A Possible Prosthetic Device for Speech Sound Disorders*

○ Takayuki Arai (Sophia Univ.)

1 Introduction

Producing speech sounds involves complicated neural and muscular controls with elaborative timings, which means any physical impairment can potentially cause communication disorders. For example, if a patient has an impairment at the larynx, it might cause voice disorders, or an abnormal voice (e.g., hoarseness) could occur due to vocal fold nodules or polyps. After laryngectomy, patients lose their voices, and an artificial larynx (such as a reed-type [1] or an electrolarynx [2]) is typically used as a voice prosthesis. The present study does not target cases with voice disorders but rather assumes that a speaker can phonate more or less normal voice.

There is another type of speech and language disorder in which a voice is normal but the speech sounds are not intelligible. According to the definition of the American Speech-Language-Hearing Association (ASHA), speech sound disorders are categorized into two types: organic and functional [3]. The latter case further contains articulation disorders and phonological disorders. For example, when a person's speech sound is "distorted" or substituted with another sound, it is considered an articulation disorder, and when problems are related to a linguistic aspect, it is considered a phonological disorder.

Organic speech sound disorders are further grouped on the basis of motor/neurological, structural, and sensory/perceptual causes [3]. Motor/neurological disorders include apraxia and dysarthria. Structural disorders are the result of structural deficits or anomalies, such as cleft palate or cleft lips. Sensory disorders include hearing loss, which often causes less intelligible speech due to the lack of auditory feedback [4].

Speech rehabilitation is also needed for patients with glossectomy due to tongue cancer [5, 6]. When the tongue is totally or partially removed by surgery, patients often exhibit deficits in both swallowing (dysphagia) and speech communication (articulation disorder). The greater the volume of the tongue that is removed, the worse the intelligibility of speech [7]. Speech-language pathologists train glossectomy patients to recover both the swallowing function and articulation.

Various speech processing and synthesis technologies support patients who have difficulty speaking [8], and there are many types of prosthetic devices utilized for different purposes to augment and/or compensate for each patient's ability. In the case of a cleft palate, for example, prosthetic management is effective after an early surgery [9]. A palatal lift prosthesis can be used to facilitate velopharyngeal movement [10], and a speech-aid prosthesis can also be applied not only to children with cleft palate but also adults [11]. A prosthodontic treatment in dentistry has been discussed on the basis of speech sound analyses [12]. The palatal augmentation prosthesis (PAP) is a prosthetic device for glossectomy patients that helps shorten the distance between the residual tongue and the palate to produce a certain set of speech sounds [13]. It is also applicable for patients with certain neurological disorders.

In the current work, we propose another type of prosthesis for people who can phonate and produce relatively normal voice but are unable to move their tongue or lips. Although continuants (including vowels and some consonants) can be produced continuously, a series of speech sounds is dynamic and must be smoothly connected. Even if each sound is less intelligible, there is still potential that a

* A possible prosthetic device for speech sound disorders, by ARAI, Takayuki, (Sophia University).

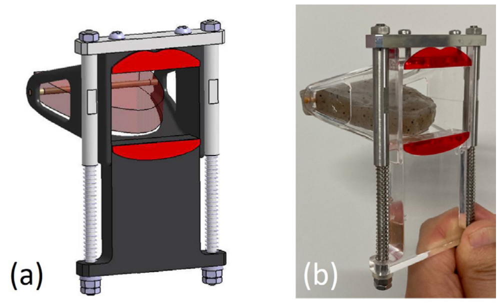

Fig. 1: Overview of proposed device: (a) schematic illustration and (b) photograph.

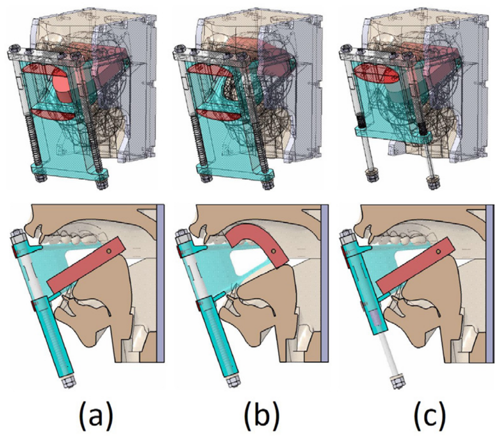

Fig. 2: Perspective (upper panel) and midsagittal (lower panel) views of proposed device with different tongue and lip configurations. The tongue is at the resting position in (a) and (c) but raised in (b). The lips are open in (a) and (b) but closed in (c).

proper sequence of articulatory motions will yield reasonable intelligibility. Therefore, we also aim to reproduce these dynamics of speech with the proposed device.

2 Proposed Device

The device we developed consists of three major components, as shown in Fig. 1: a mouth piece, a lip mechanism, and an imitation tongue. The mouth piece is made of acrylic materials and users can hold it between their teeth. Each of the inner parts has a triangular shape with the apical angle of approximately 30 degrees, which means users can maximally open their mouth with no discomfort to their jaw. The two sides of each triangular frame from the apex of the mouth piece are curved along the row of teeth. There is a shaft running between the two apexes that penetrates the end part of the imitation tongue. The tongue can be made of any flexible material, but for sanitary reasons, we utilized a jelly-type food made from Amorphophallus konjac starch.

We based the lip mechanism on our previous vocal-tract model (the BMW model) that was originally developed for research and education purposes in speech science [14]. As in the BMW model, the lip mechanism has a fixed upper lip plate and movable lower lip plate. One can close the mouth opening by manually pushing up the lower lip plate. Two springs are located at either side of the lower lip plate to facilitate smooth opening of the mouth.

3 Articulations

This section describes how to articulate vowels and consonants with the proposed device. Although the device is not limited to speech sounds in any specific language, we focus here on vowels and consonants in Japanese.

3.1 Vowels

Figure 2 shows schematic views of the proposed device with different tongue and lip configurations, where the upper and lower panels respectively show perspective and midsagittal views. In this figure, the proposed device is embedded in a vocal-tract model we previously developed [15]. The reddish brown part shows the imitation tongue, while the original tongue is shown underneath with a light brown color. Here, we estimated the vocal-tract area function for the cases shown in Fig. 2(a) and (b). First, we determined the midline of the midsagittal vocal-tract configuration, and then, the cross-sectional areas along the midline were calculated every 5 mm.

3.2 Consonants

Approximant consonants can be produced with the device as a glide sound from one vowel to another. For example, the approximant /j/ starts with vowel /i/ and gradually moves to the configuration of the following vowel. Likewise, the approximant /w/ starts with vowel /u/ with a lip closure and gradually moves to the configuration of the following vowel.

To produce stop consonants, we need to make a complete closure of the speech organs. With the proposed device, the only stop consonants that can be produced with the default lower lip plate are /p/ and /b/, as there is no hole. The other stop consonants are difficult to produce. To make nasal sounds, we have to control the velopharyngeal port opening, and to do so, we need to be able to lower the velum. If the speaker has this ability, there is a chance that nasal sounds such as /n/ and /m/ can be produced.

4 Evaluation

We recorded and analyzed five vowels, several consonants, and two phrases in Japanese to test the quality of speech sounds produced with the proposed device. Figure 3 shows the F1-F2 frequencies estimated from the recordings in this section using the proposed device with the average frequencies in [16] as reference points. The formant estimation was done by Praat [17].

Figure 4 shows a spectrogram of the recorded phrase /arigato:/, or "Thank you" in Japanese. As we can see, the formant trajectories of this utterance are more or less achieved in the vowel parts. However, consonants are sometimes substituted by other sounds or missed. The first consonant /r/ in this phrase has some abrupt interruptions in the spectrogram, and it is reflected by a rapid tongue movement. It turned out that it is intelligible as /r/ in Japanese. The second consonant in this phrase is /g/ and it was not achieved as a plosive sound. Instead, it was voiced as a fricative and is actually an allophone of voiced plosive sounds in Japanese; thus it was acceptable as /g/ to the ears of Japanese listeners. The third consonant /t/ was not achieved in this utterance, and its spectrogram only shows formant transitions from /a/ to /o/.

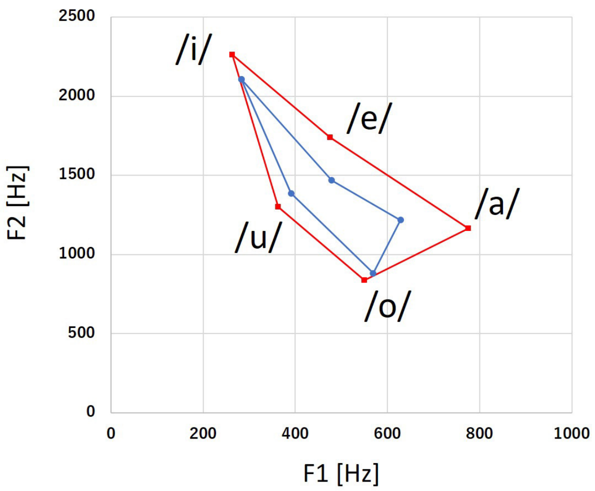

Fig. 3: F1-F2 plot of formant frequencies estimated from recordings in Section 4 using the proposed device (blue) with average frequencies based on [16] (red) as reference points.

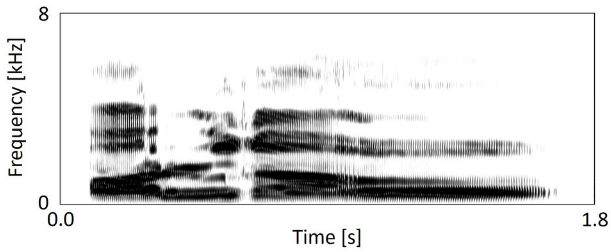

Fig. 4: Spectrogram of /arigato:/ utterance.

5 Summary

In this study, we proposed a prosthetic device for speech sound disorders and demonstrated through acoustic analyses that it can be useful for speakers who have problems articulating speech sounds for diverse reasons. We assumed that a user can produce normal phonations and can manipulate the device manually by him/herself. In this regard, there is an option to manipulate the imitation tongue by one's fingers, but for sanitary reasons, a wire can be attached to the imitation tongue, and the lower lip plate is equipped with a slit through which the wire can stick out.

Our evaluations based on the acoustic analyses indicated that the proposed device is able to produce five vowels, even if the vowel space is a little bit smaller that a normal one with average formant frequencies. Some consonants are relatively easily produced with the device, such as /b/ (/p/), /w/, /j/, /h/, and /r/, but others are substituted by similar sounds. When combined them each other with the proper timing and intonation, phrases were more intelligible than individual sounds. In terms of the phrases, we conducted an informal listening test for eleven listeners (in their twenties) in which we asked them to indicate what they heard as Japanese phrases and the results showed the both phrases were the same and 9 out of eleven listeners answered correctly. This suggests listeners are able to compensate for substituted and missing sounds by means of top-down processing.

Acknowledgments

This work was partially supported by JSPS KAKENHI Grant Numbers 21K02889 and Sophia University Special Grant for Academic Research (Research in Priority Areas).

References

[1] R. R. Riesz, "Description and demonstration of an artificial larynx," *J. Acoust. Soc. Am.*, vol. 1, pp. 273–279, 1930.

[2] H. L. Barney, F. E. Haworth, and H. K. Dunn, "An experimental transistorized artificial larynx," *Bell Syst. Tech. J.*, vol. 38, pp. 1337–1356, 1959.

[3] American Speech-Language-Hearing Association (ASHA), "Speech Sound Disorders: Articulation and Phonology," https://www.asha.org/practice-portal/clinical-topics/articulation-and-phonology/ (retrieved on Feb., 14, 2023).

[4] P. B. Denes and E. N. Pinson, *The Speech Chain: The Physics and Biology of Spoken Language*, 2nd ed., New York: W. H. Freeman, 1993.

[5] G. R. LaBlance, K. Kraus, and K. F. Steckol, "Rehabilitation of swallowing and communication following glossectomy," *Rehabilitation Nursing*, vol. 16, no. 5, pp. 266–270, Sep. 1991.

[6] K. M. Blyth, P. McCabe, C. Madill, and K. J. Ballard, "Speech and swallow rehabilitation following partial glossectomy: A systematic review," *International Journal of Speech-Language Pathology*, vol. 17, no. 4, pp. 401–410, 2015.

[7] I. Kumakura, "A Study of speech intelligibility after glossectomy: Analysis of 60 tongue cancer cases," *Japan Journal of Logopedics and Phoniatrics*, vol. 26, pp. 224–235, 1985.

[8] S. Kawahara, M. Homma, T. Yoshimura, and T. Arai, "MyVoice: Rescuing voices of ALS patients," *Acoust. Sci. Technol.*, vol. 37, no. 5, pp. 202–210, Sep. 2016.

[9] D. S. Dorf, D. J. Reisberg, and H. O. Gold, "Early prosthetic management of cleft palate. Articulation development prosthesis: A preliminary report," *J. Prosthet. Dent.*, vol. 53, no. 2, pp. 222–226, Feb. 1985.

[10] Y. Hamamura, J. Nishio, T. Matsuya, T. Goto, T, Motomura, K. Inoue, T. Fukuda, and T. Miyazaki, "Facilitation of velopharyngeal movement by palatal lift prosthesis," *Japanese Journal of Oral & Maxillofacial Surgery*, vol. 24, no. 2, pp. 253–260, 1978.

[11] D.-C. Kang, J.-H. Park, H. Seok, J.-A. Baek, D.-W. Kim, and S.-O. Ko, "Speech-aid prosthesis in velopharyngeal incompetency patient with cleft palate: Can speech aids be applicable for adult patient?," *Maxillofacial Plastic and Reconstructive Surgery*, vol. 43, no. 29, 2021.

[12] J. M. Palmer, "Analysis of speech in prosthodontic practice," *Journal of Prosthetic Dentistry*, vol. 31, no. 6, pp. 605–614, Jun. 1974.

[13] M. Marunick and N. Tselios, "The efficacy of palatal augmentation prostheses for speech and swallowing in patients undergoing glossectomy: A review of the literature," *Journal of Prosthetic Dentistry*, vol. 91, no. 1, pp. 67–74, Jan. 2004.

[14] T. Arai, "Mechanical production of [b], [m], and [w] using controlled labial and velopharyngeal gestures," in *Proc. INTERSPEECH 2016 – 17th Annual Conference of the International Speech Communication Association*, San Francisco, USA, Sep. 2016, pp. 1099–1103.

[15] T. Arai, "Flexible tongue housed in a static model of the vocal tract with jaws, lips, and teeth," in *Proc. INTERSPEECH 2018 – 19th Annual Conference of the International Speech Communication Association*, Hyderabad, India, Sep. 2018, pp. 171–172.

[16] H. Kasuya, H. Suzuki, and K. Kido, "Changes in pitch and first three formant frequencies of five Japanese vowels with age and sex of speakers," *J. Acoust. Soc. Jpn.*, vol. 24, no. 6, pp. 355–364, 1968.

[17] P. Boersma, "PRAAT: A system for doing phonetics by computer," *Glot International*, vol. 5, nos. 9/10, pp. 341–345, 2001.