

---

## 原著論文

---

### Swallowing Sound Waveform and Its Clinical Significance: Evaluation Using Ultrasonography

Tsuyoshi HONDA<sup>1)</sup>, Takuro BABA<sup>1)</sup>, Keiko FUJIMOTO<sup>1)</sup>, Kan NAGAO<sup>1)</sup>,  
Akira TAKAHASHI<sup>2)</sup>, Tetsuo ICHIKAWA<sup>1)</sup>

キーワード : Dysphagia, Swallowing sound, Ultrasonography

**Abstract : Background:** Cervical auscultation is a technique frequently used for the screening of dysphagia. However, this method is difficult to evaluate objectively and it is unclear how sound is generated during the swallowing process. The aim of this study was to analyze the waveform of swallowing sound and clarify the sound production process using recordings of swallowing sounds and ultrasound images (USI), performed simultaneously.

**Materials and Methods:** Commercial natural spring water and natural carbonated water were used in experiments 1 and 2, respectively. In experiment 1, a microphone was attached to the skin of the neck of 20 young participants and swallowing sounds were recorded and analyzed. In experiment 2, swallowing processes in three participants were recorded by a medical ultrasonography apparatus. The ultrasonic probe was placed on the skin over one of the thyroid cartilages or the thyroid gland.

**Results:** The swallowing sound wave (SSW) was divided into three sectional periods. The mean duration of the first, second, and third SSW was  $210 \pm 147$  ms,  $458 \pm 113$  ms, and  $91 \pm 61$  ms, respectively. The mean intensity ratio of the first, second, and third SSW was  $7.8 \pm 5.2$ ,  $29.2 \pm 16.5$ , and  $5.8 \pm 5.1$ , respectively. When the ultrasonic probe was placed on the skin over one of the thyroid cartilages, in the phase between the production of the second SSW and the silent period, the USI revealed an accumulation of swallowed material around the valleculae and oropharynx. In the silent period of the second SSW, the swallowed material accumulated around the hypopharynx. When the ultrasonic probe was placed on the skin over the thyroid gland, in the silent period of the second SSW, the USI revealed that the swallowed material had passed through esophagus.

**Conclusion:** Waveform and USI findings from this study suggest that swallowing sound can be divided into three sectional periods: an oral phase, a pharyngeal phase, and a repositioning phase.

#### I . Introduction

Recently, the aging rate of the global population has been increasing. Japan, in particular, is becoming a super-aged society; in 2010<sup>1)</sup> the elderly ( $\geq 65$  years old) comprised over 23% of the population<sup>2)</sup>. Dysphagia is one of the major disabilities and diseases associated with aging. It reduces the

patient's quality of life through malnutrition<sup>3, 4)</sup> and aspiration pneumonia<sup>4, 5, 6)</sup>. The serious and critical consequences of dysphagia also reduce activities of daily living. Therefore, it is important to evaluate the risk of dysphagia early. A variety of clinical examinations have been proposed as screening methods for dysphagia; among these, the repetitive saliva

---

<sup>1)</sup> Department of Oral and Maxillofacial Prosthodontics, Institute of Health Biosciences, The University of Tokushima Graduate School

<sup>2)</sup> Department of Oral Surgery, Institute of Health Biosciences, The University of Tokushima Graduate School

swallowing test<sup>7, 8)</sup>, the modified water swallow test<sup>9)</sup>, and the food test<sup>10)</sup> are commonly used. In addition, cervical auscultation, a technique for listening to and assessing swallowing sounds using a stethoscope, is a very simple and effective method also commonly utilized<sup>11-18)</sup>. However, this technique is difficult to evaluate objectively, and it is unclear how sound is generated during the swallowing process. The relationship between the swallowing sound and process has to be clarified to ensure an adequate examination of dysphagia with the cervical auscultation.

Imaging examinations, in which the movements of swallowing and bolus from the oral cavity to the esophagus can be directly detected, are also very useful for the examination of dysphagia. In particular, video-fluoroscopic examination is considered the gold-standard for evaluating dysphagia<sup>19)</sup>; however, a limitation of this method is the associated radiation risk<sup>20)</sup>. Ultrasonography, an ultrasound-based imaging technique for visualizing internal body structures, is a safe alternative method because it does not produce ionizing radiation<sup>21, 22)</sup>. There has been no study on the measurements of the swallowing process and sound using ultrasonography.

The aim of this study was to determine the swallowing sound production process by analyzing the waveform of swallowing sound recordings and ultrasound images (USI) taken simultaneously and to validate the cervical auscultation for examinations of dysphagia.

## II. Materials and Methods

### 1. Participants

Participants were randomly recruited from undergraduates and graduate students in the University of Tokushima and were selected from a group of volunteers based on the following inclusion criteria: i) normal dentition, ii) no history of stomatognathic disorder and dysphagia, iii) more than 3 of repetitive saliva swallowing test value (RSST value), iv) informed consent obtained before inclusion in this study. Twenty-three young and healthy volunteers were enrolled as participants in this study (experiment 1: 10 men and 10 women, mean age = 25.8 years; experiment 2: men and one woman, mean age = 28.3 years). This research was conducted with the approval by the clinical research ethical review board of Tokushima University Hospital (No.1406).

### 2. Swallowing materials

A Commercial, natural spring water (CRYSTAL GEYSER<sup>®</sup>, Otsuka Foods, Japan) and natural carbonated water (Minami Alps Tennensui Sparkling<sup>®</sup>, Suntory Beverage & Food, Tokyo,

Japan) at approximately 25°C were used in experiments 1 and 2, respectively.

### 3. Recording of swallowing sounds

The swallowing sound recording system is outlined in Figure 1. An ultra-small condenser microphone (AT9903, Audio-Technica, Tokyo, Japan; frequency characteristic: 30-18000 Hz) was used to detect swallowing sounds. The tip of the sensor was enlarged to 10 mm in diameter using an auto-curing resin (UNIFAST II Clear, GC, Tokyo, Japan) for ease of attachment to the skin. The microphone was attached to the skin of the neck using double-sided tape (1517, 3M Japan, Tokyo, Japan). The output signal was amplified through an amplifier (AT-MA2, Audio-Technica) and a low-pass filter (NF 3611, NF CORPORATION, Yokohama, Japan; fc: 10 kHz, -24 dB ± 2 dB dB/oct). The signals were digitized as 16 bit at a 20 kHz sampling rate using an analog-digital card (BPC-0600, Interface Corporation, Hiroshima, Japan), and analyzed using a waveform analysis software (DADiSP, CAE Solutions Corporation, Tokyo, Japan) on a personal computer.

### 4. Ultrasound imaging

The swallowing processes were recorded by a medical ultrasonography apparatus (LOGIQ Book XP Enhanced, GE Health Care Japan, Tokyo, Japan). The scanning probe (GE 8L-RS Linear Array Ultrasound Probe, Frequency: 6 MHz, Dynamic range: 90 dB, Gain: 58 dB) was placed on the skin over one of the thyroid cartilages or the thyroid gland, so as to obtain axial transverse images.

Before the recordings, anatomical structures and swallowed material in the USI were identified through video-endoscopy (FNL-10RP3, HOYA, Tokyo, Japan). Figure 2 presents the USI of three swallowing materials; carbonated water was observed as a white, bubble-like feature. The chewable, green-colored, moderately thickened water (100 ml pure water containing 2 g of a thickening agent [Tsururinko Quickly<sup>®</sup>, CLINICO, Tokyo, Japan]) which accumulated around the valleculae during chewing, was more clearly visible as a white, bubble-like feature.

Figures 3 and 4 present the USI of the thyroid cartilage and thyroid gland positions, respectively. In the thyroid cartilage position, the accumulation of swallowed material around the valleculae and oropharynx was observed as highly echogenic regions in the USI. These regions passed through the hypopharynx and were therefore found on the equivalent site on the USI. In the thyroid gland position, the esophagus was clearly observed in the USI because there was little acoustic shadow of the thyroid and cricoid cartilages. When

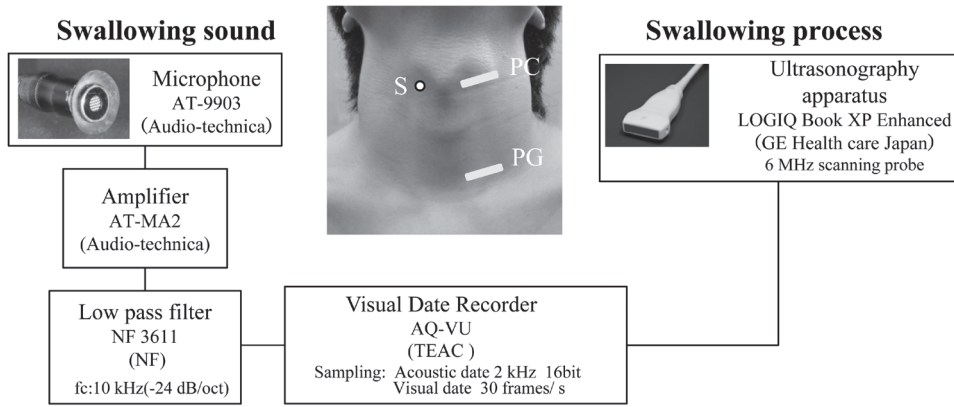


Fig. 1 An outline of the swallowing sound and ultrasound images recording system  
S: Swallowing sound recording site, PC: Probe site (thyroid cartilage), PG: Probe site (thyroid gland)

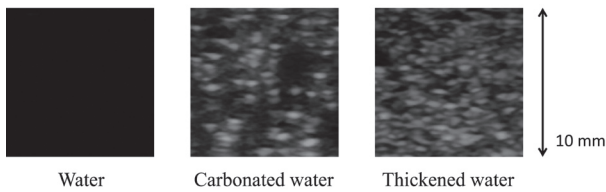


Fig. 2 Comparison of ultrasound images of three swallowing materials: water, carbonated water, and thickened water.

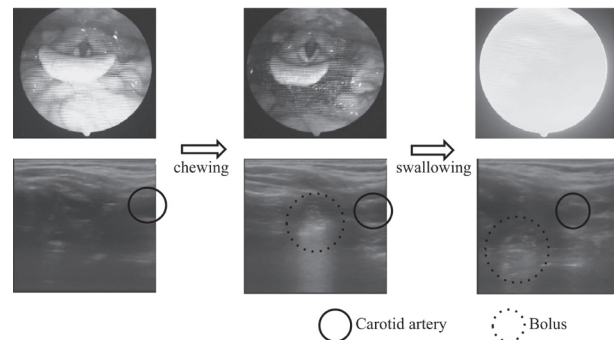


Fig. 3 Ultrasound images around the pharynx when the ultrasonic probe was positioned on the skin over the thyroid cartilage.

the swallowed material passed through the esophagus, the esophagus was observed as a highly echogenic region on the USI.

### 5. Measurement procedures

In experiment 1, participants were seated in a soundproof room, on a straight-backed chair and in an upright position. The microphone was attached to the skin over the lateral laryngeal prominence (Figure 1). Participants were asked to swallow 5 ml of water (administered into the sublingual region using a syringe) in one complete swallow, with the chin in the normal position. The recording was initiated 2 s before the investigators commanded to swallow water and stopped 2 s after completion of the swallow, as determined by the participants. Measurements for each condition were repeated three times, separated by an interval of 3 min.

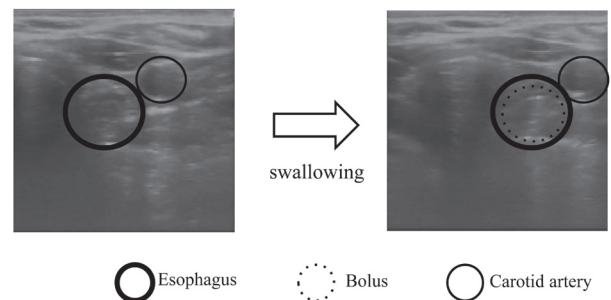


Fig. 4 Ultrasound images around the esophagus when the ultrasonic probe was positioned on the skin over the thyroid gland.

In experiment 2, the swallowing processes were recorded using a medical ultrasonography apparatus to detect swallowing sound. The scanning probe was placed on the skin, to the left, over one of thyroid cartilages or the thyroid gland. The participants were asked to swallow 5 ml of carbonated water, in a same manner as in experiment 1. The two acoustic and visual signals were transferred to a visual recorder (AQ-

VU, TEAC, Tokyo, JAPAN). The visual signals were recorded as 30 frames per second and acoustic signals were recorded as 16 bit at a 2 kHz sampling rate. Images for each swallowing sound event were evaluated.

### III. Results

Figure 5 shows a typical swallowing sound waveform

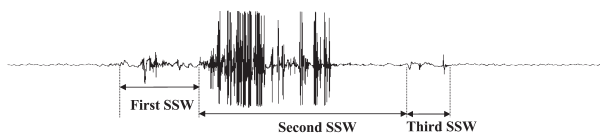
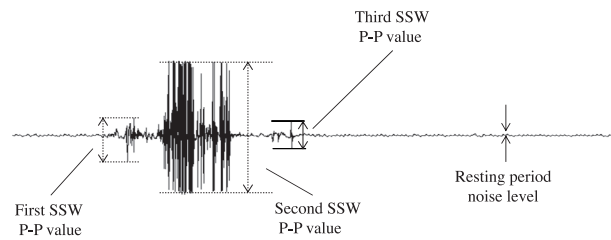


Fig. 5 A typical swallowing sound waveform and its three sectional periods.



Each SSW sound intensity ratio = Each SSW P-P value/ noise level

Fig. 6 Measurements of each SSW sound intensity ratio

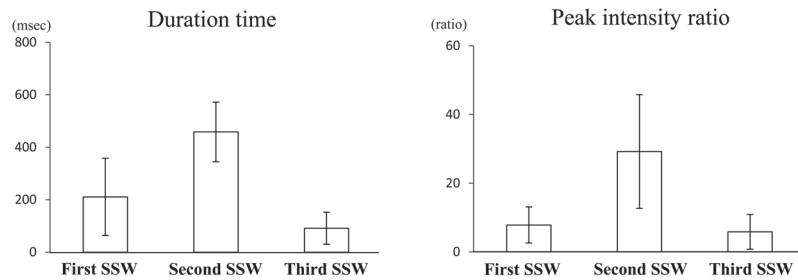


Fig. 7 Mean duration and peak intensity ratio of the first, second and third swallowing sound waves.

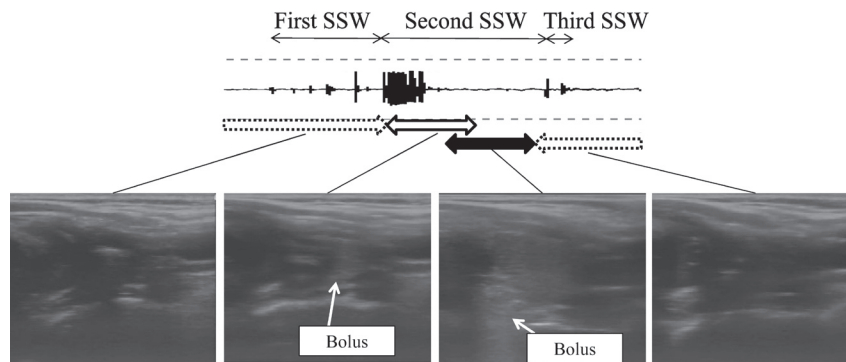


Fig. 8 Typical sonograms in the thyroid cartilage position corresponding to each swallowing sound event. In the period from the production of the second swallowing sound wave to the silent period, the swallowed material accumulated around the valleculae and oropharynx. In the latter period of the second swallowing sound wave, the swallowed material accumulated around the hypopharynx.

based on the overall data from the 20 participants. The waveform was divided into three sectional periods through the overall inspections of digital signal date. The first swallowing sound wave (first SSW) was defined as the period from the beginning of the swallowing wave (first peak) to the beginning of the following large peaks, the second SSW was the period from the beginning of the large peak to the beginning of the third peak, and the third SSW was the period from the beginning of third peak to the offset of the swallowing wave. Each first and second SSW comprised a swallowing wave production followed by a silent period in the latter part. The segmentations of each SSW were interactively generated by one examiner with a comparison to the noise level of the

resting period. Duration and peak intensities of each SSW were measured in experiment 1. The peak intensity was evaluated as the proportion of peak-to-peak value (P-P value) to the noise level of the resting period (Figure 6).

The mean duration and peak intensity ratio of the three SSWs when 5 ml water was swallowed is presented in Figure 7. The mean duration of first, second, and third SSW was  $210 \pm 147$  ms,  $458 \pm 113$  ms, and  $91 \pm 61$  ms, respectively, and the overall duration of swallowing sound was  $739 \pm 179$  ms. The mean intensity ratio of the first, second, and third SSW was  $7.8 \pm 5.2$ ,  $29.2 \pm 16.5$ , and  $5.8 \pm 5.1$ , respectively.

Figure 8 displays the typical USI for each swallowing sound waveform event which occurred when the ultrasonic

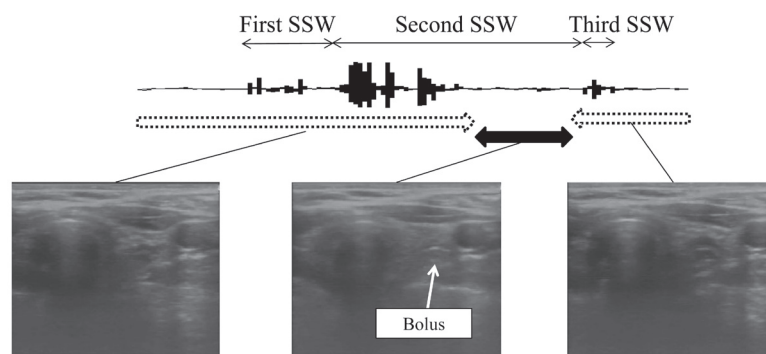


Fig. 9 Typical sonograms in the thyroid gland position corresponding to each swallowing sound event. In the latter period of the second swallowing sound wave, the swallowed material accumulated around esophagus.

probe was placed on the skin over one of the thyroid cartilages. In the phase between the production of the second SSW to the silent period, the USI revealed an accumulation of the swallowed material around the valleculae and oropharynx. The swallowed material then passed through the hypopharynx in the silent period of the second SSW. There was no swallowed material was observed on the USI in the first and third SSW.

Similar findings were found in the three participants.

Typical sonograms for each swallowing sound waveform event when the ultrasonic probe was placed on the skin over the thyroid gland are presented in Figure 9. In the silent period of the second SSW, the swallowed material passed through esophagus, while in the first and third SSW, and the production period of the second SSW, there was no swallowed material was observed on the USI. Similar findings were found in the three participants.

#### IV. Discussion

A variety of models on swallowing dynamics have been proposed, and the three-stage theory comprising oral, pharyngeal and esophageal phases is generally accepted<sup>23)</sup>. The solid bolus translation for swallowing is also understood in the same way. In addition, a five-stage model for ingestion/swallowing was reported by Leopold et al.<sup>24)</sup> as follows: pre-oral (anticipatory), preparatory, oral (propulsive), pharyngeal, and esophageal. Furthermore, Palmer<sup>25)</sup> proposed a process model in which a proportion of the solid bolus is translated around the valleculae during oral processing. In this present study, swallowing processes at the time of the swallow water command followed the five-stage process model.

The swallowing sound was divided into three sectional periods. The first SSW, with a duration of 210 ms, was the first component of the overall swallowing sound. While there

have been no previous studies on swallowing sound in the oral stage, Leonard et al.<sup>26)</sup> reported that the latency between the first movement of the bolus tip past the posterior nasal spine which leads to a swallow, and the first hyoid elevation was 210 ms. This latency corresponds to the duration of first SSW recorded in this present study. The findings from the sonogram in this period suggest that sound is generated when the swallowed material is held in the oral cavity and then transferred to the pharynx. These results indicate that posterior movements of the tongue and hyoid bone are the swallowing movements corresponding to the first SSW. Although the definition of the first SSW as a sound prior to the pharyngeal reflex may be questionable, the wave profile in addition to the USI findings, indicate that it should be interpreted as the oral stage.

The second SSW, with a duration of 460 ms, was the middle component of the swallowing sound and consisted of continuous and large waveforms in the early part, and a silent period in the latter part. In previous studies<sup>27-33)</sup>, the swallowing sound which occurs in the pharyngeal stage was found to be focused and associated with two major events: the movement of the larynx and epiglottis closure, and the passing of the bolus from the pharynx to the esophagus. In this present study, these two events could not be separated in the wave profile, despite analysis of the relevant USI. Boiron et al.<sup>31)</sup> found that the mean total duration of the pharyngeal stage was 620 ms when 5 ml water was swallowed and Hammoudi et al.<sup>33)</sup> reported a that mean total duration was  $441 \pm 150$  msec for the swallowing of solid material. These durations would be corresponded to the total duration of second SSW and third SSW in this study. Therefore, the second SSW would be generated when the pharyngeal reflex occur when the swallowing material is transferred from the pharynx to the esophagus. These findings suggest that sound production of the second SSW is related to the larynx movements beginning

with elevation of the hyoid bone, and followed by closure of the epiglottis and the passage of the swallowed material from the pharynx to the esophagus.

The third SSW, the last component of the swallowing sound, was observed after the material flowed into the esophagus; it had a mean total duration of less than 100 ms, with a weak and short waveform. Moriniere et al.<sup>28)</sup> reported that the mean duration between the descent and opening of larynx was 72 ms. The sound production of the third SSW could therefore be related to repositioning of the hyoid bone, the larynx, and the epiglottis.

Previous studies have demonstrated that the intensity of swallowing sound is dependent on age, gender, viscosity, and the volume of food bolus; however, no studies have evaluated the relative intensity of each swallowing sound component. According to Klahn et al.<sup>34)</sup>, and Perlman et al.<sup>35)</sup>, peak sound intensity occurs when the upper sphincter muscle opens completely and the bolus passes through esophageal orifice; this period corresponds to the second SSW in this present study.

## V. Conclusion

In conclusion, the waveform and USI findings from this study suggest that swallowing sound can be divided into three sectional periods which are associated with each process of a serial swallowing movements: posterior movements of the tongue and hyoid bone as an oral phase, larynx movements with the hyoid bone elevation, the epiglottis closure and the bolus passage from the pharynx to the esophagus as a pharyngeal phase, repositioning of the hyoid bone, and epiglottis as a repositioning phase. This finding could be effective to validate the cervical auscultation for a better examination.

## References

- 1) United Nations, Department of Economic and Social Affairs Population Division, Population Estimates and Projections Section: World Population Prospects: The 2012 Revision.
- 2) Paillard-Borg S1 and Strömberg L: The importance of reciprocity for female caregivers in a super-aged society: a qualitative journalistic approach. *Health Care Women* 35, 11-12 (2014)
- 3) Romco J, Warnberg J, Pozo T and Marcos A: Physical activity, immunity and infection. *Proc Nutr Soc* 69(3), 390-399 (2010)
- 4) Smith HA, O'Neill PA, Park C, England R, Renwick DS, Wyatt R, Morris J and Martin DF. North West Dysphagia Group: Can bedside assessment reliably exclude aspiration following acute stroke? *Age ageing* 29, 99-106 (1998)
- 5) Shaker R and Geenen JE: Management of dysphagia in stroke patients. *Gastroenterol Hepatol* 7(5), 308-332 (2011)
- 6) Martino R, Foley N, Bhogal S, Diamant N, Speechley M and Teasell R: Dysphagia after stroke incidence, diagnosis, and pulmonary complications. *Stroke* 36, 2756-2763 (2005)
- 7) Oguchi K, Saitoh E, Mizuno M, Baba M, Okui M and Suzuki M: The Repetitive Saliva Swallowing Test (RSST) as a Screening Test of Functional Dysphagia (1) Normal Values of RSST. *Jpn J Rehabil Med* 37(6), 375-382 (2000)
- 8) Oguchi K, Saitoh E, Baba M, Kusudo S, Tanaka T and Onogi K: The Repetitive Saliva Swallowing Test (RSST) as a Screening Test of Functional Dysphagia (2) Validity of RSST. *Jpn J Rehabil Med* 37(6), 383-388 (2000)
- 9) Saito E: An Integrated Research on Treatment and Handling of Dysphagia. General Research Report. Research Project on Aging and Health, Fiscal 1999 Health and Labour Sciences Research Grant Apr, 1-18 (2000)
- 10) Tohara H, Saito E, Mays KA, Kuhlemeier K and Palmer JB: Three tests for predicting aspiration without videofluorography. *Dysphagia* 18 (2), 126-134 (2003)
- 11) Bosma JF: Sensorimotor examination of the mouth and pharynx. *Front Oral Physiol* 2, 78-107 (1976)
- 12) Zenner PM, Losinski DS and Millis RH: Using cervical auscultation in the clinical dysphagia examination in long-term care. *Dysphagia* 10(1), 27-31 (1995)
- 13) Takahashi K, Groher ME and Michi K: Methodology for detecting swallowing sounds. *Dysphagia* 9, 54-62 (1994)
- 14) Mackwiak RC, Brenman HS and Friedman MHF: Acoustic profile of deglutition. *Proc Soc Exp Biol Med* 125, 1149-1152 (1967)
- 15) Selly WG, Flack FC, Ellis RE and Brooks WA: The Exeter Dysphagia Assessment Technique. *Dysphagia* 4, 227-235 (1990)
- 16) Hamlet SL, Patterson RL, Fleming SM and Jones LA: Sounds of swallowing following total laryngectomy. *Dysphagia* 7, 160-165 (1992)
- 17) Leslie P, Drinnan MJ, Finn P, Ford GA and Wilson JA: Reliability and validity of cervical auscultation: a controlled comparison using videofluoroscopy. *Dysphagia* 19, 231-240 (2004)
- 18) Borr C, Hielscher-Fastabend M and Lücking A: Reliability and validity of cervical auscultation. *Dysphagia* 22, 225-234 (2007)

- 19) Langmore SE, Schatz K and Olson N: Endoscopic and videofluoroscopic evaluations on swallowing and aspirations. *Ann Otol Rhinol Laryngol* 100, 678-681 (1991)
- 20) Cook IJ, Dodds WJ, Dantas RO, Massey B, Kern MK, Lang IM, Bresser JG and Hogan WJ: Opening mechanisms of the human upper esophageal sphincter. *Am J Physiol* 257, 748-759 (1989)
- 21) Morrisich KA, Stone M, Sonies BC, Kurtz D and Shawker T: Characterization of tongue shape. *Ultrasound Imaging* 6, 37 (1984)
- 22) Peng CL, Jost-Brinkmann PG, Miethke RR and Lin CT: Ultrasonographic measurement of tongue movement during swallowing. *J Ultrasound Med* 19, 15-20 (2000)
- 23) Palmer JB and Hiemae KM: Ingestion of oral and pharyngeal bolus propulsion-a new model for the physiology of swallowing. *Jpn J Dysphagia rehab* 1, 25-30 (1997)
- 24) Leopold NA and Kagel MC: Swallowing ingestion and dysphagia. *Arch Phys Med Rehabil* 64, 371-373 (1983)
- 25) Palmer JB: Bolus aggregation in the oropharynx does not depend on gravity. *Arch Phys Med Rehabil* 79, 691-696 (1998)
- 26) Leonard R and McKenzie S: Hyoid-Bolus Transit Latencies in Normal Swallow, *Dysphagia* 21(3), 183-190 (2006)
- 27) Selly WG, Ellis RE, Flack FC, Bayliss CR and Pearce VR: The synchronization of respiration and swallow sounds with videofluoroscopy during swallowing. *Dysphagia* 9 (3), 162-167 (1994)
- 28) Morinière S, Beutter P and Boiron M: Sound component duration of healthy human pharyngoesophageal swallowing: a gender comparison study. *Dysphagia* 21, 175-182 (2006)
- 29) Morinière S, Boiron M, Alison D, Makris P and Beutter P: Origin of the sound components during pharyngeal swallowing in normal subjects. *Dysphagia* 23, 267-273 (2008)
- 30) Cichero JAY and Murdoch BE: Acoustic signature of the normal swallow: Characterization by age, gender and bolus volume. *Ann Otol Rhinol Laryngol* 111, 623-632 (2002)
- 31) Boiron M, Rouleau P and Metman EH: Exploration of pharyngeal swallowing by audiosignal recording. *Dysphagia* 12, 86-92 (1997)
- 32) Youmans SR and Stierwalt JA: Normal swallowing acoustics across age, gender, bolus viscosity and bolus volume. *Dysphagia* 26, 374-384 (2011)
- 33) Hammoudi K, Boiron M, Hernandez N, Bobillier C and Morinière S: Acoustic Study of Pharyngeal Swallowing as a Function of the Volume and Consistency of the Bolus. *Dysphagia* 29 (4), 468-474 (2014)
- 34) Klahn MS and Perlman AL: Temporal and durational patterns associating respiration and swallowing. *Dysphagia* 14, 131-138 (1999)
- 35) Perlman AL, He X, Barkmeier J and Van Leer E: Bolus location associated with videofluoroscopic and respirodeglutometric events. *J Speech Hear Res* 48, 21-33 (2005)