ACOUSTICAL LETTER

Comparing the characteristics of the plate and cylinder type vocal tract models

Eri Maeda*, Nobuyuki Usuki, Takayuki Arai, Noriko Saika and Yuji Murahara

Department of Electrical and Electronics Engineering, Sophia University, 7–1 Kioi-cho, Chiyoda-ku, Tokyo, 102–8554 Japan

(*Received 1 April 2003, Accepted for publication 20 June 2003*)

Keywords: Human vocal tract, Education in acoustics, Physical model, Vowel production, Education tools

PACS number: 43.10.Sv [DOI: 10.1250/ast.25.64]

1. Introduction

The widespread use of computers along with increased processing speed has brought with it a rapid development in digital signal processing. Yet, even though advances in the digital arena (such as more general accessibility to spectrograms) are valuable prospects, we should not forget the importance of analog systems in education. Analog systems, and more specifically, physical models, are useful for teaching acoustics in the classroom. Previously, we reported on mechanical vocal-tract models we developed for intuitive learning [1–4]. We recently showed these tools to Japanese high-school students who had just finished studying acoustics, and we received positive feedback. Experience shows the tools to be particularly suitable for teaching the mechanism of vowel production [5].

Chiba and Kajiyama measured the area functions of the human vocal cavity and produced artificial vowels by using mechanical vocal tract models in the 1940's [6]. Using Chiba and Kajiyama's measurements of the oral cavity, Arai *et al.* reconstructed their vocal tract models with an acrylic cylinder to make the models suitable as the educational tool [3]. In this model, the radius at every position was approximated with a polygonal line. Their models have an opening at each end corresponding to the mouth and glottis, and are made of transparent materials so the shape of the cavity can be seen.

We have also developed several sets of square acrylic plate models [1,2]. Each plate has a hole in the center, and the holes are various diameters. A set of 16 disks simulates the vocal tract by forming an acoustic tube where the holes line up. Through rearrangement of the plates, students can learn the importance of the position of constriction for vowel production and the perturbation theory [4,7].

In the previous study [8], we made four cylinder models of the Japanese vowel /a/, which are different in step size approximations. The simulations for these 4 cylindrical models showed that the step sizes did not influence the formants of the outputs. In this paper, we compare and contrast characteristics of plate and cylinder models with the same configuration.

2. Comperative experiment

The cylinder models have external diameters of 50 mm and are 160 mm long. The plates in the plate models are made of acrylic, and each plate is 75 × 75 mm square by 10 mm thick. When all 16 plates were placed adjacent to one another, the tube measured 162 mm long. Therefore, there seems to have been a slight variation in the thickness of each plate.

Figure 1 shows vocal tract models for cylinder and plate types. Looking from left to right, the first model is the cylinder model with polygonal approximation [3], next is the plate type, which is followed by the cylinder model that has exactly same configuration as the plate model.

We compare characteristics of the plate type model with the cylinder type when in the configuration for vowel /a/. After conducting a perceptual experiment, both models conclusively produced an /a/, but there were some differences in sound quality which we wanted to investigate further. We examined the spectra of the output of the two models by linear predictive coding (LPC) analysis. Figure 2 shows two spectral envelopes: a heavy line shows the plate model, a fine black line shows the cylinder model, and a fine gray line shows the cylinder type with polygonal approximation.

3. Discussion

Using LPC analysis, we calculated the frequency and bandwidth for each type as follows:

formant frequency

$$F = \frac{\arg z_i}{2\pi \Delta T}\,[\mathrm{Hz}], \tag{1}$$

bandwidth

$$B = -\frac{\log |z_i|}{\pi \Delta T}\,[\mathrm{Hz}], \tag{2}$$

quality factor

$$Q = \frac{F}{B}, \tag{3}$$

where ΔT denotes the sampling period, and z_i is a pole of the auto-regressive (AR) model in LPC analysis.

Table 1 shows the results.

From this table, the formants in the lower frequencies, from F1 to F4, those governing vowel quality, are almost identical, although the plate models have broader bandwidths and lower quality factors. Because of this, the outputs produced from both the plate and cylinder models with the same configuration had nearly the same vowel quality.

This difference seems to result from the loss of acoustic energy when we lined up each plate. The plate type had wider bandwidths compared to the cylinder models. In the higher frequencies, plate models are also different with respect to

*e-mail: m-eri@sophia.ac.jp

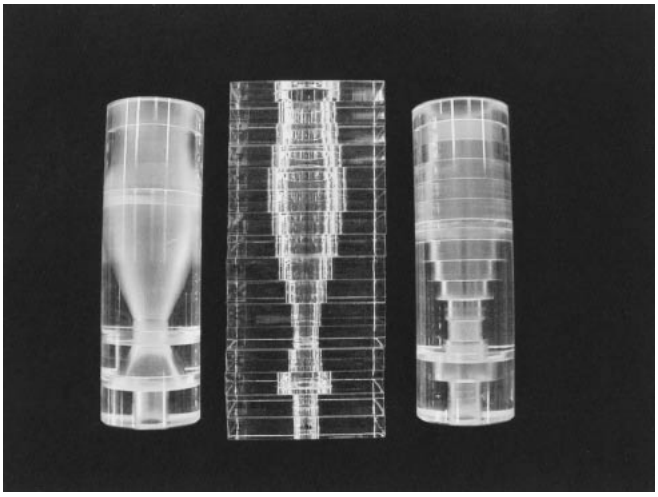

Fig. 1 Vocal Tract models for cylinder and plate models (vowel /a/).

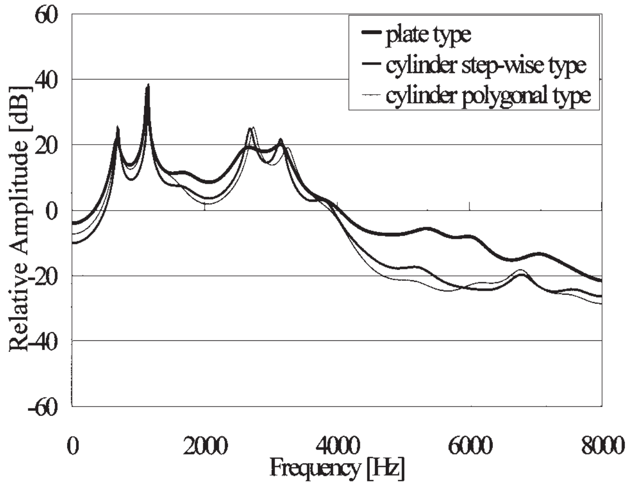

Fig. 2 Spectral envelopes of vowel /a/.

Table 1 The frequency, bandwidth and quality factor for each type.

	Formant [Hz]					Bandwidth [Hz]			Quality factor		
	Cylinder			Plate		Cylinder		Plate	Cylinder		Plate
	polygon	step-wise	(%)		(%)	polygon	step-wise		polygon	step-wise	
F1	675.3	680.1	(0.7)	647.8	(−4.0)	40.6	29.7	84.1	16.6	22.9	7.7
F2	1,127.9	1,141.1	(1.2)	1,107.6	(−1.8)	24.9	12.1	22.3	45.3	94.3	49.7
F3	2,727.1	2,674.7	(−1.9)	2,564.2	(−6.0)	84.7	92.8	361.0	32.2	28.8	7.1
F4	3,256.9	3,146.4	(−3.4)	3,127.6	(−4.0)	159.2	104.8	231.1	20.5	30.0	13.5

Note: The values in parentheses show relative errors from the cylinder models with polygonal approximations.

relative amplitude. We think these are the main factors accounting for the different acoustic impression between model types.

4. Summary

In this paper, we have emphasized the need for analog systems in education, even in a digitally oriented society. We further developed mechanical models and compared the characteristics of our two types of models: plate and cylinder. We confirmed that the vowel quality of both types were almost the same. By spectral analysis, we found that the differences were due to formant bandwidth in the lower frequency range, and relative amplitude in the higher frequencies. The plate type has the greater advantage for education in that the students can arrange them arbitrarily and experiment with them individually. They can also test perturbation theory [2] by placing a plate with a narrow hole in various locations. Because the plate-type models have sufficient characteristics, we concluded that they are useful for education in acoustics.

References

[1] T. Arai, N. Usuki and Y. Murahara, "Prototype of a vocal-tract model for vowel production designed for education in speech science," *Proc. Eurospeech*, Vol. 4, pp. 2791–2794 (2001).
[2] N. Usuki, T. Arai and Y. Murahara, "Prototype of a vocal-tract model for education in acoustics of vowel production," *Proc. Spring Meet. Acoust. Soc. Jpn.*, pp. 399–400 (2001).
[3] T. Arai, "The replication of Chiba and Kajiyama's mechanical models of the human vocal cavity," *J. Phonet. Soc. Jpn.*, **5**(2), 31–38 (2001).
[4] N. Usuki, M. Yoshida, H. A. Alwi, T. Arai, and Y. Murahara, "Usefulness of a mechanical model of the human vocal tract for education in speech sciences: Perturbation theory in vowel production," *Proc. Autumn Meet. Acoust. Soc. Jpn.*, pp. 403–404 (2001).
[5] E. Maeda, T. Arai, N. Saika and Y. Murahara, "Lab experiment using physical models of the human vocal tract for high-school students," *First Pan-American/Iberian Meeting on Acoustics*, Cancun (2002).
[6] T. Chiba and M. Kajiyama, *The Vowel: Its Nature and Structures* (Tokyo-Kaiseikan Pub. Co., Ltd., Tokyo, 1941).
[7] K. N. Stevens, *Acoustic Phonetics* (The MIT Press, Cambridge, 1998).
[8] N. Saika, E. Maeda, N. Usuki, T. Arai and Y. Murahara, "Developing mechanical models of the human vocal tract for education in speech science," *Proc. Forum Acousticum*, Sevilla (2002).