György Várallyay Jr., András Illényi, Zoltán Benyó, Zsolt Farkas, Gábor Katona:
The Speech-Chorus Method at the analysis of the infant cry.
in: Acoustic Rewiev (Akusztikai Szemle), 2005:6(2), 9-15, HU ISSN 0023-0480.

The Speech-Chorus Method at the Analysis of the Infant Cry

Csecsemősírások elemzése a Beszédkórus Módszerrel

G. Várallyay[*] Jr., A. Illényi[**], Z. Benyó[*], Z. Farkas[***], G. Katona[***]

*Dept. of Control Engineering and Information Technology, Budapest University of Technology and Economics, Hungary
**Dept. of Telecommunications and Media Informatics, Budapest University of Technology and Economics, Hungary
***Heim Pál Hospital for Sick Children, Budapest, Hungary

varallyay@bio.iit.bme.hu, illenyi@tmit.bme.hu, benyo@iit.bme.hu, drfarkaszsolt@yahoo.com, katonag@heimpalkorhaz.hu

Summary

The infant cry has been analyzed mostly in medical, pedagogical and acoustical aspects since a long time. There are several new methods in the field of signal processing which renew the problem. It proves useful to adopt the Speech-Chorus Method at the analysis of the infant cry to determine the typical attributes of the sound of the speaker (crying infant). In this article authors report their investigation about crying which was started a few years ago. As it will be shown, in case of crying with constant tonality the Speech-Chorus Method is applicable.

Összefoglalás

A csecsemősírás elemzése főleg orvosi, pedagógiai és akusztikai szempontból hosszú ideje az érdeklődés homlokterében áll. A jelfeldolgozás újabb lehetőségei a feladat megújítását eredményezik. Ezek alkalmazásánál hasznosnak bizonyul a klasszikusnak számító Beszédkórus Módszer, amelynek célja a beszélő – síró – egyed jellegzetes hangösszetevőinek kimutatása. A cikkben bemutatjuk több éves kutatásainkat és azokon belül az említett módszer mai alkalmazhatóságának lehetőségeit. Mint kiderül, állandó tonalitású sírásszakaszok esetén a Beszédkórus Módszer alkalmazható.

1. Introduction

With this study authors are to pay honor to professor Tarnóczy and his famous work. Authors' research is dealing with the analysis of the infant cry; this investigation has been started a few years ago and based on the Speech-Chorus Method from Tarnóczy (*Tarnóczy*, 1955; *Tarnóczy*, 1970).

1.1. The Speech-Chorus Method

In 1970 Tarnóczy had reported about the Speech-Chorus Method in the Acustica (Vol. 23, Heft 4). The summary of his article was the following:
"*A new method has been applied already to determine the average long time energy spectrum of speech. According to this method several persons speak different texts simultaneously and the resulting sound is recorded on a single magnetic tape. The recording can be analyzed by the methods suitable to the temporary constant signals of a continuous spectrum.*

In the present paper it is demonstrated that already a speech chorus of eight persons produces a signal having a time period of approximately 10 s. Thus averaging over a long time and on different individuals is done simultaneously.
Since speech cannot be considered a stochastic process the speech-chorus method which makes the process approach a stochastic one offers an example to extend the validity of the central limit theorem."
Authors will show how this method can be integrated into the acoustic analysis of the infant cry, as infant cry also cannot be considered a stochastic process.

1.2. The infant cry

The cry 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. This is a multimodal, dynamic behavior.
An infant cries in another way if he/she is hungry, in pain, in discomfort (*e.g.* need to be changed) or he/she is 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 (*Benyó, et. al.,* 2002; *Purhonen, et. al.,* 2001).
The infant the cry may differ if any disorder occurs. This proposes the idea of a diagnostic system based on the infant cry. In this study, infants with hearing disorders will be compared with healthy babies as an example. Authors had dealt with this question before and investigated other attributes of the infant cry (*Benyó, et. al.,* 2004; *Várallyay,* 2004; *Várallyay, et. al.,* 2005a). 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.

1.3. 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 (*Barr, et. al.,* 2000). 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. 1.** shows a universal schema of the voice-production by Gordos and Takács (*Gordos–Takács,* 1983).

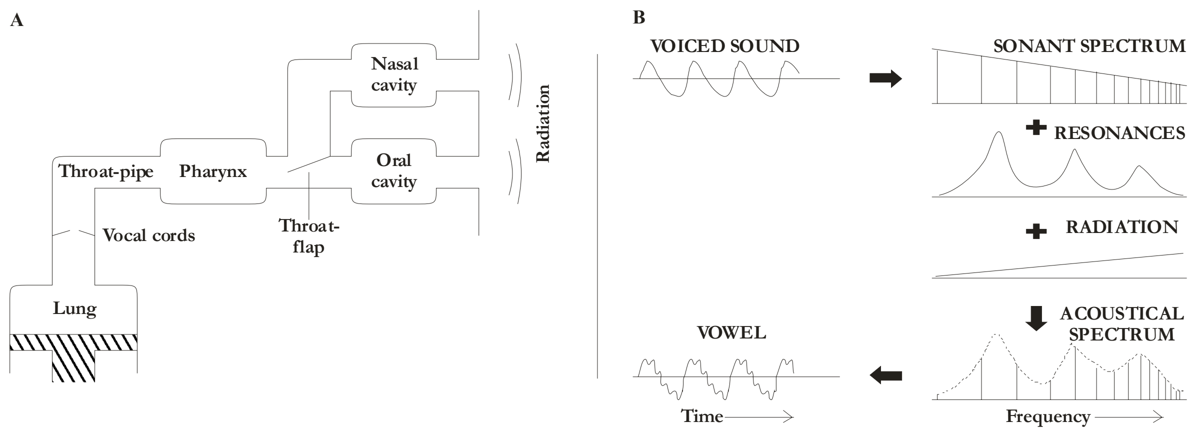

Fig. 1. A. Universal schema for the voice-production by Gordos et al.
B. Formation of a vowel in the human phonation system.

1. ábra. A. Az emberi hangképzőrendszer sematikus rajza Gordos és munkatársai szerint.
B. Az emberi hangképzőrendszer által létrehozott hangzó kialakulása.

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.
On **Fig. 2.** Golub's physioacoustic model of crying assumes three levels of central processing of the muscles contributing to the source and filters of crying (*Golub,* 1980). 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 coordinated in the act of crying.

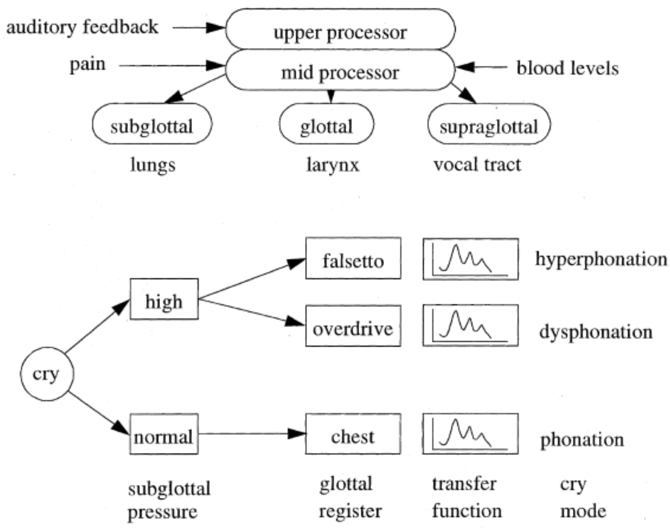

Fig. 2. Golub's physioacoustic model.
2. ábra. Golub pszichoakusztikai modellje.

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. Furthermore, the infant cry should be analyzed both in the time and in the frequency domains.

2. Analysis of the infant cry

2.1. Historical background

The sound spectrogram was the major tool for analyzing cry sounds in the 1960s and '70s (*Michelsson, et. al.,* 1977). 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.
In Hungary Hirschberg and Szende had written a complete study round *Pathological cries, stridors and coughs in infants*; which was mostly based on spectrograms in the 1980s (*Hirschberg–Szende,* 1982). 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 (*Michelsson–Michelsson,* 1999). In 1999, 14 years after the above mentioned book, they've introduced new spectrograms of infants with hypothyroidism, asphyxia or meningitis.

There are several physical 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 (e.g. color spectrograms, 3D spectrograms).

Runefors *et al.* tested the cry of 50 newborn infants after heel-prick (*Runefors, et. al.,* 2000). 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 (*Michelsson, et. al.,* 2002). They didn't find significant difference in the duration of the cries by gender.

Lind and Wermke investigated an infant from birth to the age 3 months in 2002 (*Lind–Wermke,* 2002). They used special hardware and software to record the sound and to analyze 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 (*Rothgänger,* 2003). 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 (*Benyó, et. al.,* 2002; *Gilbert–Robb,* 1996; *Wermke, et. al.,* 2002). In the change of the duration Rothgänger found increment both in crying and babbling.

2.2. Data collection and database

Recordings are mainly made in the Heim Pál Hospital for Sick Children in a quiet room. Authors use 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 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.

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 a MS Excel database (*Várallyay, et. al.,* 2004). 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.

2.3. Preprocessing

Before our 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 (*Várallyay, et. al.,* 2002).
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 analyzed. On **Fig 3.** *less important* parts are hiccough (a and f), hoarseness (c), spasmodic crying (d) and silence (e) between the crying segments (b and g).

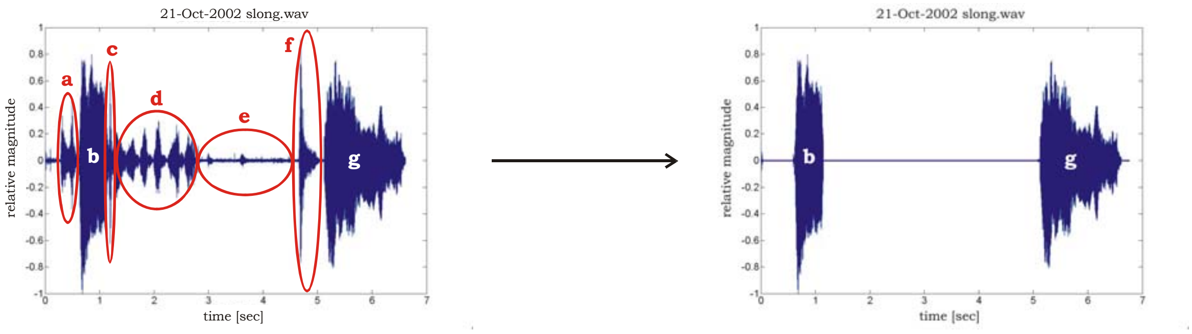

Fig. 3. A sample wave file before and after feature extraction.
3. ábra. Egy minta hang fájl a lényegkiemelés előtt és után.

There are averagely 8-10 segments in a 30 second long recording. Automatic segmentation is needed to detect the start and the end of each segment. Authors used two different short-time methods for segmentation: the short-time energy function and the short-time zero crossing rate (*Várallyay,* 2005b). These methods are well-known *e.g.* in the indexing of accompanying audio signals in movies and video programs (*Zhang–Tuo,* 2001). Authors clarify the signal by cutting the noisy parts out (*i.e.* a, c, d, e and f on **Fig. 3.**), because analyzing the whole signal could be misleading.

2.4. Adopting the speech-chorus method

In the original speech-chorus method the summarized speech of several persons (*i.e.* the speech chorus) were recorded and analyzed. Authors want to obtain an average spectrum which characterizes the frequency content of the crying segment. To adopt the speech-chorus method the crying segment are divided into windows; the length of each window is *N* (see on **Fig. 4.**). These windows are added together, as the speech of the members of the speech chorus was added in the original method.

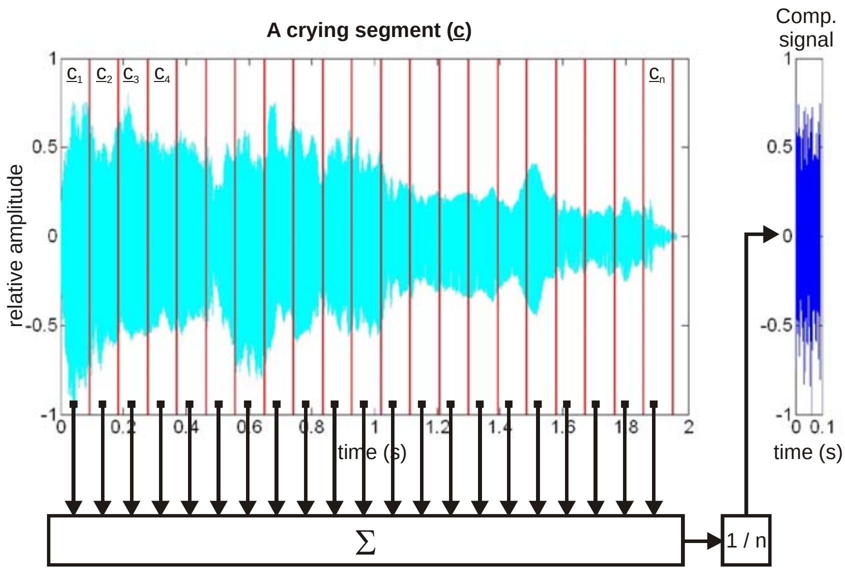

Fig. 4. Adopting the speech-chorus method.
4. ábra. A beszédkórus módszer alkalmazása.

In contradistinction to the original speech-chorus method in this case only one infant is needed. By these steps a compressed signal is created, which has only N points, while the original crying segment had $n \cdot N$ points. This compressed signal contains all the frequency information of the crying segment, thus the Fourier transform will give the average energy spectrum of the crying segment (**Fig. 5.**).

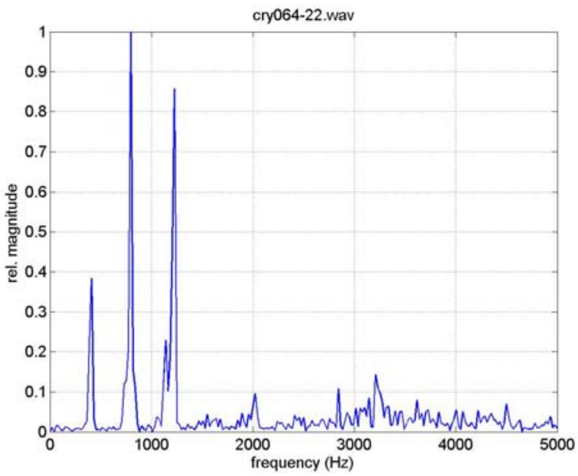

Fig. 5. The average energy spectrum of a crying segment (cry064-22.wav).
5. ábra. Egy sírásszegmens (cry064-22.wav) átlagos energiaspektruma.

This average energy spectrum shows that the crying segment is a harmonic signal. The fundamental frequency of the crying segment is approximately 400 Hz. Both the second and the third order of higher components are also visible in the spectrum. Further subharmonics have smaller intensity; a few of them can be also recognized.
Usually the value of the fundamental frequency of the infant cry is between 300 and 700 Hz (*Wermke, et. al.,* 2002; *Várallyay,* 2003; *Schönweiler, et. al.,* 1996a). The typical length of the crying segments is 1.1±0.6 s.

2.5. Mathematical explanation

A crying segment is given (see on the previous figure):

$$\underline{c} = [c(1), c(2), \ldots, c(k), \ldots c(n \cdot N)] \tag{1}$$

We split up this vector (*i.e.* the signal) into windows which are same in size:

$$\underline{c} = [\underline{c}_1, \underline{c}_2, \ldots, \underline{c}_i, \ldots, \underline{c}_n] \tag{2}$$

Where $\underline{c}_1, \underline{c}_2, \ldots, \underline{c}_n$ are N points length vectors. Now make the *compressed signal* (or *average window*) by adding these windows together, and dividing with the number of windows:

$$\langle \underline{c} \rangle = \frac{1}{n} \sum_{i=1}^{n} \underline{c}_i \tag{3}$$

The spectrum of the compressed signal is the following.

$$\langle \underline{C} \rangle = FFT\left\{ \langle \underline{c} \rangle \right\} \tag{4}$$

Where FFT means Fast Fourier Transform:

$$FFT\left\{ \underline{c}_i \right\} = C_i(k) = \sum_{j=1}^{N} c_i(j) \cdot \left[e^{(-2\pi i)/N} \right]^{(j-1)(k-1)} \tag{5}$$

By adopting the speech-chorus method authors got the *average energy spectrum* of the crying segment. With this method, only one FFT needed. The average energy spectrum could be also calculated by averaging the energy spectrums of the windows.

$$\langle \underline{C} \rangle = \frac{1}{n} \sum_{i=1}^{n} \underline{C}_i = \frac{1}{n} FFT\left\{ \underline{c}_i \right\} \tag{6}$$

If $\underline{C}_1, \underline{C}_2, \ldots, \underline{C}_n$ are known, we can calculate the average energy spectrum on this way:

$$\frac{1}{n} \sum_{i=1}^{n} \underline{C}_i = \frac{1}{n} \sum_{i=1}^{n} FFT\left\{ \underline{c}_i \right\} = FFT\left\{ \frac{1}{n} \sum_{i=1}^{n} \underline{c}_i \right\} = FFT\left\{ \langle \underline{c} \rangle \right\} = \langle \underline{C} \rangle \tag{7}$$

There are two different ways to get the average energy spectrum, but the speech-chorus method needs only one FFT. It means, by using the speech-chorus method we can obtain the average energy spectrum quicker.

3. Application

This paper deals with classical and new methods of acoustic analysis of the infant cry. The final goal is to detect disorders according to the crying at the earliest possible moment. Classical acoustic methods were reproduced and compared with solutions, which were not available before. Authors have developed several methods to analyze the infant cry and to

compare parameters of healthy and unhealthy crying (*Várallyay, et. al.,* 2004; *Várallyay, et. al.,* 2005b).
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. It is based on the harmonic structure of the spectrum and statistical algorithms. This method works with 97.99 % reliability (*Várallyay,* 2005a).
The dominant frequency is the biggest frequency component in the spectrum. The dominant frequency is one of the subharmonics of the fundamental frequency, thus the ratio of the dominant frequency and the fundamental frequency is an integer, in most cases this ratio is 2 or 3. The detection of the dominant frequency needs global maximum value detection in the whole spectrum. Authors found differences at the ratio of the dominant frequency and the fundamental frequency (*Benyó, et. al.,* 2002). If the infant has normal hearing, this ratio is 3. If the infant has hearing disease, the ratio of the dominant frequency and the fundamental frequency is 2.

4. Results

Authors have found in many cases that the average energy spectrum is not regular as it is shown on **Fig. 5**. The following figure (**Fig. 6.**) shows a few examples:

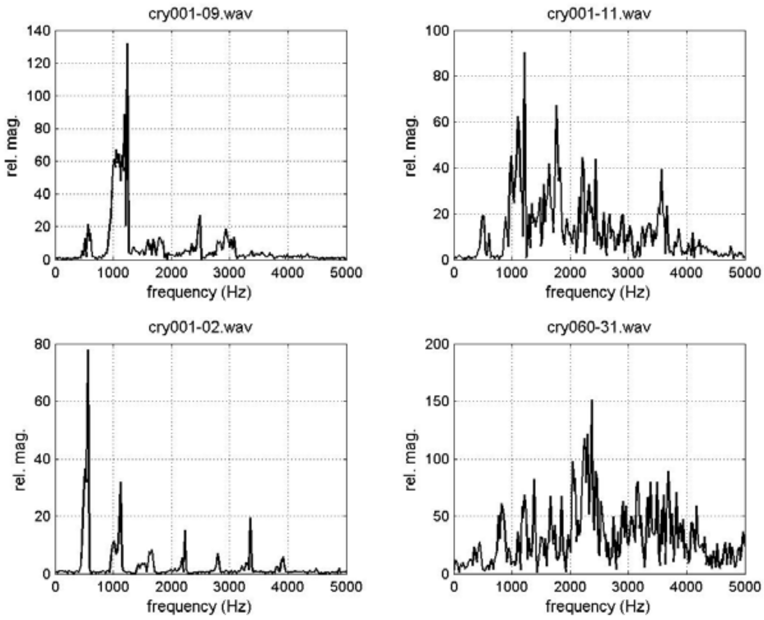

Fig. 6. Examples to the average energy spectrum of the infant cry obtained by the speech-chorus method. Two spectrums on the left recall to a harmonic signal, but the two spectrums on the right don't.
6. ábra. Példák a beszédkórus módszerrel kapott átlagos energiaspektrumra. A bal oldali két spektrum szabályosabb, a jobb oldaliak szabálytalanabb struktúrájúak.

The explanation of this observation is the following. By using the speech-chorus method the original signal is compressed and the time information is lost. The infant cry has a *melody*, which can be defined as the changing of the fundamental frequency as a function of time.

There are different melody types as flat, rising, falling, rising-falling, etc. If the fundamental frequency changes within a bigger range (*i.e.* the melody is not flat), the average spectrum gets more complex, as it is shown on the previous figure on the right hand side.
There are connections between the melody and the gender, the mood or several diseases (*Michelsson–Michelsson,* 1999; *Schönweiler, et. al.,* 1996b; *Várallyay, et. al.,* 2005c), thus the melody is an important attribute of the infant crying.

4.1. A rule to find flat melody

By using the speech-chorus method, an easy rule can be generated to find flat melody.

- Create the average energy spectrum by applying the above mentioned steps.
- If the average energy spectrum looks like a harmonic spectrum, the melody is flat.
- If the average energy spectrum doesn't look like a harmonic spectrum, we have to analyze the signal window by window, and determine the fundamental and dominant frequencies for each window.

5. Conclusion

The speech-chorus method compresses the original crying segment, and generates the average energy spectrum of the signal using only one FFT. The average energy spectrum shows all the frequency information of the whole crying segment. If the fundamental frequency doesn't change too much in the time domain (*i.e.* the melody type is flat), we get a harmonic spectrum as the average. If there is a major changing of the fundamental frequency, the average spectrum will be less regular.
The melody contour is an important attribute of the infant cry, thus it is necessary to obtain the type of the melody quickly. By using the speech-chorus method, a quick rule can be set to find flat melody. If there is flat melody, the windows of the divided segment are similar, so only one window is enough to be analyzed.

Acknowledgements

The research has been supported by Hungarian National Scientific Research Foundation (OTKA T0422990, T029830) and the Medical Research Council (ETT 089/2003).

References

Barr, R. G.–Hopkins, B.–Green, J. A.: Crying as a sign, a symptom and a signal, MacKeith Press, London, 2000.

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.

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.

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.

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.

Gordos, G.–Takács, G.: Digitális beszédfeldolgozás, Műszaki Kiadó, Budapest, 1983.

Hirschberg, J.–Szende, T.: Pathological cry, stridor and cough in infants, Akadémiai Kiadó, Budapest, 1982.

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.

Michlesson, K.–Sirviö, P. –Wasz-Höckert, O.: Pain cry in fullterm asphyxiated newborn infants correlated with late findings, Acta Paediatr Scand, 66 (1977), pp. 611–616.

Michelsson, K.–Christensson, K.–Rothganger, H.–Winberg, J.: Crying in separated and non-separated newborns: sound spectographic analysis, Acta Paediatr, 85/4 (1996), pp. 471–475.

Michelsson, K.–Michelsson, O.: Phonation in the newborn, infant cry, Int. J. Pediatr. Otorhinolaryngol., 49/1 (1999), pp. S297–S301.

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.

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.

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.

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.

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.

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.

Tarnóczy T.: Determination of the Speech Spectrum through Measurements of Superposed Samples. The Journal of the Acoustical Society of America, 28(6) (1956), pp. 1270–1275.

Tarnóczy T.: Die Sprechchor-Methode, Acustica, 23/4 (1970), pp. 173–188.

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.

Várallyay Jr., G.: A sírás és a hallás kapcsolata csecsemőkorban, In: Tavaszi Szél Konferencia 2003., Doktoranduszok Országos Szövetsége, Budapest, 2003, pp. 95–98.

Várallyay Jr., G.: Infant cry analyzer system for hearing disorder detection, Periodica Politechnica, TU Timişoara, Transactions on Automatic Control and Computer Science, 49/63 (2004), pp. 57–60.

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.

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.

Várallyay Jr., G.: Új eljárás a harmonikus jelek alapfrekvenciájának meghatározására, In: Tavaszi Szél Konferencia 2005., Doktoranduszok Országos Szövetsége, Budapest, 2005, pp. 172–175.

Várallyay Jr., G.–Benyó, Z.–Illényi, A.–Farkas, Z.–Katona, G.: An Attempt to Detect Hearing Disorders by Acoustic Features of the Infant Cry, Proc. Forum Acusticum 2005, Budapest, (2005), in press

Várallyay Jr., G.–Benyó, Z.–Illényi, A.–Farkas, Z.: Számítógépes módszerek a csecsemősírások dallamának detektálására, megjelenítésére és kiértékelésére, Proc. BUDAMED ’05, Budapest, (2005), in press

Várallyay Jr., G.: Future prospects of the application of the infant cry in the medicine, Periodica Polytechnica Ser. El. Eng. (2005), in press

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.

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.