Dysphagia 22:225–234 (2007)
DOI: 10.1007/s00455-007-9078-3

Dysphagia

Reliability and Validity of Cervical Auscultation

Christiane Borr, MA,[1] Martina Hielscher-Fastabend, HD, Dr. Phil,[1] and Andy Lücking, MA[2]
[1]Department of Clinical Linguistics, Faculty of Linguistics and Literary Studies, Bielefeld, Germany; [2]CRC 673 "Alignment in Communication," B1, Bielefeld University, Bielefeld, Germany

Abstract. We conducted a two-part study that contributes to the discussion about cervical auscultation (CA) as a scientifically justifiable and medically useful tool to identify patients with a high risk of aspiration/penetration. We sought to determine (1) acoustic features that mark a deglutition act as dysphagic; (2) acoustic changes in healthy older deglutition profiles compared with those of younger adults; (3) the correctness and concordance of rater judgments based on CA; and (4) if education in CA improves individual reliability. The first part of the study focused on a comparison of the "swallow morphology" of dysphagic as opposed to healthy subjects' deglutition in terms of structure properties of the pharyngeal phase of deglutition. We obtained the following results. The duration of deglutition apnea is significantly higher in the older group than in the younger one. Comparing the younger group and the dysphagic group we found significant differences in duration of deglutition apnea, onset time, and number of gulps. Just one parameter, number of gulps, distinguishes significantly between the older and the dysphagic groups. The second part of the study aimed at evaluating the reliability of CA in detecting dysphagia measured as the concordance and the correctness of CA experts in classifying swallowing sounds. The interrater reliability coefficient AC1 resulted in a value of 0.46, which is to be interpreted as fair agreement. Furthermore, we found that comparison with radiologically defined aspiration/penetration for the group of experts (speech and language therapists) yielded 70% specificity and 94% sensitivity. We conclude that the swallowing sounds contain audible cues that should, in principle, permit reliable classification and view CA as an early warning system for identifying patients with a high risk of aspiration/penetration; however, it is not appropriate as a standalone tool.

Key words: Cervical auscultation — Deglutition — Deglutition disorders — Stroke — Reliability — Validity.

Correspondence to: Christiane Borr, Department of Clinical Linguistics, Faculty of Linguistics and Literary Studies, Bielefeld University, Wittekindstraße 38, 33615 Bielefeld, Germany; E-mail: christiane.borr@uni-bielefeld.de

Cervical auscultation (CA) is a method of listening to the sounds of swallowing with an amplifying instrument during the pharyngeal phase to detect patients with dysphagia. Typically, a stethoscope or a microphone is placed at the lateral aspects above the cricoid cartilage in front of the sternocleidomastoid muscle and the large vessels. As opposed to videofluoroscopic swallowing study (VFSS), the alleged "gold standard," CA stands out because it is a noninvasive procedure that is simple to handle. It can be used on severely affected patients. CA is useful for checking the progress of a therapy. Finally, it can be applied flexibly because CA is not tied to a special place. So far, however, there is only little empirical data on this method. After some popularity in the 1990s, the results by Zenner et al. [2] support the use of CA as a highly sensitive and specific method of dysphagia assessment in long-term care. A clinical examination using auscultation could successfully distinguish between subjects who aspirate and those who do not in comparison to VFSS (kappa = 0.84). Moreover, the results support the clinical examination as a tool for determining appropriate dietetic treatment in long-term care when CA is used. Another research line was pursued by Takahashi et al. [3]. His investigations were aimed at determining the optimal type of acoustic detection device suited for acoustic analysis of the pharyngeal swallow, i.e., the type of adhesive to be used to attach the detector and the optimal site for sound detection of the pharyngeal swallow.

Interest in CA diminished and gave way to a growing interest in imaging techniques. However, recently CA attracted new attention as an adjunct to the clinical swallowing assessment. Kley and Biniek [4] call CA a useful supplement to imaging techniques like videography and videoendoscopy.

Stroud et al. [5] measured the inter- and intrarater reliability of five speech and language therapists (SLTs) deciding on aspiration in 16 swallow sounds. The swallow sounds were recorded simultaneously with VFSS. The authors found a merely fair agreement between the raters (kappa = 0.28). Although the outcome suggests that raters have a high true-positive rate in detecting aspiration, there is a bias for overestimation of aspiration. In the rater-based study by Leslie et al. [6], intra- and interrater reliability of judgments, agreement with the ''gold standard'' videofluoroscopy (validity), and the association between intrarater reliability and validity were measured. The authors arrived at the following results: The individual reliability outcomes varied widely and thus agreement between judgments was poor. Rater's average percentage of sensitivity was 62% and that of specificity was 66%. However, when considering the decision made by the majority of each group, the group consensus, values improved to 90% specificity and 80% sensitivity. The authors therefore conclude that improving the poor raters would improve the overall accuracy of this technique in predicting abnormality in swallowing; in principle, CA should permit reliable classification.

If we assume that CA is indeed a successful procedure, then the relevant information for deciding whether a swallow is dysphagic is to be found in its sound. That is, the swallowing sound must exhibit objective, audible properties that allow one to distinguish between pathologic and healthy swallows. In the first part of our study (''Parameterization of the Deglutition Act''), we looked for such objective acoustic properties in the temporal structuring of swallowing sounds. The second part of the study (''Reliability and External Validity of CA'') was a questionnaire-based classification task set for CA experts who had to decide whether each one of 33 swallowing sounds was produced by a dysphagic, a younger healthy, or an older healthy subject. In this part of the study the specificity and sensitivity of the experts' classifications are assessed. Reliability is defined as the degree of interrater agreement. In addition, subjects were asked to specify on what criteria their judgments are based. In sum, we address the following questions in our study:

1. Are there any objective acoustic properties that distinguish between dysphagic and nondysphagic deglutitions?
2. Are there any acoustic properties that distinguish between older and younger subjects' deglutitions?
3. To what extent do CA experts agree in classifying swallowing sounds?
4. Do the raters (experts and the layperson) classify swallowing sounds correctly?
5. Does acquaintance with CA improve classifications?
6. On which parameters do raters base their classifications?

Parameterization of the Deglutition Act

Using CA, the decision whether a given deglutition stems from a dysphagic or a healthy person is based on the sound of that deglutition. Thus, the swallowing sound should contain objective information that permits this classification. *Prima facie*, that information seems to be inherent in the acoustic profile of the swallowing sound. To verify this assumption, the structure of swallowing sounds is made explicit and annotated by parameterization. The parameters of the swallowing sounds of a dysphagic group are then compared with those of a group of older healthy individuals and those of a group of younger healthy individuals in order to flesh out differences in their acoustic profiles.

At this point it should be noted that the acoustic analysis of swallowing is presently restricted by the lack of research on the physiologic basis of the sounds and on their variation with different bolus properties and with deglutition. Some descriptive studies on throat signals recorded with a microphone have shown that the acoustic pattern of a normal swallow consists primarily of two distinct components that sound similar to ''clicks'' [cf. 9–12]. Hamlet et al. [12] assumed a close connection between the change of the deglutition sounds and the bolus flow through the upper esophageal sphincter (UES). Selley et al. [13] also acknowledged two clicks that are separated by a faint sound. Selley et al. agrees with Hamlet et al. [11] in that the first click is caused by the elevation of the larynx and the epiglottis moving down. The faint sound stems from the passage of the bolus through the UES. The second click has its origin in the retraction of epiglottis and hyoid. In contrast, Mackowiak et al. [14] distinguished three parts of the swallowing sound, α, β, and γ, intervening the silent intervals. Kley and Biniek [4] suggested a fourth part, i.e., the δ signal, which coincides with the opening of the epiglottis after swallowing. Cichero and Murdoch [15] found acoustic evidence for a release of subglottal air in the postswallow phase of nondysphagic individuals. McKaig [7] listed several landmarks that split up the course of deglutition into the following phases: two bursts followed by a glottal release, in addition, the phase of deglutional apnea, and the phase between the bursts, the so-called interburst interval. All researchers agree in proposing an acoustic structure for the deglutition act; disagreement is confined to the actual components that make up the structure. In addition, although it is agreed that the acoustic profile of a swallowing act is exclusively determined by physiologic properties of the swallowing tract, which configurations cause or shape which sounds it is not yet understood in detail.

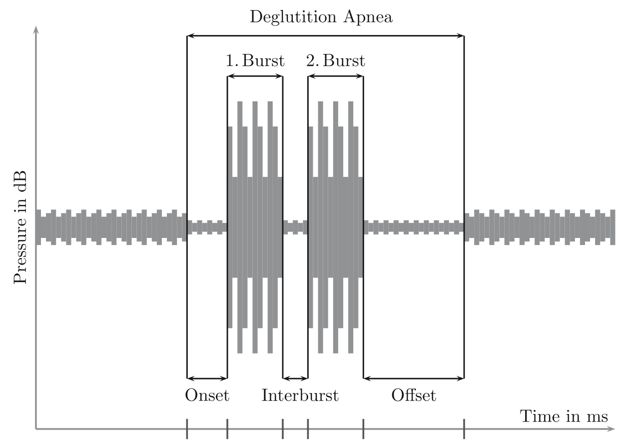

Fig. 1. Schematic depiction of a deglutition waveform. Interburst is the nearly silent section between the bursts. The phase in which respiration is interrupted is called deglutition apnea (DA). A swallow's main events are the two bursts that are presumably correlated with the transport of the bolus through the pharynx. The periods of time from initiation of DA to the leading edge of the first burst and from the trailing edge of the second burst to the end of DA are called onset and offset, respectively.

Methods

Three data sets of swallow sound recordings were used for CA analysis: those of dysphagic patients, those of younger healthy individuals, and those of older healthy individuals.

Materials

The test data consist of the swallowing sounds of 14 dysphagic patients (10 females and 4 males). The mean age of the dysphagic participants was 71.3 years (range = 44–89 years). The subjects suffered from neurogenic dysphagia resulting from anemic infarct ($n = 11$) and hemorrhage ($n = 3$). The criterion for recruitment was an acute event no more than three months post-onset. All patients showed clinical signs of dysphagia and had an aspiration risk. Diagnostics were conducted by SLTs and medical doctors.

Clinical assessment of the patients was conducted using VFSS. All patients suffered from severe dysphagia with penetration and aspiration. The time between swallow recording and clinical assessment was at most one week. The results of the videofluoroscopic examination were recorded in a clinic internal data sheet. According to the standardized penetration/aspiration scale [16], the patients have to be classified at least at level 6. Subjects with progressive neurogenic disorders and patients with chronic dysphagia were excluded.

The control data of the younger subjects comprised 25 swallowing sounds randomly chosen from 250 swallow recordings taken from subjects without a neurologic history. The younger volunteers (students and staff) were tested at the university. The random sampling yielded swallow recordings of 11 females and 14 males. The mean age of the younger subjects was 30.9 years (range = 25–44 years). The swallow recordings of 25 older subjects (13 females and 12 males) without a neurologic and dysphagic history comprised the older control group. The mean age of the older participants was 76.2 years (range = 60–97 years). The older volunteers were recruited at the medical practice of a physician. No imaging was performed because of ethical and financial reasons. The aptitude of the volunteers for participating in the study was based exclusively on their medical record.

Procedure and Apparatus

Each swallow was recorded in a controlled and prespecified way. First, the larynx was located and the so-called four-finger method, i.e., the palpation of the laryngeal elevation (cf. [17]), was applied. The four-finger method helps to obtain a "haptic overview" of the laryngeal movement and to identify the best location for placing the stethoscope. It is also a confidence-building interaction between therapist and patient. The researcher placed the stethoscope (Welch Allyn Meditron) at the lateral aspects above the cricoid cartilage in front of the sternocleidomastoid muscle and the large vessels, a position which Takahashi et al. [3] considered optimal. Then the patients were given three portions of 10 ml of nonsparkling water. The swallows were recorded via Audiorecorder (Windows XP) on a standard notebook computer (Toshiba 1900 Satellite) with a RealTek sound chip which is compliant to the AC '97 rev. 2.3 specifications. The mono line-in had a resolution of 16 bit at a sampling rate of 44.1 kHz, which was also set within the Audiorecorder software. The frequency range was from 20 Hz to 20 kHz. The stethoscope's high-pass filter was activated during recordings. The subjects were told to swallow in a normal fashion to approximate their "natural" drinking behavior.

Seven Parameter Structuring

The sound files were displayed as waveforms and were annotated by making use of the software program Soundforge 6.0 (Sonic Foundry, Inc., Madison, WI). The cardinal points for the annotation were determined following McKaig [7] (see Fig. 1). They consist of marking the beginning and the end of the deglutition apnea, the beginning of the first and the second bursts, and the time spans of the first and the second bursts. In addition to the "McKaig structuring," we recognize as a seventh parameter the number

Table 1. The parameters for structuring deglutition waveforms

Parameters	Abbreviation	Description
Onset time	ON	Period of time from initiation of deglutition apnea to leading edge of first burst
Deglutition apnea	DA	Duration of time from initiation to the end of deglutition apnea
First burst	1B	Duration of initial burst
Second burst	2B	Duration of second burst
Bolus transport signal	BTS	Duration of time from leading edge of first burst to trailing edge of second burst
Offset time	OFF	Period of time from trailing edge of second burst to end of deglutition apnea
Deglutition	D	Number of gulps used to gulp down the bolus

of gulps used to swallow the bolus. A summary of all parameters is given in Table 1. The cardinal points were annotated independently by two raters. They used a common strategy that relies on identifying the visual landmarks postulated by McKaig [7]. This approach was negotiated after an exploratory annotation session. With this strategy, Cronbach's alpha [19] was above 0.9 for all seven parameters and can be interpreted as consistent following the argument in Langer [20]. A schematic depiction of a prototypical waveform is given in Figure 1. A waveform is the visual representation of a sound's pressure over time. Every physiologic event is correlated with an acoustic signal [7] (but see the remark above). McKaig [7] divides the pharyngeal phase into five sections. The tidal breathing before the start of the gulp reflex is the first section and is called onset time. The second section is the deglutition apnea which is caused by the configuration where the pharynx is switched from the air channel into the swallow channel. The deglutition apnea continues during the entire pharyngeal phase. The first audible blasting noise is called the first burst. This acoustic event presumably correlates with the bolus penetrating the hypopharynx [7]. After a short period of time it is followed by the second burst, which marks the passage of the bolus through the hypopharynx. Thus, the first and second bursts indicate transport of the bolus; their timespan is therefore known of the bolus transport signal. The two bursts constitute the third and fourth sections of the deglutition waveform, respectively. In the quiet stripping wave (the so-called offset time) that follows (the fifth section), the hypopharynx gets emptied. The bolus is pulled toward the esophagus. The opening of the esophagus is a complex phenomenon in itself [18]. During the last phase, the sphincter closes again and muscular tonus is reestablished. If the nasopharyngeal closure recurs, the glottis opens, tidal breathing sets up again, and the pharyngeal period terminates. Sometimes a swallow contains more than one gulp. The number of gulps is an additional parameter that we called deglutition. In this case, only the first of the multiple gulps is annotated and enters into analysis. When clinicians listen to the swallowing sounds with a stethoscope, qualitative judgments are made about what is heard. Based on the ''crispness'' of the sound, the characteristic swishing double-click as the bolus passes through the pharynx and into the esophagus is judged for normalcy. A reliable definition of crispness has not been provided in physical terms, but in general it is characterized as a sharp, crackling, or brilliant sound that has high-frequency components [12].

Results

The means for the relevant seven parameters of each group are presented in Table 2. We also calculated various relative parameters from the seven basic ones. Because the very few significant outcomes we got can be reduced onto the seven parameters, we omit their discussion here (but see [1] for details). One-way analyses of variance (ANOVAs) were calculated for each parameter followed by *post hoc* Scheffé tests. The statistical results are presented in Table 3.

Younger vs. Older

Comparing the parameters annotated on the deglutition sounds of the younger and the older groups, we find two relevant differences: The duration of deglutition apnea is significant ($p = 0.047$) and the onset time ($p = 0.1$) shows a tendency. Both values are higher in the older group than in the younger one.

Younger vs. Dysphagic

There are differences between the younger group and the dysphagic group for two parameters: the duration of deglutition apnea ($p = 0.07$) and the onset time ($p = 0.03$). One additional parameter is significantly different for swallows of the younger group and the dysphagic group, namely, the number of gulps ($p = 0.001$). Dysphagics often take more than one gulp to swallow the bolus.

Table 2. The mean values (in milliseconds) for the parameters in the three groups

Parameter	Younger group		Older group		Dysphagic group	
ON	0.31	0.28	0.84	0.86	0.9	1.05
DA	1.18	0.59	1.77	0.9	2.05	1.49
1B	0.14	0.07	0.17	0.1	0.11	0.05
2B	0.2	0.17	0.2	0.12	0.2	0.09
BTS	0.35	0.14	0.44	0.21	0.44	0.27
OFF	0.52	0.55	0.49	0.39	0.72	0.89
D	1.16	0.47	1.08	0.28	1.86	0.86

The standard deviation is given in parenthesis.

Table 3. The analysis of variance with *post hoc* Scheffé test

Parameter	MSQ[a]	*F* value	*p* value	y vs. o[b]	y vs. d[b]	o vs. d[b]
ON	2.35	4.21	0.02	0.1	0.03	
DA	4	4.3	0.02	0.05	0.07	
1B	0.02	2.49	0.09			0.1
2B	0	0.02	0.98			
BTS	0.06	1.34	0.27			
OFF	0.24	0.7	0.5			
D	3.01	10.86	0.00		0.001	0.00

[a]Mean square.
[b]y = younger group (n = 25); o = older group (n = 25); d = dysphagic group (n = 14).

Older vs. Dysphagic

There are two parameters that distinguish swallowing sounds from the old group and the dysphagic group: the duration of the first burst (p = 0.1) and multiple gulps (p = 0.00). Dysphagic patients need more than one gulp to swallow the bolus and they show a slightly shorter first burst.

Summary

To sum up, there are only two parameters that might distinguish the sounds from different groups, namely, DA and onset time (see Younger vs. Older and Younger vs. Dysphagic above). Regarding the distinction between older people and dysphagics, the only significant parameter we find is the number of gulps. If we take into account that the temporal differences lie in the range of milliseconds, we conclude that dysphagic swallowing sounds are hardly to be distinguished from nondysphagic ones with respect to their temporal structure. For instance, the mean difference between the duration of the first burst in the dysphagic and the older group is 0.06 s. However, there are therapists who successfully use CA as a diagnotic tool. If they cannot base their diagnosis on the structure of the swallowing sounds, on what do they base their decision?

Reliability and External Validity of CA

The second part of the study tried to evaluate the reliability of CA in detecting dysphagia in terms of concordance and correctness. In addition, we have compared CA experts with other groups (students of clinical linguistics and layperson) in classifying swallowing sounds. We were interested in whether CA experts are able to classify the swallowing sounds used in the first part of the study. We have seen that those swallowing sounds do not contain sufficient acoustic evidence. What features do the experts use to make a diagnosis?

Method

Three groups of raters were recruited. The first comprised 20 laypeople, the second 20 students/research assistants of clinical linguistics, and the third 9 speech and language therapists who worked in seven different rehabilitation centers in Germany and Switzerland. The experience of the raters with CA and dysphagia ranged from 1 to 12 years, and CA experience alone ranged from 1 to 4 years (see Table 4 for details). The students or research

Table 4. SLTs' experience with CA and dysphagia

Expert no.	Years of working with dysphagics	CA workshop	Years of working with CA
1	7	yes	2
2	4	yes	3
3	8	no	1
4	12	yes	2
5	12	yes	4
6	6	yes	3
7	1	no	1
8	1	yes	1
9	7	yes	1

SLT = speech language therapist.

assistants had at least basic theoretical knowledge about swallowing disorders, but they were not familiar with applying CA.

The raters' task was to classify 33 sounds of deglutition into three response categories: "dysphagia," "younger healthy," or "older healthy." A questionnaire (see the Appendix) was used for the raters' answers. The raters were allowed to listen to the sound files as often as they wanted. Subsequently, the raters were asked to give reasons for their decisions: Which properties of the swallowing sounds or which additional features led them to make their classification? Regarding their acquaintance with CA, the raters were asked to give information about if and how long they had used CA, had they participated in special workshops for learning CA, about where they placed the stethoscope when using CA, and to what extent they felt CA to be a reliable method for revealing dysphagia (on a 0% to 100% scale). Finally, the raters were asked to specify other diagnostic procedures (invasive/noninvasive) they could apply. The 33 sounds were selected at random from the annotated corpus used for the first part of the study, resulting in a randomly ordered "tracklist" consisting of 10 swallow recordings from the older group, 10 from the younger group, and 13 from the dysphagic group. This tracklist was used for all raters.

Statistics

Different descriptive and inferential statistics were used to calculate reliability and external validity of the ratings, in terms of sensitivity and specificity. The level of reliability of CA was assessed by computing the interrater agreement for each single group, the intragroup concordance for short. Because of the conceptual weaknesses of the kappa statistic, we followed the proposal made by Stegmann and Lücking [21] and used the AC1 statistics developed by Gwet [22]. The assessment of validity (external validity according to Krippendorff [26]) of CA comprises two aspects: sensitivity, which is the ability to classify true dysphagics correctly, and specificity, which is the ability to classify true nondysphagics correctly. The classifications of the swallowing sounds of each rater group (experts, students, laypeople) were divided into four classes: the number of correct classifications of (i) dysphagics, (ii) older healthy, (iii) younger healthy, and (iv) the sum of i, ii, and iii, i.e., the total number of correct classifications. *t* Tests for independent measurements with heterogeneous variances for the comparison of the means of the number of correct classifications regarding all four classes were used to assess significant differences of the rating success between experts and students, experts and laypeople, and students and laypeople. Sensitivity was calculated as the percentage of correct classifications of true dysphagics. As an index for specificity, the percentage of correct classifications of nondysphagics (in this case, younger or older healthy subjects) was calculated. Reliability was calculated as the intragroup concordance of experts on the whole set of classification tasks (class iv above). For comparison, the intragroup concordances of the other groups (students and laypeople) were calculated also. To find out whether the decisions of the experts reflect the parameters used in the first part of the study, a stepwise linear regression was run on the experts' classifications of a given item using the annotation mark values as predictors. In reversal, a discriminant analysis was calculated to test whether the variables that were significant in the first study can be used to predict into which group a given sound file is classified. The statistical computation was done using the statistical programming language R [23] and SPSS v13 (SPSS, Inc., Chicago, IL).

Results

There were no significant differences between the groups of raters in the classification of older and younger healthy subjects and in the overall result. However, a significant difference between the groups of experts and that of laypeople could be found in the classification of dysphagics ($t = -2.5$; $p = 0.01$; $\alpha = 0.05$). See Figure 2 for an overview.

The mean rate of specificity for the group of experts was 70% (laypeople, 79.75%; students, 79%). The mean value for sensitivity in the expert group was 94.01% (laypeople, 85.38%; students, 91.15%). The first-order agreement coefficient AC1 for the classification of true dysphagics yielded the following results: laypeople, 0.39; students, 0.44; experts, 0.46.

A nonparametric version of the ANOVA revealed that there was no effect of the grouping (experts, students, and laypeople) on the variation of correct classifications (overall: $F = 0.75$; $p = 0.48$; dysphagics: $F = 3.09$; $p = 0.06$; older healthy: $F = 0.85$, $p = 0.44$; younger healthy: $F = 1.24$; $p = 0.31$). Only with respect to the dysphagics' swallow recordings did experts tend to show more

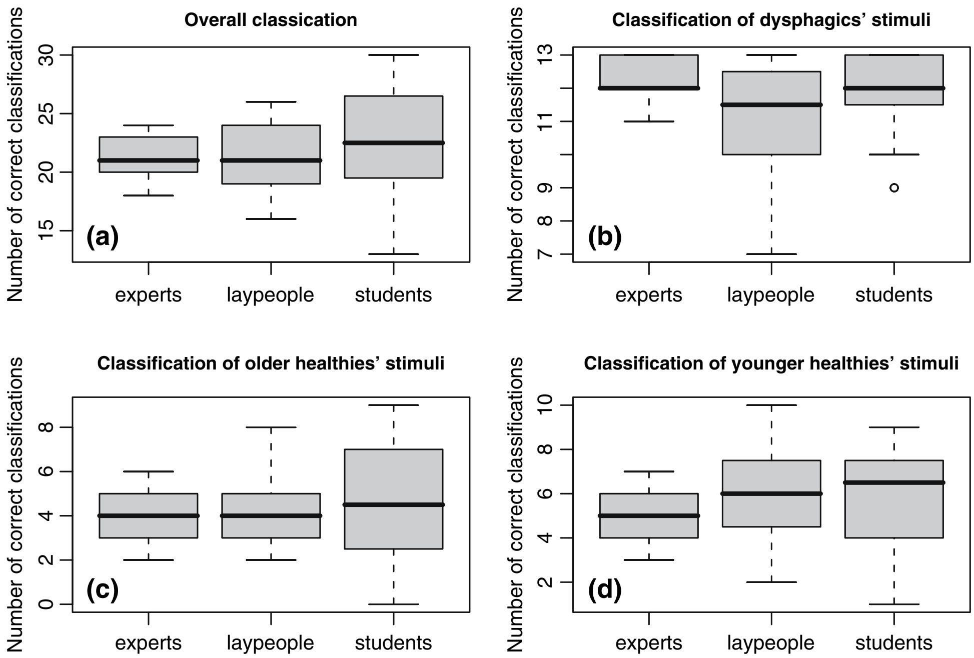

Fig. 2. Boxplots for correct outcomes. (**a**) Overview of the sum of correct classifications for each group of raters. (**b**)–(**d**) Correct classifications with respect to the three response categories.

reliable results than the other groups. Figure 3 depicts the distribution of means and their corresponding ± 1 standard errors. What are the features that the raters use to determine their ratings? The following top answers of each group are listed with the percentage of their occurrence:

Experts: respiration (100%), quality (100%), duration (66.6%), number of swallows (55.6%), cough (44.4%). Interestingly, among the experts' answers is also the frequency of bursts, a parameter that turned out to be significant in the first study.

Students: number of swallows (90%), quality (65%), duration (55%), cough (45%). Students also named a significant parameter from the first part of the study, namely, the duration of burst.

Laypeople: quality (65%), duration (60%), number of swallows (40%), loudness (40%).

The features named by all groups are "quality," "duration," and a number of swallows." The experts pay attention to a few more features of the swallowing sound, especially respiration. With regard to the first three features mentioned most frequently, there is a higher degree of concordance in the group of experts than among the students, who in turn agree about them more often than laypeople.

Turning to the results of the stepwise linear regression and the discriminant analysis very briefly: None of the seven parameters was a significant predictor of the outcome of the experts' ratings. As presumed in light of the first part of the study, those parameters are inappropriate to use to determine the group, i.e. dysphagic vs. nondysphagic, to which a swallow sound belongs.

Summary

The few parameters that turned out to be significant in the first study are indeed used as classification devices by the groups of raters in the second part of the study. Besides such quantifiable, measurable features such as number of gulps or duration of a swallow, all raters—be they experts or laypeople—take into account the "quality" of the swallowing sound.

Although speech and language therapists have a high sensitivity in detecting dysphagics, they overdetect dysphagia, i.e., they report hearing a dysphagic swallow even if it does not exist. This overestimation is not the result of guessing, rather there is a bias to label older healthy subjects as dysphagic (this finding agrees with Stroud et al. [5]).

Discussion and Conclusion

One finding from the more than 40 years of research is that deglutition sounds exhibit a specific pattern. There is, however, disagreement about how fine-grained this pattern is to be described. It is difficult to judge how convincing the evidence for the most ambitious classification, namely, that of McKaig [7], is. To arrive at well-founded conclusions, large-scale empirical studies on swallowing of

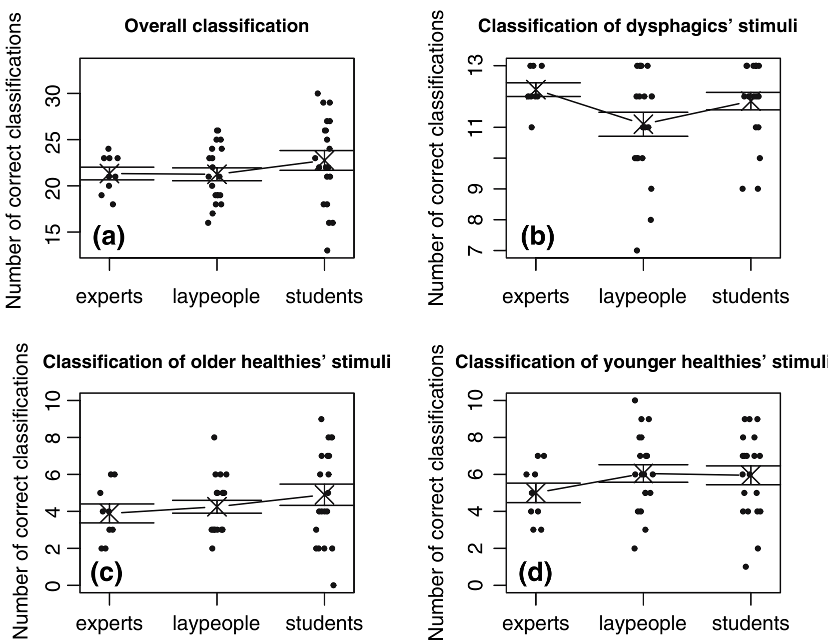

Fig. 3. Extended strip charts for the correct outcomes. The crosses indicate the mean and the bars display the range of 2 standard errors of mean.

individuals of various ages and races would have to be conducted. This research line was pursued by Youmans and Stierwalt [24]. Because of age-related changes and genetically determined variability of the anatomical basis of swallow, the swallowing sounds are subject to inter- and intrapersonal variability over time. Beyond those variations, the general pattern of deglutition remains recognizable. It is expected that this pattern gets distorted by swallowing disorders.

In the light that there is no accepted structuring of swallowing sounds yet, two acoustic parameters have been found that differentiate the swallowing sounds of older healthy subjects from those of dysphagics: First, the duration of the first burst tends to be longer in the older group than in the group of dysphagics; second, dysphagics often need more than a single gulp to swallow a small amount of water. The latter finding agrees with the results of Leslie et al. [25]. They have shown that age is not significantly correlated with multiple swallowing. Instead, they found changes with age in swallow respiration coordination that we could not replicate in our study. Turning to the first outcome—the length of the first burst—we assume that the first burst is correlated with the entrance of the bolus into the hypopharynx. The first burst of the patients is shorter than that of the other groups. A possible reason for this difference might be that sensory or motor disorders cause the bolus to enter the hypopharynx even before the swallowing reflex sets in. However, as noted already, the mapping between physiologic events and sounds is not entirely clear yet.

We also found some changes with age in swallow functions. Younger subjects showed a shorter duration of deglutition apnea (DA) and a shorter duration of onset (ON). Similarly, Leslie et al. [25] reported subtle but distinctive changes in swallow function with older age. Moreover, they also found evidence for an increase in the length of DA along with an increase in age. The parameter that seems to be the most reliable for distinguishing dysphagic from nondysphagic swallow recordings is the number of gulps. In the first part of the study we found that multiple gulps significantly differentiate dysphagics' swallow sounds from that of older and younger persons. However, all recordings from dysphagic patients show multiple gulps, whereas only one recording from the older group and no recording from the younger group exhibit this feature. That is, the dysphagic data group is almost definable exclusively in terms of parameter deglutition (D). This is evidence for deglutition being a useful diagnostic parameter, but noticing whether a person needs more than one gulp to swallow the bolus can be done without using CA!

In our view what these findings suggest is that the parameters used to investigate the time structure of deglutition sounds are not, on their own, appropriate as a guiding device for a definite diagnosis of dysphagia with a high aspiration and penetration risk. This conclusion can be corroborated with two outcomes of the second part of the study. First, all raters name "quality" when asked for the parameters on which they base their decision. The nature of the sound of a swallow clearly is not, or at least not fully, covered by the temporal structuring of that sound. Furthermore, it is unclear how the parameter "quality" can be operationalized and made accessible for measurement. Second, the agreement values are on a medium level; thus, there is some disagreement between raters (reliability). Some raters do better than others, i.e., they correctly classify more items (specificity and sensitivity). One explanation for this variability is that they are produced by the individual talents of the raters. Subjective parameters or specialized skills seem to be the ones that help the clinicians in making correct decisions. Leslie et al. [6, p. 237] also speculate in this direction. They conjecture that musically trained SLTs might be better in applying and interpreting CA than SLTs without such training.

What inferences can be drawn regarding the question of whether CA can serve as a reliable screening method for detecting dysphagia? We think that there are three worthy propositions: First, deglutition sounds carry enough information to be neatly classified—even if this information is not inherent in the temporal structure of deglutitions. Second, even layperson seems to have some naive concept of how a pathologic swallow sounds, or perhaps how a healthy one should not sound. And third, for experts, the sounds of older but healthy subjects bear some sufficient signals to alert the therapists. It is this third point that could qualify CA as an early warning system for detecting dysphagia. Here, we agree with Hamlet et al. [11] who state that "auditory or automated acoustic analysis of swallowing sounds could become a useful noninvasive aid in alerting a clinician to the presence of swallowing dysfunction or tracking its course in dysphagia treatment" [11, p. 749]. At least in our study and for Hamlet et al. [11] it must be admitted that it cannot be taken for granted that those sounds are indeed problematic.

What does that mean for the clinical applicability of CA? With regard to what we know about CA, we are viewing this method with skepticism. The clinical significance of CA is too weak for CA to be used as a stand-alone method. Therefore, in a future study we will evaluate the reliability and validity of CA connected with parameters of CSE (Clinical Swallow Examination) in detecting geriatric patients with a high risk of penetration/aspiration.

Acknowledgments. The authors thank the patients and control subjects and the SLT teams of the Aatalklinik, Wünnenberg, and the Edith-Stein-Klinik, Bad Bergzabern, the team of the medical practice in Mühlhausen/Thüringen, and the raters from Germany and Switzerland. Special thanks go to Petra Dünger and Stephan Braun. The authors are very grateful to Sönke Stanschus, T. Neil McKaig, Heike Borr, Michael Borr, and Seong-Ji Kim for their general support. Finally, they thank Max Sichelschmidt and the anonymous reviewers for their helpful comments.

Appendix

Instructions for raters

- Please listen to the 33 tracks on the CD. For listening you can use the PC or a standard CD player. The swallowing sounds are played ordered from 1 to 33.
- Listen to each track as often you feel is necessary.
- Please don't confer with your colleagues.
- Check the appropriate box to rate each swallow as dysphagia, aspiration, younger healthy, older healthy.
- After listening to the CD please complete the questionnaire, which is related to the basis for your classification.
- Please answer every question. Do not leave any question blank.

Questions relating to raters' decision and their experience in the field of dysphagia and CA

- Please specify the reasons for your classification. Which parameters did you use to diagnose?
- How many years have you been working with patients suffering from dysphagia?
- How many years have you been performing cervical auscultation?
- Where do you place the stethoscope?
- Have you participated in a special cervical auscultation workshop?
- How reliable is the method in your opinion?
- Which other diagnostic methods are used in your institution?

References

1. Borr C: Zervikale Auskultation zur Diagnostik neurogener Dysphagien bei Erwachsenen. Master's Thesis, Bielefeld University, 2003
2. Zenner P, Losinski D, Mills R: Using cervical auscultation in the clinical dysphagia examination in long-term care. *Dysphagia 10:*27–31, 1995

3. Takahashi K, Groher M, Michi K: Methodology for detecting swallowing sounds. *Dysphagia 9:*54–62, 1994
4. Kley C, Biniek R: Dysphagie: Sind Schluckgeräusche diagnostisch nutzbar? *Nervenarzt 76:*1495–1505, 2005
5. Stroud A, Lawrie B, Wiles C: Inter- and intra-rater reliability of cervical auscultation to detect aspiration in patients with dysphagia. *Clin Rehabil 16:*640–645, 2002
6. Leslie P, Drinnan M, Finn P, Ford G, Wilson J: Reliability and validity of cervical auscultation: a controlled comparison using videofluoroscopy. *Dysphagia 19:*231–240, 2004
7. McKaig T: Auskultation Gú Zervikal und Thorakal. In: *Methoden in der klinischen Dysphagiologie*. S Stanschus (ed), 2002, vol 1, chap 3, Idstein: Schulz-Kirchner Verlag, pp 111–1378
8. McKaig T: http://www.grapeescape.com/mckaignet/ [accessed 6 October 2006]
9. Lear C, Flanagan J, Moorress C: The frequency of deglutition in man. *Arch Oral Biol 10:*83–99, 1965
10. Logan W, Kavanaugh J, Wornall A: Sonic correlates of human deglutition. *J Appl Physiol 23:*279–284, 1968
11. Hamlet S, Nelson R, Patterson R: Interpreting the sounds of swallowing. *Ann Otol Rhinol Laryngol 99:*749–752, 1990
12. Hamlet S, Penney D, Formolo J: Stethoscope acoustics and cervical auscultation of swallowing. *Dysphagia 9:*63–68, 1994
13. Selley W, Ellis R, Flack F, Bayliss C, Pearce V: The synchronization of respiration and swallow sounds with videofluoroscopy during swallowing. *Dysphagia 9:*162–167, 1994
14. Mackowiak R, Brenman H, Friedman M: Acoustic profile of deglutition. *Proc Soc Exp Biol 125:*1149–1152, 1967
15. Cichero J, Murdoch B: What happens after the swallow? introducing the glottal release sound. *J Med Speech Lang Pathol 11(1):*31–41, 2003
16. Rosenbek J, Robbins J, Roecker E, Coyle J, Wood J: A penetration-aspiration scale. *Dysphagia 11:*93–98, 1996
17. In: Perlman A, Schulze-Delrieu K (eds.) *Deglutition and its disorders*. San Diego, CA: Singular Publishing Group, 1997
18. Mu L, Sanders I: Neuromuscular organisation of the human upper esophageal sphincter. *Ann Otol Rhinol Laryngol 107:*370–377, 1998
19. Cronbach L: Coefficient alpha and the internal structure of tests. *Psychometrika 16:*297–333, 1951
20. Langer W: Methoden V: Explorative Methodenanalyse, 1999, http://www.soziologie.uni-halle.de/lauger/lisrel/skripten/reliabil.pdf [accessed June 6, 2003]
21. Stegmann J, Lücking A: Assessing reliability on annotations (1): Theoretical considerations. Tech Rep, Bielefeld University, SFB 360, 2005
22. Gwet K: *Handbook of inter-rater reliability*. Gaithersburg, MD: STATAXIS Publishing Company, 2001
23. R Development Core Team: R: *A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, 2005 [http://www.R-project.org]
24. Youmans S, Stierwalt J: An acoustic profile of normal swallowing. *Dysphagia 20:*195–209, 2005
25. Leslie P, Drinnan M, Ford G, Wilson J: Swallow respiratory patterns and aging: presbyphagia or dysphagia? *J Gerontol 60A:*391–395, 2005
26. Krippendorff K: *Content Analysis: An introduction to its methodology*. The Sage COMMMTEXT Series. vol. 5. Beverly Hills, CA: Sage Publishing, 1980