Baseline characteristics of cervical auscultation signals during various head maneuvers

I.Jestrović[a], J. M. Dudik[a], B.Luan[a], J. L. Coyle[b], E.Sejdić[1,*]

[a]*Department of Electrical and Computer Engineering, Swanson School of Engineering, University of Pittsburgh, Pittsburgh, PA, USA*
[b]*Department of Communication Science and Disorders, School of Health and Rehabilitation Sciences, University of Pittsburgh, Pittsburgh, PA, USA*

Abstract

Cervical auscultation (CA) is an emerging method of assessing swallowing disorders that is both non-invasive and inexpensive. This technique utilizes microphones to detect acoustic sounds produced by swallowing activity and characterize its behavior. Though some properties of swallowing sounds are known, there is still a need for a complete understanding of the baseline characteristics of cervical auscultation signals as well as how they change due to the patient's head motion, age, and sex. In order to examine these parameters, data was collected from 56 healthy adult participants that performed six different head movement tasks without swallowing. After preprocessing the signal, features were extracted. Dependent variables were time domain, frequency domain and time-frequency domain features. Statistical tests showed that only the skewness and peak frequency were not statistically different for all tasks. The peak frequency results indicate that head movement does not significantly affect the microphone signal, and that it is unnecessary to filter out the lowest frequency components. No sex differences were observed on the extracted features, but several features exhibited age dependence.

Keywords: Cervical auscultation, baseline, head motions, signal characteristic.

*Corresponding author
Email addresses: ivj2@pitt.edu (I.Jestrović), jmd151@pitt.edu (J. M. Dudik), bol12@pitt.edu (B.Luan), jcoyle@pitt.edu (J. L. Coyle), esejdic@ieee.org (E.Sejdić)

1. Introduction

Swallowing difficulties, or dysphagia can occur for many reasons [1]. Almost half of stroke patients suffer from dysphagia [2]. In stroke, dysphagia can impair one or more of the oral, pharyngeal, or esophageal contributions to the transfer of a swallowed bolus of food or liquid into the digestive system. Given that the upper aerodigestive tract is a single tube that is shared by the digestive system and the respiratory system, the numerous events that occur during this alternating access to the mechanism must be precisely executed to prevent swallowed material from entering the airway. Sensorimotor events that occur during a single swallow include oral and pharyngeal activities that systematically transfer intrabolus pressure from the mouth to the esophagus: oral propulsion of the bolus into the pharynx while preventing its leakage into the nasal cavity, and transfer of the bolus from the mouth, through the pharynx into the esophagus while ensuring that the airway is closed and the upper esophageal sphincter is adequately opened. Since the duration of a pharyngeal swallow is very brief (about one second) and several biomechanical events occur during this duration, the mistiming of the sequenced components of a swallow, or other sensorimotor impairments affecting swallowing can produce adverse events such as misdirection of a swallowed bolus into the airway leading to aspiration, in which food and fluids enter the trachea and into upper airways and lungs. Aspiration is a serious immediate consequence of dysphagia in many stroke patients and contributes significantly to the morbidity and mortality of these patients by producing pneumonia, airway obstruction, and other chronic pulmonary consequences [3]. Likewise, dysphagia after stroke is responsible for important adverse clinical outcomes such as malnutrition, dehydration, and impaired quality of life [4, 5]. In fact patients with new onset of stroke who develop pneumonia after onset, have a three-fold increased relative risk of death compared to those that do not develop pneumonia after stroke onset [6]. Hence, early diagnosis of dysphagia is very important for patient safety and health [7]. One current gold standard for diagnosis is the videofluoroscopic swallowing study (VFSS), which is an imaging technique that uses X-rays for recording biomechanical events that occur during swallowing, and visualize the path of swallowed foods and fluids [8, 9]. However, VFSS instrumentation

is not always available for immediate access when needed by patients residing in settings in which it is not available, and in the absence of gold-standard instrumentation, reasonably sensitive screening methods are needed to identify patients with elevated likelihoods of aspiration.

Cervical auscultation (CA) is a screening method that has received much attention in the past two decades [10]. CA uses sensors attached to the patient's neck that record the acoustic sounds that occur during swallowing [11]. Clinically it is implemented with an ordinary stethoscope. Previous contributions have shown that sounds of normal and abnormal swallows are different [12], though identification of specific impairments, or characterization of the nature of the abnormal swallows has yet to be identified using CA. However if the value of CA can be raised by improving its ability to detect specific and clinically important biomechanical errors or bolus misdirection, it would become more useful in screening and possibly diagnosis. Because of its low cost, noninvasiveness and accessibility, an objective, valid and reliable method for clinically using CA would be attractive [13, 14]. Current CA implementation methods utilize a subjective perceptual assessment by the test administrator to make a judgment regarding the clinical relevance of the sounds that occur during the swallowing event, and several prior studies have found such judgment methods to produce relatively low overall accuracy in identifying specific biomechanical and bolus-flow abnormalities [15]. However, developing algorithms for instrumented analysis may prove to produce a more objective characterization of swallowing events from acoustic data than perceptual judgment, and significantly increase the sensitivity and specificity of the examination.

A number of studies have investigated swallowing sounds (e.g., [16, 17, 11]). However, none of the studies have investigated the baseline characteristics of cervical auscultation signals (i.e., when swallows are not present). Understanding such characteristics is important for several reasons. First, cervical auscultation signals can potentially contain signal components that are present even when no swallowing is performed, such was the case for swallowing accelerometry signals [18]. Second it is possible that, like for accelerometry signals, the baseline characteristics of swallowing sounds could be affected by the patient's

movement or head position. For example, Sejdić *et al.* [19] showed that head motions can severely impact swallowing accelerometry components, and later developed an algorithm for removing those components [20]. Similarly, microphones can sometimes record sounds associated with skin displacement (e.g., [21, 22]). Therefore, the same procedure should be investigated for swallowing sounds to reduce the effect of acoustic artifact on CA signals that may mask data that reflect physiologic events. Also relationships between the different head positions or compensatory postures used during swallowing, age, gender and other diagnostic differences among patients, may affect the resultant acoustic sounds of the swallowing event; if these could be accounted for and subtracted from the overall CA product signal, the resultant signal would have more value as diagnostic data.

To address these open questions, we examined baseline characteristics of cervical auscultation signals in time domain, frequency domain and time-frequency domains. In particular, these signals were examined while participants completed several tasks in head neutral position and the chin-tuck (head/neck flexion) positions which is a common compensatory posture used with some dysphagic patients to mitigate specific aspects of dysphagias. Sex and age dependence were also examined.

2. Methodology

2.1. Data Acquisition from Participants

56 people, aged from 18 to 65, with no previous self-reported history of neurological diseases, swallowing disorder, head, neck or spinal trauma, neck, brain or mouth cancer or abnormal brain activity, participated in the data acquisition process. Each subject provided written consent and provided basic demographic information such as their age. The study was approved by Institutional Review Board at the University of Pittsburgh.

We recorded sounds with a contact microphone (AKG C411L, AKG Acoustics GmbH, Vienna, Austria) that had a frequency response from 10Hz to 18kHz. We also recorded accelerometry signals using a dual-axis accelerometer (ADXL322, Analog Devices, Norwood, MA, USA). However, swallowing accelerometry signals were not considered in the current

manuscript. All signals were recorded using LabView software Signal Express (National Instruments, Austin, TX, USA) which provided 40kHz sampling rate, and recorded data was saved to a hard drive.

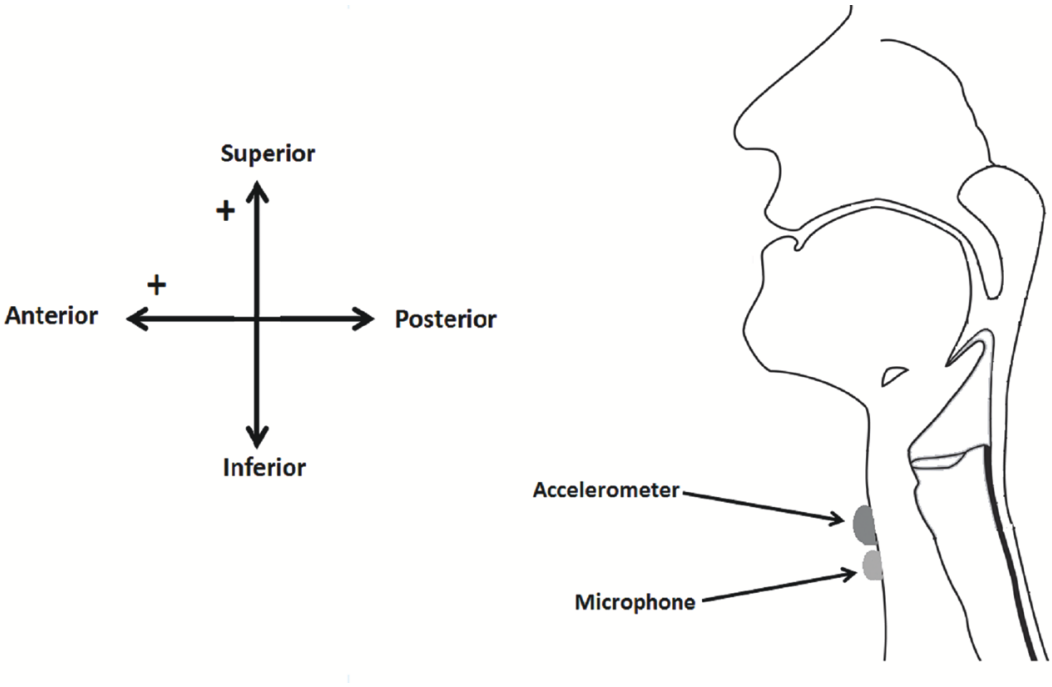

Figure 1: Position of accelerometer and microphone

The sensors were attached to the subject's anterior neck with double sided tape. The accelerometer was positioned below the thyroid cartilage as shown in Figure 1 and the microphone was positioned far enough from the accelerometer such that the two sensors would not come into contact. After the placement of sensors, the subject was asked to complete six different tasks. First the resting state was recorded, where the subject was asked to refrain from moving, talking, or swallowing for one minute. Next, the subject was instructed to hold their head in the head neutral position and hold their breath for 10 seconds while again refraining from moving, talking, or swallowing. The next four tasks each consisted of subjects tilting their heads then returning to the neutral starting position ten times in one of four directions: in the sagital plane (flexion, extension), and in the coronal plane (right and left lateral flexion) (Figure 2). During these tasks, the subjects were asked

again to refrain from talking or swallowing.

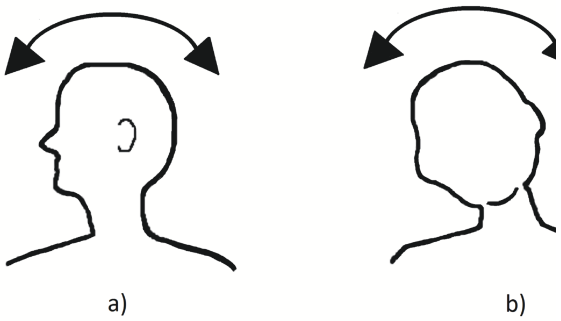

Figure 2: Different head motion: a) flexion and extension, b) right and left lateral flexion

2.2. Pre-processing

The raw signals were pre-processed by the algorithm reported in previous studies (e.g., [18]). In order to annul effects from the recording devices, a finite impulse response filter, was created using AR coefficients from 18 baseline recordings, a method described in [18]. After filtering, the signals were denoised with 10-level discrete wavelet decomposition using the discrete Meyer wavelet with soft-thresholding. The global denoising threshold as proposed in [23] was used for wavelet denoising.

As the frequency response of the microphone which was used in experiment is from 10 Hz to 18 kHz, the pre-processed signals were filtered with a 4 order, infinite impulse response Butterworth high pass filter with cut off frequency of 10Hz to eliminate any spurious frequency components that may present as a side effect of pre-processing steps.

The subsequent feature extraction was performed on the pre-processed signals.

2.3. Feature Extraction

Considered here are the time domain, frequency domain and time-frequency domain features. Under the assumption, that each microphone signal is represented as an array of time elements, of length n, $M = \{m_1, m_2, ..., m_n\}$, each subsections below are describes the

computation of each feature. The same formulas are used for features extraction for both axis of the accelerometer signal.

2.3.1. Time Domain Features

- Standard deviation describes spread of amplitude probability distribution of signal, and it is computed by

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(m_i - \mu_m)}, \tag{1}$$

where μ_m is mean value of the signal amplitude.

- Skewness describes symmetry of the probability distribution curve [24], and it is calculated by:

$$\nu = \frac{\frac{1}{n}\sum_{i=1}^{n}(m_i - \mu_m)^3}{(\frac{1}{n}\sum_{i=1}^{n}(m_i - \mu_m)^2)^{1.5}}. \tag{2}$$

- The kurtosis describe the "peakedness" of the probability distribution curve. A high value means that curve is sharp and narrow, while a low value describes a flat distribution peak. It is computed by:

$$\varpi = \frac{\frac{1}{n}\sum_{i=1}^{n}(m_i - \mu_m)^3}{(\frac{1}{n}\sum_{i=1}^{n}(m_i - \mu_m)^2)^{2}}. \tag{3}$$

- The entropy rate [25, 26] describes regularity of the signal. To calculate entropy rate, first the signal M should first be normalized to zero mean and unit variance. The normalized signal is than quantized to 10 equally spaced levels ranging from minimum to maximum, $\hat{M} = \{\hat{m}_1, \hat{m}_2, ..., \hat{m}_n\}$. In the next step, the signal $\hat{M}$ is coded with U consecutive points

$$s_i = \hat{m}_{i+U-1} \cdot 10^{U-1} + ... + \hat{m}_i \cdot 10^0, \tag{4}$$

where $i = 1, 2, ..., n - U + 1$, and $S_i = \{s_1, s_2, s_{n-U+1}\}$ are coded integers. The entropy is estimated using the Shannon entropy formula by

$$E(U) = -\sum_{k=1}^{10^{U-1}} P_{S_u}(k) \cdot lnP_{S_U}(k), \tag{5}$$

where P_{S_u} is probability of observing k in S_u. The entropy $E(U)$ is normalized with

$$\widehat{NE(U)} = \frac{E(U) - E(U-1) + E(1) \cdot \alpha}{E(1)}, \tag{6}$$

where α is the percentage of the coded integers in S_i that occur only once. Finally, the entropy rate or regularity index is calculated by:

$$\rho = 1 - min\widehat{NE(U)}. \tag{7}$$

- The Lempel-Ziv complexity (L-Z) [27] describes predictability of the signal. To calculate L-Z complexity, first signal M is quantized to 100 equally spaced levels ranged from minimum to maximum, $A_1^n = \{a_1, a_2, ..., a_n\}$. In the next step, the signal A_1^n is decomposed into L different blocks, which are defined as $\Psi = \{a_j, a_{j+1}, ..., a_l\}$. The length of the blocks are $l - j + 1$. Now the signal can be represented by:

$$A_1^n = \{\psi_1, \psi_2, ..., \psi_n\} \tag{8}$$

The first block is equal to first symbol, $\Psi_1 = a_1$, while others are computed by

$$\Psi_{m+1} = A_{h_m+1}^{h_{m+1}}, m \geq 1, m \in \mathbb{Z}^+ \tag{9}$$

where h_m is ending index for ψ_m. At the end, the formula for calculating the L-Z complexity is given by:

$$LZ = \frac{L \log_{100} n}{n} \tag{10}$$

2.3.2. Frequency Domain Features

- The peak frequency is the frequency which contains most of the energy of the signal, defined as

$$f_p = \underset{f\in[0,f_{max}]}{\operatorname{argmax}} |F_M(f)|^2, \tag{11}$$

where f_{max} is the highest available frequency in a signal and F_M is the Fourier transform of the signal.

- The centroid frequency is the frequency at which the center of mass of the signal lies [18] and is computed by

$$f_c = \frac{\int_0^{f_{max}} f|F_M(f)|^2 df}{\int_0^{f_{max}} |F_M(f)|^2 df}. \tag{12}$$

- Bandwidth represents spectral spread and it is defined as

$$BW = \sqrt{\frac{\int_0^{f_{max}} (f-f_c)^2 |F_M(f)|^2 df}{\int_0^{f_{max}} |F_M(f)|^2 df}}. \tag{13}$$

2.3.3. Time-Frequency Domain Features

- The relative energy describes the concentration of the signal's energy at various frequency levels. It is computed using a 10-level discrete wavelet decomposition of the signal with the Meyer wavelet [24, 28, 29, 30, 31, 32]. The signal decomposition is written as $W_M = \{a_{10}, d_{10}, d_9, ..., d_1\}$, where a_{10} is approximation signal and and d_i is detail signal.

  The energy at each decomposition level is computed as

$$E_{a_{10}} = ||a_{10}||^2, \tag{14}$$

$$E_{d_i} = ||d_i||^2, \tag{15}$$

where $|| \bullet ||$ is the Euclidean norm of decomposition coefficient vectors.

$$E_T = E_{a_{10}} + \sum_{i=1}^{10} E_{d_i} \tag{16}$$

is the total energy of the signal, and the percent of relative energy contribution from each decomposition level is

$$E_{t_{a_{10}}} = \frac{E_{a_{10}}}{E_T} \times 100\%, \tag{17}$$

$$E_{t_{d_i}} = \frac{E_{d_i}}{E_T} \times 100\%. \tag{18}$$

- The wavelet entropy describes the spread of relative energy and it is computed as

$$WE = -\frac{E_{t_{a_1 0}}}{100} \cdot log_2 E_{t_{a_1 0}} - \sum_{i=1}^{10} \frac{E_{t_{d_i}}}{100} \cdot log_2 E_{t_{d_i}} \tag{19}$$

where $E_{t_{a_1 0}}$ and $E_{t_{d_i}}$ are relative energy calculated above from each wavelet decomposition level.

2.4. Data Analysis

The statistical differences between all different conditions were tested using the Kruskal-Wallis test [33]. Next, the Wilcoxon rank-sum test [34] was used for determining pairwise statistical differences between similar head motions. Namely, we examined the statistical differences between the 1 minute baseline and 10 seconds breath holding segments, between the positive and negative sagittal head tilts, and between the positive and negative coronal head tilts. Due to the clinical significance of the chin-tuck position, statistical differences between the 1 minute baseline and the positive sagittal position was examined as well. The Wilcoxon rank-sum test was also to examine sex effects. To examine the age effects on features, we employed a standard linear regression [35].

3. Results

In this section, we summarize the obtained results. As a starting point, we provide graphical representations of sample signals obtained in our experiment.

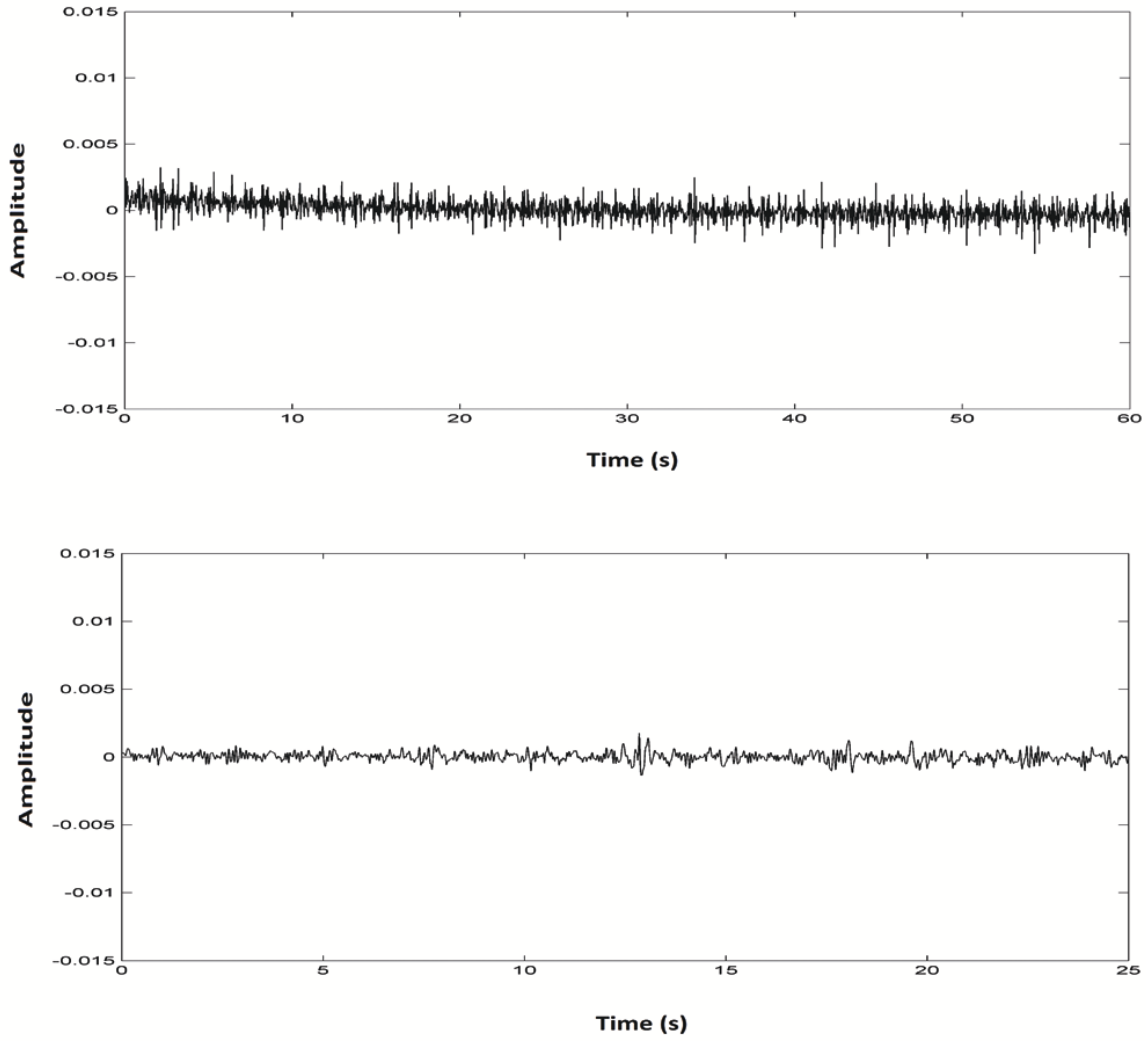

Figure 3: Example of the 1 minute baseline task (the upper signal) and tilt forward task (the lower signal)

Tables 1 and 2 summarize mean values of the different features, expressed as mean $\pm$ standard deviation. We first examined the effects of different tasks on all considered features. The Kruskal-Wallis test showed that only skewness (ν), peak frequency (f_p) and

$d1$ relative energy level did not exhibit significant statistical differences between different tasks ($p > 0.05$).

Table 1: Time domain features for cervical auscultation signals. * denotes multiplication by 10^{-2}, ** denotes multiplication by 10^2.

Feature	**1 minute baseline**	**10 sec hold breath**	**tilt forward**	**tilt backward**	**tilt right**	**tilt left**
s^*	0.04 ± 0.01	0.04 ± 0.01	0.19 ± 0.04	0.14 ± 0.05	0.11 ± 0.02	0.12 ± 0.04
ν	-0.53 ± 0.22	-0.98 ± 0.93	-1.21 ± 2.42	-0.79 ± 0.75	0.13 ± 0.09	1.04 ± 0.71
ϖ^{**}	50.4 ± 29.1	7.91 ± 4.52	26.2 ± 8.95	16.3 ± 7.74	15.1 ± 5.75	6.51 ± 2.81
ρ	0.99 ± 0.01	0.99 ± 0.01	0.99 ± 0.01	0.99 ± 0.01	0.99 ± 0.01	0.99 ± 0.01
LZ^*	0.81 ± 0.08	1.86 ± 0.12	0.66 ± 0.08	0.96 ± 0.11	1.14 ± 0.09	1.14 ± 0.09

Pairwise comparisons between right and left lateral flexion did not reveal significant differences for any of the features ($p > 0.05$). Pairwise comparison between flexion and extension did not show statistical differences for standard deviation (σ), skewness, and entropy rate (ρ), as well as for all of the frequency domain features ($p > 0.05$). Flexion experienced a higher mean value for kurtosis (ϖ) ($p = 0.02$) and a lower mean value for the Lempel-Ziv complexity ($p = 0.04$) than the extension. While performing pairwise comparisons for the time-frequency domain features, we observed that the wavelet entropy (WE) and most of the relative energy levels were not affected by head motion ($p > 0.05$). The relative energy distribution was only statistically different between flexion and extension for the levels $d4$ and $d3$ ($p < 0.05$).

Table 2: Frequency domain features for cervical auscultation signals.

Feature	**1 minute baseline**	**10 sec hold breath**	**tilt forward**	**tilt backward**	**tilt right**	**tilt left**
f_p	14.8 ± 0.79	16.5 ± 0.97	20.6 ± 3.96	15.2 ± 1.44	16.7 ± 3.03	16.8 ± 3.71
f_c	131 ± 40.8	84.2 ± 27.3	434 ± 80.1	287 ± 56.9	325 ± 87.4	218 ± 83.3
BW	556 ± 151	239 ± 69.7	963 ± 154	782 ± 127	325 ± 158	499 ± 116

The pairwise comparison between 1 minute baseline and 10 seconds breath holding were not statistically different the standard deviation, skewness and entropy rate ($p > 0.05$).

The 1 minute baseline showed higher mean value for kurtosis ($p << 0.01$) and lower mean value for L-Z complexity ($p << 0.01$) than the breath holding segments. The observation of frequency domain features for the same pairwise comparison shows that peak frequency was not statistically different between the 1 minute baseline and breath holding segments. However, the 1 minute baseline segment had higher mean value of centroid frequency (f_c) and bandwidth (BW) than the 10 seconds breath holding segments.

Figure 4 present results for the mean relative energy per the composition band. Wavelet entropy showed mean value of 0.93 ± 0.06 for 1 minute baseline, 0.97 ± 0.06 for 10 sec hold breath, 1.85 ± 0.11 for tilt forward, 1.67 ± 0.09 for tilt backward, 1.44 ± 0.09 for tilt right and 1.25 ± 0.09 for tilt left. The wavelet entropy and the relative energy in levels $a10$, $d3$ and $d2$, were not statistically affected ($p > 0.05$). Significant differences were found between the 1 minute baseline and 10 seconds breath holding segments for relative energy in the levels $d10$, $d9$, $d8$, $d7$, $d6$, $d5$, $d4$ and $d1$.

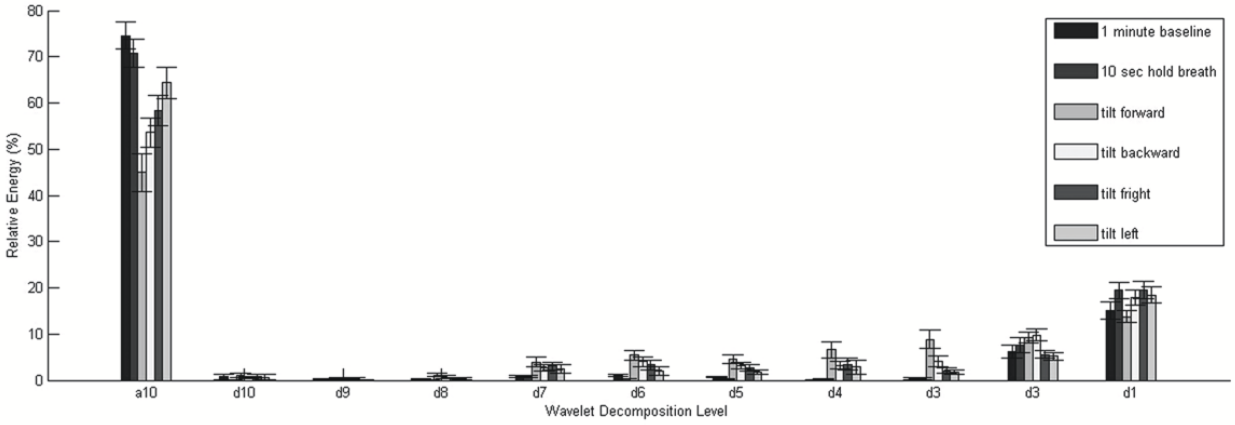

Figure 4: Mean relative energy per decomposition band.

Pairwise comparison between 1 minute baseline and head flexion did not show statistical difference for skewness, L-Z complexity, entropy rate, peak frequency and $d1$ relative energy level ($p > 0.05$), while all other feature show significant difference ($p << 0.01$).

Sex differences were not present for most of the features except for the skewness during the 1 minute baseline ($p = 0.02$) and kurtosis during left lateral flexion ($p = 0.02$).

According to the results of linear regression, frequency and time frequency domain features do not depend on the subject's age for all of the head motions ($p > 0.05$). There was an observed age dependence of skewness and kurtosis for 10 seconds breath holding, extension, right and left lateral flexion ($p < 0.02$) tasks. Standard deviation was affected with age for extension, right and left lateral flexion, while L-Z complexity were affected for extension and right lateral flexion ($p < 0.02$).

4. Discussion

In this paper, we extracted a number of features in different acoustic signal domains during various head maneuvers. Understanding relationships between these maneuvers and associated features will help us to further understand swallowing sounds and the mechanism that generates these sounds. Since head movements could be present during swallowing, these presented findings are important for future investigations of swallowing sounds.

4.1. The Effects of Head Maneuvers on Cervical Auscultation Signals

A lower mean value for the Lempel-Ziv complexity implies higher predictability. In this case, as we completed a pairwise comparison between flexion and extension, a lower values denotes that flexion produces a more predictable signal than tilting in the extension. Also, kurtosis describes "peakness" of the amplitude probability distribution of the signal. A higher mean value of the kurtosis for the flexion task than for the extension task means that extension contains more variant amplitudes in the sound signal. However, the behavior of the sensors on the skin during motion needs to be considered. During extension, the sensor moves with the skin over the laryngeal framework and produces a sound. This behavior likely explains higher kurtosis and higher predictability of flexion compared to extension.

A lower kurtosis value for breath holding segments than for the 1 minute baseline task implies that the 1 minute baseline task contains less components of the different amplitudes (loudness) than that the baseline task. The microphone attached on the subject's neck can also record sounds from the carotid artery [36]. Studies have shown that the heart rate increases and becomes more prominent while holding breath in comparison to the resting

state [37]. These heart rate changes can potentially provide more signal components, which can explain the results for kurtosis, as well as the lower result of L-Z complexity for 1 minute baseline to the 10 seconds holding breath task. A lower mean value of the Lempel-Ziv complexity for the 1 minute baseline implies that task tends to be a more well defined pattern than the 10 seconds breath holding task, which is expected since the cervical auscultation signal for the 10 seconds breath holding task can change with time as the heart rate increases. A higher value for the centroid frequency during the 1 minute baseline than during the 10 second breath holding tasks can be attributed to higher bandwidth values during the 1 minute baseline task. Given the effects of the heart rate on swallowing sounds, future studies should investigate if these effects can be annulled via filtering operations.

The comparison between the 1 minute baseline and flexion is a clinally important question. A similar study has been done with the accelerometer signal [38], which showed that the flexion contain low frequency components which contaminate the signal information. The significant influence of head motion can be observed even by visual inspection of the accelerometer signals in time domain for this two tasks, and an algorithm was developed that removes low frequency components produced by head motions from the swallowing accelerometry signals [20]. In the case of swallowing sounds, statistical differences for most of the features between the 1 minute baseline and the flexion task are expected due to the clearly different behaviors. Of particular interest for this pairwise comparison is the peak frequency, which denotes the frequency component with the greatest energy. Statistically different peak frequency values would mean that the flexion contains dominant frequency components different from the dominant frequency component during the baseline signals, resulting in motion-based artifacts found in accelerometer signals. The presented result did not show statistical difference for the peak frequency for these two tasks. Visual inspection of the swallowing sounds also did not find any significant differences (Figure 3). Hence, we can conclude that there is no need for removing signal components associated with head movements.

We anticipate that the observed age effects are due to the behavior of the skin in older

subjects. With the age, skin loses the collagen and elastin which are supportive connectivity for the tissue of the skin. These changes causes wrinkling, laxity and sagging of the skin [39]. The attached sensor on the subject's neck should record sounds next to cricoid cartilage through the skin. Due to the sagging skin on the neck, it is possible that the microphone does not directly sit at top of the laryngeal framework in various head position. Consequently, some information is lost or artifacts are introduced (for example, a sound produced by touching the skin with cricoid cartilage when the head moves backward).

Males had significantly higher mean values for skewness during the 1 minute baseline task and for kurtosis during the tilting left task. We anticipate that these sex differences are not of any importance, as there is no theoretical reasons for these features to differ between genders during our passive recording tasks (e.g., [40] and references within).

4.2. Remarks

In the same way that background noise influences the acoustic signal perceived by humans in a sound field, head movement can introduce extraneous artifact into recorded swallowing acoustic data. This problem also exists with imaging studies such as videofluoroscopy, and with nonimaging myoelectric signals such as those recorded with surface electromyography. The primary goal of this study was to investigate potential acoustic artifacts produced only by common head movements performed by dysphagic patients using compensatory maneuvers while swallowing. The current manuscript does not make any inferences about muscle activity or other biomechanical events associated with swallowing or with head movements as our study were not geared towards such activities. In future studies we hope to be able to subtract this artifact from the total acoustic signal to derive more specific information regarding physiologic events occurring during swallowing.

5. Conclusion

In this paper, the baseline characteristic of non-swallowing cervical auscultation signals and the effects of head movements on their characteristics were analyzed. Signals were collected from 56 participants and 10 different features were considered. Statistical differences

between clinically relevant head movements were examined. Any age and gender affects on a signal were also observed and discussed. We found that head tilting forward (flexion) and tilting backward (extension) influences some features, but these head movements do not affect the peak frequency and so it is not necessary to remove them from the signal. However, the study also showed that certain features exhibited age dependence. These findings may indicate that sex, head position and possibly other variables may influence swallowing acoustics. Further exploration of these findings may generate methods that increase the diagnostic value of CA. For CA to be eventually make its way into clinical usefulness as a valid and reliable screening or diagnostic method for dysphagia, abnormalities in swallow physiology need to be very reliably attached to specific acoustic signals that can be discerned either perceptually or with instrumentation.

References

[1] C. Wiles, "Neurogenic dysphagia," *Journal of Neurology, Neurosurgery and Psychiatry*, vol. 54, no. 12, p. 1037, 1991.

[2] D. Gottlieb, M. Kipnis, E. Sister, Y. Vardi, and S. Brill, "Validation of the 50 ml3 drinking test for evaluation of post-stroke dysphagia," *Disability and Rehabilitation*, vol. 18, no. 10, pp. 529–532, 1996.

[3] P. E. Marik, "Aspiration pneumonitis and aspiration pneumonia," *New England Journal of Medicine*, vol. 344, no. 9, pp. 665–671, 2001.

[4] H. M. Finestone, N. C. Foley, M. G. Woodbury, and L. Greene-Finestone, "Quantifying fluid intake in dysphagic stroke patients: a preliminary comparison of oral and nonoral strategies," *Archives of Physical Medicine and Rehabilitation*, vol. 82, no. 12, pp. 1744–1746, 2001.

[5] D. Ambrosi, R. Goldstein, M. Horn, and U. Bogdahn, "Predictors of survival after severe dysphagic stroke," *J Neurol*, vol. 252, pp. 1510–1516, 2005.

[6] I. Katzan, R. Cebul, S. Husak, N. Dawson, and D. Baker, "The effect of pneumonia on mortality among patients hospitalized for acute stroke," *Neurology*, vol. 60, no. 4, pp. 620–625, 2003.

[7] D. J. Ramsey, D. G. Smithard, and L. Kalra, "Early assessments of dysphagia and aspiration risk in acute stroke patients," *Stroke*, vol. 34, no. 5, pp. 1252–1257, 2003.

[8] J. A. Logemann, *Evaluation and treatment of swallowing disorders.* Austin, TX: PRO-ED Inc, 1998.

[9] A. Tabaee, P. Johnson, C. Gartner, K. Kalwerisky, R. Desloge, and M. Stewart, "Patient-controlled comparison of flexible endoscopic evaluation of swallowing with sensory testing FEESST and videofluoroscopy," *The Laryngoscope*, vol. 116, no. 5, pp. 821–825, 2006.

[10] P. Leslie, M. Drinnan, P. Finn, G. Ford, and J. Wilson, "Reliability and validity of cervical auscultation: a controlled comparison using videofluoroscopy," *Dysphagia*, vol. 19, no. 4, pp. 231–240, 2004.

[11] J. Cichero and B. Murdoch, "Acoustic signature of the normal swallow: characterization by age, gender, and bolus volume," *The Annals of Otology, Ohinology and Laryngology*, vol. 111, no. 7 Pt 1, p. 623, 2002.

[12] S. Hamlet, R. Nelson, and R. Patterson, "Interpreting the sounds of swallowing: fluid flow through the cricopharyngeus." *The Annals of Otology, Rhinology and Laryngology*, vol. 99, no. 9 Pt 1, p. 749, 1990.

[13] P. Zenner, D. Losinski, and R. Mills, "Using cervical auscultation in the clinical dysphagia examination in long-term care," *Dysphagia*, vol. 10, no. 1, pp. 27–31, 1995.

[14] J. Lee, S. Blain, M. Casas, D. Kenny, G. Berall, and T. Chau, "A radial basis classifier for the automatic detection of aspiration in children with dysphagia," *Journal of NeuroEngineering and Rehabilitation*, vol. 3, no. 1, pp. 14–1–17, 2006.

[15] J. L. Coyle, C.-S. BRS-S, G. A. Ford, and J. A. Wilson, "Cervical auscultation synchronized with images from endoscopy swallow evaluations," *Dysphagia*, vol. 22, no. 4, pp. 290–298, 2007.

[16] K.-i. Michi, "Methodology for detecting swallowing sounds," *Dysphagia*, vol. 9, no. 1, pp. 54–62, 1994.

[17] S. Youmans and J. Stierwalt, "An acoustic profile of normal swallowing," *Dysphagia*, vol. 20, no. 3, pp. 195–209, 2005.

[18] E. Sejdić, V. Komisar, C. Steele, and T. Chau, "Baseline characteristics of dual-axis cervical accelerometry signals," *Annals of Biomedical Bngineering*, vol. 38, no. 3, pp. 1048–1059, 2010.

[19] E. Sejdić, C. Steele, and T. Chau, "Segmentation of dual-axis swallowing accelerometry signals in healthy subjects with analysis of anthropometric effects on duration of swallowing activities," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 4, pp. 1090–1097, 2009.

[20] E. Sejdić, C. Steele, and T. Chau, "A method for removal of low frequency components associated with head movements from dual-axis swallowing accelerometry signals," *PLoS One*, vol. 7, no. 3, p. e33464, 2012.

[21] D. T. Barry, T. Hill, and D. Im, "Muscle fatigue measured with evoked muscle vibrations," *Muscle and Nerve*, vol. 15, no. 3, pp. 303–309, Mar. 1992.

[22] N. Alves and T. Chau, "The design and testing of a novel mechanomyogram-driven switch controlled by small eyebrow movements," *Journal of NeuroEngineering and Rehabilitation*, vol. 7, pp. 22–1–10, May 2010.

[23] D. Donoho, "De-noising by soft-thresholding," *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 613–627, 1995.

[24] J. Lee, C. Steele, and T. Chau, "Time and time-frequency characterization of dual-axis swallowing accelerometry signals," *Physiological Measurement*, vol. 29, no. 9, p. 1105, 2008.

[25] A. Porta, S. Guzzetti, N. Montano, R. Furlan, M. Pagani, A. Malliani, and S. Cerutti, "Entropy, entropy rate, and pattern classification as tools to typify complexity in short heart period variability series," *IEEE Transactions on Biomedical Engineering*, vol. 48, no. 11, pp. 1282–1291, 2001.

[26] A. Porta, G. Baselli, D. Liberati, N. Montano, C. Cogliati, T. Gnecchi-Ruscone, A. Malliani, and S. Cerutti, "Measuring regularity by means of a corrected conditional entropy in sympathetic outflow," *Biological Cybernetics*, vol. 78, no. 1, pp. 71–78, 1998.

[27] A. Lempel and J. Ziv, "On the complexity of finite sequences," *IEEE Transactions on Information Theory*, vol. 22, no. 1, pp. 75–81, 1976.

[28] O. Rosso, S. Blanco, J. Yordanova, V. Kolev, A. Figliola, M. Schurmann, and E. Basar, "Wavelet entropy: a new tool for analysis of short duration brain electrical signals," *Journal of Neuroscience Methods*, vol. 105, no. 1, pp. 65–76, 2001.

[29] E. Sejdić, I. Djurović, and J. Jiang, "Time-frequency feature representation using energy concentration: An overview of recent advances," *Digital Signal Processing*, vol. 19, no. 1, pp. 153–183, Jan. 2009.

[30] S. Pittner and S. Kamarthi, "Feature extraction from wavelet coefficients for pattern recognition tasks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 1, pp. 83–88, 1999.

[31] G. Yen and K. Lin, "Wavelet packet feature extraction for vibration monitoring," *IEEE Transactions on Industrial Electronics*, vol. 47, no. 3, pp. 650–667, 2000.

[32] S. Stanković, I. Orović, and E. Sejdić, *Multimedia Signals and Systems*. New York, NY: Springer US, 2012.

[33] W. H. Kruskal and W. A. Wallis, "Use of ranks in one-criterion variance analysis," *Journal of the American Statistical Association*, vol. 47, no. 260, pp. 583–621, 1952.

[34] V. DePuy, V. Berger, and Y. Zhou, "Wilcoxon-Mann-Whitney test," *Encyclopedia of Statistics in Behavioral Science*, 2005.

[35] D. C. Montgomery, E. A. Peck, and G. G. Vining, *Introduction to linear regression analysis*. Wiley, 2012, vol. 821.

[36] J. F. Stapleton and M. M. El-Hajj, "Heart murmurs simulated by arterial bruits in the neck," *American Heart Journal*, vol. 61, no. 2, pp. 178–183, 1961.

[37] P. M. Gross, B. J. Whipp, J. T. Davidson, S. N. Koyal, and K. Wasserman, "Role of the carotid bodies in the heart rate response to breath holding in man," *Journal of Applied Physiology*, vol. 41, no. 3, pp. 336–340, 1976.

[38] E. Sejdić, C. Steele, and T. Chau, "The effects of head movement on dual-axis cervical accelerometry signals," *BMC Research Notes*, vol. 3, no. 1, p. 269, 2010.

[39] S. Imayama and I. Braverman, "A hypothetical explanation for the aging of skin. chronologic alteration of the three-dimensional arrangement of collagen and elastic fibers in connective tissue." *The American Journal of Pathology*, vol. 134, no. 5, p. 1019, 1989.

[40] D. Hung, E. Sejdić, C. M. Steele, and T. Chau, "Extraction of average neck flexion angle during swallowing in neutral and chin-tuck positions," *Biomedical Engineering Online*, vol. 8, no. 1, p. 25, 2009.