, *Masahiro Kuroda* * and *Yoshihiko Ogawa**
**Faculty of Engineering, Hokkaido University*
** *Institute de la Communication Parlee*

ABSTRACT

The variation of the vocal tract transfer function (VTTF) is discussed for the vowels with pitch variation. Even in the isolated vowels, we show the fluctuation of frequency and bandwidth in the higher formants of the upper third formant; this fluctuation pattern corresponds to small pitch variations. We try to compare the VTTF variation between Japanese and Swedish speakers.

1 INTRODUCTION

It is well known that the accent in Japanese speech is controlled by pitch variation. In the traditional theory of speech production, it has been assumed that the vocal tract transfer function (VTTF) is not changed by pitch variation in steady state articulation. Recently, however, it is known that the position of the vocal fold moves up or down a few cm in pitch control for Japanese vowels even in steady state articulation [3]. If the size of the vocal tract is varied in the order of cm, the corresponding transfer function should be varied in the phonation of steady state vowels. From this point of view, we tried to estimate the VTTF variation from the speech signal by using our short-time speech analysis algorithm (M-algorithm)[1], and we pointed out that this variation is associated with the variation of the 3-dimensional figure of the laryngeal part caused by pitch control [2]. In order to estimate the VTTF from the cut speech signal corresponding accurately to the glottal closure, we record speech signals and EGG signals simultaneously using a DAT. From the estimated results, we show the fluctuation of formant frequencies even in the isolated vowels or the accentuated vowels. Moreover we try to compare the VTTF variation between Japanese and Swedish speakers.

2 DIRECT MEASUREMENT OF VOCAL TRACT TRANSFER FUNCTION

Our interest is to find the relation between the laryngeal movement including the glottal height and the variation of VTTF, and we used the VTTF measurement system of ICP [4]. In this system the shaker on the neck is driven by a white noise sequence; the response at the lips is picked up by a microphone, and is digitized by a PC. We compare the difference in the estimated VTTFs of the vowel under the two conditions of glottal control;

(a): The subject utters a high pitch vowel, closes the glottis for few seconds, utters again with low pitch, and then closes the glottis. The subject keeps the articulation as neary the same as possible during this measurement. The

VTTF is measured in two parts of the glottal closure, and the difference is compared for high and low pitch.

(b): The subject utters a vowel, closes the glottis for few seconds, and then relaxes the glottis. The VTTF is measured in the closure part and the relaxed part.

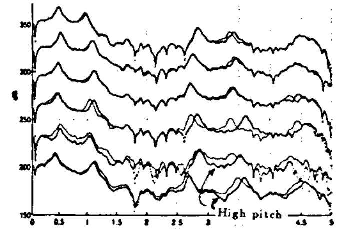

Fig.1 Results for experiment (a)

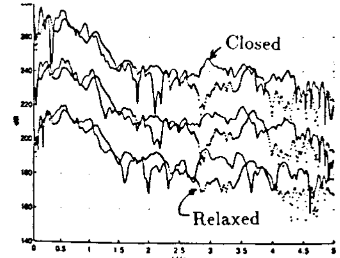

Fig.2 Results for experiment (b)

The experimental results are shown in Fig.1 and 2. As in Fig.1, the difference of VTTFs appears, but is smaller than that of Fig. 2. In the relaxed condition the position of the glottis is lower than in ordinal phonation, and we see the influence of the subglottal system on the VTTF. In the above experiments, the measured VTTFs are average frequency responses in the several hundreds mm seconds under steady state glottal conditions. We see that the higher formants shift to high frequency directions according to height position condition of the glottis in both figures. From these data, we see that the VTTF (formants) can be varied even in the same vowel by pitch control.

3 EVALUATION OF THE VARIATION OF VTTF BY USING A MODEL

In order to evaluate the formant shift with the movement of glottal position, we compute the two kinds of VTTF for high or low pitch by using the vocal tract model and the 3-D data of the vocal tract from MRI data. The difference of VTTF is shown in Fig. 3; the glottis moves up 0.5 cm and the laryngeal part near the glottis shrinks about 15 % . From this figure we can conclude that the formant shift is larger in the higher formants, but the shift frequency is small.

We can suppose that during utterance the glottal movement is from a relaxed position to an upper position, and back again to the relaxed position. Since this small movement influences the VTTF and gives a dynamic shift on the formants, we need to evaluate this dynamic shift by using a short-time analysis algorithm.

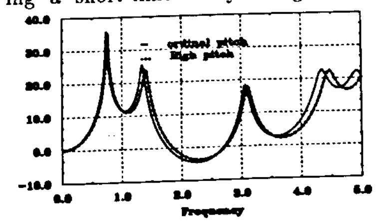

Fig.3 VTTFs computed by the model

4 ESTIMATION OF THE VARIATION OF VTTF IN ISOLATED VOWELS

Using a detection algorithm of peaks of the EGG signal, we obtain a short-time interval data cut from the closed phase of vowels. Since the obtained data have discontinuous points, the data are mapped to continuous waveforms with the Fejer kernel, and are analyzed to find accurate VTTFs (formants) by using the M-algorithm [1]. As in the first experiment, the subject uttered a vowel with pitch control from low to high frequency during 400 - 600 mm sec. In each glottal closure, the VTTF was estimated with the M-algorithm. Fig.4 shows the difference of VTTF of /a/ uttered by a Japanese speaker, and we see that almost all the formants shift from low to high frequency according to the pitch frequency. In Fig.5, the difference is shown for a Swedish speaker; the frequency shift of formants is smaller than in Japanese. It may be caused from different articulation for /a/.

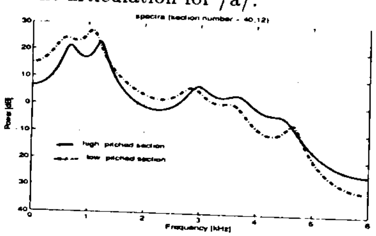

Fig.4 VTTF of Japanese /a/

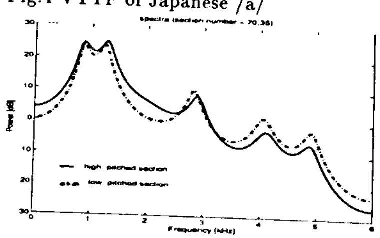

Fig.5 VTTF of Swedish /a/

As the second experiment, we estimate the formant movement for the Japanese vowel /a/ with pitch control from high to low frequency, and the result is shown in Fig.6 (a) and (b). The estimated formants are not of steady state frequency but of variable pattern. Around 2 kHz we see the false formant sequence which may be caused by the subglottal coupling (incomplete closure), and in this duration the 5th formant of the upper 4kHz range is blurred. The estimated formants vary with pitch frequency, and the subglottal coupling appears in some durations. From these experiments, we see that the VTTF can be varied with pitch frequency even in the isolated vowels.

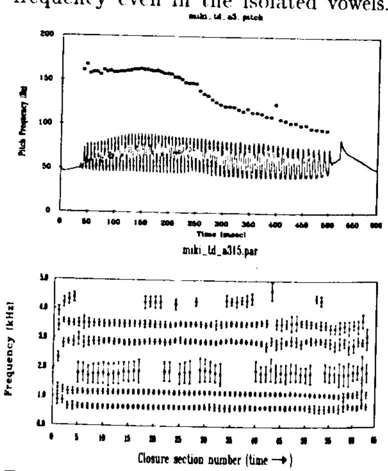

Fig.6 Pitch frequency and traced formants

5 ESTIMATION OF THE VARIATION OF VTTF IN VCV

It is interesting to compare the two VTTF for vowels in the sequence VCV (Vowel Consonant Vowel). Here we show the estimated VTTFs for the two /a/s in /aká/ uttered by Japanese and Swedish. Although the these VTTFs are influenced by the articulation of /k/,we see the formant shift from the first vowel to the second vowel; the formant shift for Japanese is larger than for Swedish, and the influence of /k/ for Japanese is also larger.

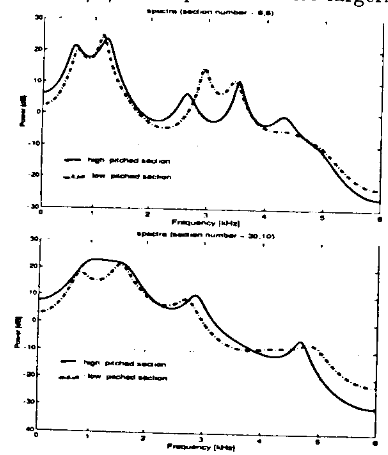

Fig.7 VTTF for /a/ in /aká/

In Fig.8, the estimated formant traces are shown for the second vowel /a/ in the VCV /áka/ (upper graph) and /aká/ (lower graph). The upper trace corresponds to the case of low pitch frequency, and the lower trace is high. From these experiments, we see that the pattern of the formant trajectories is different in the two graphs, and the formants are influenced much by the pitch frequency or the accent. The fluctuation of formant trajectories is larger for the low pitch (non-accent) than the high pitch (accent) vowel.

As our conclusion, since the pitch accent causes movement of the glottal position even in the same articulation effort for the vowel, the VTTF or formant can be shifted by accent.

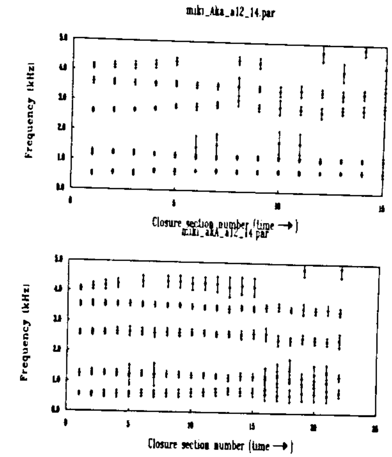

Fig.8 Traced formants for /a/ part in the second vowel in /áka/ (upper) and /aká/ (lower)

Acknowledgment We are grateful to Dr. Y. Pham Thi Ngoc, Dr. I. Karlsson, and Dr. C.H. Shadle for there help in our data aquisition.

References

[1] Miki N., Takemura K., and Nagai N., "A Short-time Speech Analysis Method with Mapping Using the Fejer Kernel," IEICE Trans. Fundamentals, E77-A, pp.792-799 (1994)

[2] Miki N., Badin P., Pham Thi Ngoc Y., and Ogawa Y. (1994)," Vocal Tract Model and 3-Dimensional Effect of Articulation," ICSLP94, 1,pp167-170.

[3] Hirai H. and Honda K., "Analysis of magnetic resonance images on the physiological mechanisms of fundamental frequency control," J. Acoust. Soc. Jpn. vol.50, 4,pp.296-304 (1994) (in Japanese)

[4] Badin P.(1991), "Fricative consonants: acoustic and X-ray measurements," J. Phonetics, 19, pp397-408