PERIODICA POLYTECHNICA SER. EL. ENG. VOL. 50, NO. 1–2, PP. 47–62 (2006)

FUTURE PROSPECTS OF THE APPLICATION OF THE INFANT CRY IN THE MEDICINE

György VÁRALLYAY JR.

Department of Control Engineering and Information Technology
Budapest University of Technology and Economics
H–1117 Budapest, Magyar Tudósok krt. 2, Hungary
Phone +36 1 463-4027
e-mail: varallyay@iit.bme.hu

Received: April 28, 2005

Abstract

Babies cry for the same reason adults talk: to let others know about their needs or problems. The infant cry contains, besides, more information about the baby, particularly, information about the health of the infant (*e.g.* airway affections cause different sounds from the original). Thus, we can conclude on diseases from the modified cry signal.

In this study 35 infants with hearing disorders and 35 healthy babies are tested. The author compares several attributes of the cry in the time domain between the two groups.

Keywords: infant cry, non-invasive diagnosis, acoustic signal analysis.

1. The Infant Cry

Cry is a multimodal, dynamic behaviour. This is the first tool of communication and the sign of life at birth. It involves characteristic vocalizations, facial expressions and limb movements, all of which change over time.

Infants cry in another way if they are hungry, in pain, in discomfort (*e.g.* need to be changed) or if they are happy. Cry is individual; mothers (and experienced nurses) recognize the sounds of their own infants. Babies cry for the same reason adults talk: to let others know about their needs or problems. The infant cry is in the most sensitive range of the human auditory sensation area [2, 15].

If any disorder occurs with the infant the cry may differ. This proposes the idea of a diagnostic system based on the infant cry. In this study, infants with hearing disorders are compared with healthy babies. The author had dealt with this question before and investigated other attributes of the infant cry [3, 24, 25]. The main goal of this project is to bring on a knowledge base about the infant cry and connections with several disorders. The analysis of the infant cry is evaluated by regenerating former methods, and developing new ones.

1.1. Historical Overview

1.1.1. Spectrography

In the 1960s and '70s, the sound spectrogram was the major tool for analysing cry sounds. Produced by an analog device, a spectrogram plots time on the x axis and frequency on the y axis, and encodes the darkness of the frequency lines (*Fig. 1).*

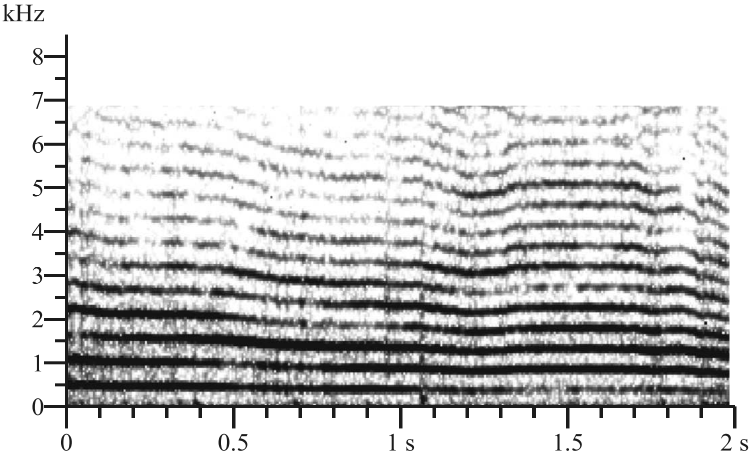

Fig. 1. A (digitally-produced) spectrogram of an infant cry.

In the 1980s HIRSCHBERG and SZENDE wrote a complete study round *Pathological cries, stridors and coughs in infants*; which was mostly based on spectrograms [7]. In their work they showed numerous examples to several diseases (e.g. Down's syndrome, Cri du chat syndrome) and summarized the history of the investigation of the infant cry. They had recorded 109 samples and attached them in a gramophone record to the book.

MICHELSSON et al. also defined healthy and unhealthy cry types by spectrography [12]. In 1999, 14 years after the above mentioned book, they introduced new spectrograms of infants with hypothyroidism, asphyxia or meningitis.

There are several attributes of the cry which can be obtained from a spectrogram, as length of the cry, spectral components, the shape of the melody contour, etc. Nowadays spectrography is still a general tool in the analysis of the infant cry, although we use, of course, digitally-produced spectrograms.

1.1.2. The Importance of Time in the Analysis

In 1999, MÖLLER and SCHÖNWEILER have reported of subjective classification of the infant cry and the reliability of this method [14]. They found that nurses, who work with several infants every day, had the most consentaneous answers in the subjective classification; questions were: if the cry sounds pleasant/unpleasant, aggressive/reconciled, high/low, monotonous/melodic, screaming/harmonic, etc. A few decades before the subjective (*i.e.* listening the sound) diagnostic at oto-rhino-laryngologists was a more typical way to determine diseases than today. The before-mentioned book of HIRSCHBERG and SZENDE, was based on subjective experiences.

SCHÖNWEILER *et al.* investigated the cry of hearing impaired infants [19], [20]. They found differences between the duration of the cry signal between 3 healthy infants and 4 babies with hearing diseases.

MICHELSSON *et al.* tested the cry of 50 newborn infants after heel-prick [17]. In their pain cries authors investigated the duration of the first five cries. They found decrement in the duration from the first cry to the fifth. In 2002 MICHELSSON *et al.* reported on the cry of 172 newborn infants [13]. They did not find significant difference in the duration of the cries by gender.

LIND *et al.* investigated an infant from birth to the age of 3 months in 2002 [11]. They used special hardware and software to record the sound and to analyse it. Authors defined a time limit (0.8 s), and divided the cry signals into two categories according to this limit. They compared these two groups; no significant differences were found

In 2003, ROTHGÄNGER reported on the change of the duration and the fundamental frequency in crying and babbling over a year [16]. He found increment in the fundamental frequency of crying and decrement of the babbling. The change of the fundamental frequency over a year can also be decrement or flat [2, 4, 27]. In the change of the duration ROTHGÄNGER found increment both in crying and babbling.

Studies, working in the time domain, mostly reported about the duration of the infant cry. But there are several other time parameters which also should be analysed. In this study *segment density*, *specific segment length*, *average pause length* and *average segment length* are analysed and compared between healthy babies and infants with hearing disorders (these parameters will be defined in 3.2.2).

1.2. Relation between Deafness and the Infant Cry

In Hungary, HIRSCHBERG spent decades with finding the differences between the crying of an infant with normal hearing and hard of hearing [8, 9, 10]. *Fig. 2* shows a block diagram as a theoretical background of this investigation.

In this simplified model, the human brain works with dual feedback. The brain instructs the organ of speech to create sound. The organ of speech executes the request and gives a mechanical feedback to the brain. At normal hearing the organ of hearing gives the second feedback to the brain. If this feedback is missing, the created sound may differ from the original [18, 21, 23].

Hence other diseases or disorders may affect the produced cry signal. By ob-

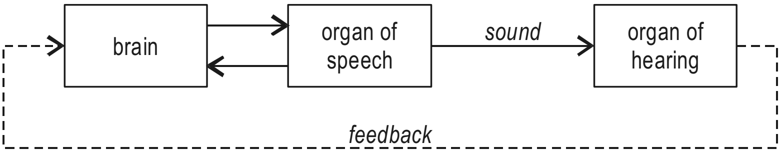

Fig. 2. The role of the organ of hearing at the phonation. If the feedback is missing, the created sound may differ from the original.

taining and applying these differences a simple and non-invasive diagnostic system could be created.

1.3. Cry Production

Several models of cry production have been theorized. The theory that underlies most acoustic analyses of cry sounds is the sound-filter theory [1]. This suggests that the waveform that impinges upon the listener's ear is a function of the characteristics of the source (*i.e.* the vibrating vocal cords) and its filters (*i.e.* the resonances of the supraglottal vocal tract and the radiation characteristics from the lips). Fig. 3 shows a universal schema of the voice-production by GORDOS et al. [6].

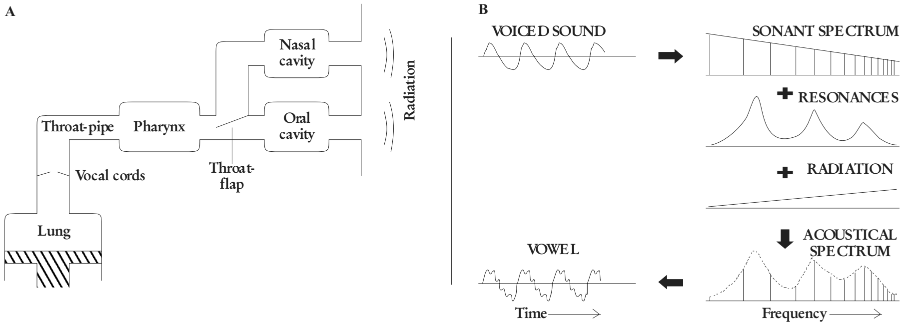

Fig. 3. A. Universal schema for the voice-production by Gordos et al. B. Formation of a vowel in the human phonation system.

In the first graphic (A) a schematized model of the human phonation system is shown. The other graphic (B) illustrates the way of an acoustic signal from the vocal cords until the radiation.

GOLUB's physioacoustic model of crying assumes three levels of central processing of the muscles contributing to the source and filters of crying [5]. These three levels are identified as the upper, middle and lower processors. The upper processor is implicated in determining the state of the infant (*e.g.* fussiness). The middle processor is involved with the infant's vegetative states, such as swallowing, coughing, digestion and crying. The lower processor involves control of many muscle groups, including the subglottal, supraglottal, glottal and facial muscles. These muscle groups are co-ordinated in the act of crying.

According to the above-mentioned conceptions we can declare that crying is an important source of information about the infant. This information can be obtained by using suitable analyses.

2. Data Acquisition

The elaborated method has a 2-stage procedure. The first stage is the recording of the sound signals in several hospitals. At the same time further information about the infants (*e.g.* age, gender) are also recorded. The second stage (signal analysis) is executed in the Laboratory of Medical Informatics; it has three main modules: preprocessing, processing and evaluation (*Fig. 4*). These steps are performed in Matlab environment, with own developed software.

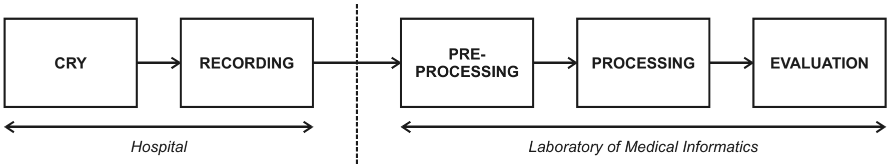

Fig. 4. The process of the investigation.

2.1. Sound Recording

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

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

The cries are collected during manual audiometry; the doctor looks into the ears of the infant. The procedure is painful and the baby starts to cry. The cries selected for analysis are, whenever possible, chosen from the start of the cry sequences. The cries for each baby are, however, usually very similar both auditively and on the spectrograms.

2.2. Database

The infant cries are transferred into a PC and converted to wave (.wav) files. All the further information about the infants, as name, age, etc. is stored in an MS Excel database [26]. This .xls file contains:

- *Information of identity:* name, date of birth, gender, address or telephone number of the parents.
- *Predetermined auditory diagnosis:* degree of deafness, type of deafness, other diseases existing.
- *Details about the sound recording:* date of recording, place of recording, length of the cry signal, sampling frequency, type of the recording device, filename of the cry signal.
- *Circumstances of the recording:* background noise, echo, overdriving, etc.

3. Analysis

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 cry 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 [22].

The second role of preprocessing is to distinguish *important* and *less important* contents in the cry signal from each other. Only the *important* parts (named crying segments) should be analysed. On *Fig. 5 less important* parts are hiccough (a and f), hoarseness (c), spasmodic crying (d) and silence (e) between the crying segments (b and g).

There are averagely 8-10 segments in a 30 second long recording (*Fig. 6/A*). Automatic segmentation is needed to detect the start and the end of each segment. For this task there are two short-time methods described below: the short-time

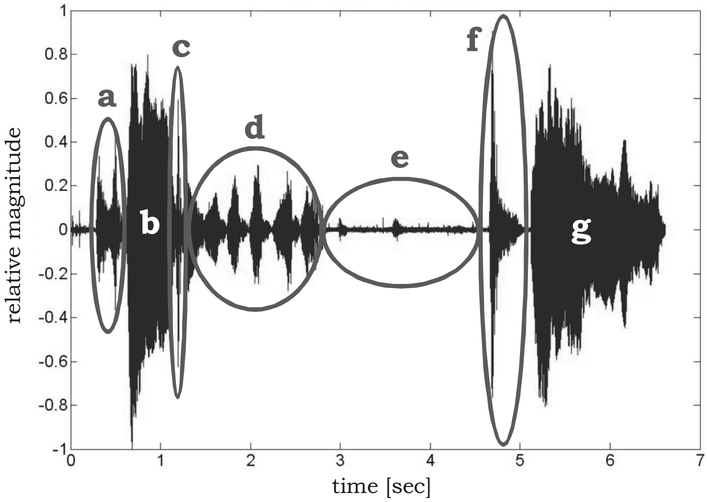

Fig. 5. Sample .wav file before feature extraction.

energy function and the short-time zero crossing rate. These methods are well-known *e.g.* in the indexing of accompanying audio signals in movies and video programs.

3.1.1. Short-time Energy Function

The short-time energy function of an audio signal is defined as:

$$E_n = \frac{1}{N} \sum_{m} \left[x\,(m) \cdot w\,(n-m) \right]^2 \tag{1}$$

where $x(m)$ is the discrete time audio signal, n is time index of the short-time energy, and $w(m)$ is a rectangle window, *i.e.*

$$w\,(n) = \begin{cases} 1, & 0 \le n \le N-1, \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

It provides a convenient representation of the amplitude variation over the time. It is fact that values of E_n for the unvoiced (*i.e.* coughing, silence, etc.) components are in general significantly smaller than those of the voiced (*i.e.* the real crying) components (*Fig. 6/B*). It can be used as the measurement to distinguish audible sounds from silence when the signal-to-noise ratio is high [28].

3.1.2. Short-time Average Zero-crossing Rate

In the context of discrete-time signals, a zero-crossing is said to occur if successive samples have different signs. The rate at which zero-crossings occur is a simple measure of the frequency content of a signal. The short-time average zero-crossing rate (ZCR_n) is defined as:

$$ZCR_n = \frac{1}{2}\sum_{m} \left|\operatorname{sgn}\left[x\left(m\right)\right] - \operatorname{sgn}\left[x\left(m-1\right)\right]\right| \cdot w\left(n-m\right) \tag{3}$$

where

$$\operatorname{sgn}\left[x\left(n\right)\right] = \begin{cases} 1, & x\left(n\right) \geq 0, \\ -1, & x\left(n\right) < 0, \end{cases} \tag{4}$$

and $w(n)$ is a rectangle window of length N. The average zero-crossing rate can be used as another measure for making distinction between voiced and unvoiced parts of the whole cry signal, because unvoiced components normally have much higher ZCR_n values than voiced ones (*Fig. 6/C*).

3.1.3. Boundaries of the Cry Segments

From E_n or ZCR_n the boundaries of the cry segments can be determined by using a satisfactory threshold. For example, on *Fig. 6/B* the value of an optimal threshold is approximately 0.01, but this number may be different in case of other cry signals. The value of the threshold always depends on the cry signal and its energy or frequency content.

Another method for detecting vocal signals was applied by ZHANG *et al.* [28]. Their technique is based on observing the spectrum of the cry signal in certain times. If harmonic spectral components are found, the investigated signal is part of a cry segment.

From the 70 infants a total of 556 cry segments can be detected, their distribution between the two groups is the following: 271 segments in the group of infants with hearing disorders and 285 segments in the healthy group. Thus, there are 556/70 $\approx$ 8 segments in a whole cry signal on the average.

3.2. Processing

After preprocessing, the cry signal is filtered and useful information (*i.e.* the cry segments) is extracted. This signal is ready for the signal processing and analysis. The analysis can be evaluated in the time domain as well as in the frequency domain. This study deals with several parameters of the infant cry only in the time domain.

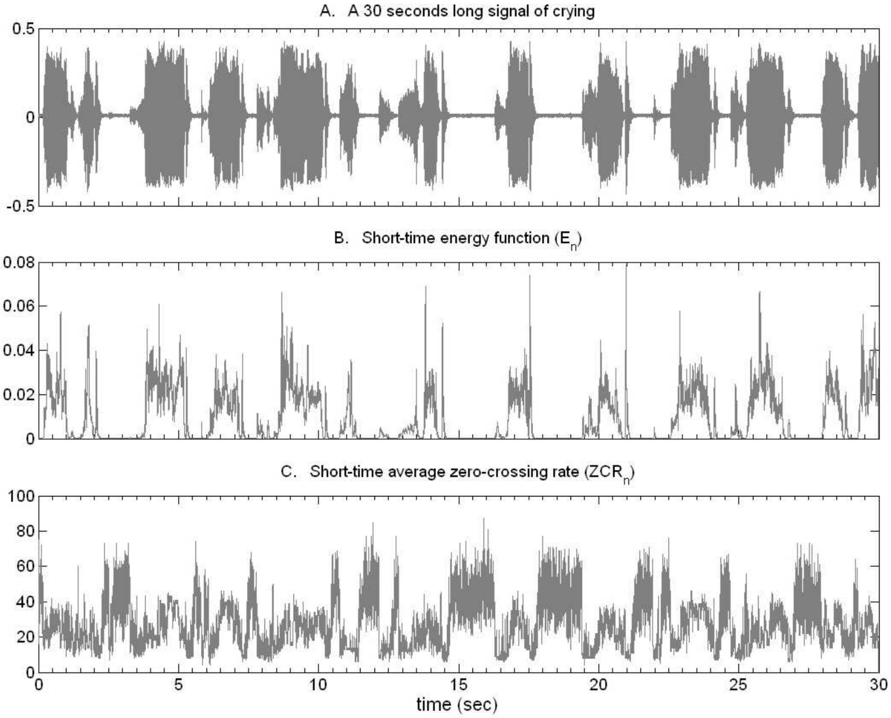

Fig. 6. Example for the automatic segmentation: crying signal (A), short-time energy function (B), short-time average zero-crossing rate (C).

3.2.1. Segment Length

After automatic segment detection, the length of a segment can be determined as the difference of the detected end and start points. The distribution of the duration of the detected 556 segments is shown on the following histogram (*Fig. 7*).

The mean value of the segment lengths is 1.1003 sec, the standard deviation is 0.6066 sec, the median is 0.9750 sec.

3.2.2. Determining the Four Attributes of the Cry

Attributes are determined from the following three parameters:

- *Total time:* the length of the whole cry signal in seconds (T_{tot}).
- *Segment number:* the number of the segments inside a whole cry signal (N).
- *Segment time:* the length of each cry segments ($t_{i,seg}$), where $i = 1, 2, \ldots, N$.

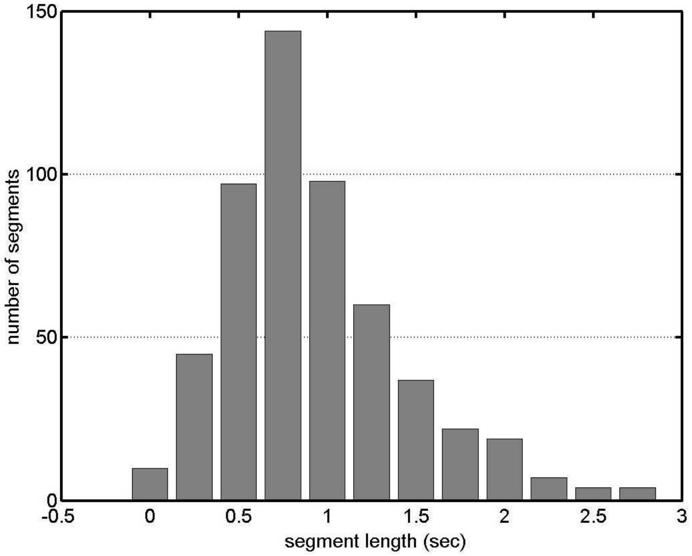

Fig. 7. Histogram of the segment lengths for the investigated 70 infants.

These three parameters are illustrated on *Fig. 8*.

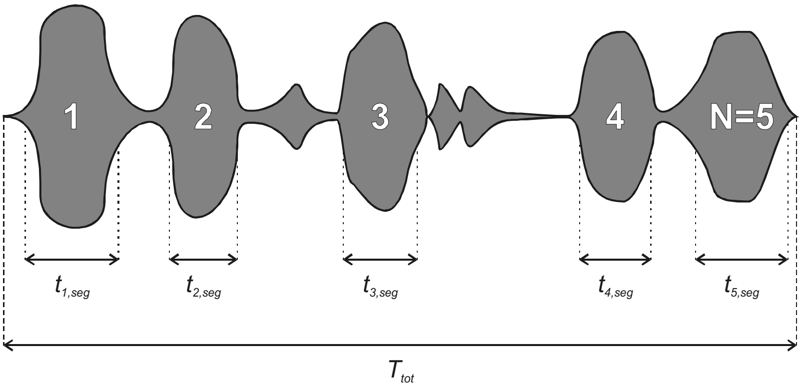

Fig. 8. Illustrating total time (T_{tot}), segment number (N) and segment time ($t_{i,seg}$).

The total segment time (T_{seg}) can be calculated by summing the segment times.

$$T_{seg} = \sum_{i=1}^{N} t_{i,seg} \tag{5}$$

The total segment time is always less than the total time, because there are not just segments inside a whole cry signal, but pauses. The total length of these pauses can be obtained from the total time and total segment time:

$$T_{pause} = T_{tot} - T_{seg} \tag{6}$$

As the foregoing parameters are not relative values they should not be compared ($e.g.$ N or T_{tot} does not depend on the infant itself, but the length of the recording). From these parameters (*i.e.* absolute values) the following relative values (in our case the expected attributes) can be calculated:

- *Segment density* (1/s). It indicates how many segments are included in a one second cry signal on the average.

$$n = \frac{N}{T_{tot}} \tag{7}$$

(If there are 8-10 segments in a 30 seconds long cry signal, the value of the segment density is approx. 0.25-0.3 1/s.)

- *Specific segment length* (%). It shows the relation between the total time and the total segment time.

$$\tau_{seg} = \frac{T_{seg}}{T_{tot}} \cdot 100 \tag{8}$$

(If there are 8-10 segments in a 30 seconds long cry signal, and the segment length is typically around 1 sec, the value of the specific segment length is approx. 30-35 %)

- *Average pause length* (s). It characterizes the time between two neighbouring segments on the average.

$$\langle t_{pause} \rangle = \frac{T_{pause}}{N} \tag{9}$$

(If there are 8-10 segments in a 30 seconds long cry signal, and the segment length is typically around 1 sec, the value of the average pause length is approx. 2.25-2.5 s)

- *Average segment length* (s). It shows the length of a segment on the average.

$$\langle t_{seg} \rangle = \frac{1}{N} \sum_{i=1}^{N} t_{i,seg} = \frac{T_{seg}}{N} \tag{10}$$

(The value of the average segment length is approx. 0.5-1.25 s)

4. Results

A total of 556 cry segments were analysed from 70 babies. From a majority of the babies 8 segments were analysed. 4 cry attributions (n, τ_{seg}, $< t_{pause} >$, $< t_{seg} >$) were calculated and compared between the two groups of infants: Group D includes 35 infants with hearing loss or deafness; in Group H there are 35 healthy infants. In addition, the relation between these parameters and gender was also tested.

4.1. Segment Density

The mean segment density was 0.3095 ± 0.1339 1/s and the median 0.2875 1/s (*Table 1*). The mean segment density was somewhat higher for Group H, 0.3199 ± 0.1361 1/s, than for Group D, 0.2992 ± 0.1328 1/s; the difference was not statistically significant (d.f. = 69, F = 0.41, p = 0.5232). Neither were any significant differences noted according to gender (d.f. = 79, F = 0.07, p = 0.7989).

Table 1. Segment density (mean and median) according to hearing and gender.

Hearing	Gender	Infants	Segment density (n), 1/s			
			mean	SD	median	range
Group D	boys	24	0.2939	0.1351	0.2627	0.0873-0.6053
	girls	11	0.3108	0.1331	0.2581	0.1779-0.5999
	total	35	0.2992	0.1328	0.2602	0.0873-0.6053
Group H	boys	16	0.3365	0.1481	0.3029	0.1422-0.7048
	girls	19	0.3058	0.1274	0.2943	0.1167-0.5687
	total	35	0.3199	0.1361	0.2968	0.1167-0.7048
Total	boys	40	0.3110	0.1402	0.2824	0.0873-0.7048
	girls	30	0.3076	0.1273	0.2922	0.1167-0.5999
	total	70	0.3095	0.1339	0.2875	0.0873-0.7048

4.2. Specific Segment Length

The mean of the specific segment length was 33.66 ± 16.10 % and the median 32.05% (*Table 2*). There were no significant differences noted according to hearing (d.f. = 69, F = 8.42e-5, p = 0.9927), and what is more the high value of p indicates that there is no difference between the two groups in respect of the specific segment length. Neither were any significant differences noted according to gender (d.f. = 79, F = 0.35, p = 0.5543).

Table 2. Specific segment length (mean and median) according to hearing and gender.

Hearing	Gender	Infants	Specific segment length (τ_{seg}), %			
			mean	SD	median	range
Group D	boys	24	30.17	14.30	29.55	3.90-54.96
	girls	11	41.32	15.32	41.99	17.79-66.23
	total	35	33.67	15.32	31.57	3.90-66.23
Group H	boys	16	36.05	17.76	36.42	7.43-61.63
	girls	19	31.61	16.67	31.87	12.20-62.19
	total	35	33.64	17.07	32.23	7.43-62.19
Total	boys	40	32.52	15.82	30.37	3.90-61.63
	girls	30	35.17	16.61	32.59	12.20-66.23
	total	70	33.66	16.10	32.05	3.90-66.23

4.3. Average Pause Length

The mean of the average pause length was 2.7641 ± 1.8930 s and the median 2.4182 s (*Table 3*). There were no significant differences according to hearing (d.f. = 69, F = 0.43, p = 0.5124) or gender (d.f. = 79, F = 0.63, p = 0.4292).

Table 3. Average pause length (mean and median) according to hearing and gender.

Hearing	Gender	Infants	Average pause length (t_{pause}), s			
			mean	SD	median	range
Group D	boys	24	3.2230	2.4463	2.7824	0.8146-11.0091
	girls	11	2.2390	1.1358	1.9519	0.8010-4.6203
	total	35	2.9137	2.1547	2.6076	0.8010-11.0091
Group H	boys	16	2.3566	1.4444	2.1969	0.6797-6.2703
	girls	19	2.8316	1.7415	2.4970	0.6717-7.4489
	total	35	2.6145	1.6074	2.3999	0.6717-7.4489
Total	boys	40	2.8764	2.1252	2.4182	0.6797-11.0091
	girls	30	2.6143	1.5529	2.3993	0.6717-7.4489
	total	70	2.7641	1.8930	2.4182	0.6717-11.0091

4.4. Average Segment Length

The mean of the average segment length was 1.1338 ± 0.5032 s and the median 1.6083 s (*Table 4*). There were no significant differences according to hearing (d.f. = 69, F = 0.59, p = 0.4441) or gender (d.f. = 79, F = 1.27, p = 0.2639).

Table 4. Average segment length (mean and median) according to hearing and gender.

Hearing	Gender	Infants	Average segment length (t_{seg}), s mean	SD	median	range
Group D	boys	24	1.0475	0.3773	0.99774	0.4465-1.9634
	girls	11	1.4700	0.7283	1.3328	0.7240-2.8805
	total	35	1.1802	0.5403	1.0534	0.4465-2.8805
Group H	boys	16	1.1134	0.5477	1.0228	0.3233-2.4657
	girls	19	1.0655	0.3998	1.0967	0.2698-1.9283
	total	35	1.0874	0.4665	1.0866	0.2698-2.4657
Total	boys	40	1.0738	0.4477	0.9977	0.3233-2.4657
	girls	30	1.2138	0.5670	1.1266	0.2698-2.8805
	total	70	1.1338	0.5032	1.0683	0.2698-2.8805

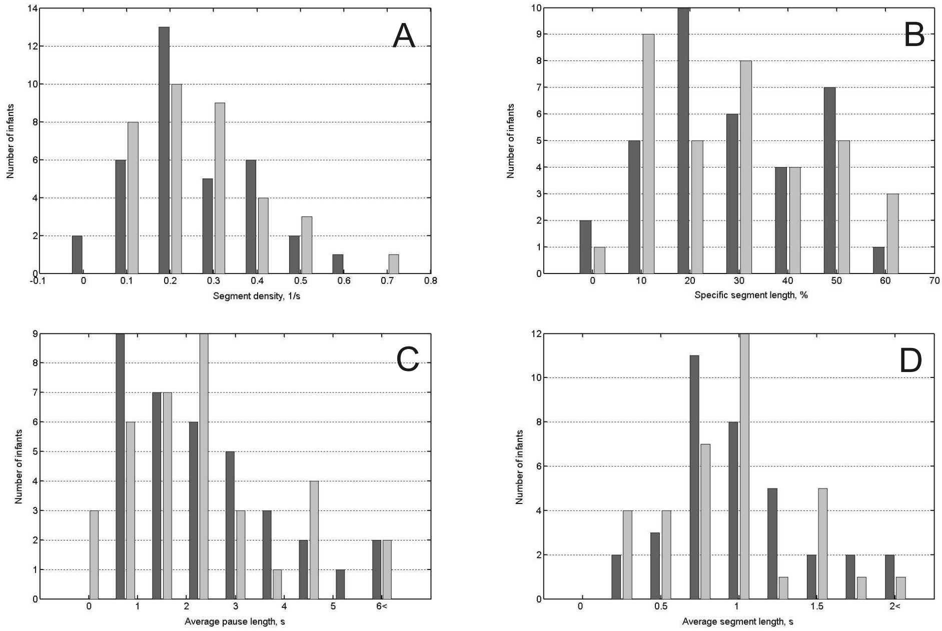

Fig. 9. Histograms for the above mentioned attributes of 70 infants.
Infants with hearing disorders: dark grey. Healthy infants: light grey.
A: segment density (1/s); B: specific segment length (%);
C: average pause length (s); D: average segment length (s).

All the results from the investigated 4 attributes are shown on *Fig. 9*. The group of infants who have hearing diseases (Group D) is shown with dark grey, the healthy group (Group H) is shown with light grey.

5. Conclusions and Future Work

There were no significant differences noted according to hearing or gender from the investigated 70 infants.

One of the investigated attributes, namely the specific segment length, had a very high similarity in the two groups. The further three parameters should be analysed in respect of more circumstances (*e.g.* how these parameters relate to the age of the infant).

If the attributes of the cry change over the time (as it was proved with the mean of the fundamental frequency [16]), the speed of these changes could be different in the two groups. To testify this idea the number of the investigated infants should be increased. In another future work respiratory data (*e.g.* frequency of breathing) will be involved.

Acknowledgements

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

Special thanks to Zoltán Benyó (BUTE, Dept. of Control Engineering and Information Technology, Laboratory of Medical Informatics), András Illényi (BUTE, Dept. of Telecommunications and Media Informatics, György Békésy Acoustic Laboratory) and Zsolt Farkas, MD (Heim Pál Hospital for Sick Children, Budapest).

References

[1] BARR, R. G.–HOPKINS, B.–GREEN, J. A., *Crying as a sign, a symptom and a signal,* MacKeith Press, London, 2000.

[2] BENYÓ, Z.–FARKAS, Z.–ILLÉNYI, A.–KATONA, G.–VÁRALLYAY JR, G., Csecsemők hangelemzése különös tekintettel a hallásképességükre. In: Beszédkutatás 2002., M. Gósy ed., Budapest, Oct 2002, MTA Nyelvtudományi Intézet, Budapest, pp. 100–117.

[3] BENYÓ, Z.–FARKAS, Z.–ILLÉNYI, A.–KATONA, G.–VÁRALLYAY JR, G., Information Transfer of Sound signals. A case study: The infant cry. Is it a noise or an information?, *Proc. 33rd International Congress and Exposition on Noise Control Engineering, Prague, Aug 2004*, pp. 528/1–8.

[4] GILBERT, H. R.–ROBB, M. P., Vocal Fundamental Frequency Characteristics of Infant Hunger Cries: Birth to 12 Months, *Int. J. Pediatr. Otorhinolaryngol.*, **34** (1996), pp. 237–243.

[5] GOLUB, H. L., A Physioacoustic Model of the Infant Cry and its Use for Medical Diagnosis and Prognosis, Ph.D. Thesis, Massachusetts Institute of Technology, MA, USA, 1980.

[6] GORDOS, G.–TAKÁCS, G., *Digitális beszédfeldolgozás*, Műszaki Kiadó, Budapest, 1983.

[7] HIRSCHBERG, J.–SZENDE, T., *Pathological cry, stridor and cough in infants*, Akadémiai Kiadó, Budapest, 1982.

[8] HIRSCHBERG J.–DEJONCKERE, P. H.–HIRANO, M.–MORI, K.–SCHULTZ-COULON, H.-J.–VRTICKA, K., Voice disorders in children, *Int. J. Pediatr. Otorhinolaryngol.*, **32** (1995), pp. S109–S125.

[9] HIRSCHBERG, J., Dysphonia in infants, *Int. J. Pediatr. Otorhinolaryngol.*, **49/1** (1999), pp. S293–S296.

[10] HIRSCHBERG, J., *A foniátria és a Magyar Fonetikai, Foniátriai és Logopédiai Társaság története. A kommunikáció, a hangképzés és a beszéd zavarainak kezelése,* Rétiron Kft nyomdaüzeme, Budapest 2003. pp. 163–207.

[11] LIND, K.–WERMKE, K., Development of the Vocal Fundamental Frequency of Spontaneous Cries during the first 3 Months, *Int. J. Pediatr. Otorhinolaryngol.,* **64** (2002), pp. 97–104.

[12] MICHELSSON, K.–MICHELSSON, O., Phonation in the Newborn, Infant Cry, *Int. J. Pediatr. Otorhinolaryngol.*, **49/1** (1999), pp. S297–S301.

[13] MICHELSSON, K.–EKLUNKD, K.–LEPPANEN, P.–LYYTINEN, H., Cry Characteristics of 172 Healthy 1- to 7-day-old Infants, *Folia. Phoniatr. Logop.*, **54** (2002), pp. 190–200.

[14] MÖLLER, S.–SCHÖNWEILER, R., Analysis of infant cries for the early detection of hearing impairment, *Speech. Commun.*, **28** (1999), pp. 175–193.

[15] PURHONEN, N.–PAAKKONEN, A.–YPPARILA, H.–LEHTONEN, J.–KARHU, J., Dynamic Behavior of the Auditory N100 Elicited by a Baby's Cry, *Int. J. Psychophysiol.*, **41/3** (2001), pp. 271–278.

[16] ROTHGANGER, H., Analysis of the Sounds of the Child in the First Year of Age and a Comparison to the Language, *Early Hum. Dev.*, **75** (2003), pp. 55–69.

[17] RUNEFORS, P.–ARNBJORNSSON, E.–ELANDER, G.– MICHELSSON, K., Newborn Infants' Cry after Heel-prick: Analysis with Sound Spectrogram. *Acta. Paediatr.*, **89/1** (2000), pp. 68–72.

[18] SCHÖNWEILER, R.–SCHÖNWEILER, B.–SCHMEIZEISEN, R., Hearing function and language skills of 417 children with cleft palates, *Int. J. Pediatr. Otorhinolaryngol.*, **32/3** (1995), pp. 278–279.

[19] SCHÖNWEILER, R.–KAESE, S.–MÖLLER, S.–RINSCHEID, A.–PTOK, M., Neuronal Networks and Self-organizing Maps: New Computer Techniques in the Acoustic Evaluation of the Infant Cry, *Int. J. Pediatr. Otorhinolaryngol.*, **38** (1996), pp. 1–11.

[20] SCHÖNWEILER, R.–KAESE, S.–MÖLLER, S.–RINSCHEID, A.– PTOK, M., Classification of Spectrographic Voice Patterns Using Self-organizing Neuronal Networks (Kohonen maps) in the Evaluation of the Infant Cry with and without Time-delayed Feedback, *Int. J. Pediatr. Otorhinolaryngol.*, **38/2** (1996), pp. 181.

[21] SCHÖNWEILER, R.–PTOK, M.–RADÜ, H.-J., A Cross-Sectional Study of Speech- and Language-abilities of Children with Normal Hearing, Mild Fluctuating Conductive Hearing Loss, or Moderate to Profound Sensoneurinal Hearing Loss, *Int. J. Pediatr. Otorhinolaryngol.*, **44** (1998), pp. 250–258.

[22] VÁRALLYAY JR., G.–BENYÓ, Z.–ILLÉNYI, A.–FARKAS, Z.–KATONA, G.–CSÁKÁNYI, ZS.–MAJOROS, T.–BENEDEK, P.–PATKÓ, T., Csecsemősírás számítógépes elemzése. In: BUDAMED '02 Előadás kivonatok, Z. Pintér ed., MATE, Budapest, 2002. pp. 60–61.

[23] VÁRALLYAY JR., G., A sírás és a hallás kapcsolata csecsemőkorban, In: Tavaszi Szél Konferencia 2003., Doktoranduszok Országos Szövetsége, Budapest, 2003, pp. 95–98.

[24] VÁRALLYAY JR. G., Infant Cry Analyser System for Hearing Disorder Detection, *Periodica Politechnica, TU Timiºoara, Transactions on Automatic Control and Computer Science*, **49/63** (2004), pp. 57–60.

[25] VÁRALLYAY JR., G.–BENYÓ, Z.–ILLÉNYI, A.–FARKAS, Z.–KOVÁCS, L., Acoustic Analysis of the Infant Cry: Classical and New methods, *Proc. 26th Conf. IEEE Engineering in Medicine and Biology, San Francisco, CA, Sept 2004,* pp. 313–316.

[26] VÁRALLYAY JR., G.–BENYÓ, Z.–ILLÉNYI, A.–FARKAS, Z., Methods for the Analysis of Acoustic Biomedical Signals. Proc. IASTED International Conference on Biomedical Engineering (BioMED 2005), Innsbruck, Feb 2005, pp. 434–438.

[27] WERMKE, K.–MENDE, W.–MANFREDI, C.–BRUSCAGLIONI, P., Developmental Aspects of Infant's Cry Melody and Formants, *Med. Eng. Phys.*, **24/7–8** (2002), pp. 501–514.

[28] ZHANG, T.–KUO, C. C. J., Audio Content Analysis for Online Audiovisual Data Segmentation and Classification, *IEEE Trans. on Speech and Audio Processing,* **9/4** (2001), pp. 441–457.