

Articles

Swallowing Sound Analysis System for Quantitative Evaluation of Dysphagia Rehabilitation

KANI Masaya¹, FURUTAKI Rika¹, YAMAUCHI Shinobu²,
MOTOHASHI Yuka², SATO Toshio*,
AGISHI Tetsuzo³ and TAKAHASHI Tsuneo⁴

(Received Date: March 17, 2018)

Key Words: Swallowing sound, Time-frequency analysis, Swallowing function, Rehabilitation

I. Introduction

In Japan, the examinations available for investigating the eating and swallowing functions are videofluorography and videoendoscopy. These examinations enable detailed visual assessment of the function and dynamics of swallowing. However, these methods have disadvantages, such as radiation exposure during the examinations; require technicians to be trained in specific skills; and are not simple. Meanwhile, screening tests of the swallowing function have been proposed and include repetitive saliva swallowing test (RSST), modified water swallowing test (MWST), and cervical auscultation (CA). Although RSST has a high sensitivity of 0.98, it has a specificity of only 0.66 and was reported to have a high frequency of false positives. Although the MWST has relatively high sensitivity and specificity, its possibility to overlook silent aspiration that does not present

with choking or coughing remains. In addition, evaluation by CA depends on the subjective judgment of the person performing it. Therefore, no detailed and simple screening methods have been established.

Meanwhile, head lift exercises (i.e., shaker exercise) and ice massages are being performed for eating and swallowing rehabilitation in patients with impaired swallowing function. Although some studies on triggering the swallowing reflex with cold carbonated water have been conducted in recent years ^{1),2)}, few studies have assessed the efficacy of using this method in a standardized manner.

In our previous studies, we performed time-frequency analyses to investigate the efficacy of measuring the swallowing sounds at the neck for the assessment of swallowing function ³⁾⁻⁵⁾. In the present study, we attempted to increase the reproducibility of measuring the swallowing sounds by identifying the optimal locations for obtain-

* SATO Toshio: Professor, Graduate School of Engineering; Faculty of Biomedical Engineering, Toin University of Yokohama. 1614 Kurogane-cho, Aoba-ku, Yokohama 225-8503, Japan

¹ KANI Masaya and FURUTAKI Rika: Department of Medical Engineering, Graduate School of Engineering, Toin University of Yokohama

² YAMAUCHI Shinobu and MOTOHASHI Yuka: Lecturer, Faculty of Biomedical Engineering, Toin University of Yokohama

³ AGISHI Tetsuzo: Faculty of Biomedical Engineering, Toin University of Yokohama

⁴ TAKAHASHI Tsuneo: former Professor, Department of 3D Imaging Anatomy, Graduate School of Kanagawa Dental University

ing the swallowing sounds. We also investigated the potential applicability of this newly proposed method of obtaining swallowing sounds using cold carbonated water for rehabilitation.

II. Experimental Methods

1. Measurement of swallowing sounds with the Bio Sound Analyzer

We previously developed a device called the Bio Sound Analyzer (BSA) through collaboration among the industry, academia, and medicine. Although it is currently not yet approved by the Pharmaceutical Affairs Law, this device can non-invasively measure biological sound signals at bedside and converts them into wavelets in order to screen biological functions (*Fig.1*). This device can quantitatively analyze the biological sounds of interest in real-time. In the field of blood purification therapy, the BSA was developed to measure and analyze the sound of blood flow, which

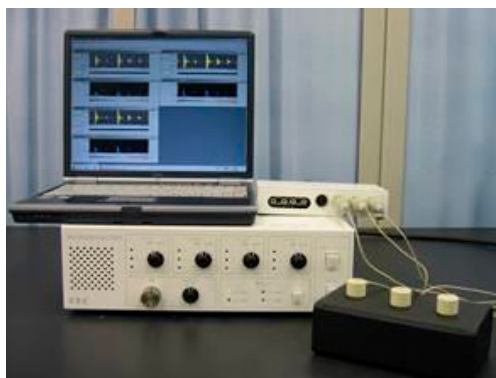


Fig.1 Bio Sound Analyzer (BSA)

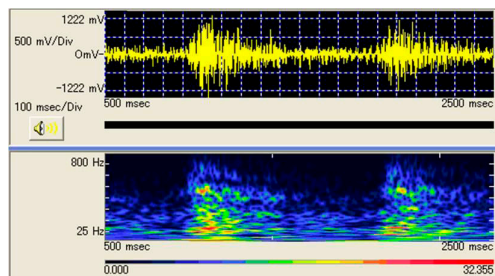


Fig.2 Analysis example of shunt murmur of vascular access

is referred to as shunt murmur, and uses phonocardiographic sensors to obtain biological sounds. It can simultaneously measure biological sounds from a maximum of 3 channels with a 44.1-kHz sampling frequency and 16-bit resolution. *Figure 2* shows an example of the results of a shunt murmur analysis taken from the vascular access of a hemodialysis patient. The upper graph shows the shunt murmur signals measured with the BSA, with time on the horizontal axis and signal voltage on the vertical axis. The lower graph is the result of wavelet conversion of the shunt murmur signals from the upper graph, with time on the horizontal axis and frequency on the vertical axis; the size of the frequency spectrum is represented in a color map from blue to red, and Gabor wavelets were used for the mother wavelets during wavelet conversion. In the present study, as an example of BSA application, we measured swallowing sounds with the same system that was used for the shunt murmurs in an attempt to establish a non-invasive, quantitative method for assessment of swallowing function.

2. Quantitative assessment of swallowing function

Based on our past results on simultaneously measuring swallowing sounds and the videofluorographic images of swallowing, we determined that a swallowing sound comprises 3 different components: type I corresponds to laryngeal elevation during swallowing; type II corresponds to the movement of a food bolus; and type III corresponds to the laryngeal descent. In people without impaired swallowing, this sequence of actions happens quickly and correctly, in combination with the swallowing reflex, through the activity of the nerves involved in eating and swallowing. However, with aging, the muscles related to swallowing atrophy, which leads to insufficient laryngeal movement and inability to increase the swallowing pressure. This results in problems, such as

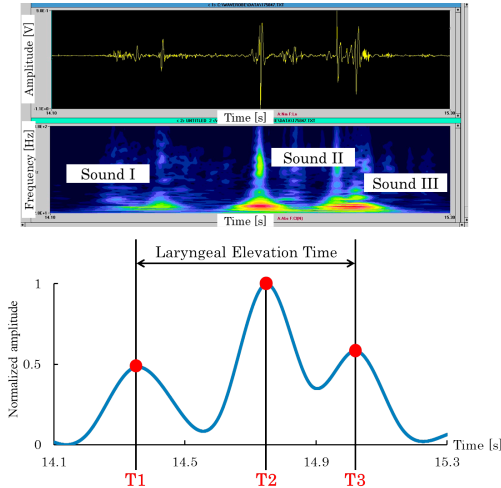


Fig.3 Envelope of swallowing sound signals

the inability to properly send the food bolus down the esophagus, which can trigger aspiration. Based on these, the temporal relationship of the 3 sounds may provide useful information on the swallowing dynamics for the assessment of older adults and individuals with impaired eating and swallowing functions. Therefore, we attempted to quantitatively assess the swallowing function by creating an envelope of swallowing sound signals using cubic spline approximation.

As an example, **Figure 3** shows the results of the conversion of the swallowing sound signals to wavelets and the cubic spline approximation in unimpaired individuals in their twenties who were asked to swallow 5 mL of room-temperature water. The envelope was created using a dedicated BSA analysis software (Wavelet Bmp Analyzer: WBA) that we created. The horizontal axis shows the time and the vertical axis shows the amplitude standardized to values from 0 to 1. When assessing the swallowing function using this envelope, the positions of the 3 amplitude peaks for the 3 sounds were defined as T1, T2, and T3 on the temporal axis. The time from the sound I to the sound III (T3-T1) corresponded to the duration of laryngeal elevation. The ratio of the duration between the sound I and sound II (T2-T1) over the duration

of all 3 sounds $P = (T2-T1) / (T3-T1)$ was then calculated to define a parameter that represented the swallowing function of the subject.

3. Measurement of swallowing sounds with multiple sensors

The acceleration sensors used in the present study were of the floating type with an outside diameter of 20 mm, thickness of 16 mm, and weight of 41 g. Nominally, they have frequency characteristics of 20–600 Hz, a resonant frequency of at least 350 Hz, and a sensitivity of 0–12 dB (150 Hz), when 0 dB = 0.1 V/ (m/s²). Takahashi *et al.*⁶ found that the optimal location of auscultation for measuring the swallowing sounds with an acceleration pick-up and a microphone was lateral to the trachea just inferior to the cricoid cartilage. In our study, therefore, we used this site for attachment of the acceleration sensor onto the subject using a double-sided tape and measured the swallowing sounds. On this spot, the sound II was very clear,

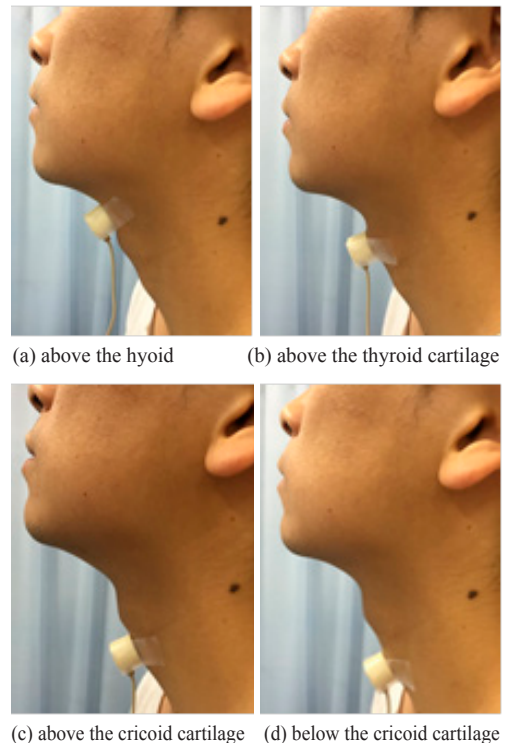


Fig.4 Spots for measurement of swallowing sounds

but the sound I and sound III could not be obtained sometimes or the amplitudes of those sounds were far too small. Therefore, there was a tendency that the laryngeal movement at the time of the swallowing reflex cannot be determined on this spot. In an attempt to identify the sensor attachment position/s that will enable acquisition of clear sound I and sound III, we measured the swallowing sounds at 4 spots, including (a) just above the hyoid, (b) just above the thyroid cartilage, (c) just above the cricoid cartilage, and (d) just below the cricoid cartilage along the medial line of the neck as shown in *Figure 4*. We also aimed to establish a reproducible method of obtaining swallowing sounds on these 4 spots and on a spot lateral to the trachea just inferior to the cricoid cartilage.

4. Test of the swallowing reflex triggered by carbonated water

Currently, as part of the rehabilitation of patients with delayed swallowing reflex, cold pressure stimulation and ice massage are being performed to trigger the swallowing reflex by combining cold and pressure stimulation. In particular, recent attention has been on the use of cold carbonated water as a way to trigger the swallowing reflex both physically, with cold and effervescence, and chemically, with carbonic acid²). However, the use of carbonated water for eating and swallowing rehabilitation had been investigated by only few studies and is not currently performed in clinical practice. In this section, in order to test the effectiveness of such procedure, we asked the subjects to swallow carbonated water in varying carbonic acid concentrations and temperatures and calculated the duration of laryngeal elevation and the P.

First, we attached the sensors on 2 places: (1) the spot identified in section II.3 as the best place for obtaining sound I and sound III and (2) the spot lateral to the trachea just inferior to the cricoid cartilage, which was the original location for sensor

attachment. Next, we measured the swallowing sounds by having the subjects swallow normally the samples twice during a 20-second period; the timing for swallowing was 5 and 15 seconds after the start of measurement. The swallowing samples were made using a soda machine to mix water and carbonated water to create weak, moderate, and strong effervescence. A thermostatic device was used to adjust the temperature of each sample to 10, 20, and 30 °C. For the measurement conditions, each swallowing sample was set to 5 mL and the subjects were asked to be seated while swallowing. After measuring the swallowing sounds for each sample, a composite signal was created from the swallowing sound signals obtained from each of the 2 sensors and the respective envelope curve was created. Then, the duration of laryngeal elevation and the P were calculated from the envelope created. As an ethical consideration, the study protocol was approved by the research ethics committee of Toin University of Yokohama.

III. Experimental Results

1. Identification of the sensor attachment position

Figure 5 shows the results of the swallowing measurement and the wavelet conversion at the 4 spots along the medial line of the neck. The results of this analysis suggested the possibility to clearly obtain sound I and sound III at the spots just above the hyoid and just above the thyroid cartilage. In contrast, at the spots just above and below the cricoid cartilage, the measurements did not provide sufficient sound I and sound III and there was a constant and frequent tendency for the noise that was likely the sound of carotid pulsation to be mixed with the actual swallowing sound. There were also anatomical differences between sexes in the degree of protrusion of the thyroid cartilage (i.e., more pronounced in men than in women).

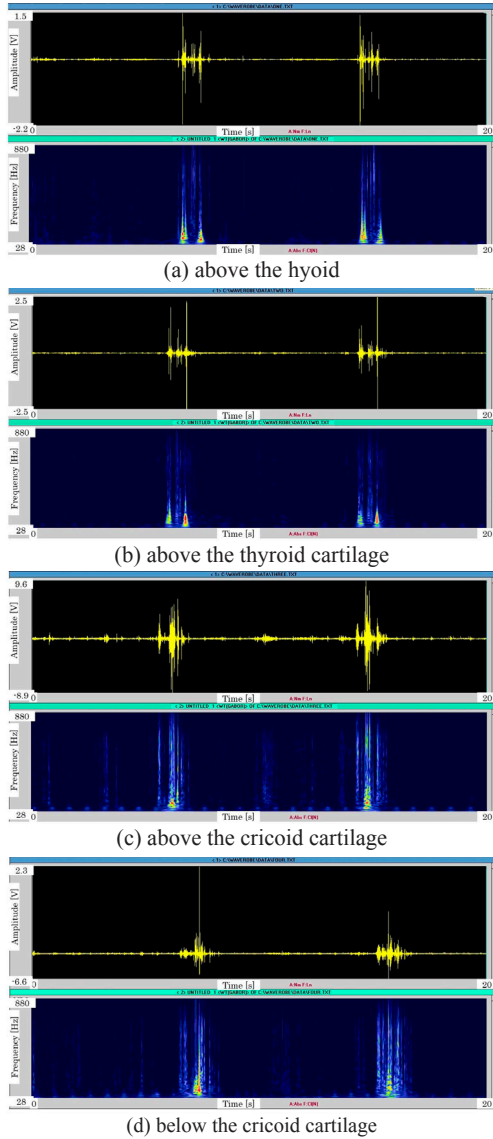


Fig.5 Analysis results of swallowing sound signals

This anatomical difference may affect the measurement of the swallowing sounds at the spot just above the thyroid cartilage, suggesting that the spot just above the hyoid would be more appropriate for attaching the sensor to obtain sound I and sound III.

The swallowing sounds were measured simultaneously just above the hyoid and lateral to the trachea just inferior to the cricoid cartilage, and the amplitudes of the swallowing sound signals

obtained from each of the 2 sensors were added to yield a composite swallowing sound signal. We then attempted to create an envelope by wavelet conversion of that signal and cubic spline approximation. **Figure 6** shows one example of the composite swallowing sounds that we obtained. When the swallowing sounds were combined, the amplitude and frequency spectra of the 3 swallowing sounds became clear. Moreover, the swallowing sounds obtained with this method tended to be more stable than those obtained with the original method.

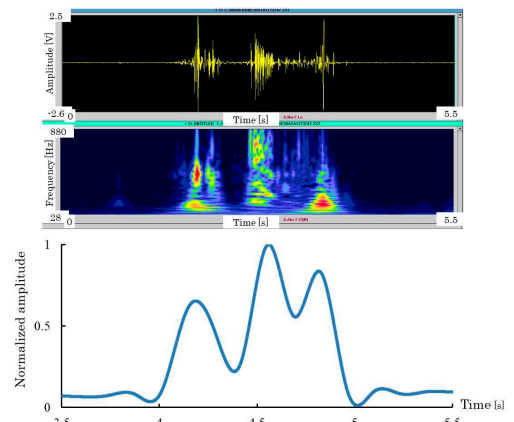


Fig.6 Composite swallowing sounds

2. Effects of carbonated water temperature and effervescence on the swallowing function

Figure 7 shows the results of creating composite swallowing sound signals from the swallowing sounds obtained when subjects swallowed water at each temperature and carbonated water with 3 different levels of effervescence, and calculation of the laryngeal elevation duration and P from the envelope obtained. When the subjects were asked to drink the 20– °C samples, the duration of laryngeal elevation tended to be shorter when drinking carbonated water than when drinking plain water (**Fig.7a**). Furthermore, the P for each sample tended to be smaller when swallowing carbonated water than when swallowing plain water at 20 and 30 °C (**Fig.7b**).

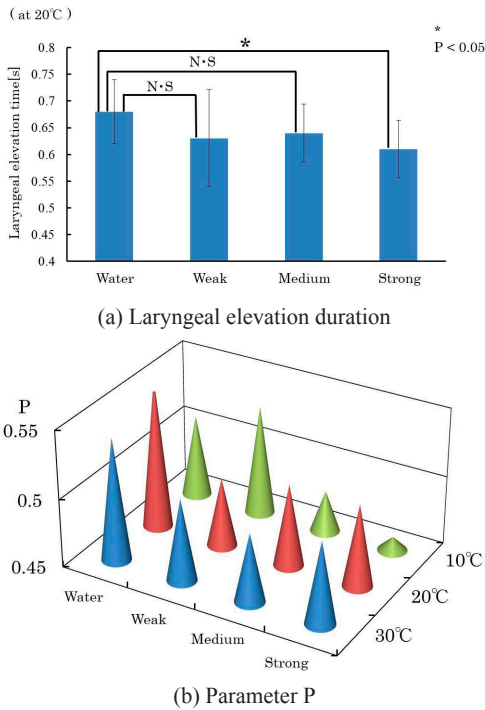


Fig. 7 Calculation results

IV. Discussion

In a study on the mechanisms underlying the swallowing sounds, Hamlet *et al.*⁷⁾ found 3 acoustic components when they attached small accelerometers to spots about 2 cm anterior and posterior to the left mandibular angle in unimpaired adults who were asked to swallow 10 mL of diluted liquid barium. They reported that the first component was related to laryngeal elevation and movement of the food bolus into the pharynx; the second component temporally matched the moment when the food bolus passed into the hypopharynx and through the upper esophageal sphincter (UES); and the third component was related to the laryngeal descent at the end of swallowing. Moreover, they noted that the first and third acoustic components were weak and not always present. In an attempt to clarify the sites of production of swallowing sounds and the corresponding acoustic

characteristics on those sites, Nakayama *et al.*⁸⁾ attached acceleration transducers to the spot lateral to the trachea just inferior to the cricoid cartilage on unimpaired adults and analyzed the 96 swallowing sounds obtained after swallowing 5 mL of angiographic agent. The swallowing sounds were divided into 5 passing sounds from the tongue base, epiglottis, initial UES, middle UES, and final UES. The occurrence rates of each of the sounds measured, in synchrony with the obtained video-fluorographic images of swallowing, were 63.5%, 97.9%, 90.6%, 70.8%, and 53.1%, respectively. Comparison of these occurrence rates showed that the sounds measured at the beginning and end of swallowing were likely to be difficult to obtain.

In the present study, we found that the spot just superior to the hyoid was the position for sensor attachment that enabled detection of laryngeal movement with high sensitivity and stable sampling of sound I and sound III, which are especially difficult to obtain among the 3 swallowing sounds. Moreover, our results suggested the possibility to obtain highly reproducible swallowing sounds (i.e., each of the 3 sound components is clearly distinguishable) by measuring the swallowing sounds simultaneously at a point lateral to the trachea just inferior to the cricoid cartilage, where sound II can be clearly detected, and by creating a composite swallowing sound signal from the swallowing sound signals obtained from the 2 sensors. This implied the ability to stably calculate the P and the duration of laryngeal elevation, which are defined as indicators of swallowing function, offered an effective method for obtaining swallowing sounds.

Furthermore, prevention of aspiration pneumonia requires not only screening of the eating and swallowing functions, but also rehabilitation. In particular, some of the risk factors for aspiration pneumonia in older patients include age-related changes in the structure and functioning of the organs for eating and swallowing. One example is

the relatively lower position of the larynx at rest, which can cause prolonged laryngeal elevation, insufficient laryngeal closure, and incomplete dilation of the UES. Additionally, sarcopenia, which is the reduction in muscle mass that is associated with aging, can affect the tongue and the muscles involved in swallowing and consequently decreases the ability to send a food bolus down the esophagus and increases the risk of aspiration. Aspiration triggered by impaired eating and swallowing functions can be classified as before, during, and after swallowing, depending on the timing of occurrence. Aspiration after swallowing is triggered by incomplete dilation of the UES or reduced ability to send a food bolus down the esophagus. This type of aspiration accounts for over half of all cases of aspiration and is a well-known major issue in clinical practice. Therefore, prevention of aspiration after swallowing is very important.

In the context of these issues, we focused on carbonated water and attempted to test its effectiveness in the assessment of swallowing. A total of 12 different types of carbonated water, including plain water, at different temperatures and levels of effervescence, were used as swallowing samples to calculate the duration of laryngeal elevation and the P from the swallowing sound signals obtained. Although no differences were seen among the types of carbonated water at 20 °C, the duration of laryngeal elevation tended to be shorter when swallowing carbonated water than when swallowing plain water. To test the effects of carbonated beverages in triggering the swallowing reflex, Morishita *et al.*¹⁾ used the duration of laryngeal elevation as an indicator of the ease of sending the beverage down the esophagus. Eight older inpatients were asked to swallow 30 mL of carbonated beverage that had been cooled to between 5 and 10 °C. The duration of laryngeal elevation was reportedly shorter when they swallowed the carbonated beverage than when they swallowed water. Although the temperature of the swallowed

samples differed from that used in the present study, the trend was similar between the 2 studies. Moreover, we found that the P tended to become smaller with increasing level of effervescence and decreasing temperature. This P parameter that we are proposing as an indicator of swallowing function represents the ratio of the duration of onset of a sound II over the duration of laryngeal elevation. Accordingly, a drop in the P when swallowing would indicate faster transit of a food bolus down the esophagus after laryngeal elevation for cold carbonated water than for plain water; this finding can be linked to prevention of aspiration after swallowing. Furthermore, one study showed that injections of ≤ 1 mL of carbonated water at 12 to 14 °C on the tongue base triggered the swallowing reflex in patients with cerebrovascular diseases. Another study on 17 people with dysphagia found that aspiration occurred in 5 subjects when drinking 5 ml of plain water, but it did not occur when drinking 5 ml of carbonated water. These findings indicated that even small amounts of cold carbonated water may trigger the swallowing reflex and may be safe to use for eating and swallowing rehabilitation.

V. Conclusion

Using the BSA and phonocardiographic sensors, we were able to obtain swallowing sounds to test the effects of cold carbonated water at different levels of effervescence and temperatures in triggering the swallowing reflex. In this article, we first proposed a method for obtaining reproducible swallowing sounds using multiple sensors. This method showed a tendency to improve the ability to obtain highly reproducible and clear swallowing sounds with 3 components. Further analysis suggested that the use of cold carbonated water may be a novel tool for eating and swallowing rehabilitation.

[References]

- 1) M. Morishita *et al.*, “Sustained swallowing- improving effect of carbonated beverage in elderly inpatients”, *Jurnal of Kibi International University Research Institute of Health and Welfare*, No.16, pp.7–12, 2005.
- 2) E. Michou *et al.*, “Examining the role of carbonation and temperature on water swallowing performance: a swallowing reaction-time study”, Vol.37, pp.799–807, 2012.
- 3) T. Sato *et al.*, “Evaluation of swallowing by time-frequency analysis of swallowing sounds with wavelet transform”, *Therapeutics & Engineering*, Vol.17, No.4, pp.181–187, 2005.
- 4) M. Kani *et al.*, “Time-frequency analysis of swallowing sounds in non-invasive evaluation of swallowing function”, Abstracts of ASAIO 62nd Annual Conference, p.3, 2016.
- 5) A. Murakami *et al.*, “Noninvasive biofunctional diagnosis based on time-frequency analysis of various biosounds”, *IEICE Technical Report*, Vol.109, No.135, pp.15–20, 2009.
- 6) K. Takahashi *et al.*, “Methodology for detecting swallowing sounds”, *Dysphagia*, Vol.9, pp.54–62, 1994.
- 7) SL. Hamlet *et al.*, “Sounds of swallowing following total laryngectomy”, *Dysphagia*, Vol.7, pp.160–165, 1992.
- 8) H. Nakayama *et al.*, “Evaluation of the sites where swallowing sound produces and the acoustic characteristic of swallowing sound in healthy adults”, *The Journal of Showa University Dental Society*, Vol.26, No.2, pp.163–174, 2006.