ANALYSIS OF INFANT CRIES FOR THE EARLY DETECTION OF HEARING IMPAIRMENT

S. Möller*, R. Schönweiler**

*Institute of Communication Acoustics, Ruhr-Universität Bochum, D-44780 Bochum, Germany
Tel. +49 234 700 3979, Fax +49 234 709 4165, E-mail: moeller@ika.ruhr-uni-bochum.de

**Department of Phoniatrics and Pedaudiology, Hannover Medical School, D-30623 Hannover, Germany
Tel. +49 511 532 9104, Fax +49 511 532 4609

ABSTRACT
Auditory and instrumental analysis of cries from normally hearing and profoundly hearing-impaired infants (2-11 month) is presented. Results from listening experiments lead to the assumption that differences exist between cries from the two infant groups. Attributes expressing the difference are related to the emotional state of the infant, to prosodic features, and to voice quality. Signal analysis of the cries confirms these findings showing statistically significant differences for spectral parameters and those describing the melody contour of the cries. The usability of neural networks for an automatic classification and discrimination of cries is discussed. If the tendencies shown here hold true for other data sets, the findings can be used to develop a new screening method detecting hearing impairment and auditory perception disorders at a very early age.

1. INTRODUCTION
Infant cries both of healthy and impaired children have widely been investigated in the past (cf. e.g. [9] for a review). Besides the scientific interest in the cry production process, the aim of most of the studies was to use the infant cry as a diagnostic tool for different diseases. Although there exist some very typical cry features (e.g. the cry-du-chat syndrome, showing a very high-pitched cry related to a chromosome defect), to our knowledge the cry is currently not used in clinical diagnostics.

Disappointingly, little is known about the cries of profoundly hearing-impaired children. Most studies aim at the canonical babbling onset [1, 3]. However, reactions of newborn infants to the delayed presentation of their own crying have been reported [2]. This shows an influence of auditory feedback on infant sound production, far before the babbling stage. The lack of investigations in this area may be due to the fact that hearing-impairment is diagnosed in general at a very advanced age (> 2 years). However, in many cases an early diagnosis and provision with hearing aids can enable nearly normal language development.

Based on spectrographic analysis, Truby and Lind [11] identified three major types of cries: the basic phonation cry, a turbulent dysphonation cry and a very high-pitched hyperphonation cry. The most complete model for the production of these cry types was developed by Golub [4], separating the cry production into subglottal, glottal and supraglottal production areas, and relating muscle activities to each cry type. A schematic presentation of the model is given in Figure 1.

However, it must be admitted that little is known about the control of the cry production process, as it regards the ‘control organizer’ depicted in the upper part of the figure. Concerning the tests reported here, results should therefore be seen not only with regard to the development of a diagnostic tool. They could also help to improve our knowledge of the cerebral organization and development in language acquisition.

Our study is based on the assumption that the cries of hearing-impaired children differ from those of their normally hearing counterparts due to the lack of auditory feedback. This intuition, which relies on information from clinical experts, is investigated by means of auditory experiments with both, naive and expert listeners. The listening test results are then interpreted in the context of data obtained by means of signal analysis of the cries. This analysis regards the spectral, temporal and melodical structure of the cries. Finally, a possible application of the findings for an early detection of hearing impairment is discussed.

2. AUDITORY TESTS

2.1 Recording of Infant Cries
Cries from seven healthy, normally hearing infants and seven profoundly hearing-impaired infants have been recorded during visits in hospital. The cries were triggered by putting on uncomfortable headphones, which provided auditory feedback to the infants in the limits of their hearing capacity.
It is important to mention that these cries reflect the children’s uneasiness, and not a maximum response like pain. We think that cries stimulated in this way show more influence of the auditory feedback than, e.g., pain cries.

In the first auditory test sequences of several cries (approx. 7 sec.) from 10 children were presented to the listening subjects. In the second auditory test, as well as for signal analysis, single cries were used.

2.2 Listening Subjects

Listening subjects were divided into different groups according to their experience with infants cries: 10 midwifes and pediatric nurses, 7 audiologists, 7 mothers or fathers of younger infants, and 10 naive listeners. Former experiments suggest that the degree of experience has a large influence on the classification ability, and on the cry perception in general.

2.3 Test Procedures

In the first test, the listening subjects had to classify the cries regarding the children’s expected hearing capacity. They were then given time to comment on their impressions and on the classification scheme they used. This phase of so called ‘intermitting loud thinking’ [10] helps to interpret the subjects’ answers and to relate them to perceptual and acoustic cry features.

In a second test, subjects had to describe the cries with regard to 10 attribute scales, according to the semantic differential method [8]. The attributes that we used relate to voice and melodic cry characteristics as well as the cry perception by the listener.

3. SIGNAL ANALYSIS

A number of 153 cry signals of all 14 infants were analysed by various algorithms. The following features related to different aspects of the cry production were extracted:

Temporal parameters

Besides the duration of the single cries, the rhythm was investigated by spectral analysis of the cry-and-pause signal of longer sequences.

Spectral parameters

The absolute energy as well as the energy distribution in frequency bands were analysed. Additionally, mel-scaled cepstral coefficients were calculated.

Vocal tract parameters

Contours of the first three formants were analysed, and mean, minimum and maximum values were calculated.

Voice source parameters

Many of the known fundamental frequency (F0) extraction algorithms fail for the high-pitched, variable infant cries. A new algorithm was implemented, working both in the time and autocorrelation domain. Mean, minimum and maximum F0-values and the F0 jitter were calculated. Additional parameters describing the type of voice source (percentage voiced excitation, glottal-to-noise excitation ratio [7]) were extracted.

Melody parameters

The melody contour plays an important role in cry perception. In order to achieve a parametric representation of the F0 contour which otherwise would be difficult to handle, an LPC based model described in [6] as well as a model based on a simplified Fujisaki approach were implemented.
4. RESULTS AND DISCUSSION

4.1 Results from Auditory Tests

In the first auditory test, discrimination between normally hearing and hearing-impaired infants proved to be impossible for non-expert listeners (see boxplot in Fig. 2). Classification accuracy around the statistically motivated 50% are usual for naive listeners as well as parents and even audiologists. Only the midwifes’ and nurses’ group which is in very frequent contact with both healthy and hearing-impaired infants shows significantly higher classification scores. In the free description, however, the cries of hearing-impaired infants were described as more whining, ailing, monotonous and weaker. The voice was described as flat and guttural.

The direct cry description by means of the semantic differential method shows a large spread in the judgement. As a consequence, significant differences do not become very clear. A better description can be obtained by factor analysis of the auditory judgements, which reveals two factors related to voice characteristics, and one factor related to prosodic cry features. Presenting the cries in the rotated factor space (Fig. 3), it can be seen that the cries of the normally hearing infants 1, 3, 4 and 5 cluster in an other area of the factor space then the cries obtained from the hearing-impaired infants. Considering the little amount of data material used for the auditory tests, an auditory discrimination of the cries seems to be possible.

![Figure 3: Cries from auditory test 2 in the rotated factor space (numbers refer to infants: 1-5 norm. hearing, 6-10 hearing imp.)](image)

4.2 Results from Signal Analysis

A statistical comparison of the extracted parameters for the two infant groups should be done with care. One reason is the limited amount of data material obtainable so far. Additionally, for some features it is not clear whether they are cry-specific or individual-specific, and consequently whether parameters from single cries or infant means should be compared. We therefore only refer to findings which show the same tendency for single cries and infant means.

Main statistically significant differences were found in the distribution of energy in different frequency bands, the cry duration, and some melody parameters. Cries of hearing-impaired infants are longer in duration and more complicated in their melodic structure, i.e. the number of $F_0$ maxima and minima. Frequency analysis shows less energy between 2-4 kHz and 6.4-8.5 kHz.

The analysis of the cry-and-pause rhythm was carried out on less data material, as longer cry sequences had to be found. In this data material, the basic rhythm frequency is much lower for the hearing-impaired children. This correlates with the more whining and monotonous cry impression of the listening subjects.

4.3 Discussion

The results obtained from auditory tests as well as signal analysis congruently show that the cries of normally hearing infants differ from those of profoundly hearing-impaired. Main differences are the spectral characteristics, which influence the ‘sound’ perceived, and this can be used by experienced subjects to discriminate cries. Other differences regard melody and duration.

Generalization of the results, however, is restricted due to the limited amount of data material obtained so far. It will be necessary to repeat the investigations with other data sets. In literature, it has been shown that several central nervous system (CNS) diseases coincide with a higher fundamental frequency in the infant cry. For newborn infants a modified $F_0$ fine structure has been reported, even though the CNS disease for those infants was detected at a more advanced age [12].

Our findings suppose that a missing auditory feedback may be manifested in a modified spectral shaping and global cry characteristics, e.g. melody contour. This correlates with effects in later language development. The canonical babbling is very restricted and delayed for hearing-impaired infants [1, 3]. If the results of our experiments can be repeated with other data sets, they could be used to develop a very early applicable and easy-to-handle diagnostic tool.

5. APPLICATION FOR DIAGNOSIS

In order to develop a diagnostic tool, automatic classification of the data by means of self-organizing feature maps has been investigated. The network type used is a two-dimensional Kohonen LVQ card, which allows supervised learning and a simple visualization of
the classification results [5]. Among the input features tested there are short time and long time mel spectra, linear and mel-scaled cepstra and LPC PARCOR coefficients.

Some classification results are shown in Table 1. The classification accuracy was calculated as a mean for several cries from the same infant, and for several training and test sets. Children classified as ‘uncertain’ (i.e. between 33 and 66%) were excluded from this statistic. It should be noted that these results refer to input vectors from infants whose cries have not been used for the training, so these infants were completely ‘unknown’ to the feature map.

![Table 1: Classification results for different input vectors](image)

<table>
<thead>
<tr>
<th>input vectors</th>
<th>classification accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>mel-scaled short time spectra</td>
<td>63 %</td>
</tr>
<tr>
<td>mel-scaled long time spectra</td>
<td>68 %</td>
</tr>
<tr>
<td>linear scaled cepstral coefficients</td>
<td>75 %</td>
</tr>
<tr>
<td>mel-scaled cepstral coefficients</td>
<td>82 %</td>
</tr>
<tr>
<td>LPC PARCOR coefficients</td>
<td>73 %</td>
</tr>
</tbody>
</table>

It can be seen from Table 1 that with all kind of input vectors the classification accuracy is higher than the statistically motivated 50%. Although the absolute numbers should not be generalized as they reflect the limited data set to a high degree, these results confirm the findings of the auditory tests and signal analysis. Improvements should be possible by incorporating melodic and long-term cry characteristics.

6. CONCLUSION

It has been shown that an auditory as well as signal analysis based discrimination of the cries of hearing-impaired infants from those of normally hearing infants is possible. Restrictions in generalization have to be made due to the little amount of data material available for the tests until now. The use of self-organizing feature maps opens up the possibility for an application for the early detection of hearing-impairment. Interpretation of the results in the context of cry and language production can also improve our knowledge on the cerebral control of these features.

ACKNOWLEDGEMENTS

This study is supported by the German Federal Research Fund, DFG (Deutsche Forschungsgemeinschaft).

REFERENCES