ABSTRACT
In this study the first communication signal of life (i.e. the infant cry) will be analyzed. Authors show methods for the acoustic analysis of the infant cry and ways how to obtain the melody of crying. Previously several research groups had reported about the development of the infant cry, now authors are to present the development of the melody. The results are based on an infant cry database containing 2640 crying samples. If we know what kind of developmental chain we can expect from a healthy infant, we could check if an unhealthy infant followed this development or not. This study mostly deals with the classification of melody shapes, and the development of the fundamental frequency and the duration of the melodies.

KEYWORDS
Infant cry, development, fundamental frequency, melody, and acoustic analysis

1. Introduction
After birth, infants start an extremely high mental, physical, physiological, etc. development. Many stimuli have betided the infants, which ones the infants try to process, to memorize and to answer. There are well known developmental chains during the infancy, e.g. social, emotional, language and moving. For example, the development of moving is the following: in the first months of life the baby doesn’t use his limbs to change place, later he starts creeping, walking on all fours, and finally the baby makes the first move on his own feet. The development itself is not only important because of the developed attribute. If we find that our baby doesn’t follow the expected, normal development we think that there must be some disorder with the infant, so we draw a conclusion from the lack of the normal development.

In this study authors want to obtain the development of the melody of the infant cry. If the developmental chain of the human sound production system is known, several disorders could be diagnosed (as hearing disorders, central nervous system disorders, etc.) by the analysis of the produced sound. In case of hearing disorders the infant cannot hear his own voice partly or totally, in this way we expect that the crying melody development of hard-of-hearings differs from the development of normal hearing infants.

Several research teams have published papers about the development of attributes of the infant cry. Michelsson et al. tested approx. 1800 crying samples from 172 infants during the first 7 days of life; they tested the development of the duration and the mean fundamental frequency (F₀) but no significant differences were found [1]. By the analysis of the hunger cry from 4 infants over a year Gilbert & Robb found increasing mean F₀ values from 400 to 500 Hz [2]. As results of unhealthy development Hirschberg & Szende summarized the sounds from over 100 infants with disorders a few decades before [3]. Lind & Wermke investigated 658 spontaneous cries during the first 3 months [4]. They divided the recorded sounds by duration into 2 groups; the group of durations over 0.8 seconds had increasing F₀ values, while the other group had stagnation during the first 3 months. Wermke et al. tested the development of formant frequencies and the melodies of crying between the age of 8 and 24 weeks [5]. They found that the typical rising-falling melody shape is multiplying as the infants get older.

2. Subjects and Methods
2.1 Data Collection
As the summary of previous studies shows, the number of infants in these investigations was between 4 and 172, and the number of the analyzed crying samples was up to 1800. For this study a database containing 2640 crying samples from 320 infants was created as a result of a 5 years long data collection in several hospitals and homes in Hungary. All the sound recordings were made in quiet places, but not in special silence rooms. During the recordings we tried to diminish the background noises as much as possible. Besides the crying samples the database
contains information about the infant (date of birth, gender, etc.), about the recording (date of recording, place of recording, recording device, length of recording, etc.) and some additional information about medical observations.

Out of the 320 infants there were 147 boys and 173 girls with a mean age at 194.73 days. As there were recordings from several maternity wards, a part of the infants were between the age of 0 and 7 days.

There were several recording devices applied during these 5 years as minidisk recorder (SONY MZ-R55), digital video camera (SONY DCR-TRV25), digital dictaphone (SONY ICD-P28) and PC sound card with several microphones (SONY ECM-MS907, AKG D55S) attached.

As the goal of this study is to obtain and to analyze the melody of crying we don’t have to bother that the applied devices are different with different transfer functions, because the basic information we need (i.e. the melody) is carried wholly in every type of recording.

The typical duration of a recording was 25–30 seconds in the hospitals and 60–80 seconds in homes. There were 8–9 crying samples in each of the recordings on the average. The reason of crying was pain in most cases (evoked crying); there were only a few infants, at whom other reasons were found (hunger crying, boredom crying, etc.).

The digitalization of the recorded crying sounds was applied at 44.1 kHz or 48 kHz (depending on the recording device), each sample was assigned to 16 bits; as there was no need for stereo recordings in this research, to spare storage capacity stereo recordings were converted into mono ones. Finally, all the recorded sounds were saved to PC as separate wave (.wav) files.

2.2 Getting the Melody from the Crying Samples

The following steps, applied methods and algorithms are executed in Matlab 7.1. It provides flexible and precise environment for signal analysis.

First, the recordings are segmented into useful and useless parts. By this segmentation the crying samples (i.e. the useful parts) are separated from the original recording. Simplifying this task, this step can be performed by analyzing the amplitude (or energy) and the spectral information of the recording [6]. There is a crying sample inside the recording where high energy and harmonic structure is detected.

After determining the crying samples from the recordings, the second step is to detect the fundamental frequency (F₀) of the crying in short (typically 50 ms) distances. The fundamental frequency (it is often called pitch) is the lowest, useful frequency component of the spectrum. In case of harmonic structure (as in human vowels, singing or crying) the harmonics of F₀ are also presented. There are several methods to obtain the fundamental frequency automatically: the Smoothed Spectrum Method [7], the Harmonic Product Spectrum Method [8], the Cepstrum Analysis [9], the Autocorrelation Function [9], etc.

Authors obtain the melody of crying by representing the detected F₀ values as a function of time. Authors developed the Five Line Method (FLM) for visualizing the melody in an advantageous way: the melody curve is visualized with five guidelines similarly to the five lines on a music paper, but the frequency values of the lines are marginally different [10]. As the typical values of the fundamental frequency of crying is between 300 and 700 Hz according to the results of several teams [1, 11] and this study as well, the Five Line Method with its fixed range of 200–1000 Hz provides a suitable interval for the visualization of the melodies of the infant cries. The following figure (Fig. 1) shows examples of the obtained melodies represented by the FLM.

These randomly chosen examples show as well how diverse the melodies of crying are: there are shorter or longer melodies (Duration), there are melodies in higher or lower frequencies (Mean value of F₀) and there are several kinds of melody shapes (Melody classification). In the following these three topics will be discussed, and the development of each of them will be shown in the first 18 months of life.

As the infants involved this investigation are different in ages, authors divided the 2640 crying samples into 12 isometric groups (named G01–G12). The groups are in ascending order by age: the first group contains the youngest 220 infants; the second group contains the next 220 young infants, and so on.

By this division, however, authors haven’t created fully homogenous groups by age, though they are able to check the development of the melody of the infant cry from the age of 0 to 18 months. The contents of the 12 groups are shown on Table 1.
Table 1. Contents of the 12 groups.

<table>
<thead>
<tr>
<th>group</th>
<th>number of infants</th>
<th>minimum age (days)</th>
<th>maximum age (days)</th>
<th>median age (days)</th>
</tr>
</thead>
<tbody>
<tr>
<td>G01</td>
<td>220</td>
<td>0</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>G02</td>
<td>220</td>
<td>3</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>G03</td>
<td>220</td>
<td>5</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>G04</td>
<td>220</td>
<td>8</td>
<td>15</td>
<td>10</td>
</tr>
<tr>
<td>G05</td>
<td>220</td>
<td>15</td>
<td>27</td>
<td>20</td>
</tr>
<tr>
<td>G06</td>
<td>220</td>
<td>27</td>
<td>45</td>
<td>38</td>
</tr>
<tr>
<td>G07</td>
<td>220</td>
<td>46</td>
<td>64</td>
<td>57</td>
</tr>
<tr>
<td>G08</td>
<td>220</td>
<td>64</td>
<td>101</td>
<td>81</td>
</tr>
<tr>
<td>G09</td>
<td>220</td>
<td>101</td>
<td>154</td>
<td>117</td>
</tr>
<tr>
<td>G10</td>
<td>220</td>
<td>154</td>
<td>222</td>
<td>196</td>
</tr>
<tr>
<td>G11</td>
<td>220</td>
<td>222</td>
<td>341</td>
<td>250</td>
</tr>
<tr>
<td>G12</td>
<td>220</td>
<td>341</td>
<td>499</td>
<td>413</td>
</tr>
</tbody>
</table>

3. Analysis and Results

3.1 Duration

As it was shown above, there are shorter and longer melodies of crying as well. Fig. 2 represents the common distribution of the durations of the 2640 melodies.

![Fig. 2. Common distribution of duration](image)

The common distribution of the durations is bell-shaped with a maximum value at 0.6 s. The most typical value of the duration of a crying sample is around 0.4–0.8 s; the 95% of the durations are between 0 and 2 s. By checking not the common, but the separated distribution of each groups authors have found the following results (see Table 2).

Authors have found significant differences by statistical comparison (one-way analysis of variance – ANOVA) between the distribution of durations at G01 and G12 (d.f. = 439, F = 17.01, p = 0.00004).

While crying samples having a shorter duration (between 0 and 1 s) were in vast majority in the first 2 months of life (G01–G07), later this difference between the shorter and the longer durations decreased (G08–G12). From another side, the mean value of the durations has been increased from around 0.8 to 1.0 s, after the first 2 months of life.

This kind of development of the duration can be explained by other developmental effects. During the first 2 months of life the infants cry more spontaneously; later they start communicating with their environment more consciously. While spontaneous crying has a shorter, conscious (or voluntary) crying has a longer duration. We can declare:

- After the first 2 months of life infants start treating the crying as a tool, the crying is not just spontaneous any more.
- The older the infant is, the longer the crying samples appear.

3.2 Mean value of F0

The melody is the changing of the fundamental frequency as a function of time. On Fig. 1 it was shown that there were many kinds of melodies; by the five guidelines of FLM we could see that some of the melodies were on higher frequencies, while others were on lower frequencies.

An obvious way to feature this difference with a scalar is to calculate the mean value of F0 for each of the 2640 crying samples. The following figure (Fig. 3) represents the common distribution of the 2640 melodies.

The common distribution of the mean value of F0 is bell-shaped; the maximum of the distribution is at 400 Hz. The 96% of the mean values are between 300 and 600 Hz.
By testing the distribution of these mean values from group by group, the following results had been found: the center of the mean values of the distributions at the 12 groups was varying between 404 and 464 Hz without any significant tendency.

There were a few reports about the development of the mean value of the fundamental frequency at crying, and some of them had resulted decreasing, while others had resulted increasing; e.g. Rothganger reported about a positive regression from 450 to 500 Hz over a year [12]. Those reports were mostly case studies, testing individuals. We can declare that:

- The development of the fundamental frequency of infant cry melodies can be decreasing, stagnating or increasing individually.
- By analyzing 2640 melodies we find that there is no significant decreasing or increasing, the mean value of the fundamental frequency of crying is varying only between 404 and 464 Hz in the first 18 months of life.

3.3 Melody classification

As we are able to compare the durations or the mean values of F0 between the melodies of crying we can see that these two parameters are not enough to describe the melody. The melody shapes should be categorized to be able to compare them and to test the development of the melodies. In general we can say that there are several kinds of melody shapes among the 2640 melodies.

Schönweiler et al. classified the melody shapes of crying [13]. Their categories were: rising, falling, rising-falling, falling-rising, flat and glottal plosive. As this last category is rather a sound phenomenon, than a melody type authors use the first 5 types for the current study. In the following these categories will be referred as ‘classical categories’. On Fig. 1 we can recognize a few of these categories; for example ‘CR0186-28’ has a falling shape, and ‘CR0127-07’ has a rising-falling shape. But what is the situation with e.g. ‘CR0134-20’ or ‘CR0004-15’? It seems the 5 classical categories are not enough to cover all of the melodies.

The Idea of Determining New Categories: Authors defined three fundamental units: falling (-1), flat (0) and rising (I). All the melodies can be treated as a sequence of some of these fundamental units. The idea of this method is to create categories according to the obtained sequences. The length of the fundamental units individually wasn’t adverted, only the order of the units was important.

For the automatic classification the melodies were partitioned into various number of parts (depending on the number of the local extremums), then all of these partitions were classified as a falling unit, a flat unit or a rising unit. By this way from the 2640 melodies authors determined 2640 sequences containing -1, 0 and 1. After assorting these sequence codes authors found the following results.

Results of the New Classification Method: A total of 77 different sequence codes were found from the easier ones (having only one unit) to the more complex ones (having 6–8 units). Out of the 77 categories there were 20 which include the 95% of the 2640 melodies. The distribution of these “top 20” categories and the schemes of their shapes are shown on Table 3.

The most typical category of this classification system was the [1 -1] category with 933 cases. It can be clearly seen that the all simplest shapes (including 1 or 2 units) are present in the top 20; the most complex shapes have 4 units.

Table 3. Illustrating the 20 most important melody shapes.
Typically the starting partition of the melodies of crying was a rising unit or a flat unit, while the typical ending partition is a falling unit, or a flat unit. It might have happened because before and after sound production the tightness of the vocal cords was in static conditions.

Comparing these results with the classical melody shapes, authors find that there are more than 5 important categories of the melodies. These important melody shapes can be put together from 1 or 2 units. However, the 5 classical categories can be found in the top 20 of the new classification system, but not in the top 5, because the falling-rising category (i.e. \([-1 \ 1]\)) is only the 13th on this new list.

Before testing the development of these categories we can declare the following remarks about this new classification system:

- The five classical categories are not suitable to make a precise classification of the melodies of crying.
- The new classification system partitions the melody into fundamental units, and the information about the shape of the melody is carried in the order of these units; this method enables to make a precise categorizing of all the melody shapes.
- The simpler the melody shape is (i.e. containing less units), the bigger number of cases is present in the distribution of the melody shapes.
- The typical starting unit of a sequence is rising (or flat), while the typical ending unit is falling (or flat).

**The Development of the Obtained Categories:** To determine the development of the obtained categories authors used the same technique as it was shown before in the case of the duration and the mean value of \(F_0\).

There were five main categories which ones presented a more active varying than the others; these categories were \([1 \ -1], [1], [-1], [0] \) and \([0 \ -1]\). The development of these five categories is shown on Table 4 and on Fig. 4.

By the last column of the table it is turned out that the development of the dominancy of the top 5 categories starts with a short increasing section from 80% to 88% at G1–G5 and finishes with a decreasing section from 88% to 60% between G05 and G12. There is a considerable supereminence at \([1]\) between G03 and G07, and relatively smaller ones at \([1\ -1]\) between G04–G05 or G07–G09.

There is an increasing tendency in the number of the flat (0) units with age which means: the older the infants are the more sustained parts are in the melody of crying.

From the results authors emphasize the followings:

- In the first months after birth, infants cry mostly with simple melody shapes, but the dominancy of these simple categories is continuously decreasing during the following months.
- Out of the 77 categories there are 57 which ones represent the 5% of the 2640 melodies; these “rare shapes” appear mostly at the higher numbered groups (G06–G12).

### 4. Conclusion

In this study the development of the melody of the infant cry was investigated. After the recognition of developmental chains of the healthy infants some disorders can be detected by the lack of these developments. As the crying recordings were digital sound files and the methods applied in this study were functions in Matlab environment the obtained results were reproducible and precise.
In the field of the classification of the melodies authors elaborated a new objective method based on partitioning the melody into fundamental units; then the melody shapes could be described by the order of these units. Authors proved that the 5 classical categories were not enough to categorize the melodies. The new classification system, however, resulted much too much categories, nevertheless it offered a precise resolution; the top 20 categories covered 95% of the melodies.

By using statistical methods to analyze and to classify the obtained melodies authors were able to find differences, tendencies between the 12 groups. The obtained results about the development of the melody of the infant cry were the followings:

- In the first few months of life infants cry with mostly shorter, simpler melodies, while later the durations increase and the melody shapes get more complex. Emotional reason: as the infants are getting older, they try to express their needs better and better. Physiological reason: as the infants are getting older they have an increased capacity of their lungs, so they are able to produce longer melodies of crying.

- During the first 18 months of life, although, there can be increasing/decreasing tendencies in the changing of F0 individually, no significant tendency were found statistically. The mean value of the fundamental frequency of crying has only a small varying around 404–464 Hz.

**Future Work:** This study has left some open questions which should be answered in a future work. In a later study not only the order but the duration of the fundamental units should be also treated. Another task is to focus on the “rare melody shapes”.

Out of the three tested attributes (Duration, Fundamental Frequency and Melody Classification) the first and the third proved to be important in the investigation of development of the melody.

The development of more parameters (as formant structure) of crying will be also recognized. Authors want to compare the development of crying melodies from healthy and unhealthy infants; a data collection has just started to get crying samples (thus melodies) from infants with hearing disorders.

**Acknowledgements**

This research has been supported by Hungarian National Scientific Research Foundation (OTKA T0422990, T029830) and Medical Research Council (ETT 089/2003). Authors thank for all the hospitals involved in the data collection: Heim Pál Hospital for Sick Children (Budapest), Szent István Hospital (Budapest), Schöpf-Merei Hospital (Budapest), and Borsod County Hospital (Miskolc). Special thanks for their help to Zsolt Farkas and Gábor Katona chief doctors from the Heim Pál Hospital, and Zsolt Szabó chief doctor from the Borsod County Hospital.

**References**