A stucy on estimation of acoustic characteristics of three-dimensional vocal tract shapes by Finite Element Method

Tomiyama Yoshihide

School of Information Science,
Japan Advanced Institute of Science and Technology

February 15, 2000

Keywords: Finite Element Method (FEM), 3-D Vocal Tract, Formant Frequency, Misarticulation.

1 Introduction

The purpose of this paper is to find handholds about relationship between vocal tracts and their sound spectra by simulating the spectra from the vocal tract data using Finite Element Method (FEM). Each sound spectrum is different together with speaker even when the same phoneme is uttered. It is important to find the relationship between the vocal tracts and the sound spectra.

Figure1 shows two different shapes of the vocal tracts, normal and abnormal vocal tracts when uttering Japanese vowel /i/. The cross section of the abnormal vocal tract is more intricate than the normal one. Therefore, it is necessary that we consider influence of the three-dimensionally intricate shape on the spectrum. FEM is an effective analysis method to simulate such intricate shape.

In this work, FE models are constructed from one normal and four abnormal vocal tract data uttering Japanese vowel /i/, and the spectra are simulated using the FE model. The purpose is to find the relationship between the vocal tract and the sound spectrum.

2 Relationship between element division and calculation precision of FEM

Straight acoustic tubes whose element size is different on each FE model simulated to follow relationship between the element division and the calculation precision using

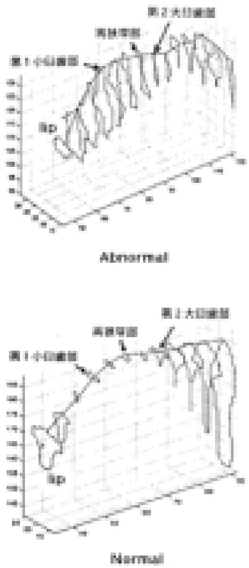

Figure 1: Vocal Tract Shape of Nomal and Abnormal

FEM analysis. The length of the straight acoustic tube is 17 cm, and the sound velocity is 340 m/s. In this case, the theoretical values of the formant frequencies are 500,1500,2500...[Hz]. The simulated results show that the simulated values of formant frequencies using FE model whose element sizes 0.50 and 0.25 cm are almost same as the theoretical values in the range below 5000 Hz.

3 Relationship between vocal tract around the glottis and spectrum

When the sound spectrum of the abnormal vocal tract uttering Japanese vowel /i/ was compared with normal one, formant frequencies of the sound spectra were greatly different. Whereas, when the abnormal vocal tract uttering Japanese vowel /i/ was compared with normal one, structural characteristics of vocal tracts around tongue positions were greatly different. Therefore, it is considered that the vocal tract around tongue position affected the sound spectrum from the abnormal vocal tract. Then,area functions around the glottis and vocal tract lengths are standardized in the normal and abnormal vocal tract data and relationship between vocal tract shapes around tongue position and sound spectra are discussed.

Table 1: Results of Speach Analysis uttering Japanese vowel /i/

	F1	F2	F3
Normal-1	312	2421	3398
Abnormal-1	312	1992	2343
Abnormal-2	351	2304	3281
Abnormal-3	312	1640	2734
Abnormal-4	312	1953	2578

4 Relationship between vocal tract around the tongue and spectrum

There were area function data of vocal tracts around tongue positions of normal (Normal-1) and abnormal (Abnormal-1, 2, 3, 4) from MR images. Additionally, there were analyzed results of speech uttered Japanese vowel /i/ (Table.1). However, there were no data of vocal tracts around glottis and no data of vocal tract lengths. Then, the proposed models adopted vocal tract shape around the glottis and the vocal tract length . In this work, these proposed models were one-dimensional models, which only the area functions on the cross-section are taken into consideration.

As for Normal-1, the structural characteristics of the vocal tract were that the narrow field around the tongue was formed enough from the back tongue to the front tongue. Results of speech analysis were that the first formant frequency (F1) and the second (F2) separated each other and that F1 was lower and F2 was higher. Then, the aspect of FEM analysis result was the same as the result of speech analysis.

As for Abnormal-1, the structural characteristics of the vocal tract were that the narrow field around the tongue was formed at the front tongue and not formed at the back tongue comparing to Normal-1. Results of speech analysis were that F2 and F3 were lower than those of Normal-1 and approached each other. Then, the aspect of FEM analysis result was the same as the result of speech analysis.

As for Abnormal-2, the FEM analysis result cannot approximate the result of speech analysis. In this case, two factors were considered: (1) the area function of the proposed model around the glottis differed from the actual vocal tract function because of no data around glottis and no data of the vocal tract length, (2) three-dimensional characteristics affected to the spectrum because the actual vocal tract shape on cross-section was more intricate than other one (Show Fig.2).

As for Abnormal-3, the structural characteristics of the vocal tract were that the narrow field around the tongue was formed at the back tongue and not formed at the front tongue comparing to Normal-1. Results of speech analysis were that F2 and F3 separated each other. Then, the aspect of FEM analysis result was the same as the result of speech analysis.

As for Abnormal-4, results of speech analysis were that F2 and F3 approached each other. On the contrary, the results of FEM analysis were that F2 and F3 separated each other. The same factors as Abnormal-2 shoud be related.

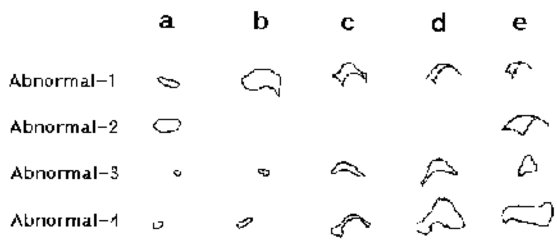

Figure 2: Vocal Tract Shape of Abnormal (a:front tongue - e:back tongue)

By comparing the vocal tracts around the tongue of Abnormal-1 and Abnormal-3, the narrow field positions are different. By comparing the speech analysis results of Abnormal-1 and Abnormal-3, formant frequency positions were different. Moreover, the FEM analysis results of Abnormal-1 and Abnormal-3 can approximate the results of speech analysis. Therefore, the results indicate that "the narrow field" greatly influenced on the spectrum.

5 The modification of the model

When the vocal tract shape on the cross-section around the tongue was intricate like Abnormal-2, the one-dimensional model is not valid. It is considered that three-dimensional characteristics affect the spectrum. The proposed model was constructed, in which three-dimensional characteristics like the actual vocal tract shape were taken into consideration (Fig.3 shows this model).

In these results of FEM analyses, the formant frequencies of FEM analysis results could not approach those of speech analysis. However, when the intricate models like these proposed models was analyzed, three-dimensional shapes of vocal tracts influenced on formant frequencies of the FEM analysis results. Therefore, when the intricate shape of vocal tract model is analyzed, it is necessary that we do not take only area functions on cross-section but also three-dimensional vocal tract shapes into consideration.

6 Conclusion

The results of experiments conclude that:

1. Vocal tract shapes around tongue position uttering Japanese vowel /i/ influence on formant frequency positions than those around glottis.

2. The narrow field position in the vocal tract around tongue position is important to utter Japanese vowel /i/.

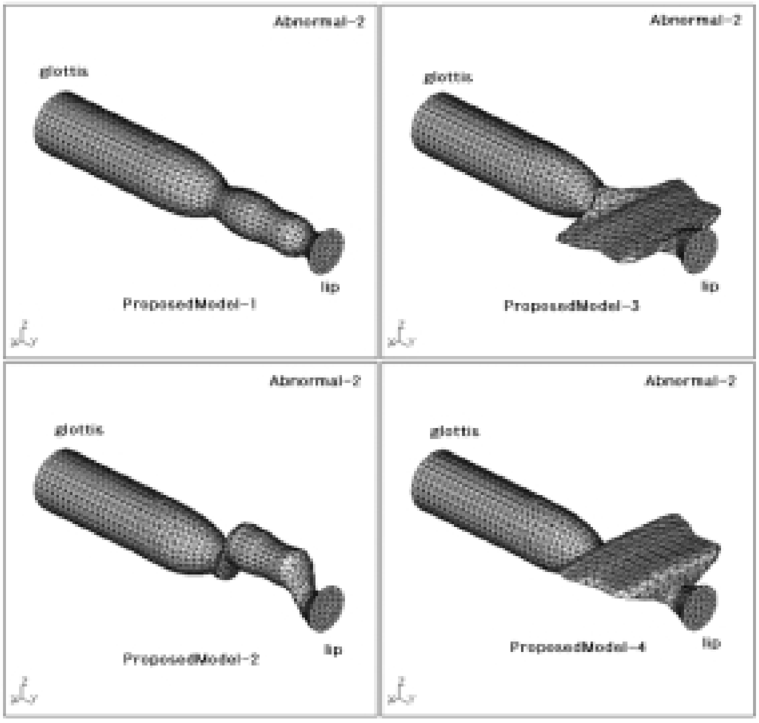

Figure 3: Proposed Model (Abnomal-2)

3. When vocal tract shapes around tongue are intricate, only the area functions cannot simulate that formant frequencies is affected by three-dimensional characteristics.

This paper shows that relationship between vocal tracts and their sound spectra is led by this simulation when aspect of FEM analysis results uttered Japanese vowel /i/ are same as speech analysis results. Moreover, it shows that sound spectra are influenced by three-dimensional characteristics of intricate shapes when FEM analysis results cannot approximate to speech analysis results.