ACOUSTICAL LETTER

Gel-type tongue for a physical model of the human vocal tract as an educational tool in acoustics of speech production

Takayuki Arai*

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

(*Received 17 August 2007, Accepted for publication 3 September 2007*)

Keywords: Gel, Tongue, Vocal tract model, Vowel production, Speech science, Education in acoustics

PACS number: 43.70.Bk, 43.10.Sv [doi:10.1250/ast.29.188]

1. Introduction

Recently, Arai (2001) [1] replicated Chiba and Kajiyama[2]'s mechanical models of the human vocal tract from transparent materials and showed that the models are useful as educational tools. Since then, we have continued to develop educational tools, because we feel that physical models of the human vocal tract should be widely used for education, especially in speech sciences [3–12].

Arai (2001) [1] proposed two types of models: the cylinder-type model with the precise reproduction of the original vocal-tract shapes of Chiba and Kajiyama [2], and the plate-type model, a stepwise approximation of these original shapes. Arai's models are effective tools because they help learners understand the relationship between vocal tract shape and speech output. In both models, a vowel-like sound is produced at the lip end when a sound source excites the glottis end.

To visualize how the vocal tract is positioned in the head, we designed a series of head-shaped models in our education system [3,11,12]. There are two major types of head-shaped models: fixed tongue model and manipulable tongue model. For the both types of the head-shaped models, we had two goals: 1) to make the tongue shape visible from the outside, and 2) to produce actual vowel sounds from the models. We achieved the first goal with transparent acrylic plates, so that the tongue shape is visible from the outside. For the second goal, we obtained a proper area function for each vowel to form a three-dimensional model.

In the "fixed tongue" head-shaped model, the vocal tract shape was fixed, so one had to use different models for different vowels. For the "manipulable tongue" model, a learner can change the vocal tract shape by manipulating the tongue in a single model, and thereby better understand the dynamics of speech production [3,11,12]. The tongues in the models described in [3,11,12], are made of acrylic resin, a bent rubber tube, and felt. In each case, the tongue position may be changed by sliding the tongue between the two acrylic plates in the rest of the head-shaped model.

In this paper, an extended version of the "flexible tongue" model in [3] is proposed (Fig. 1). Unlike other "flexible-tongue" models, such as [13], the tongue in this version is manipulable by hand.

*e-mail: arai@sophia.ac.jp

2. Flexible tongue model by gel-type material

The model shown in Fig. 1 consists of the gel part, the transparent part with acrylic plates, and the non-transparent part with metal (Fig. 2). The gel part corresponds to the tongue, the bottom of the oral cavity, and the anterior pharyngeal wall. The transparent part corresponds to the both sides of the oral cavity. The non-transparent part corresponds to the palate and the posterior pharyngeal wall. The gel material used was a polyethylene-styrene copolymer. We tested two different degrees of the hardness: 4 (slightly softer than the actual tongue) and 15 (slightly harder than the actual tongue) in ASKER-C hardness. We decided to use a hardness of 4 because it is more flexible and easier to manipulate. The dimensions of the vocal tract model are shown in Fig. 3.

The entire body of the tongue is not necessarily reproduced in this model. The reproduced parts are mainly the upper body of the tongue, such as, the tongue tip, the upper part of the tongue body including the tongue dorsum, and the surface of the tongue root. This is due to a result of pursuing the manipulability. Thus, one can change the shape of the tongue manually by manipulating it from the bottom of the tongue, so that the constriction in the vocal tract can easily be moved front-back and high-low directions.

Figure 4 shows pictures when /i/, /a/ and /u/ vowels are produced. In case of vowel /i/, the tongue attaches to the hard palate (Fig. 4, left). There is a semicircular groove (approximately 13 mm in diameter) on the top surface of the tongue. When the tongue is in the high position, this groove with the palate forms a long (approximately 50 mm) narrow constriction, and vowel /i/ is produced. For high-front vowels, the sound wave propagates through the narrow constriction and should not leak out of the sides of the tongue. To achieve that, the ridges are designed to close the lateral pathways as shown in Fig. 5. In this figure, the photograph was taken from the back side by removing the rear panel temporarily. When the degree of constriction is less severe than vowel /i/, /e/ is produced. In case of vowel /a/, the tongue body is lowered whereas (the surface of) the tongue root is pushed backward (Fig. 4, middle). For vowel /u/, one can push back and up the middle of the tongue toward the curve of the vocal tract (Fig. 4, right).

Because the tongue is connected to the bottom of the oral cavity, they move almost in unison as shown in Fig. 4. In other words, for high vowels, the tongue is located at the high position while the bottom of the oral cavity is also positioned

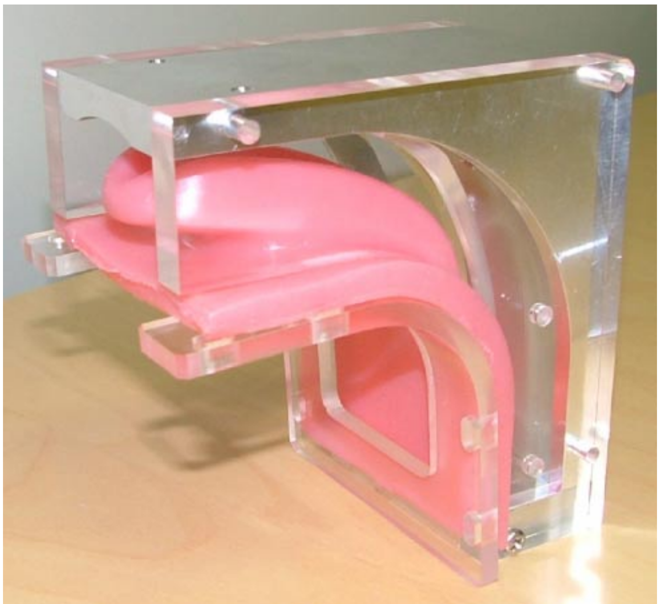

Fig. 1 Physical model of the human vocal tract with a gel-type tongue.

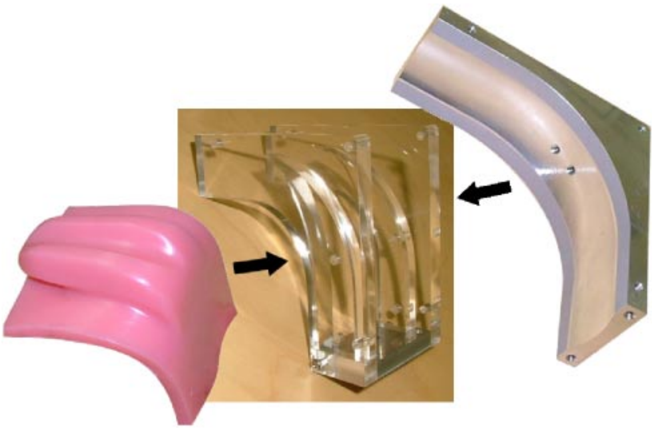

Fig. 2 The model consists of the gel part, the transparent part with acrylic plates, and the non-transparent part with metal.

relatively high. For low vowels, the tongue is located at the low position, and at the same time, the tongue pushes down the bottom of the oral cavity. This model can not be configured so that the tongue is high and the bottom is low. However, we do

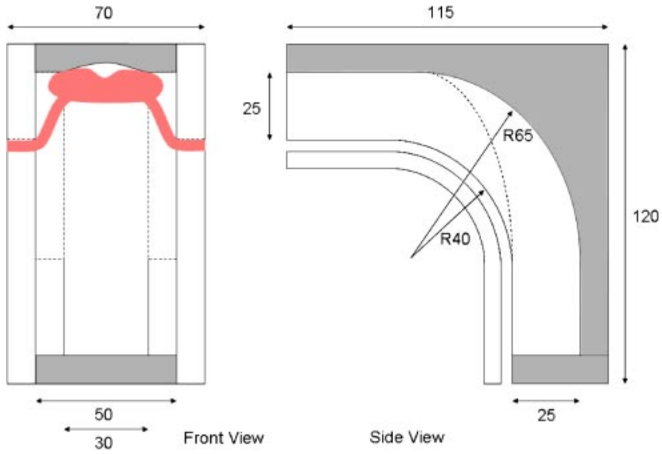

Fig. 3 The dimensions of the proposed vocal tract model (the unit is in mm).

not produce a sound with this kind of configuration in natural speech. Thus, in some sense, it is reasonable that the tongue and the bottom of the oral cavity move together.

3. Discussions

Because the proposed model is transparent as other models in our education system [3], the location of constriction is visible to the naked eye, as is the overall shape of the cavity. This transparent design helps observers intuitively associate the quality of a vowel with the location of the constriction in the model.

Figure 6 shows the vowel spectra obtained from the measurement for the vowels /i/, /a/, and /u/. In this measurement, a whistle-type artificial larynx [3] was used as a sound source. The microphone was placed approximately 20 cm in front of the model's lip end, and produced sounds were digitally recorded through a USB digital audio interface (SE-U33GX, Onkyo). Finally, all signals were stored with sampling frequency of 16 kHz. The spectrum analysis was based on 20th order linear predictive coding (LPC) with a Hamming window. As shown in these plots, the lower formants were clearly observed, and the vowel quality was also clear when listening to each vowel.

This model can also produce a part of consonantal sounds. Sonorants are relatively producible. Consonant /j/ is the easiest one to produce. Consonant /l/ is barely producible but less intelligible. If we place an obstacle, such as lower incisors, sibilant sounds can be produced.

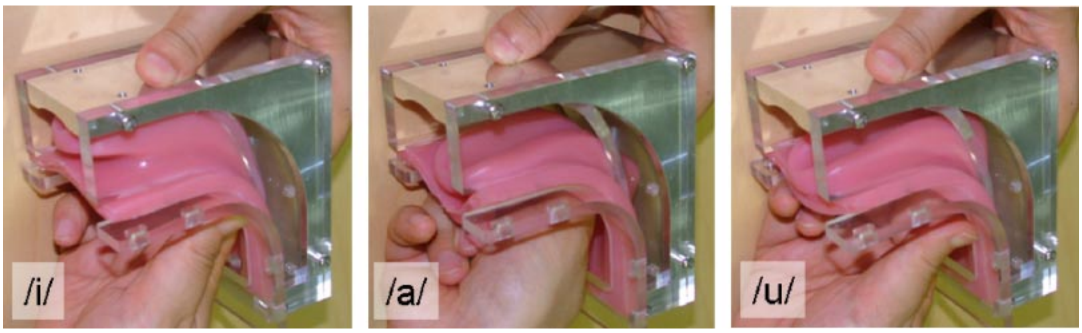

Fig. 4 How to produce different vowels; /i/ (left), /a/ (middle), and /u/ (right).

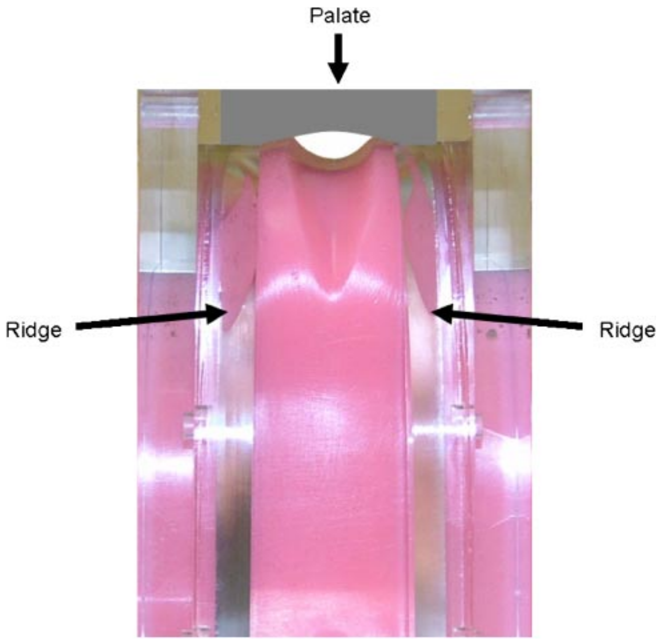

Fig. 5 The model from the back side when producing vowel /i/ (the rear panel is temporarily removed).

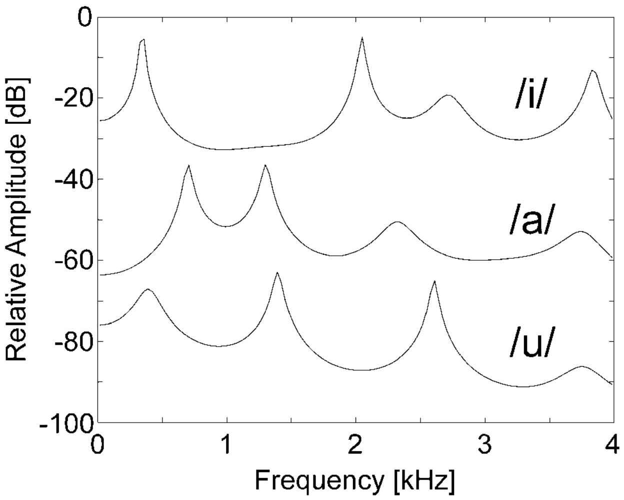

Fig. 6 Measured spectra for vowels /i/, /a/, and /u/.

The author is teaching acoustics at Sophia University not only to technical students but also to students majoring in Linguistics, Psychology, and Speech Pathology, and this model should be useful for those non-technical students. One of the benefits of these models is that students are able to see and hear the change from one sound to another in real time. This model is valuable when measuring how sensitive sound quality is to changes in configuration. From our teaching experience for non-technical students, we confirmed that these models are particularly useful when teaching the dynamic aspects of speech production. Furthermore, this model is attractive for children. A science workshop is one of the appropriate places where we can demonstrate the model effectively for them. In any case, a learner can manipulate the tongue by hand, so that he/she can intuitively learn how we produce speech sounds. The combination of the simultaneous sensations of tactility, somatosensory and auditory perception helps learners understand the phenomenon more naturally and easily.

4. Conclusions

The gel-type flexible-tongue model is proposed as an extended version of the physical model of the human vocal tract. With this model, a learner can manipulate the tongue shape and understand vowel production by seeing the shape and listening to the produced sound. In the proposed model, the entire body of the tongue is not necessarily reproduced. This is due to a result of pursuing the manipulability. Coexisting the real shape of the tongue and the optimal shape for the manipulability is going to be one of our future works.

Acknowledgments

This research was supported in part by Grants-in-Aid for Scientific Research (KAKENHI 16203041; 17500603 and 19500758), and Sophia University Open Research Center from MEXT.

References

[1] T. Arai, "The replication of Chiba and Kajiyama's mechanical models of the human vocal cavity," *J. Phonet. Soc. Jpn.*, **5**(2), pp. 31–38 (2001).

[2] T. Chiba and M. Kajiyama, *The Vowel, Its Nature and Structure* (Tokyo-Kaiseikan, Tokyo, 1942).

[3] T. Arai, "Education system in acoustics of speech production using physical models of the human vocal tract," *Acoust. Sci. & Tech.*, **28**, 190–201 (2007).

[4] T. Arai, "An effective method for education in acoustics and speech science: Integrating textbooks, computer simulation and physical models," *Proc. Forum Acusticum Sevilla* (2002).

[5] T. Arai, E. Maeda, N. Saika and Y. Murahara, "Physical models of the human vocal tract as tools for education in acoustics," *Proc. 1st Pan-American/Iberian Meet. on Acoustics*, Cancun (2002).

[6] T. Arai, "Incorporating more intuitive acoustic education into speech science," *Proc. Spring Meet. Acoust. Soc. Jpn.*, pp. 1219–1220 (2002).

[7] T. Arai and E. Maeda, "Acoustics education in speech science using physical models of the human vocal tract," *Trans. Tech. Comm. Education in Acoustics, Acoust. Soc. Jpn.*, EDU-2003-08, pp. 1–5 (2003).

[8] T. Arai, "Physical and computer-based tools for teaching Phonetics," *Proc. Int. Congr. Phonetic Sciences*, Vol. 1, pp. 305–308, Barcelona (2003).

[9] T. Arai, "Education in Acoustics using physical models of the human vocal tract," *Proc. Int. Congr. Acoustics*, Vol. III, pp. 1969–1972, Kyoto (2004).

[10] T. Arai, "Visualizing vowel-production mechanism using simple education tools," *J. Acoust. Soc. Am.*, **118**, Pt. 2, 1862 (2005).

[11] T. Arai, "Lung model and head-shaped model with visible vocal tract as educational tools in Acoustics," *Proc. Spring Meet. Acoust. Soc. Jpn.*, Vol. 1, pp. 273–274 (2005).

[12] T. Arai, "Lung model and head-shaped model with visible vocal tract as educational tools in acoustics," *Acoust. Sci. & Tech.*, **27**, 111–113 (2006).

[13] N. Sakakibara, E. Shintaku, A. Shimomura, K. Fukui, Y. Ishikawa, A. Takanishi and M. Honda, "Three dimensional tongue model of anthropomorphic talking robot," *Proc. Autumn Meet. Acoust. Soc. Jpn.*, pp. 341–344 (2007).