Showa Univ J Med Sci **27**(4), 271～284, December 2015

Original

Identifying the Timing of Swallowing Sounds Using Videoendoscopy Findings in Healthy Adults

Nanae FURUYA*[1)], Kaoru YOKOYAMA[1)], Koji TAKAHASHI[1)], Madoka YAMASHITA[1)], Yuka NAKAMICHI[1)], Yoshiaki IHARA[1)] and Michael E. GROHER[2)]

Abstract: Cervical auscultation is a useful tool for detecting dysphagia; however, the sites where swallowing sounds are produced are unknown. In this study, we investigated the relationship between swallowing sounds and videoendoscopy (VE) images in healthy adults to identify the timing of swallowing sounds. Fifteen healthy young adults participated in the study. Each participant was seated in an upright position while a stethoscope probe with an inserted microphone was placed at the center of his or her lower neck to detect swallowing sounds during the VE. The detected sounds were recorded simultaneously with the VE images while the subjects swallowed 4 g of liquid or jelly. Swallowing duration, swallowing sound duration, and VE findings at the beginning and end of swallowing sounds were analyzed. One hundred and thirty-four sound samples produced by a single swallowed bolus were obtained and analyzed. The mean swallowing duration for each material ranged from 1.25 to 2.39 s. Swallowing duration was significantly longer for jelly compared with liquids ($p < 0.01$). Swallowing sound duration was approximately 0.5 s in all samples, and there were no significant differences between materials. Most swallowing sounds started during velopharyngeal closure (109/134, 81.3%), and most swallowing sounds ended during velopharyngeal closure (98/134, 73.1%). For all materials, swallowing sounds did not start when the materials flowed into the pyriform sinuses, and very few sounds corresponded with epiglottic movements. These results show that many movements associated with physiologic events—including hyoid bone and laryngeal excursion, and opening of the upper esophageal sphincter—may be involved in the production of swallowing sounds.

Key words: videoendoscopy, swallowing sounds, acoustic analysis, deglutition, cervical auscultation

Introduction

Cervical auscultation (CA) is a non-invasive clinical tool that is frequently used for assessing dysphagia[1, 2)]. Borr *et al* reported that CA had 70% specificity and 94% sensitivity in detecting dysphagia[3)]. Takahashi *et al* reported that the percentage agreement when diagnosing dysphagia using acoustic analysis and videofluorography (VF) was 77.3%[4)]. Thus, swallowing sounds are an

[1)] *Department of Special Needs Dentistry, Division of Oral Rehabilitation Medicine, Showa University School of Dentistry,* 2-1-1 Kitasenzoku, Ohta-ku, Tokyo 145-0062, Japan.
[2)] *Truesdail Center for Communicative Disorders, University of Redlands.*
* To whom corresponding should be addressed.

important clue in detecting dysphagia. However, the anatomic sites where swallowing sounds are produced are unknown.

Research aimed at identifying the sites where swallowing sounds are produced has involved comparison of images from VF of swallowing to swallowing sounds detected by CA. However, visualization with VF is limited when the contrast agent is too thin or too small, and it is difficult to detect slight spillage of materials into the pharynx that might generate sound. Videoendoscopy (VE) of swallowing provides a direct view of the pharynx and larynx and can detect slight spillage. In addition to observing the movement of soft tissue, VE enables researchers to directly observe residue of saliva and secretion products, movement of the base of the tongue and velopharyngeal closure, contraction of the pharyngeal side wall, shapes of the pyriform sinus and epiglottis, and movement of arytenoid cartilage. When VF is used, contraction of the pharyngeal side wall cannot be observed by the lateral view, and contraction of pharyngeal side wall is difficult to observe by the frontal view because the maxilla bone overlaps the pharyngeal side wall.

During velopharyngeal closure, bolus movements and swallow-related organs cannot be observed by VE. However, VE can detect slight spillage of materials into the pharynx, pharyngeal residue in great detail, slight penetration of materials or saliva into the larynx, aspiration before and after swallowing, and the cough reflex[5–7]. In addition, VE enables examinations with usual diets, flavors, and properties because the procedure does not use barium. Without the use of barium, severely dysphagic patients can be examined more safely by VE than by VF[5]. Moreover, VE equipment is compact and portable, so evaluation location is not restricted.

The purpose of this study was to investigate the relationship between swallowing sounds and VE images in healthy adults, and differences in swallowing sounds for various materials. From these data, insight into the source and timing of swallowing sounds can be obtained.

Materials and methods

Participants

Fifteen healthy young adults (eight men, seven women; mean age: 26.5±3.0 years) were recruited to participate in the study. The procedure was explained to them, and they all provided written informed consent.

Ethics approval

The study was approved by the Ethics Committee of Showa University School of Dentistry (acceptance number: 2011-010).

VE examination of swallowing

Before the VE experiments, three dentists who had more than 4 years of experience in managing dysphagia assessed participants' laryngopharyngeal structures to confirm that they were normal. In addition, they checked whether participants experienced any discomfort and whether they could swallow as usual with transnasal insertion of a fiberscope. Participants were seated

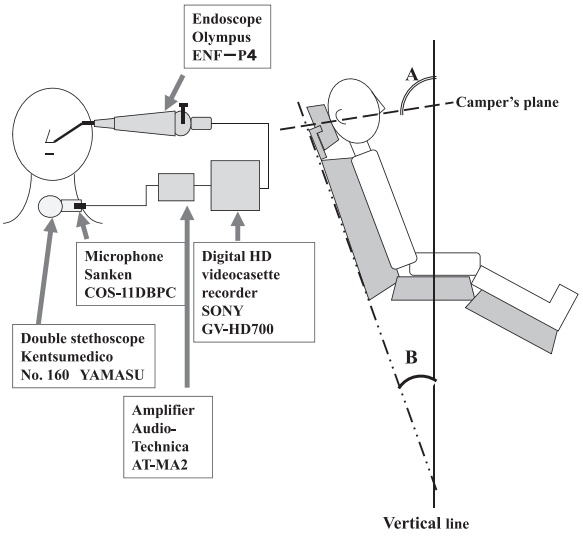

Fig. 1. Detection and recording of swallowing sounds during videoendoscopy (VE)
The system for detecting and recording swallowing sounds during VE is shown on the left. The detected swallowing sound signals were amplified and recorded simultaneously with the VE images using a digital high-definition video cassette recorder. On the lateral view, the angles between the vertical line and Camper's plane (A), and the vertical line and the backrest of the dental chair (B) were measured, as shown on the right.

in a dental chair in a comfortable, upright position. A flexible fiberoptic endoscope (ENF-P4, Olympus, Tokyo, Japan) was introduced transnasally directly below the inferior turbinate, and the tip of the endoscope was positioned at the level just below the uvula during the swallow sequence. On the lateral view, the angles between the vertical line and Camper's plane and between the vertical line and the backrest of the dental chair were measured (Fig. 1). To perform the examination while the participants were in a relaxed position, participants' head and body were unfixed; however, participants were instructed to keep their head and body as still as possible during the examination.

Black and white colored liquids were used for contrasting pharyngeal mucosa to detect slight spillage of materials into the pharynx (coffee [Roots Aroma Impact, JT Co., Tokyo, Japan] and milk [Hokkaido Gyuunyuu 3.7%, Eight Cooperative Buying Co., Tokyo, Japan]; 4 g each). Each liquid was injected onto the floor of the oral cavity using a syringe.

Two kinds of jelly (4 g each) were used: Jelly A, low adhesion (V CRESC Jelly, Nutri Co., Mie, Japan) and Jelly B, high adhesion (Okunosu Nutrition Support Dessert Azuki, Horika Foods, Niigata, Japan). Each jelly was placed on the dorsum of the tongue using a spoon. The properties of the jellies were measured by the method described by Kayashita *et al*[8].

Three sets of each material were prepared for each participant and kept in a refrigerator (at 5–6°C) until just before the experiment. Liquid and jelly were then given to participants at the same temperature. The participants were asked to swallow three samples each of coffee, milk, Jelly A and Jelly B, resulting in a total of 180 swallowing samples for the 15 participants. To analyze volitional swallows, the participants were instructed to swallow the materials whenever they were ready.

The number of masticatory cycles, number of swallows, and the timing of swallows varied across the participants' volitional swallows. We selected single swallows to exclude any influence that might be caused by multiple swallows. Multiple swallows could indicate segmentation of the 4 g of material into small quantities—that is, a change in the bolus amount. In multiple swallows, total swallow duration and swallowing sound duration might be extended, and acoustic analysis of swallowing sound might be changed. We therefore excluded 46 multiple swallow samples, and 134 single swallow samples were analyzed (coffee, 39; milk, 39; Jelly A, 28; Jelly B, 28).

To detect swallowing sounds, a microphone (Lavalier Microphone COS-11DBPC, Sanken Microphone, Tokyo, Japan) was inserted into a stethoscope probe (double-faced stethoscope, No. 160, Yamasu, Kentsumedico, Saitama, Japan). The probe was placed with double-sided tape (Naisutakku, NW-K25, Nichiban, Tokyo, Japan) at the center of the lower neck just above the sternum. This location was designated as Site 6, located on the midline of the neck, immediately superior to the jugular notch. It was considered to be the most appropriate site by Takahashi *et al*[9] for recording a volume of swallowing sound similar to that at Site 11, which is over the lateral border of the trachea immediately inferior to the cricoid cartilage. Takahashi *et al* recorded swallowing sounds at 24 sites on the neck and found that Site 11 was the optimal site for detecting swallowing sounds because this site showed the greatest signal-to-noise ratio with the smallest variance[9].

The detected swallowing sound signals were amplified (AT-MA2 microphone amplifier, Audio Technica, Tokyo, Japan) and recorded simultaneously with the VE images using a digital high-definition video cassette recorder (GV-HD700, Sony, Tokyo, Japan).

Measurement of total swallow duration and swallowing sound duration

Recorded VE images and swallowing sounds were fed to a computer (PC-VJ25AAU5HJR9, NEC, Tokyo, Japan) at a frame rate of 30 frames per second and a sampling rate of 48 kHz. The total swallowing duration and the duration of each sound generated were analyzed by a system for synchronous analysis of swallowing sounds and video images (Dimagic, Tokyo, Japan).

Total swallow duration taken from the VE images started on the frame where the bolus entered the oropharynx or whenever velopharyngeal closure began (i.e. "whiteout"; the tip of the endoscope was placed just above the uvula, so we referred to "whiteout" as velopharyngeal closure), and ended with a frame where the epiglottis returned to its resting position. Total swallow duration was divided into three parts: (i) from when the bolus entered the oropharynx to the beginning of velopharyngeal closure (Fig. 2 A, B); (ii) velopharyngeal closure duration, from the beginning to the end of velopharyngeal closure (Fig. 2 B-D); and (iii) from the end of velopharyngeal closure to the return of the epiglottis to its resting position (Fig. 2 D-F). Swallowing sound duration was measured on the screen for monitoring the time waveform. The criteria for beginning and end of swallowing sounds were determined by three dentists who had 4 years of experience managing dysphagia.

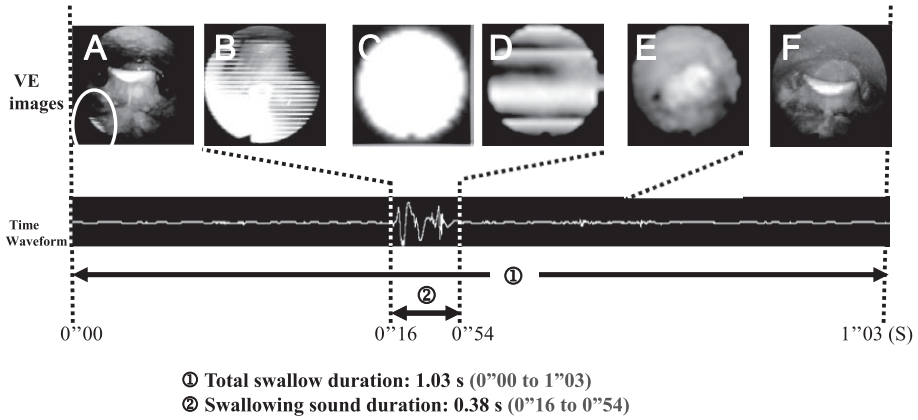

Fig. 2. Examples of total swallow duration and swallowing sound duration
Series of VE images taken during swallowing of milk: (A) bolus, indicated by the oval, entering the oropharynx; (B) beginning of velopharyngeal closure; (C) during velopharyngeal closure; (D) end of velopharyngeal closure; (E) epiglottis rotation; (F) epiglottis returning to its resting position. Total swallow duration ① is taken from the VE images starting on the frame where the bolus entered the oropharynx or whenever velopharyngeal closure began, and ending with a frame where the epiglottis returned to its resting position. Swallowing sound duration ② was measured on the time waveform showing the beginning and end of the swallowing sound.

Relationship between beginning and end of swallowing sounds and VE findings

To clarify the relationship between the beginning and end of swallowing sounds and VE findings, VE images representing the beginning and end of swallowing sounds were analyzed for each material. VE findings at the beginning of swallowing sounds were classified into three image categories: (i) before velopharyngeal closure, (ii) beginning of velopharyngeal closure, and (iii) during velopharyngeal closure. VE findings at the end of swallowing sounds were classified into four image categories: (i) during velopharyngeal closure, (ii) end of velopharyngeal closure, (iii) epiglottis rotation, and (iv) epiglottis returning to its resting position. VE findings at the beginning and end of swallowing sounds were evaluated using a six-point VE score: 1, before velopharyngeal closure; 2, beginning of velopharyngeal closure; 3, during velopharyngeal closure; 4, end of velopharyngeal closure; 5, epiglottis rotation; 6, epiglottis returning to its resting position.

Statistical analysis of total swallow duration and swallowing sound duration

To assess the differences between the materials for total swallow duration, swallowing sound duration, and the beginning and end of swallowing sounds, we used the Kruskal–Wallis one-way analysis of variance by ranks. In cases where statistically significant differences between materials were detected, we used Games–Howell multiple comparison tests to assess differences between specific pairs of materials. Analyses were conducted using SPSS Version 15.0. Data are presented as mean ± SD.

Analysis of frequency characteristics of swallowing sounds

For the entire swallowing sound duration, sound samples were extracted using a 1,024-point Hamming window with a 512-point overlap, and a time-averaged frequency waveform was

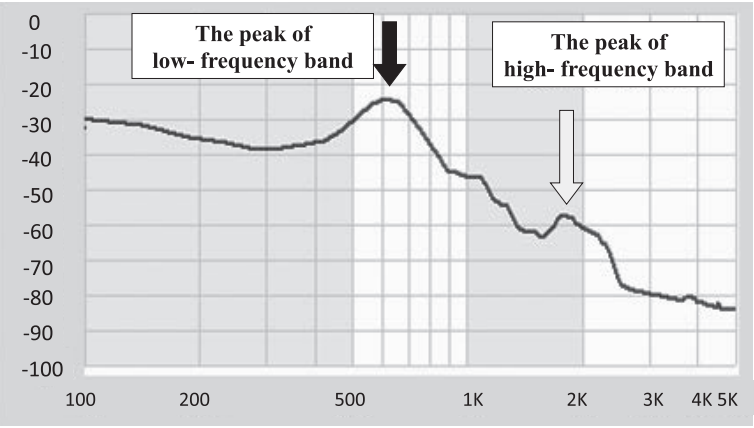

Fig. 3. Frequency characteristics of swallowing sounds
We set the low-frequency band under 1 kHz and the high-frequency band from 1 kHz to 5 kHz, and analyzed the peak frequency of each material.

obtained with fast Fourier transform analysis. We set the low-frequency band under 1 kHz and the high-frequency band from 1 kHz to 5 kHz, and analyzed the peak frequency of each material (Fig. 3).

Results

Posture of participants

The mean angle between the vertical line and Camper's plane (A) was 83.6±8.7°, and the mean angle between the vertical line and the backrest of the dental chair (B) was 16.6±5.2° (Fig. 1). This confirmed that the participants remained close to an upright posture during the VE examination.

Total swallow duration

For the 134 single swallow samples included in the analysis, the mean total swallow duration for each material was 1.25±0.33 s for coffee, 1.25±0.41 s for milk, 2.39±1.34 s for Jelly A, and 2.35±1.25 s for Jelly B (Fig. 4a). There was no significant difference between coffee and milk, or between Jelly A and Jelly B. Compared to liquids, jellies showed significantly longer total swallow duration ($p < 0.01$). The results for each section of the total swallow duration are shown in Fig. 4.

The mean duration of bolus entry to velopharyngeal closure beginning was 0.08±0.15 s for coffee, 0.15±0.32 s for milk, 1.25±1.26 s for Jelly A, and 1.19±1.18 s for Jelly B. For liquids, there was almost no time-lag between bolus entry into the oropharynx and the beginning of velopharyngeal closure. For jellies, however, there was a mean time-lag of 1.22 s. The difference between liquids and jellies was statistically significant ($p < 0.01$) (Fig. 4b).

The mean duration of velopharyngeal closure was 0.64±0.13 s for coffee, 0.68±0.15 s for milk, 0.65±0.09 s for Jelly A, and 0.67±0.10 s for Jelly B. There was no significant difference between

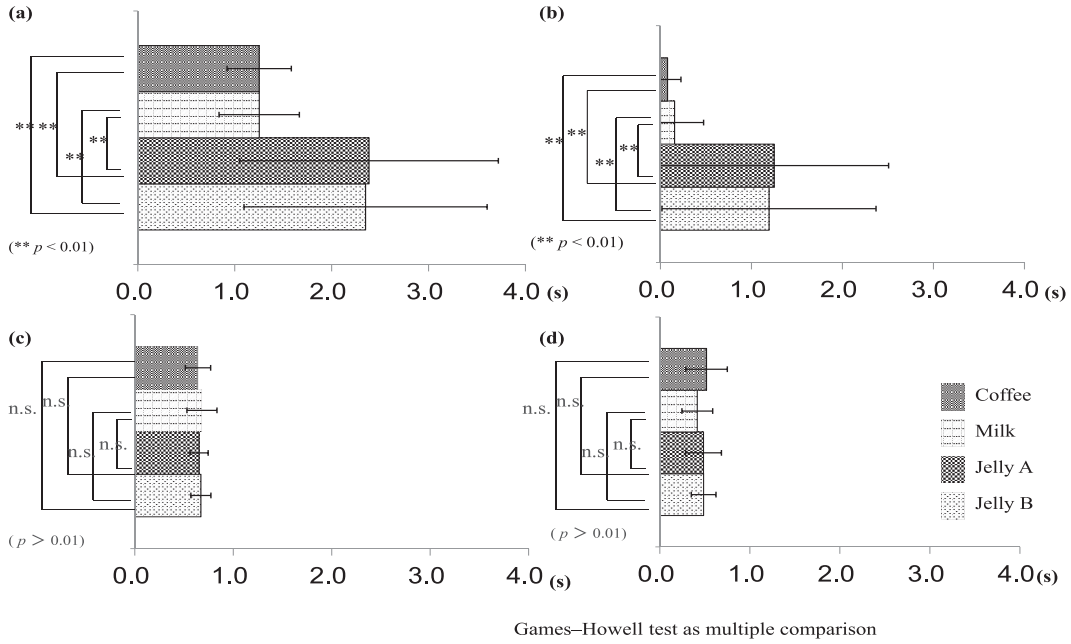

Fig. 4. Total swallow duration*

(a) There were significant differences in mean swallow duration between materials ($p < 0.05$; Kruskal–Wallis one-way analysis of variance by ranks). The mean swallow duration for jellies was significantly longer than that for liquids (coffee v Jelly A: $p = 0.001$; coffee v Jelly B: $p = 0.001$; milk v Jelly A: $p = 0.001$; milk v Jelly B: $p = 0.001$; Games–Howell multiple comparison test).

(b) There were significant differences in duration from bolus entering the oropharynx to beginning of velopharyngeal closure between materials ($p < 0.05$; Kruskal–Wallis one-way analysis of variance by ranks). The mean swallow duration for jellies was significantly longer than for liquids (coffee v Jelly A: $p < 0.001$; coffee v Jelly B: $p < 0.001$; milk v Jelly A: $p = 0.001$; milk v Jelly B: $p < 0.001$; Games–Howell multiple comparison test).

(c) There were no significant differences in velopharyngeal closure duration (from the beginning to end of velopharyngeal closure) between materials ($p = 0.827$; Kruskal–Wallis one-way analysis of variance by ranks).

(d) There were no significant differences in duration from the end of velopharyngeal closure to the epiglottis returning to its resting position between materials ($p = 0.130$; Kruskal–Wallis one-way analysis of variance by ranks).

*Error bars indicate SD. n.s : not significant.

the materials (Fig. 4c).

The mean duration from the end of velopharyngeal closure to the return of the epiglottis to its resting position was 0.52 ± 0.23 s for coffee, 0.42 ± 0.17 s for milk, 0.48 ± 0.20 s for Jelly A, and 0.49 ± 0.14 s for Jelly B. There was no significant difference between the materials (Fig. 4d).

Swallowing sound duration

The mean swallowing sound duration was 0.55 ± 0.10 s for coffee, 0.54 ± 0.11 s for milk, 0.50 ± 0.11 s for Jelly A, and 0.50 ± 0.11 s for Jelly B. There was no significant difference between the materials.

Relationship between beginning and end of swallowing sounds and VE findings

For all the materials, swallowing sounds began most frequently during velopharyngeal closure

Table 1. Videoendoscopy findings at the beginning and end of swallowing sounds*

	Beginning of swallowing sound			End of swallowing sound			
	before VC	beginning of VC	during VC	during VC	end of VC	epiglottis rotation	epiglottis returning
Coffee	5/39 (12.8)	5/39 (12.8)	29/39 (74.4)	27/39 (69.2)	2/39 (5.1)	4/39 (10.3)	6/39 (15.4)
milk	1/39 (2.5)	4/39 (10.3)	34/39 (97.4)	28/39 (71.8)	2/39 (5.1)	2/39 (5.1)	7/39 (17.9)
jelly A	5/28 (17.9)	0/28 (0.0)	23/28 (82.1)	21/28 (75.0)	2/28 (7.1)	1/28 (3.6)	4/28 (17.8)
jelly B	3/28 (10.7)	2/28 (7.2)	23/28 (82.1)	22/28 (78.5)	1/28 (3.6)	1/28 (3.6)	4/28 (14.3)
Total	14/134 (10.4)	11/134 (8.2)	109/134 (81.3)	98/134 (73.1)	7/134 (5.2)	8/134 (6.0)	21/134 (15.7)

*Data are numerator/denominator (percentage). VC : velopharyngeal closure.

Table 2. Videoendoscopy scores at beginning and end of swallowing sounds*

material	n	VE scores at beginning of swallowing sounds		VE scores at end of swallowing sounds	
		Mean	SD	Mean	SD
Coffee	39	2.5	0.8	3.7	1.2
Milk	39	2.9	0.4	3.6	1.2
Jelly A	28	2.6	0.8	3.6	1.1
Jelly B	28	2.7	0.7	3.5	1.1
total	134	2.7	0.7	3.6	1.1

*A six-point VE score was used: 1, before velopharyngeal closure; 2, beginning of velopharyngeal closure; 3, during velopharyngeal closure; 4, end of velopharyngeal closure; 5, epiglottis rotation; 6, epiglottis returning to its rest position. There was no significant difference between materials for beginning or end of swallowing VE scores (beginning: $p = 0.41$, end: $p = 0.88$; Kruskal–Wallis one-way analysis of variance by ranks).

(74.4%-87.2%), and swallowing sounds began during velopharyngeal closure for 81.3% of all samples (Table 1). For all the materials, the end of swallowing sounds finished most frequently during velopharyngeal closure. For 73.1% of all samples, end of swallowing sounds were during velopharyngeal closure (Table 1). The VE scores at the beginning and end of swallowing sounds are shown in Table 2. There was no significant difference between the materials in the beginning and end VE scores (beginning: $p = 0.41$, end: $p = 0.88$; Kruskal–Wallis one-way analysis of variance by ranks).

Acoustic analysis of swallowing sounds

Most swallowing sounds exhibited two peaks on the frequency waveform analysis (120/134 samples, 89.6%) (Fig. 5). One peak was at a low-frequency band and the other peak was at

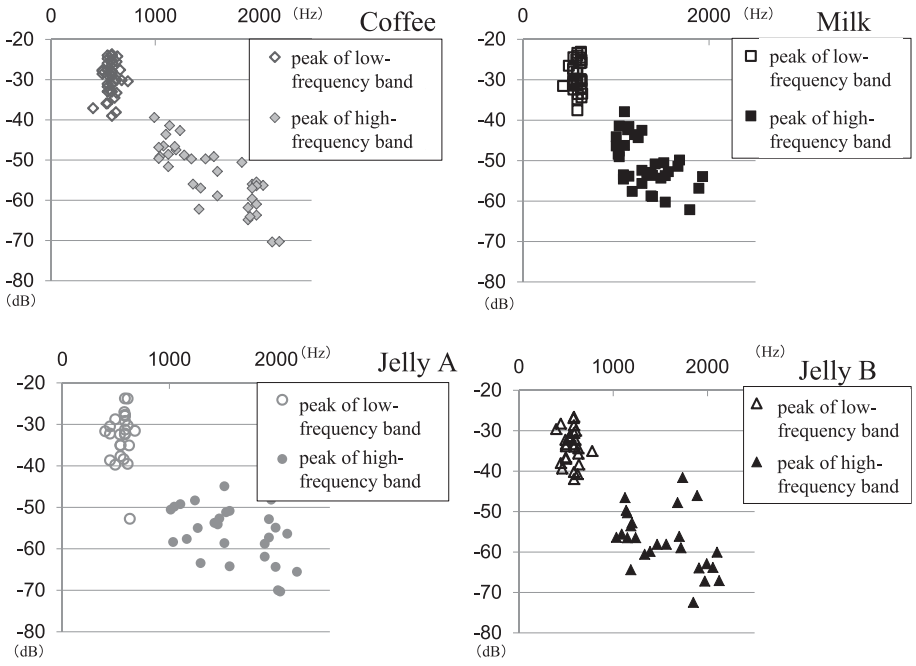

Fig. 5. Acoustic analysis of swallowing sounds
(a) Coffee (b) Milk (c) Jelly A (d) Jelly B
The adhesiveness, firmness, and cohesiveness for the jellies, measured by the method described by Kayashita *et al*[8], were 74 J/m^3, 4,845 N/m^2, and 0.42 for Jelly A, and 807 J/m^3, 12,780 N/m^2, and 0.41 for Jelly B.

a high-frequency band. The peak at the high-frequency band had a lower sound pressure level than the peak at the low-frequency band. The mean peak frequency was 569.7 Hz for the low-frequency band and 1487.8 Hz for the high-frequency band. With respect to each material, the mean peak frequency of the low-frequency band was 576.6 H z for coffee, 580.4 Hz for milk, 556.0 Hz for Jelly A, and 554.9 Hz for Jelly B. The mean peak frequency of the high-frequency band was 1550.6 Hz for coffee, 1319.2 Hz for milk, 1580.3 Hz for Jelly A, and 1510.6 Hz for Jelly B.

Multiple swallows

We investigated multiple swallows across sex and materials. Across sex, four of seven women and two of eight men swallowed multiple times. The capacity of the mouth and pharynx is smaller in women than in men, so women might have a greater tendency to swallow more than once. Across materials, multiple swallows were noted for five of 45 samples of coffee, six of 45 samples of milk, 13 of 45 samples of Jelly A, and 16 of 45 samples of Jelly B. There were more multiple swallows for jelly than liquid, probably due to smoother flow of liquid than jelly.

Mastication cycles

We counted the number of mastication cycles by observing downward motion of participants' mentum from video images recorded during the experiment. The mean number of mastication

cycles was 7.00±2.62 for Jelly A and 5.97±2.65 for Jelly B.

Discussion

We investigated the relationship between swallowing sounds and videoendoscopy (VE) images in healthy adults using 4 g of liquid or jelly. Our technique was similar to that of the "food test"—one of the most widely used assessment methods for screening swallowing function in Japan—which uses 4 g of pudding[10]. We found that many movements associated with physiologic events may be involved in the production of swallowing sounds.

Total swallow duration

Palmer *et al* reported a process model for the chew–swallow complex, stating some part of the masticated food is transferred into the oropharynx during stage II transport[11, 12]. They also reported differences in timing and location of bolus transport for "volitional swallow" versus "command swallow"[12]. In our study, participants were asked to swallow the material in their usual manner when they were ready. For the jellies, participants were asked to chew and swallow volitionally; for the liquids participants were asked to swallow volitionally without chewing. We analyzed "volitional single swallows" to exclude any influence of multiple swallows. From our results, the total swallow duration for jellies was longer than that for liquids because the jellies were transferred into the pharynx by the tongue, by stage II transport. The duration from when the bolus entered the pharynx to the beginning of velopharyngeal closure was also prolonged with jellies.

There were fewer mastication cycles for Jelly B compared to those for Jelly A. It should be noted that the total sample volume of jelly could increase during mastication because saliva was mixed with the jelly. Moreover, the properties of jelly, solids, and semi-solids could be changed during mastication because of the influence of saliva mixing and morphologic changes of materials during mastication. However, the mean number of mastication cycles was low for both jellies, so possible changes in volume and morphology are likely to be limited. It is very difficult to measure the properties of a bolus that has passed into the pharynx. One of the greatest challenges of swallowing studies is investigating the influence of changes in bolus properties.

Temperature in the mouth may have also changed the properties of the jellies during mastication. We used samples that were refrigerated before the experiment. The samples were encased in polysaccharide thickener, which might be less affected by temperature than gelatin.

There was no difference between liquids and jellies in velopharyngeal closure duration, or the duration from the end of velopharyngeal closure to the return of the epiglottis to its resting position. The results suggest that the swallow duration for jellies without a period of stage II transport was almost equal to that for liquids.

Swallowing sound duration

Takahashi *et al* investigated swallowing sounds and reported that the mean duration of swallowing sounds in adults using 5 ml of water was 0.52 s[13]. Youmans and Stierwalt investigated

swallowing sounds using 5 ml of various textured foods, and the mean duration across bolus types was 0.53 s[14]. Cichero and Murdoch indicated that there was no significant difference in the mean durations of swallowing sounds for different volumes of liquid[15]. Moriniere *et al* reported the swallowing sounds for liquids using VF[16]. Results from these previous studies are similar to those of our study. Youmans and Stierwalt reported that aging, increased bolus viscosity and increased volume of bolus prolonged swallowing sound durations[7]. In their study, they requested that participants swallowed materials when ready; however, it was unclear if they swallowed the bolus in one swallow attempt. Using solid food, the total swallow duration may be affected by command versus swallow-when-ready instructions or by single versus multiple swallows. These conditions should be clearly defined in studies that measure total swallow duration. In our study, materials with different textures were used and sound samples from single swallows were analyzed. The total swallowing sound duration did not differ significantly between materials. This might be because the adhesiveness, firmness, and cohesiveness of the jellies became closer to that of a liquid owing to mastication, even if the possible changes in volume, properties, and morphology of materials were limited.

Relationship between beginning and end of swallowing sounds and VE findings

Using sound and VF images, Lear *et al* suggested that swallowing sounds may arise from the apposition and parting of the mucous membrane while the bolus flows into the pyriform sinus[18]. In our investigation, the beginning of swallowing sounds did not correspond with the time of the bolus reaching the pyriform sinus. Our results showed that swallowing sounds almost always began during the initial period of velopharyngeal closure, and almost always ended before completion of velopharyngeal closure. Several other studies have investigated the origin of swallowing sounds, with the common assumption that swallowing sounds were associated with laryngeal elevation, passing of the bolus through the upper esophageal sphincter (UES) and laryngeal descent[19–21].

Logemann *et al* investigated normal swallowing using VF and VE[22]. They found that both procedures could be used to visualize the duration of pharyngeal delay and laryngeal elevation. They also found that swallowing events were most often seen 1–2 video frames (0.03–0.06 s) later on VF than on VE. This suggests that the timing of swallowing sounds versus swallow events observed using VE might be faster than those observed using VF.

However, few studies have investigated swallowing sounds using VE images. Leslie *et al* used simultaneous VE and respiration monitoring to analyze the relationship between respiration and epiglottic movement[23]. They suggested an association between swallowing sounds and epiglottic movement, but found no evidence of a causal link.

In our investigation, swallowing sounds did not correspond with epiglottic movement (rotation and returning). Hamlet *et al* also reported that swallowing sounds in patients who had undergone total laryngectomy did not show acoustic differences for liquid versus paste swallows, as found in control subjects[19]. Thus, epiglottic movement might be unrelated to the occurrence of swallowing sounds.

VF cannot be used to visualize lateral pharyngeal wall contraction by the lateral nor frontal view, or to observe whether or not velopharyngeal closure occurs completely. Therefore, the relationships between swallowing sounds and velopharyngeal closure are still unclear, but we believe that VE is suitable for investigating these relationships. In our study, swallowing sounds for all materials were produced most frequently during velopharyngeal closure, suggesting that velopharyngeal closure was unrelated to the occurrence of swallowing sounds.

Physiologic events including laryngeal excursion and opening of the UES may be involved in production of swallowing sounds, but we could not confirm this in our investigation. Simultaneous VE and VF investigation with a multiple microphone system that can detect swallowing sounds from several sites on the neck may help verify the mechanisms involved in production of swallowing sounds.

Acoustic analysis of swallowing sounds

Hamlet *et al* reported an acoustic analysis of sounds produced by swallowing 10 ml of liquid using an accelerometer and simultaneous VF imaging[24]. They found a spectral shift in the normal swallow. The earlier spectra had an energy peak in the low frequencies (an average center frequency of 556 Hz), and the spectra 40 ms later displayed an energy peak in the high frequencies (an average center frequency of 1,384 Hz). Hamlet *et al* also reported that there were three components in a normal swallow and that spectral shift occurred in the second component[19]. They noted that the first, weaker component was generally associated with laryngeal elevation and bolus flow through the pharynx, and the third, weaker component was associated with laryngeal descent after the swallow. The first and third components were not always present. The strong second component corresponds temporally with flow in the hypopharynx and through the UES.

In our study, most swallowing sounds (89.6%) had two peaks on frequency waveforms. Regardless of the swallowed materials, the peak was around 570 Hz for the low-frequency band, and around 1,500 Hz for the high-frequency band. The ranges of these peaks were similar to those of Hamlet *et al*[24]; however, these peaks were observed at almost the same time in the windowed segment of the time wave and no spectral shift was observed. These results might be due to the differences in bolus volume, or the manner of swallow and posture; however, further research using new procedures, as described above, is needed to verify the acoustic characteristics of swallowing sounds.

These results show that many movements associated with physiologic events—including hyoid bone and laryngeal excursion, and opening of the upper esophageal sphincter—may be involved in the production of swallowing sounds.

Acknowledgements

This study was supported by a Grant-in-Aid for Young Scientists from the Japan Society for the Promotion of Science, KAKENHI (24292361). There are no conflicts of interest to disclose.

Conflict of interest disclosure

The authors declare no conflict of interest.

References

1) Vice FL, Heinz JM, Giuriati G, *et al*. Cervical auscultation of suckle feeding in newborn infants. *Dev Med Child Neurol*. 1990;**32**:760-768.
2) Zenner PM, Losinski DS, Mills RH. Using cervical auscultation in the clinical dysphagia examination in long-term care. *Dysphagia*. 1995;**10**:27-31.
3) Borr C, Hielscher-Fastabend M, Lucking A. Reliability and validity of cervical auscultation. *Dysphagia*. 2007;**22**:225-234.
4) Takahashi K, Uyama R, Hirano K, *et al*. Accuracy of cervical auscultation for detecting dysphagia in head and neck cancer patients. *Head Neck Cancer*. 2001;**27**:198-203.
5) Bastian RW. Videoendoscopic evaluation of patients with dysphagia: an adjunct to the modified barium swallow. *Otolaryngol Head Neck Surg*. 1991;**104**:339-350.
6) Bastian RW. The videoendoscopic swallowing study: an alternative and partner to the videofluoroscopic swallowing study. *Dysphagia*. 1993;**8**:359-367.
7) Kelly AM, Drinnan MJ, Leslie P. Assessing penetration and aspiration: how do videofluoroscopy and fiberoptic endoscopic evaluation of swallowing compare? *Laryngoscope*. 2007;**117**:1723-1727.
8) Kayashita J, Kanaya S, Yamagata Y, *et al*. Dysphagia diets pyramid-lebel food on market 250. Tokyo: Ishiyaku Shuppan; 2008. (in Japanese).
9) Takahashi K, Groher ME, Michi K. Methodology for detecting swallowing sounds. *Dysphagia*. 1994;**9**:54-62.
10) Saito E. Food test. In Annual report of Grant-in-Aid for Scientific Research (KAKENHI), Comprehensive Research on Aging and Health. Tokyo: Ministry of Health and Welfare; 2000. pp 1-17. (in Japanese).
11) Palmer JB, Hiiemae KM. Integration of oral and pharyngeal bolus propulsion: a new model for the physiology of swallowing. *Jpn J Dysphagia Rehabil*. 1997;**1**:15-30.
12) Palmer JB, Hiiemae KM, Matsuo K, *et al*. Volitional control of food transport and bolus formation during feeding. *Physiol Behav*. 2007;**91**:66-70.
13) Takahashi K, Groher ME, Michi K. Symmetry and reproducibility of swallowing sounds. *Dysphagia*. 1994;**9**:168-173.
14) Youmans SR, Stierwalt JA. An acoustic profile of normal swallowing. *Dysphagia*. 2005;**20**:195-209.
15) Cichero JA, Murdoch BE. Acoustic signature of the normal swallow: characterization by age, gender, and bolus volume. *Ann Otol Rhinol Laryngal*. 2002;**111**:623-632.
16) Moriniere S, Boiron M, Alison D, *et al*. Origin of the sound components during pharyngeal swallowing in normal subjects. *Dysphagia*. 2008;**23**:267-273.
17) Youmans SR, Stierwalt JA. Normal swallowing acoustics across age, gender, bolus viscosity, and bolus volume. *Dysphagia*. 2011;**26**:374-384.
18) Lear CS, Flanagan JB Jr, Moorrees CF. The frequency of deglutition in man. *Arch Oral Biol*. 1965;**10**:83-100.
19) Hamlet SL, Patterson RL, Fleming SM, *et al*. Sounds of swallowing following total laryngectomy. *Dysphagia*. 1992;**7**:160-165.
20) Boiron M, Rouleau P, Atipo B, *et al*. Esophageal swallowing phase assessed by audiosignal recording: relationship with manometry in gastroesophageal reflux disease patients. *Dig Dis Sci*. 1999;**44**:529-535.
21) Sarraf-Shirazi S, Baril JF, Moussavi Z. Characteristics of the swallowing sounds recorded in the ear, nose and on trachea. *Med Biol Eng Comput*. 2012;**50**:885-890.
22) Logemann JA, Rademaker AW, Pauloski BR, *et al*. Normal swallowing physiology as viewed by videofluoroscopy and videoendoscopy. *Folia Phoniatr Logop*. 1998;**50**:311-319.

23) Leslie P, Drinnan MJ, Zammit-Maempel I, *et al*. Cervical auscultation synchronized with images from endoscopy swallow evaluations. *Dysphagia*. 2007;**22**:290–298.
24) Hamlet SL, Nelson RJ, Patterson RL. Interpreting the sounds of swallowing: fluid flow through the cricopharyngeus. *Ann Otol Rhinol Laryngol*. 1990;**99**:749–752.

[Received May 8, 2015 : Accepted October 8, 2015]