Most of the hearing disorders might be treatable in the early age of life. There are variant methods to test the hearing abilities of an infant; still none of them cover the whole auditory system by itself. Thus a comprehensive audiometry with high hit probability is required. Crying is the first tool of communication for an infant. Emotions, state of health or several organ functions can be obtained from it. Normally, an auditory feedback works between the produced sound and the brain. When this feedback has problems, crying may differ from the original one. Authors are to find these differences by acoustic analysis. This paper deals with acoustic features of the infant cry in order to detect hearing impairments.

1 Introduction

1.1 The infant cry

Crying is an important information source about the infant. It is a doubtless sign of life at birth. Investigations dealing with the infant cry can be classified into three main categories: investigations of the reason of crying, investigations of the development of crying and investigations of the connection between diseases and crying.

Infant cry is a mean of communication. The sound of crying is in the most sensible range of the human auditory sensation area, in order to let the babies express their needs. Crying is different if the baby is hungry, sleepy, bored or need to be changed. Several research teams had been investigated the reason of crying, they tested mostly acoustical and psychological respects [1-4].

The second main category is the investigation of the development of crying. The sound of crying has a development in the same way as the infant itself [5-7]. The cry of preterm infants is acoustically different from the normal term babies in the first three months of life [8]. According to Zeskind et al. the sound of crying is characteristic to the birth process [9]. If there are complications at birth the sound of crying is much sharper and shorter; the time of the pause between crying blocks is longer than they are at normal birth [10]. Crying can also refer to emergency. E.g. normal term infants kept in incubator observably decrease the volume of their crying, while normally it should be increased [11].

Last but not least, the infant cry can be an important aid in medical diagnostics. Experienced otorhinolaryngologist paediatricians use their impressions from the movements and crying of the infant in the diagnostics. E.g. infants suffers from Mongolism have a very weak sound of crying [12, 13].

1.2 Hardness of hearing

Approximately 0.2-0.3% of the infants is born with at least a medium grade permanent hearing loss (the hearing threshold is above 40 dB) [14]. A baby learns to speak by hearing its environment. It is essential for the correct development of speech to have good hearing in the critical period. This critical period (i.e. in the case of the development of the organ of speech and organ of hearing) is between the first 6 and 24 months of life. There are further repercussions of deafness in the physical, psychic or intellectual development.

The hearing system is a complex system, which has numerous elements cooperating together. The main elements are: the pinna, the ear canal, the ear drum, the auditory ossicles (malleus, incus and stapes), the oval and round windows, the Eustachian tube, the cochlea, the nerve fibers and the central nervous system. If any of these elements has functional disorders, hardness of hearing will occur.
Hearing disorders can be cured in many cases, if they are detected in the first few months of life.

1.3 Basic hypothesis

There is an auditory feedback between the organ of speech and the brain, which instructs the organ of speech to create sound [15]. If this feedback is missing, the created sound may differ from the normal sound. Our basic hypothesis is the following: infants who have hearing disorders or hearing loss cry unlike healthy infants.

Our primary goal is to detect deafness by analyzing the sound of crying. By means of the above mentioned auditory feedback the whole hearing system could be tested on this wise. There are several parameters which can be obtained from crying, as length of the crying blocks, as fundamental frequency, as spectral components, as melody contour of crying, as rising time, etc [16].

Authors have already tested several parameters of crying, in this study the rising time of crying will be investigated between healthy and deaf (partly or totally) infants.

The melody contour (which correlates with the rising time) is an important parameter to determine the mood of the infant [17, 18]. Authors want to examine, if there is any connection between deafness and the rising time (which is part of the melody). The definition of the rising time will be given later.

2 Data collection

Recordings were mainly made in the Heim Pál Hospital for Sick Children in a quiet room. Authors used a digital camera (SONY DCR-TRV25) in order to recognize the infant and the circumstances of crying. The sampling frequency for audio recording was 48 kHz; there were 16 bits assigned to each sample. No image processing was used. During the recording the infants were sitting on their mothers’ lip, the distance between the microphone and the mouths of the infants was 1-2 meters.

For this study, cries from 70 infants were recorded; the length of the cry signals is 27.49 s on the average. The hearing function of all the infants was assessed by means of subjective and objective audiometry. 35 of the babies had hearing loss or deafness; the other 35 had normal hearing. The mean age was 11 months, with a standard deviation of 8 months. There were boys (40) as well as girls (30) within the group.

The cries were collected during manual audiometry; the doctor looked into the ears of the infant. The procedure was painful and the baby started to cry. The cries selected for analysis were, whenever possible, chosen from the start of the cry sequences. The cries for each baby were, however, usually very similar both auditorily and on the spectrograms.

3 Method

Authors used their own developed Matlab software to evaluate preprocessing, feature extraction and fundamental frequency detection, and to plot the melody contour and calculate the rising time.

3.1 Preprocessing

Before the acoustic analysis starts, preprocessing is necessary to eliminate the technical defects of the recording and to find the important parts of the whole crying signal. Two main tasks are defined, they are filtering and feature extraction.

The recorded sounds may contain unwanted effects, as background noises, echo, etc. Some of these effects can disadvantageously affect the results of the analysis. In the frequency domain the lowest component of the infant cry is not less than 250-300 Hz, thus using a high-pass filter, with a frequency cut-off at 250 Hz is a good solution to reduce most of the background noise [19].

The feature extraction is also called as Automatic Infant Cry Recognition [20]. The basis of this procedure is to find the useful crying signals among pauses and disturbances in the whole signal. Crying signals (see in Figure 1) can be identified by their frequency content or energy [21]. In a typical crying process crying signal is the first signal after respiration.

The automatic identification has found 556 crying signals from the 70 infants.

3.2 Fundamental frequency detection

The fundamental frequency ($F_0$) is the smallest, useful frequency component of the spectrum. In the case of crying the range of the fundamental frequency is between approximately 300 and 750 Hz [22-24].
Each crying signals are divided into successive windows, which are same in size. $F_0$ is calculated to each window. Authors chose such a small window size to be able to ignore the changing of the frequency components within a window.

Because crying is a harmonic signal (except crying with disturbances, as hoarseness), the fundamental frequency can be determined by the regular structure of the spectrum. Authors developed an algorithm (Smoothed Spectrum Method) to detect $F_0$ in each window. It is based on the harmonic structure of the spectrum and statistical algorithms. This method works with 97.99 % reliability [25].

3.3 Plotting the melody contour

The melody contour can be obtained by imaging the fundamental frequency as a function of time.

There are typical types of the melody contour, as rising, rising-falling, falling, and flat, etc [26]. Crying melodies can be categorized by duration as well. In a previous study it was proved that the duration doesn’t differ between the group of infants of normal hearing and hard of hearing [27].

Because the range of the fundamental frequency comes near the frequency range of the music paper, authors represented the melody contour on a music paper as it was used in the eighties [28]. But the original music paper, which contains five lines (namely E4, G4, H5, D5, F5), has an unequal scaling, so the original music paper cannot be used as a correct frequency scaling. Authors developed the Five Line Method (FLM) to visualize the melody contour, which i.a. applies logarithmic scaling to obtain equal ratio between the neighboring lines [29].

The difference of the five lines between the music paper and the Five Line Method is shown in the following Matlab code.

```matlab
>> lin_old=[329.63,392.0,493.88,587.33,698.46];
>> lin_old(2:5)./lin_old(1:4)
ans = [1.1892, 1.2599, 1.1892, 1.1892]
>> lin_new=[329.63,397.7,479.82,578.91,698.46];
>> lin_new(2:5)./lin_new(1:4)
ans = [1.2065, 1.2065, 1.2065, 1.2065]
```

Figure 2 shows an example to the FLM. There is the time in seconds on the X-axis. The Y-axis is a frequency axis, containing the (round) values of the new five lines.

3.4 Calculating the rising time

The rising time ($t_{rising}$) can be only defined if the melody of the crying signal starts with rising; it is the duration of the first rising period.

In the following examples, we use the Matlab code of a crying signal.

```matlab
>> cry009-02.wav
```

The rising time is shown with the vertical dashed line.

**Figure 2:** an example to the Five Line Method.

E.g. on the previous figure (Figure 2) the melody starts with a rising period, the duration of that period (i.e. the rising time) is approximately 0.3 s.

To obtain the rising time the following steps are to evaluated.

\[ M(t) = F_0(y(t)) \] (1)

Where $M(t)$ is the melody contour of the $y(t)$ crying signal. $M(t)$ is a function of time. $F_0\{\}$ means the detection of the fundamental frequency.

To find the rising period, the authors check the derived function of the melody contour.

\[ \frac{dM(t)}{dt} |_{0<t<t_{rising}} > 0 \] (2)

If the derived function is positive at starting, it means the melody starts with a rising period. In this case the rising time is interpretable, and the value of $t_{rising}$ can be obtained by finding the first zero value at the derived function of the melody contour.

The next figure (Figure 3) shows a melody contour and its derived function. It can be clearly seen that the rising time is at the zero point of the derived function (in this case it is 0.27 s).

**Figure 3:** The melody contour and its derived function. The rising time is shown with the vertical dashed line.
The following Matlab code shows how to obtain the rising time from a crying signal in practice. This code contains the before mentioned methods as \textit{preproc}, \textit{ssm} and \textit{flm} functions.

\begin{verbatim}
% input is the filename and the window size
% output will be the auto_detected rising time
%--------------------------------------------
% importing the wave file and the sampling
% frequency into variables
[y,fs] = wavread(filename);
% preprocessing (filtering, etc.)
y = preproc(y);
% number of windows
win_number = floor(length(y)/win_size);
%--------------------------------------------
for i = 1:win_number
    actual_window = y((i-1)*win_number+1:i*win_number);
    % calculating the fundamental
    % frequency to the actual window by
    % the Smoothed Spectrum Method
    F0 = ssm(actual_window,fs);
    melody(i) = F0;
end
% imaging the melody with the Five Line Method
flm(melody,fs,win_size)
%--------------------------------------------
% auto-calculating the rising_time
start_point = 0;
finish_point = ...
(min(find(diff(melody)<0))-1)*win_size/fs;
rising_time = finish_point - start_point;
\end{verbatim}

The graphical user interface of the rising time detection software is shown below (Figure 4).

The auto-calculated rising time can be manually corrected if the (experienced) user find a different start time or finish time.

The program shows the determined melody contour and calculates the rising time. The start and the finish points are highlighted from the points of the melody.

If the user finds both the start time and the finish time true, by clicking the OK button, the calculated rising time will be saved and the next crying signal will appear.

In questionable situations the user can play the crying signal, and listen to the melody to testify the results.

If there is no rising part in the beginning of the crying signal or it is not obvious according to the plotted melody contour, the user can choose not to save rising time for the current sample and to get the next crying signal.

\section{Results}

From the 556 crying signals a total of 265 rising periods (\textit{i.e.} 265 crying signals started with rising period) were detected. There were 128 crying signals from infants with hearing disorders, and 137 signals from healthy babies.

The average of the rising times was 0.3282±0.1487 s, and the median was 0.2987 s. The distribution of the rising times is shown in Figure 5.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{distribution.png}
\caption{The distribution of the rising time.}
\end{figure}

\subsection{Comparison}

Comparing the two groups (\textit{i.e.} the group of infants with hearing disorders and the group of healthy babies) no significant differences were found by the rising time (from the well-known ANOVA method: d.f. = 273, \(F = 1.31, p = 0.253\)). At the first group the mean of the rising time was 0.3178±0.1206 s, and the median was 0.2887 s. At the group of healthy infants the mean value was 0.3380±0.1701 s, and the median was 0.2987 s. The results are presented on the following table (Table 1).
Table 1: The mean, the deviation and the median values of the rising time for both groups.

<table>
<thead>
<tr>
<th>Hearing</th>
<th>Gender</th>
<th>No. of signals</th>
<th>Rising time (s)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>mean</td>
<td>SD</td>
</tr>
<tr>
<td>Disorders</td>
<td>boys</td>
<td>69</td>
<td>0.2951</td>
<td>0.1056</td>
</tr>
<tr>
<td></td>
<td>girls</td>
<td>59</td>
<td>0.3442</td>
<td>0.1321</td>
</tr>
<tr>
<td></td>
<td>total</td>
<td>128</td>
<td>0.3178</td>
<td>0.1206</td>
</tr>
<tr>
<td>Healthy</td>
<td>boys</td>
<td>42</td>
<td>0.2853</td>
<td>0.1303</td>
</tr>
<tr>
<td></td>
<td>girls</td>
<td>95</td>
<td>0.3612</td>
<td>0.1807</td>
</tr>
<tr>
<td></td>
<td>total</td>
<td>137</td>
<td>0.3380</td>
<td>0.1701</td>
</tr>
<tr>
<td>Both</td>
<td>boys</td>
<td>111</td>
<td>0.2914</td>
<td>0.1151</td>
</tr>
<tr>
<td></td>
<td>girls</td>
<td>154</td>
<td>0.3547</td>
<td>0.1636</td>
</tr>
<tr>
<td></td>
<td>total</td>
<td>265</td>
<td>0.3282</td>
<td>0.1484</td>
</tr>
</tbody>
</table>

Authors found differences between the rising time of boys and girls (d.f. = 307, F = 17.02, p < 0.01). At boys the mean of the rising time was 0.2914±0.1151 s, and the median was 0.2560 s. Girls had a 0.3547±0.1636 s mean value and a 0.3251 s median value. The difference is shown on Figure 6.

![Figure 6: The distribution of the rising time between boys and girls.](image)

5 Discussion

As experienced otorhinolaryngologists could distinguish the crying of infants with hearing disorders from the crying of healthy infants a few decades before, these differences may still exist. By using a suitable parameter obtained from the crying signal, this separation could be successful.

This method would cover the whole auditory system, contrary to the existing objective audiometry methods, which test just a part of the whole system. There are further advantages of the infant cry based method, as it is a non-invasive method and sound analysis could be evaluated by low cost instruments.

6 Conclusions

The melody contour was determined to each crying signal and the rising time was calculated from the melody. Authors had found 265 melodies which started with rising period. No significant differences were found between the rising time of infants with hearing disorders and healthy babies.

Differences were found, although, between boys and girls. By using the one-way analysis of variance a very small (< 0.01) p-value was determined between the rising time of boys and girls.

Acknowledgements

The research has been supported by Hungarian National Scientific Research Foundation (OTKA T0422990).

References


[7] H. Rothganger, ’Analysis of the sounds of the child in the first year of age and a comparison to


