

DEVELOPING MECHANICAL MODELS OF THE HUMAN VOCAL TRACT FOR EDUCATION IN SPEECH SCIENCE

PACS: 43.10.Sv

Saika, Noriko; Maeda, Eri; Usuki, Nobuyuki; Arai, Takayuki; Murahara, Yuji

Dept. of Electrical and Electronics Eng., Sophia University

7-1 Kioi-cho, Chiyoda-ku

Tokyo, 102-8554

Japan

Tel: +81-3-3238-3417

Fax: +81-3-3238-3321

E-mail: s-noriko@splab.ee.sophia.ac.jp

ABSTRACT

Formerly, Arai (2001) proposed two types of mechanical models of human vocal tract for education in speech science [1]. In the first, several acrylic plates with holes formed an acoustic tube. In the second, a hollowed out cylinder created an acoustic resonator. Vowels were produced for both models with an electrolarynx or other sound source. In the current study, we newly made four cylinder models that were different in step sizes of approximation and compared their characteristics. We then developed another new model for nasalized vowels by attaching a side branch to the oral cavity. By means of experiments, we confirmed that these models are powerful tools for education in speech science.

INTRODUCTION

By using mechanical models of vocal tract, we can associate the configuration of the vocal cavity with the acoustic output easily. Therefore, many vocal-tract models are proposed until now. Chiba and Kajiyama (1941) measured area functions of the human vocal cavity and produced artificial vowels by using the mechanical vocal-tract models [2]. They then examined characteristics of the vocal-tract models and the acoustic theory. Umeda and Teranishi (1965) developed mechanical vocal-tract models different in the length of the vocal tract by sex and age [3]. They also made a nasal cavity attaching to the models.

From an educational point of view, Arai [1] and Arai et al. [4] developed two types of the mechanical vocal tract model of the human vocal cavity. One model consists of several acrylic plates with holes formed an acoustic tube. Another one consists of a hollowed out cylinder created an acoustic resonator. They used their models during a lecture in a speech science classroom and found that students were able to experience the acoustic phenomena of the vowel production.

In this study, we extend Arai's models and 1) investigate how the difference in step sizes of approximation affects formants for the cylinder type; and 2) develop a new model for nasalized vowels by attaching a side branch to the oral cavity.

FOUR CYLINDER MODELS

There are two types of mechanical models of human vocal tract of Japanese vowel /a/ designed by Arai (2001) [1]. One is several acrylic plates, the other is hollowed out cylinder. For the cylinder type, we newly made four models that were different in step sizes of step-wise approximation of Arai's model. The step-sizes are 80 mm (C80), 10 mm (C10), 5 mm (C5), and polygonal line approximation (CL). The length of all vocal tracts was 160 mm each.



Fig. 1.- Four Cylinder Models

We recorded the acoustic outputs of CL, C5, C10 and C80 using the electrolarynx and measured their spectra. Table 1 shows the formant frequencies (F1-F4) obtained from the measured spectra. As a result, there were no huge differences in formant frequencies up to F4 among four models. According to an informal listening test, C80 sounded slightly differently from others, although all of models produced /a/-like sounds.

Simulation

First, the electric equivalent four-pole network were considered for each segment by each step of four cylinder models based on Flanagan (1965) [5]. Each of R, L, C, G can be computed by the following formulas:

$$R = \frac{S}{A^2} \sqrt{\frac{\omega r m}{2}}; \quad L = \frac{r}{A}; \quad C = \frac{A}{r^2}; \quad G = S \frac{h-1}{r^2} \sqrt{\frac{I \omega}{2 c_p r}};$$

where A is tube area, S is diameter, r is air density, c is sound velocity, m is viscosity coefficient, I is coefficient of heat condition, h is adiabatic constant, and c_p is specific heat of air at constant pressure. Then we simulated the spectra of the acoustic output for each model. Formant frequencies (F1-F4) obtained from the spectra are shown in Table 1. CL, C5, and C10 have nearly close values on F1, F2 and F3. However, C80 has slightly higher F1 and very low F4 comparing to the others. Because crucial formants for vowel quality are F1, F2 and F3, the gap of F4 frequencies does not contribute much for perceptual difference in quality.

Table 1.- Formant frequencies from measurement and simulation (in Hz).

	Measurement				Simulation			
	CL	C5	C10	C80	CL	C5	C10	C80
F1	717	655	702	671	740	732	741	866
F2	1138	1123	1154	1123	1300	1283	1324	1343
F3	2636	2729	2682	2636	3040	3058	3002	3073
F4	3025	3136	3150	3103	4460	4407	4345	3549

VOCAL TRACT MODEL FOR NASALIZED VOWELS

A nasalized vowel is a vowel that is nasalized by a process in which the soft palate is lowered, opening the velum, which allows air to go through both the nasal and oral cavities. The nasal tract plays an important role in shaping appropriate nasal spectra. Several investigations have been conducted to nasal acoustic characteristics and results show that nasal cavity changes the spectrum in the vicinity of the low frequency of the transfer function (House and Stevens, 1956 [6]).

Sinus

Takeuchi et al. (1977) and Maeda (1982) estimated the volume of the paranasal cavity and discussed the resonance properties of the nasal cavity with sinus. Masuda (1992) measured sinus and demonstrated acoustic characteristics by obstructing the ostia of the sinus [9]. The demonstration showed that there was a strong pole-zero pair at about 700Hz in the spectra of nasal sounds, although no pole-zero pair appeared in the frequency range when the ostia were obstructed.

Dang and Honda (1994) investigated the relevance of morphological measurements of the nasal and paranasal cavities and the acoustic properties of the human vocal tract [10]. Their results suggested that paranasal cavities play an important role in shaping the spectral characteristics of human nasal sounds.

Design

To confirm the acoustic effects of nasals, we created a new model for nasalized vowels by attaching a side branch to the oral cavity of /a/ based on Arai [1]. The dimensions of the nasal cavity was based on the measurements of human speech organs with sinus by Dang et al. [10], Stevens [11] and Chen [12]. Fig. 2 shows the side view of the model. The materials of this model was acrylic resin.

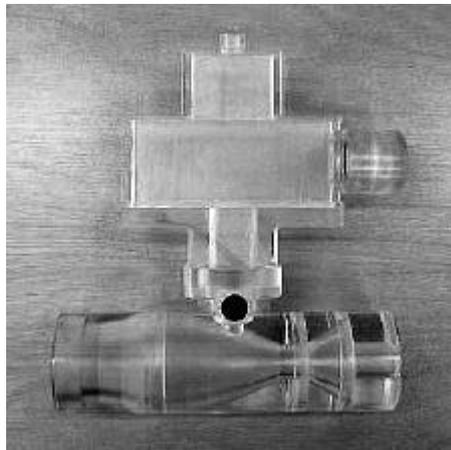


Fig. 2.- Vocal tract model for nasalized vowels

The upper part of Fig. 2 represents the nasal cavity and the bottom represents the oral cavity. For convenience, we created only the one cavity in the nasal cavity. The dimension of the oral cavity is basically the same as the one used in the previous section. The connection between oral and nasal cavities consists of an elastic tube. By tightening up a tension screw, we can press the tube and change the cross sectional area of the opening between oral and nasal cavities. Sinus is removable, so that we can compare the two conditions: with and without the sinus.

Table. 2.- Measured first and second formant frequencies and bandwidths (in Hz)

Coupling	Sinus	F1	B1	F2	B2
No	w/o	609	62	1109	47
Full	w/o	718	250	1125	109
Full	W/	625	187	1140	78

The Effect of Velopharyngeal Coupling

We examined the effect of the velopharyngeal coupling. By adjusting the screw, we changed the cross sectional area between oral and nasal cavities and obtained two degrees of coupling: 1) no coupling, and 2) full coupling. The difference between two acoustic signals with full and no coupling was apparent, and the one with full coupling sounded a nasalized vowel /a/.

Table 2 shows the first and second formant frequencies and bandwidths of the acoustic signals by driving the mechanical models with an electrolarynx. We confirmed that the bandwidth of F1 was increased as well as frequency shifts, when coupled.

The Effect of The Sinus

Next, we examined the effect of the sinus. Sinus is connected to the nasal cavity through a small opening. We recorded the sounds with and without sinus and evaluated them with spectral analysis (Table 2). Although there was no distinct difference when we listened to the acoustic outputs, we observed a change due to sinus in lower frequency range, especially F1. By taking the ratio of spectra with and without sinus, we confirmed the evidence of the sinus resonance in the vicinity of the low frequency region.

CONCLUSIONS

We extended Arai's models [1] as tools for education and developed two types of vocal-tract models for education in speech science. We confirmed that 1) formants are not much affected by the difference in step sizes of approximation for the cylinder type; 2) by attaching a side branch to the oral cavity, we confirmed that nasalized vowel is produced; and 3) by attaching sinus to the nasal cavity of the model, changes take place in the vicinity of the low frequency of the transfer function. Thus, each model has potential as an educational tool in speech science and acoustics.

BIBLIOGRAPHICAL REFERENCES

- [1] T. Arai, "The replication of Chiba and Kajiyama's mechanical models of the human vocal cavity," *J. Phonetic Soc. Japan*, Vol.5, No.2, pp.31-38, 2001.
- [2] T. Chiba and M. Kajiyama, *The Vowel: Its Nature and Structures*, Tokyo-Kaiseikan Pub. Co., Ltd., Tokyo, 1941.
- [3] N. Umeda and R. Teranishi, "Phonemic feature and vowel feature: Synthesis of speech sounds, using an acoustic model of vocal tract," *J. Acoust. Soc. Jpn.*, Vol.22, No.4, pp.195-203, 1966.
- [4] T. Arai, N. Usuki and Y. Murahara, "Prototype of a vocal-tract model for vowel production designed for education in speech science," *Proc. of Eurospeech*, Vol.4. pp.2791-2794, 2001.
- [5] J.L. Flanagan, *Speech Analysis Synthesis and Perception*, Springer-Verlag, pp.23-35, 1965.
- [6] A. House and K. Stevens, "Analog studies of the nasalization of vowels," *J. Speech Hear. Disord.*, Vol.21, No.2, pp.218-232, 1956.
- [7] S. Takeuchi, H. Kasuya and K. Kido, "A study on the effects of nasal and paranasal cavities on the spectra of nasal sounds," *J. Acoust. Soc. Jpn.*, Vol.33, pp.163-172, 1977 (in Japanese).
- [8] S. Maeda, "The role of the sinus cavities in the production of nasal vowels," *Proc. of IEEE ICASSP*, Vol. 2, pp. 911-914, 1982.
- [9] S. Masuda, "Role of the maxillary sinus as a resonant cavity," *J. Otolaryng. Jpn.*, pp.95-70, 1992 (in Japanese).
- [10] J. Dang and K. Honda, "Morphological and acoustical analysis of the nasal and the paranasal cavities," *J. Acoust. Soc. Am.*, Vol.96, No.4, pp.2088-2100, 1994.
- [11] K. Stevens, *Acoustic Phonetics*, MIT Press, Cambridge, MA, 1998.
- [12] M.Y. Chen, "Acoustic correlates of English and French nasalized vowels," *J. Acoust. Soc. Am.*, Vol.102, No.4, pp.2360-2370, 1997.