ACCELERATION SENSOR MEASUREMENTS OF VIBRATIONS OF THE LARYNX IN PATIENTS WITH VOCAL FOLD ADDUCTION DEFICIENCIES

Wolfgang Wokurek*, Manfred Pützer†

*Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart, Deutschland
†Klinische Phonetik, Institut für Phonetik, Universität des Saarlandes, Saarbrücken, Deutschland
wokurek@ims.uni-stuttgart.de, puetzer@coli.uni-saarland.de

Abstract: In this study we investigate voice productions of normal speakers and patients with varying vocal fold adduction deficiencies using a non-invasive method to measure spatial vibrations of the larynx. The current version ACCV4 of our acceleration sensor device was used. The study’s primary goal is to find out whether lesions of vocal folds lead to additional spatial vibration modes compared to a non-disordered voice. Our results allow to assume that the composition of spatial vibration modes at the skin over the cricothyroid ligament may depend on the symmetry of vocal fold movements.

Keywords: acceleration sensor, vocal fold adduction deficiencies

I. INTRODUCTION

We use the current version ACCV4 of our acceleration sensor device to measure spatial vibrations of the skin of the neck covering the larynx. These vibrations are driven in part by the subglottal sound pressure and also indirectly by the vocal folds. We suspect that the movement of each vocal fold is passed on via their own arytenoid cartilage to the left or right part of the cricoid cartilage, respectively. The cricoid cartilage is situated at the lower part of the cricothyroid ligament [3]. Therefore, a path of vocal fold vibrations to the skin over this ligament can be assumed. Both the subglottal sound pressure and the symmetric vocal fold movements result in skin and sensor movements in the ventral and dorsal direction. These movements were attributed to the subglottal sound pressure alone in our previous studies [1], [4], [5], [6], [7]. Any deviations of vocal fold movements from symmetry could cause additional skin and sensor movements in lateral and/or cranial and caudal direction. We attempt to quantify the amount of asymmetry of the skin and sensor movement by an analysis of the spatial sensor movement.

II. METHODS

A. Acceleration Sensor

In this study the spatial vibrations of the current version ACCV4 of the acceleration sensor device are recorded simultaneously with the nasal and oral sound pressure signal captured separately through a Rothenberg mask. Preceeding versions of the acceleration sensor device were presented in [4], and [5]. The relevant aim in this study is its ability to track the spatial movement of the body tissue with a high bandwidth.

Figure 1: Acceleration sensor device

The acceleration sensor device is shown in Fig. 1. The tip T is mounted at the acceleration sensor A.

© 2011 Firenze University Press
Both are held by spiral springs in the suspension ring S. The suspension ring is glued to the handle H that also contains electrical connectors to conduct the analog signals to an external preamplifier by a cable.

The acceleration sensor A consists of three ADXL202E two axis microelectromechanical acceleration sensors that are glued to different planes of an aluminium cube. The tip T is a plastic screw fixed at the cube by a counter nut. The electronic components are soldered to a flexible printed circuit board (PCB). The force of the tip T to the neck is about 0.2 N – which proved to be strong enough to keep tissue contact but is hardly noticed by the speaker.

The arrangement of the ADXL202E chips tracks the acceleration along each spatial direction at two different points of the cube. Hence, this six signals are sufficient to compute the spatial vibrations of the body tissue.

B. Sensor Placement

The glottis is located in the larynx and separates the supraglottal from the subglottal cavity. It lies behind the thyroid cartilage. A soft tissue – the cricothyroid ligament – connects the lower end of the thyroid cartilage to the cricoid cartilage. The vocal fold vibration is passed on through the thyroid, arytenoid, cricoid cartilage, and the cricothyroid ligament to our sensor, respectively.

![Figure 2: Sensor at neck](image)

The cricothyroid ligament can be found by touching the larynx with a finger and searching for a small soft gap in the elsewhere hard larynx structure. The sensor tip T is placed perpendicular to the neck and pressed gently to the soft gap until the suspension ring S touches the skin as shown in Fig. 2. Now the speaker is asked to speak. The correct placement of the sensor is immediately seen in the amplitude display of the six accelerometer channels. The amplitude of the two channels corresponding to the tip axis rise to high levels, the other four stay at low levels. In many cases this situation holds for several minutes. Sometimes the perpendicular position of the tip to the neck is lost and the signal amplitude distributes over more than two channels. In that case the session is paused and the sensor is arranged correctly again. In our recordings usually the sensor device was held by an assistant.

C. Speech sounds

Speech sounds are recorded via two electret microphones mounted in the oral and nasal section of a Rothenberg mask. In this study both sounds are added and used for labelling the short and long vowels. The lower part of the mask is visible in Fig. 2. The mask was held by the speaker.

D. Recordings

The recordings were made in a consultation room that was not sound treated. Eight channels were recorded simultaneously, six channels of the acceleration sensor as well as the oral and the nasal sound of the Rothenberg mask. The first order 5 kHz RC-lowpass recommended by the ADXL202E data sheet was implemented by analog hardware. All channels were digitized with a sampling rate of 48 kHz. The RME soundcard offers only AC coupling, hence no static acceleration signals like the gravitation vector are available as a direction reference in the evaluation.

As speech material sustained vowels (e.g. [i:], [a:] [u:]) produced at the subject’s normal pitch were used for this study.

E. Spatial analysis

A segment stable for about a second is located manually in the sound of one of the sustained vowels. The temporal sample indices \( n = 1, \ldots, N \) correspond to that segment. To study the modes of vibrations of the sensor, the samples \( a_i(n) \) of the six sensor channels are arranged in columns

\[
a_i = (a_i(1), \ldots, a_i(N))^*, \quad i = 1, \ldots, 6
\]

where \(^*\) denotes transposition. The columns are put together to form the \( N \times 6 \) matrix of acceleration data

\[
A = (a_1, \ldots, a_6)
\]
Each row of A may be viewed as a sample of the six dimensional vector valued sequence of acceleration measurements.

Whereas the mechanical sensor axes are not perfectly aligned parallel and perpendicular, no measurement of these deviation was done and hence no correction can be attempted. According to the data sheet of the sensor chip ADXL202E, the cross axis sensitivity of each sensor chip is ±2% or −34dB. It stems from axis misalignments and inherent sensor errors. Sensor assembly misalignments of 1 degree would result in additional cross sensitivity of about −39dB.

To find independent modes of vibration in the acceleration vector sequence A, the $6 \times 6$ correlation matrix is computed

$$ R_A = A^* A $$

The eigen decomposition of this correlation matrix

$$ R_A = V^* \Lambda V $$

results in the diagonal matrix

$$ \Lambda = \text{Diag}(\lambda_1, \lambda_2, \ldots, \lambda_6) $$

with non-negative eigenvalues

$$ \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_6 \geq 0 $$

and in the orthogonal matrix

$$ V = (v_1, v_2, \ldots, v_6) $$

containing the eigenvectors $v_i$ as columns.

According to Ineq.(6), the eigenvalues are arranged in the order of descending magnitude starting with the largest eigenvalue $\lambda_1$. The eigenvalue $\lambda_1$ represents the energy of the major vibration mode. The direction of the major vibration mode is given by the corresponding eigenvector $v_1$. Similarly, the second vibration mode is given in energy and direction by $\lambda_2$ and $v_2$. Due to the symmetry of the correlation matrix, the eigenvectors corresponding to different eigen-values are always orthogonal. The case of equal or multiple eigenvalues with its associated subspace is not considered here further, since it never appeared in our measurements and it is very unlikely due to measurement noise.

The orientation of the major vibration mode is basically perpendicular to the skin at the neck, along the ventral and dorsal direction. The second mode vibrates along a line in the plane spanned by the lateral and the cranial and caudal direction. In order to quantify the symmetry of the sensor movement we propose the ratio between the energies of the major and the second vibration mode

$$ \sigma = \frac{\lambda_1}{\lambda_2} $$

A large $\sigma$ corresponds to a dominant major vibration mode and a weak second vibration mode. In this situation the vibration of the cricothyroid ligament in lateral and/or cranial and caudal direction is weak—a highly symmetric vibration. On the other hand a stronger vibration mode in lateral and/or the cranial and caudal direction reduces $\sigma$ and corresponds to a more asymmetric vibration.

Since energy ratios may result in large figures the logarithmic decibel scale

$$ \sigma_{dB} = 10 \log \sigma $$

is more familiar and often preferred. Both versions of the proposed symmetry measure $\sigma$ and $\sigma_{dB}$ will be shown in Tab.1.

F. Speakers

We investigated normal voices produced by two speakers with no known speaking or hearing problems as a control group. Additionally, three patients with varying vocal fold adduction deficiencies resulting from unilateral and bilateral paralysis of the recurrent nerve were considered [2]. This kind of pathology is a frequent cause of deficient vocal fold adduction. Patients compensate or do not use compensatory strategies for the adduction deficiency. Our three patients cover a wide range of physiological constellations. They were classified on the basis of the observed vocal fold adduction, judged from laryngoscopic and videostroboscopic recordings of their vocal folds during phonation by an experienced ENT physician. The clinical judgements were made during consultation.

III. RESULTS

The acceleration sensor device was previously used to get indirect access to the subglottal sound pressure and to measure the resonance parameters of the sub-glottal cavity [5]. It records the spatial components of the acceleration of its moving part. The analysis is based on the eigen decomposition of the correlation matrix of the acceleration vector. The projection to its main component was assumed to be driven mainly by the subglottal sound pressure. Now the strength of the second largest component is compared to the strength of the main component.
Table 1: Symmetry measure \( \sigma \) for normal adduction behaviour and different vocal fold adduction deficiencies.

<table>
<thead>
<tr>
<th></th>
<th>norm. voice</th>
<th>unilat. uncomp.</th>
<th>unilat. comp.</th>
<th>bilat. comp.</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma )</td>
<td>15</td>
<td>23</td>
<td>4.7</td>
<td>12</td>
</tr>
<tr>
<td>( \sigma_{dB} )</td>
<td>12dB</td>
<td>14dB</td>
<td>7dB</td>
<td>11dB</td>
</tr>
</tbody>
</table>

Tab. 1 shows the resulting ratios of the underlying different vocal fold adduction behavior. Normal voices (norm. voice) having relative regular adduction behaviour (first column in Tab. 1) show a ratio of 15 and 23. The first vibration modes for these individuals are 12dB and 14dB stronger than their second modes.

Our first patient with uncompensated unilateral vocal fold paralysis (unilat. uncomp.; second column in Tab. 1) produces a much stronger second vibration mode which is only 7dB weaker than the first mode. This result may indicate non-symmetric vocal fold movements.

Our second patient with compensated unilateral vocal fold paralysis (unilat. uncomp.; third column in Tab. 1) offers results very closed to those of normal voices. Consequently, compensation of vocal fold paralysis may restore symmetric vocal fold movements.

Finally, our third patient with compensated bilateral vocal fold paralysis (bilat. comp.; fourth column in Tab. 1) shows a very weak second vibration mode which may be caused by a high degree of symmetry in the vocal fold movement.

A closer look to the eigenvectors of our speakers confirms that the major mode vibrates, as conjectured, in the ventral and dorsal direction. The second vibration mode turns out to have always a component in the cranial and caudal direction. An additional lateral component in the second vibration mode is only seen with the second patient, not as expected with the first one.

IV. DISCUSSION

In the present study voice productions of normal speakers and patients with varying vocal fold adduction deficiencies were investigated. Instrumentally the current version ACCV4 of our acceleration sensor device was used. In extension to our previous approaches the spatial capabilities of the sensor were made use of. To measure the amount of symmetry of the vocal fold vibration the energy ratio of the first and second vibration mode was proposed and evaluated. It seems to mirror the symmetry condition of the vocal fold vibration despite of the underlying complex coupling path via arytenoid cartilage, cricoid cartilage, and the cricothyroid ligament.

V. CONCLUSIONS

Phonation behavior of patients with vocal fold adduction deficiencies resulting from unilateral and bilateral paralysis of the recurrent nerve show varying degrees of symmetry in their vocal fold vibration. The proposed symmetry measure of the energy ratio of the first and second vibration mode properly represents this situation. These observations encourage a further look at other phonation qualities to find out whether this symmetry measure is still applicable.

REFERENCES