Available online at www.sciencedirect.com

ScienceDirect

Speech Communication 58 (2014) 91–100

SPEECH COMMUNICATION

www.elsevier.com/locate/specom

Infant cry reliability: Acoustic homogeneity of spontaneous cries and pain-induced cries

Tanja Etz [a,b,*], Henning Reetz [b,1], Carla Wegener [a,2], Franz Bahlmann [c,3]

[a] *Fresenius University of Applied Sciences Idstein, Limburger Straße 2, 65510 Idstein, Germany*
[b] *Department of Phonetics, Goethe University of Frankfurt, Senckenberganlage 31, 60325 Frankfurt, Germany*
[c] *Bürgerhospital Frankfurt am Main, Nibelungenallee 37-41, 60318 Frankfurt, Germany*

Received 12 April 2013; received in revised form 22 October 2013; accepted 14 November 2013
Available online 28 November 2013

Abstract

Infant cries can indicate certain developmental disorders and therefore may be suited for early diagnosis. An open research question is which type of crying (spontaneous, pain-induced) is best suited for infant cry analysis. For estimating the degree of consistency among single cries in an episode of crying, healthy infants were recorded and allocated to the four groups spontaneous cries, spontaneous non-distressed cries, pain-induced cries and pain-induced cries without the first cry after pain stimulus. 19 acoustic parameters were computed and statistically analyzed on their reliability with Krippendorff's Alpha. Krippendorff's Alpha values between 0.184 and 0.779 were reached over all groups. No significant differences between the cry groups were found. However, the non-distressed cries reached the highest alpha values in 16 out of 19 acoustic parameters by trend. The results show that the single cries within an infant's episode of crying are not very reliable in general. For the cry types, the non-distressed cry is the one with the best reliability making it the favorite for infant cry analysis.

Keywords: Infant cry; Reliability; Acoustic analysis

1. Introduction

For many years, acoustic analysis has been used to explore the infant cry. Infant cry research has been looking for differences between the cry of normally developing infants as well as the cry of infants with several diseases and developmental disorders. Differences in acoustic parameters have been related to disturbances of the vocal neuromuscular maturation (Lind et al., 1967; Golub and Corwin, 1982; Fort and Manfredi, 1998), to brain disorders (Sirviö and Michelsson, 1976; Fisichelli and Karelitz, 1966; Wasz-Höckert et al., 1968; Karelitz and Fisichelli, 1962), to central nervous system insults (Lester et al., 2002; Corwin et al., 1992; Blinick et al., 1971; Nugent et al., 1996; Michelsson et al., 1977; Verduzco-Mendoza et al., 2012), to various developmental disorders like hearing impairment (Möller and Schönweiler, 1999; Arch-Tirado et al., 2004; Várallyay, 2007; Etz et al., 2012) or autism (Esposito et al., 2013; Sheinkopf et al., 2012), and to genetic defects like the down syndrome (Fisichelli and Karelitz, 1966; Lind et al., 1970), morbus crabbe (Thoden and Michelsson, 1979) and the cri-du-chat syndrome (Wasz-Höckert et al., 1968; Vuorenkoski et al., 1966). Many of those studies aimed at using the infant cry as an early, non-invasive diagnostic instrument to test for various diseases by analyzing several acoustic parameters.

While differences between healthy cries and those related to disorders are already explored, only little research about the *reliability* of the infant cry has been conducted (Robb et al., 1997; Lind and Wermke, 2002; Green et al., 1998).

* Corresponding author. Address: Müllerwies 14, 65232 Taunusstein, Germany. Tel./fax: +49 6128 210 764.

E-mail addresses: tanja.etz@hs-fresenius.de (T. Etz), reetz@em.uni-frankfurt.de (H. Reetz), wegener@hs-fresenius.de (C. Wegener), f.bahlmann@buergerhospital-ffm.de (F. Bahlmann).

[1] Tel.: +49 69 798 25031; fax: +49 69 798 23774.

[2] Tel.: +49 6126 935 2913; fax: 49 6126 935 2174.

[3] Tel.: +49 69 1500 853.

0167-6393/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.specom.2013.11.006

For this study, an *episode of crying* was defined as the total period of continuous crying activity (Grau et al., 1995). In an episode of crying, an infant produces multiple *cries*. Cries are extracted from the episode of crying for analysis. In the context of infant cry analysis, we define "reliability" as the homogeneity of all cries in an episode of crying of an infant. Therefore, reliability provides information about the reproducibility of acoustic analyses when analyzing multiple cries of an infant. This becomes especially important when aiming to automatically classify infant cries, e.g., for diagnostic purposes. Here, all cries of an infant must be similar enough for a classification model to be able to predict the same (diagnostic) result for each cry. Otherwise, for one infant its cries are classified ambiguously and the infant may not be allocated to one (diagnostic) group.

Research on infant cry reliability especially lacks of studies about which *type of crying* is the most reliable one. Cry types that have often been used in infant cry research are spontaneous cries (e.g., by Wermke et al., 2002, 2011; Manfredi et al., 2009) and pain-induced cries (e.g., by Branco et al., 2007; Runefors et al., 2000; Runefors and Arnbjörnsson, 2005). For *spontaneous cries*, the researcher has to wait until the infant starts crying without intervention. Here, the cry can have various causes like hunger, mood, desires or indisposition. Often, it is not traceable which of those might be the actual cause for crying. For *pain-induced cries*, infants are recorded when they start crying due to a pain stimulus. For ethical reasons, pain stimuli are necessarily independent of the cry analysis. Vaccinations or blood withdrawals in the context of regular screenings are often used as pain stimulus. Here, the cause of the cry can clearly be related to the pain stimulus. Although pain-induced cries being assumed to be more standardized because of the known cause, they are said to be more biased due to the high energy of the cry (Thoden and Koivisto, 1980). In contrast, spontaneous cries seem to be less standardized because of their unknown reason. Which type of crying is suited best to be used in diagnostic instruments is not answered conclusively.

In this study we analyzed the reliability of healthy infant cries for spontaneous cries and for pain-induced cries. Furthermore we generated subgroups for each of these two groups. For both sub-groups, we analyzed their reliability and compared them to the full spontaneous and pain-induced group. In this paper, we provide new insights about which acoustic parameters are consistent over multiple cries of an infant. In addition, we give lead about which type of crying is most reliable and therefore might be suited best for infant cry research.

2. Method

2.1. Subjects

In this study, 68 infants were included. 268 spontaneous cries were recorded from 35 infants (14 female, 21 male). 236 pain-induced cries (after heel prick) were recorded from 33 infants (15 female and 18 male).

All infants had no complications during birth. Their age, birth weight and gestational age were without pathological findings (Table 1). APGAR scores ("Appearance, Pulse, Grimace, Activity, Respiration", (Apgar, 1953)) were documented after 1, 5 and 10 min. For all infants, the APGAR scores were 9/10/10.

All infants were found to be healthy by paediatricians at postpartum examination. They had no indication of neurological diseases or further anomalies or any diagnosis that might influence normal development. The hearing function of all infants was assessed for both ears by otoacoustic emissions. No limitation of the hearing function was found. According to paediatricians, there was no indication for an existing cold at the time of recording for all infants.

All parents of the infants were native speakers of German and gave written informed consent to participate in this study. The study was approved by the Ethic Review Committee of the Fresenius University of Applied Sciences.

2.2. Data acquisition

Infant cries were recorded with a sampling rate of 48 kHz and 24 bit digital resolution on a Zoom H2n recorder with its built-in microphone. The microphone was held about 30 cm from the infant's mouth. Recordings were made in similar environments. For each infant, one full episode of crying was recorded. Recordings started with the first cry of the infant (using the H2n's pre-recording functionality). Recordings were stopped after the last cry of an infant when there was a 15 s pause with no crying. All episodes of crying were recorded in a supine position of the infants. One recording lasted about 10 to 30 s. For acoustic analysis, single cries were extracted from the episodes of crying.

2.3. Grouping of cries by type

Participants in this study were divided into two main groups – spontaneous cries and pain-induced cries – as those two general groups are often used in infant cry analysis. For the first group (*SP* group), recordings were started when infants began crying spontaneously. None of these cries were pain-related or triggered by any known cause. To exclude causes like sleepiness, hunger or discomfort, it was assured the infant was awake, properly fed (but not right after feeding) and had dry diapers. For the second group (*PI* group), cries were recorded during the

Table 1
Statistical parameters for the subjects ($N = 68$).

Parameters	Mean	SD	Range
Birth weight (g)	3320.85	354.51	2710–4120
Gestational age (weeks)	39.25	1.24	37–42
Age (days)	2.01	0.77	1–3

phenylketonuria screening (PKU) as part of the routine newborn screening. During the PKU, a blood sample was drawn by heel prick. If the heel was not warmed by socks before, the heel was warmed with warm water to achieve a good circulation of the blood. A Microtainer lancet was used for the heel prick. Recordings were started before the prick to ensure that the first cry after the pain stimulus was not missed.

In addition to those two general groups, we tried to extract one especially homogeneous subgroup from each of the two main groups to explore if reliability is higher when focusing on special cries within each main group.

Within the spontaneous group, a special kind of spontaneous cry was identified by acoustic analysis: the *non-distressed cry*. The characteristic of this type is a harmonic structure of the signal with a continuous contour for F0 as well as for the intensity. Both contours are without shifts and breaks (Truby and Lind, 1965). Furthermore, this type has a clear rising in intensity at the beginning and a clear falling at the end. Additionally, Lester (1976) described a reduced intensity compared to other cry types. In contrast to other subtypes of spontaneous cries, this cry type can be identified by spectral analysis and therefore is clearly recognizable. Non-distressed cries were assigned to a third group (*ND* group). This group contained non-distressed cries only, whereas the SP group contained non-distressed cries as well as other spontaneous cries.

For the pain-induced cries, a subgroup was created by removing the first cry after the pain stimulus. Runefors et al. (2000) described the first cry after a painful stimulus being different compared to the following cries. Pain-induced cries without the first cry were assigned to a fourth group (*PI w/o 1st* group).

Summarizing, four groups were defined after this grouping process: (1) the *SP* group, containing 268 spontaneous cries, (2) the *ND* subgroup, containing only spontaneous, non-distressed cries ($N = 115$), (3) the *PI* group, containing 236 pain-induced cries, (4) and the *PI w/o 1st* subgroup, containing 203 pain-cries without the first cry after the pain stimulus.

2.4. Acoustic analysis

Infant cries were analyzed on their acoustic parameters with the phonetic Software Praat 5.3.39 (Boersma and Weenink, 2013b). Included were *cries* during the expiratory phase that lasted at least 0.4 s. All valid cries were extracted from the episode of crying that was recorded for each infant. Depending on the duration of the recording, 3 to 12 cries were extracted per infant, summing up to a total of 532 cries. To guarantee sufficient quality of the recordings, only recordings with more than 30 dB intensity between the minimum intensity within an episode of crying (corresponding to the noise level) and the maximum intensity within the episode were included in the study. No recordings had to be excluded from this study.

For each cry, 19 acoustic parameters were computed with Praat software. Algorithms and settings are described in the following paragraphs.

Fundamental frequency (F0). The fundamental frequency was computed with Praat's autocorrelation algorithm (Boersma, 1993). The algorithm was parameterized to use a Gaussian window; the pitch floor was set to 100 Hz and the pitch ceiling to 1000 Hz. According to various studies, spontaneous cries as well as pain-induced cries range from 200 Hz to 600 Hz (Crowe and Zeskind, 1992; Michelsson et al., 2002; Wolff, 1969; Furlow, 1997; Sirviö and Michelsson, 1976). Porter et al. (1986) described fundamental frequencies of pain-induced cries not exceeding 730 Hz. Cries that were auditorily perceived as high-pitched were analyzed in an oscillogram to validate that the fundamental frequency was below the pitch ceiling of 1000 Hz. The frame duration was left to be selected automatically by the autocorrelation algorithm $\left(\frac{6}{\text{pitchfloor[Hz]}} = \frac{6}{100\ \text{Hz}} = 0.06\ \text{s}\right)$.

For the fundamental frequency, the median was computed as robust parameter for the central tendency of F0. The 10th percentile (P_{10}) and the 90th percentile (P_{90}) were computed as lower and upper bounds of F0. In analyses of the spectrogram, the 10th and 90th percentile proved to be more robust against outliers in F0 than minimum and maximum. The interquartile range ($IQR = P_{75} - P_{25}$) was used as measure for the dispersion of F0.

Intensity. The intensity's median, the 10th and 90th percentile and the interquartile range were computed. Constant noise levels that might have been introduced by the microphone (DC offset) were subtracted by Praat's intensity algorithm automatically. In addition, the algorithm filters pitch-synchronous intensity variations (Boersma and Weenink, 2013a).

Cry duration. Infant cries were extracted manually from the episode of crying for each infant. Boundaries of cries were identified by spectrographic analysis, based on the intensity contour as well as the waveform and the spectrum. Cry duration was then computed as the duration of the extracted cry.

Formants. The first six formants (F1–F6) as references to frequency ranges with high spectral intensities were computed with Praat's Burg algorithm (Andersen, 1974; Childers, 1978; Press et al., 2002). Here, the formants are used as a *spectral smoothing technique* rather than a modeling of the baby's vocal tract. Therefore, more than the usual first two formants were computed to model key properties of the cry signal. In addition to the fundamental frequency, we found six formants to be appropriate for approximating the cry signal sufficiently. The ceiling of the formant search range was set to 8000 Hz. For each formant, the median was computed. For the dataset, computing six formants provided sufficient spectral smoothing while keeping a reasonable spectral resolution.

Micro-variability of vocal folds. For analyzing the micro-variability of the vocal folds, Praat's Waveform-Matching algorithm (Boersma, 2009) was used and the local jitter and shimmer values were computed.

In Praat, the local jitter is defined as the mean absolute difference between the duration of succeeding periods, divided by the mean duration of periods:

$$jitter_{local} = \frac{jitter_{absolute}}{meanPeriod}$$

with

$$jitter_{absolute} = \frac{\sum_{i=2}^{N} T_i - T_{i-1}}{N-1}$$

$$meanPeriod = \frac{\sum_{i=1}^{N} T_i}{N}$$

where T_i is the duration of the ith period and N is the number of periods. A period is defined to be the smallest interval after a signal recurs. Analogously, the local shimmer is defined as the mean absolute difference between the amplitudes of succeeding periods divided by the mean amplitude of the cry. Formulas are similar to the local jitter; only the duration of a period T_i is replaced by the amplitude of a period A_i.

Even though not typical for infant cry research, jitter and shimmer were included in the set of parameters. According to Barr et al. (2000), jitter and shimmer are features of the infant cry worth to be analyzed. In contrast to older algorithms like peak picking, Praat's waveform matching algorithm is robust against additive noise (Boersma, 2009) and therefore seems to be suited for infant cry analysis.

Harmonics-to-noise ratio. For computing the harmonics-to-noise ratio (HNR), Praat uses a forward cross-correlation analysis (Boersma and Weenink, 2013a). The HNR mean and the standard deviation were computed as Praat does not support computing the median or any percentile on harmonicity values.

2.5. Statistical analysis

For statistical analysis, the software SPSS Statistics 19 (IBM, 2011) was used. To compute how consistent each of the acoustic parameters are over multiple cries of an infant, Krippendorff's Alpha was computed. To test if there are significant differences in the Krippendorff's Alpha values between the groups, a Kruskal–Wallis test was used.

2.5.1. Infant cry reliability: Krippendorff's Alpha

The homogeneity of the infant cries can be seen as a form of reliability. For each infant, an acoustic parameter, e.g. the fundamental frequency median, is computed for each of the infant's cries. Comparing those acoustic parameter values to each other allows an estimation about how reproducible the values for an infant are. Because the computation algorithms for acoustic parameters are perfectly reproducible (for identical signals, the algorithms compute always the same results), reliability estimation analyzes the reliability of the cry production itself.

Because of those considerations, an algorithm for computing inter-rater reliabilities was chosen to quantify the extent of agreement (i.e., the reliability) among the single cries. Krippendorff's Alpha (Krippendorff, 2003; Hayes and Krippendorff, 2007) is a coefficient used in content analysis to compute inter-rater reliabilities (IRR). The inter-rater reliability measures for given events (called units), how exactly multiple observers (called raters) rate the given units. If all observers give similar ratings, the IRR is high and it can be assumed that the rating results are reliable. If the observers give completely different ratings, the IRR is low and it can be assumed that the ratings are given more randomly and therefore are unreliable.

For analyzing the similarity of infant cries, Krippendorff's Alpha was adapted. For each acoustic parameter, one Krippendorff's Alpha value is computed. Here, "units" are the infants. Each cry is a "rater" for the real value of the infants' acoustic parameter. By this, Krippendorff's Alpha computes the consistency of one acoustic parameter over all cries of an infant, averaged over all infants.

To allow a better understanding of the adaption and to provide scientific transparency to the validity of this approach, algorithmic details of Krippendorff's Alpha computation are presented. For that purpose, the Krippendorff's Alpha algorithm as used for infant cry research is explained on an example. To compute the alpha coefficient, four steps are to be performed (Krippendorff, 2003):

(a) Construct the reliability matrix.
(b) Tabulate coincidence within units.
(c) Compute difference between values.
(d) Compute the α-coefficient.

In the following, the steps are described in detail.

Construct the reliability data matrix. First, the reliability data matrix is computed (Matrix (1)).

$$\begin{array}{lcccl}
\hline
\text{Infant:} & \text{I1}_{F_0} & \text{I2}_{F_0} & \text{I3}_{F_0} & \\
\hline
Cry & & & & \\
1 & 334 & 479 & 497 & \\
2 & 373 & 360 & 492 & \\
3 & 345 & 378 & . & \\
\hline
m_i & 3 & 3 & 2 & \sum_i m_i = 8 \\
\hline
\end{array} \tag{1}$$

For an infant, the acoustic parameter for each of its cries is noted; in the example, this is the median of the fundamental frequency F_0. The matrix has as many lines as cries were recorded per infant at most. In the example, 3 cries was the maximum number of recorded cries per infant. If no 3 cries were recorded for an infant, the F0 median values for the unrecorded cries are marked as missing values (marked as "·"). In the bottom line the number m_i of valid (not missing) cries for infant i is noted. On the far right in this line, the overall number of cries is summed up.

Tabulate coincidences within infants. For each *value* occurring in the reliability data matrix, the observed coincidence is computed. The observed coincidence is the probability for a value to appear together with the other values as observed in the dataset (For numeric values, the distance between two values will be computed, see Section 2.5.1).This information will later be used to determine, if the fundamental frequency values for an infant may be on chance or not.

To construct the coincidence for all possible pairs of values, a coincidence matrix is calculated (Matrix (2)).

$$
\begin{array}{c|ccccc|c}
F_0 & 1 & \cdots & w & \cdots & & \\
\hline
1 & O_{11} & \cdots & O_{1w} & \cdots & & N_1 \\
\vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\
v & O_{v1} & \cdots & O_{vw} & \cdots & & N_v = \sum_w O_{vw} \\
\vdots & \vdots & \cdots & \vdots & \vdots & & \vdots \\
\hline
 & N_1 & \cdots & N_w & \cdots & & N = \sum_{v,w} N_{vw}
\end{array}
\tag{2}
$$

The rows (v) and columns (w) in this matrix represent all F0 median values occurring in the dataset. Each entry O in the matrix at point (v, w) is computed as

$$
O_{vw} = \sum_i \frac{\text{Number of } (v,w) \text{ pairs in infant } i}{m_i - 1} \tag{3}
$$

with v and w as F0 median values and m_i as number of cries for infant i.

For the given example, the coincidence matrix shown in Matrix (4) is computed.

$$
\begin{array}{c|cccccccc|c}
F_0 & 334 & 345 & 360 & 373 & 378 & 479 & 492 & 497 & \\
\hline
334 & . & 0.5 & . & 0.5 & . & . & . & . & 1 \\
345 & 0.5 & . & . & 0.5 & . & . & . & . & 1 \\
360 & . & . & . & . & 0.5 & 0.5 & . & . & 1 \\
373 & 0.5 & 0.5 & . & . & . & . & . & . & 1 \\
378 & . & . & 0.5 & . & . & 0.5 & . & . & 1 \\
479 & . & . & 0.5 & . & 0.5 & . & . & . & 1 \\
492 & . & . & . & . & . & . & . & 1 & 1 \\
497 & . & . & . & . & . & . & 1 & . & 1 \\
\hline
 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 8
\end{array}
\tag{4}
$$

Compute distance matrix. To determine *how different* two F0 median values are, a distance function is used. Krippendorff's Alpha uses diverse distance functions according to the level of measurement of the data. It supports nominal, ordinal, interval and ratio scale for distance computation. For a comprehensive description of all distance functions see Krippendorff's book on content analysis (Krippendorff, 2003).

In this example, data are of ratio scale. The appropriate distance function between the ratio variables v and w is defined as:

$$
{}_{\text{ratio}}\delta^2_{vw} = \left(\frac{v - w}{v + w}\right)^2 \tag{5}
$$

Distances between all (v, w) pairs are calculated in a distance matrix. For the example the distance matrix is:

$$
\begin{array}{c|cccccccc}
\delta^2_{vw} & 334 & 345 & 360 & 373 & 378 & 479 & 492 & 497 \\
\hline
334 & .000 & .000 & .001 & .003 & .004 & .032 & .037 & .038 \\
345 & .000 & .000 & .000 & .002 & .002 & .026 & .031 & .033 \\
360 & .001 & .000 & .000 & .000 & .001 & .020 & .024 & .026 \\
373 & .003 & .002 & .000 & .000 & .000 & .015 & .019 & .020 \\
378 & .004 & .002 & .001 & .000 & .000 & .014 & .017 & .018 \\
479 & .032 & .026 & .020 & .015 & .014 & .000 & .000 & .000 \\
492 & .037 & .031 & .024 & .019 & .017 & .000 & .000 & .000 \\
497 & .038 & .033 & .026 & .020 & .018 & .000 & .000 & .000
\end{array}
\tag{6}
$$

Compute alpha-Coefficient. Finally, Krippendorff's Alpha coefficient is computed as the ratio between the observed disagreement D_o among infant cry parameters and the disagreement D_e one would expect when the parameters are attributable to chance instead to the properties of the cry:

$$
\alpha = 1 - \frac{D_o}{D_e} \tag{7}
$$

For ratio values, Krippendorff's Alpha coefficient is defined as:

$$
{}_{\text{ratio}}\alpha = 1 - (N - 1)\frac{\sum_v \sum_{w>v} O_{vw\,\text{ratio}}\delta^2_{vw}}{\sum_v \sum_{w>v} N_v N_{w\,\text{ratio}}\delta^2_{vw}} \tag{8}
$$

where N is the number of cries, O_{vw} is the coincidence for the pair (v, w) of the coincidence matrix (Matrix (4)), ${}_{\text{ratio}}\delta^2_{vw}$ is the distance between both items of the pair as noted in the distance matrix (Matrix (6)), N_v and N_w are the number of times the items v and w occur. For other levels of measurement, only the distance computation changes.

Inserting the corresponding values computes the alpha coefficient (the formula is abbreviated for readability reasons):

$$_{\text{ratio}}\alpha = 1 - (8-1)\frac{0.5\cdot 0 + 0.5\cdot 0.003 + \ldots + 1\cdot 0}{1\cdot 1\cdot 0 + 1\cdot 1\cdot 0.001 + \ldots + 1\cdot 1\cdot 0} = 0.634$$

We chose Krippendorff's Alpha as IRR coefficient for infant cry reliability for the following reasons. For infant cry reliability, algorithms that are able to cope with multiple units (acoustic parameter) and multiple raters (the measured value for the acoustic parameter. Each rater is one cry in an episode of crying) are required. In addition, the algorithm must be able to handle missing values (not all infants had the same number of cries in an episode of crying, so some "ratings" were missing). Finally, acoustic parameters are interval-scaled data which required the algorithm to support this level of measurement. Given those criteria, many inter-rater reliability IRR) coefficients were excluded (see Hayes' comparison of IRR coefficients for a discussion of IRR coefficient properties: Hayes and Krippendorff (2007)). Krippendorff's Alpha and the Intra-Class Coefficient (ICC: Shrout and Fleiss, 1979) were considered as appropriate algorithms. We decided to use Krippendorff's Alpha as it allows to compute inter-rater agreement for all levels of measurement while the ICC is fixed to metric data. In this paper, we only used interval-scaled acoustic parameters. However, infant cry research already explored nominal properties of cries (e.g., "is bi-phonetic or is not bi-phonetic"). So using Krippendorff's Alpha would allow us to extend reliability analysis and compare results on nominal data with the results in this paper.

Interpretation of the Krippendorff's Alpha Coefficient. For interpreting inter-rater reliability coefficients like Krippendorff's Alpha, the best known conventions are those proposed by Landis and Koch (1977). They categorize reliability coefficients into six ranges as shown in Table 2. Values less than 0.0 have a poor agreement and can be interpreted as having a great statistical spreading and being very unequal to each other. Values less than 0.2 can be interpreted as a slight agreement between the cries. Values between 0.2 and 0.4 are said to have fair agreement. A moderate agreement can be assumed at values up to 0.6. Values between 0.6 and 0.8 can be interpreted as having a substantial agreement. Values between 0.8 and 1.0 can be interpreted as a perfect agreement between the single cries.

Table 2
Interpretation of alpha coefficients according to Landis and Koch (1977).

Alpha	Interpretation
$\alpha \leqslant 0.0$	Poor
$0.0 > \alpha \geqslant 0.2$	Slight
$0.2 > \alpha \geqslant 0.4$	Fair
$0.4 > \alpha \geqslant 0.6$	Moderate
$0.6 > \alpha \geqslant 0.8$	Substantial
$0.8 > \alpha \geqslant 1.0$	Perfect

For the research field of content analysis (where Krippendorff's Alpha originates from), Krippendorff declared alpha values above 0.8 as good reliability and values above 0.667 as acceptable ones (Krippendorff, 2003). For interval-scaled data, especially in medical and language studies, alpha values higher than 0.4 are considered adequate (Artstein and Poesio, 2008; Rietveld and Hout, 1993).

In infant cry research, we assume some degree of dispersion within the cries of an infant as normal. For this reason, we propose using a relaxed interpretation of alpha values. In this study, we declared alpha values above 0.4 as acceptable reliability and alpha values above 0.667 as good reliability.

2.5.2. Differences in reliability between cry types

To identify, if the overall distribution of Krippendorff's Alpha values were significantly different between groups, a Kruskal–Wallis test was performed. In a new dataset, the group and Krippendorff's Alpha value were defined as variable. For each group, 19 Krippendorff's Alpha values from the acoustic parameters were added as items. This distribution of alpha values was then compared between groups. A non-parametric test was chosen, because a Shapiro–Wilk test revealed that the alpha values for the acoustic parameters were not normally distributed.

2.5.3. Validation of the Krippendorff's Alpha approach

To explore the validity of the Krippendorff's Alpha results, reliability of infant cries was computed with a second algorithm for inter-rater reliability; the intraclass correlation coefficient (ICC). As the ICC cannot deal with missing values, they were replaced by the group mean. For all acoustic parameters and the four groups, an intraclass correlation coefficient type $ICC(3,1)$ was computed according to Shrout and Fleiss (1979). Spearman's correlation coefficient was calculated to analyze the similarity of Krippendorff's Alpha and ICC.

3. Results

3.1. Acoustic parameters

Table 3 provides an overview about the results of the acoustic analysis.

3.2. Reliability of acoustic parameters

Krippendorff's Alpha was computed for all 19 acoustic parameters. Table 4 summarizes the results of Krippendorff's Alpha computation.

For the spontaneous cries (*SP*) 2 out of 19 acoustic parameters had good alpha values ($\alpha > 0.667$): the intensity median and the 90th percentile (P90) of intensity. In the non-distressed group (*ND*), the F0 median, F0 P90, as well as the intensity median and intensity P90 reached good reliability values. For the pain-induced cries (*PI*), good reliability was achieved for intensity median and HNR

Table 3
Mean (averaged over all cries within a group) and standard deviation of acoustic parameters over groups.

Parameter	Group			
	SP $N = 268$	ND $N = 115$	PI $N = 236$	PI w/o 1st $N = 203$
	Mean ± SD	Mean ± SD	Mean ± SD	Mean ± SD
Cry duration (ms)	1056.88 ± 525.70	855.15 ± 364.46	1103.78 ± 562.73	1021.87 ± 460.48
F0 P10 (Hz)	371.95 ± 107.55	427.93 ± 77.43	344.95 ± 111.85	346.45 ± 113.37
F0 median (Hz)	468.57 ± 87.28	473.25 ± 79.43	456.87 ± 112.35	451.38 ± 108.82
F0 IQR (Hz)	79.06 ± 79.94	39.95 ± 24.87	106.56 ± 100.42	104.89 ± 95.86
F0 P90 (Hz)	527.19 ± 93.05	501.74 ± 81.46	538.77 ± 125.93	529.88 ± 117.20
F1 median (Hz)	1288.36 ± 231.50	1213.33 ± 225.75	1221.22 ± 265.86	1210.97 ± 245.41
F2 median (Hz)	2434.31 ± 440.93	2305.13 ± 430.86	2317.51 ± 343.73	2304.43 ± 341.74
F3 median (Hz)	3635.89 ± 393.02	3525.46 ± 387.83	3563.16 ± 415.65	3549.79 ± 399.69
F4 median (Hz)	4913.86 ± 235.78	4810.99 ± 234.86	4778.42 ± 354.76	4772.86 ± 377.11
F5 median (Hz)	6220.62 ± 279.48	6213.82 ± 272.80	5964.41 ± 428.29	5966.38 ± 415.30
F6 median (Hz)	7237.30 ± 247.48	7240.01 ± 249.42	7173.16 ± 226.14	7174.18 ± 210.38
Intensity P10 (dB)	66.50 ± 5.33	66.81 ± 5.01	55.85 ± 6.89	56.34 ± 6.31
Intensity median (dB)	71.99 ± 6.09	71.72 ± 4.98	60.40 ± 6.72	60.69 ± 5.13
Intensity IQR (dB)	4.67 ± 2.49	4.27 ± 2.18	4.63 ± 2.71	4.46 ± 2.70
Intensity P90 (dB)	75.21 ± 5.32	74.57 ± 5.18	64.30 ± 6.53	64.43 ± 6.00
Jitter (local) (%)	0.68 ± 0.36	0.49 ± 1.01	1.05 ± 0.51	1.04 ± 1.01
Shimmer (local) (%)	4.75 ± 2.27	3.87 ± 2.18	9.32 ± 3.82	9.39 ± 3.88
HNR mean (dB)	14.70 ± 5.19	18.35 ± 5.11	11.80 ± 5.14	11.70 ± 3.49
HNR mean SD (dB)	5.83 ± 1.55	4.66 ± 1.37	4.94 ± 1.68	4.83 ± 1.46

Table 4
Results of Krippendorff's Alpha for the acoustic parameters grouped by the type of cry.

Parameter	Kripp. Alpha for group			
	SP	ND	PI	PI w/o 1st
Cry duration	0.368	0.385	0.337	0.379
F0 P10	0.350	0.558	0.269	0.266
F0 median	0.544	0.727	0.312	0.349
F0 IQR	0.229	0.257	0.223	0.234
F0 P90	0.631	0.708	0.406	0.489
F1 median	0.492	0.550	0.370	0.418
F2 median	0.530	0.548	0.490	0.507
F3 median	0.578	0.634	0.440	0.494
F4 median	0.392	0.488	0.339	0.336
F5 median	0.475	0.512	0.433	0.470
F6 median	0.184	0.053	0.456	0.463
Intensity P10	0.580	0.624	0.554	0.575
Intensity median	0.702	0.773	0.698	0.736
Intensity IQR	0.201	0.219	0.142	0.149
Intensity P90	0.728	0.779	0.663	0.718
Jitter (local)	0.439	0.582	0.529	0.544
Shimmer (local)	0.454	0.592	0.655	0.652
HNR mean	0.416	0.518	0.681	0.687
HNR mean SD	0.339	0.395	0.279	0.319

mean. The pain-induced cries without the first cry (*PI w/o 1st*) reached in HNR mean and intensity median as well as in intensity P90 good reliability values.

3.3. Differences in reliability between cry types

To test if one of the cry types had significantly better reliability values over all acoustic parameters, a non-parametric Kruskal–Wallis test was computed. No significant differences ($p = 0.92$) were found between the four groups when including the Krippendorff's Alpha values for all 19 acoustic parameters.

However, Krippendorff's Alpha values were visualized in a box diagram to identify trends between the groups (Fig. 1). By trend, the non-distressed cry (*ND* group) has the most reliable alpha values in 16 of 19 acoustic parameters. The acoustic parameters cry duration, all parameters of both F0 and intensity, formants 1 to 5, as well as the jitter and the HNR mean SD have their highest alpha values in the non-distressed group. The remaining three parameters F6, shimmer and HNR mean in the non-distressed group had alpha values below those in the other groups. For F6 and the HNR mean, the alpha values in the pain-induced cries without the first cry (*PI w/o 1st*) were higher. For the shimmer, the pain-induced group (*PI*), reached the highest values.

When only exploring the three acoustic parameters for which high differences in Krippendorff's Alpha between groups occurred (F0 P10, F0 median and F0 P90), significant differences can be verified between the ND group and the PI group ($p = 0.005$) as well as the ND group and the PI w/o 1st group ($p = 0.010$).

3.4. Validation of the Krippendorff's Alpha approach

Fig. 2 compares the Krippendorff's Alpha values with the intraclass correlation coefficient for all 19 parameters and all 4 groups. Spearman's correlation coefficient revealed significant, moderate correlation ($R = 0.598$, $p = 0.00$) between Krippendorff's Alpha and ICC.

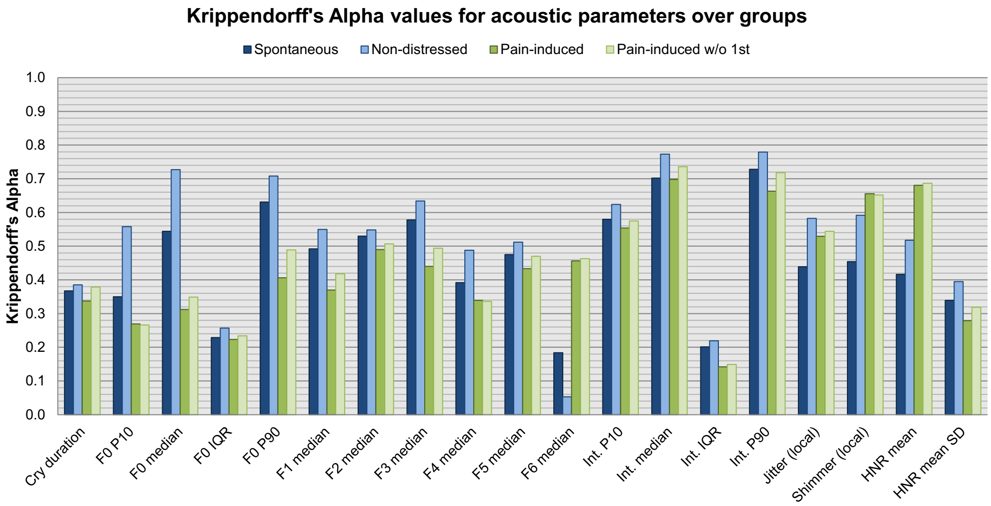


Fig. 1. Comparison of Krippendorff's Alpha values for the different cry types.

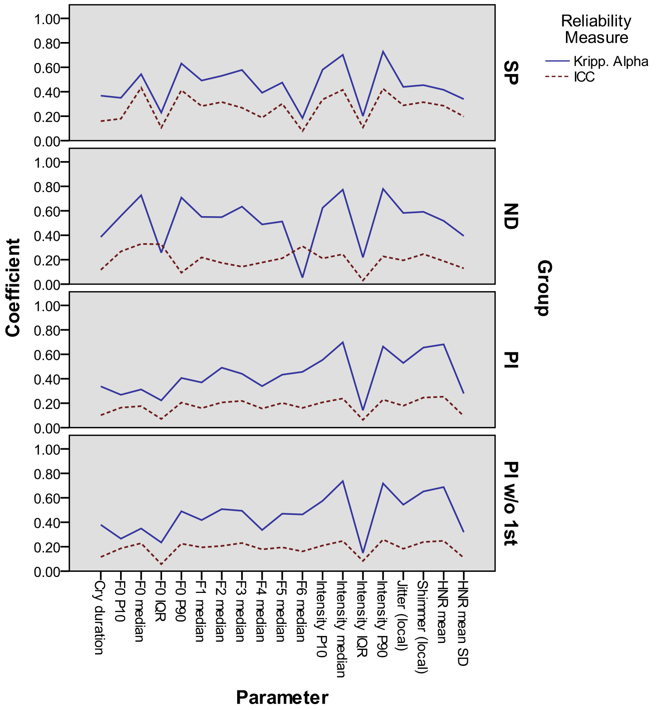


Fig. 2. Comparison of Krippendorff's Alpha and intraclass correlation coefficient for all 19 acoustic parameters over the four groups of cry types.

4. Discussion

As described in Section 3, Krippendorff's Alpha values over all groups were not very high for many of the acoustic parameters. As for all alpha coefficients, the threshold for acceptable similarity must be defined with respect to the research context and the research goals. For developing diagnostic instruments based on the infant cry, it is still to be defined which alpha values are satisfying and which are not. We decided to use a threshold of 0.4 for acceptable alpha values and a threshold of 0.667 for good alpha values. It is still to be evaluated, if those threshold values prove to be appropriate, when developing diagnostic instruments based on the infant cry.

When exploring which cry type has the most similar cries, statistically significant differences could not be found when including all parameters. However, some conclusions might be drawn by trend (Fig. 1).

First of all, comparing spontaneous and pain-induced cries, the spontaneous cries are more reliable in 15 out of 19 parameters. This seems to refute the expectation that pain-induced cries might be more reliable as the trigger of cries is more standardized than for spontaneous cries.

Considering the pain-induced cries (*PI*) and the subgroup of the pain-induced cries without the first cry (*PI w/o 1st*), the *PI w/o 1st* subgroup reached better values in 16 out of 19 parameters by trend. However, in most of the cases, the differences between both groups are only marginal. By these marginal differences, we could not confirm that excluding the first cry after the pain stimulus clearly improves reliability of pain-induced cries as was expected due to the findings of Runefors et al. (2000) about the differences between the first cry and the remaining ones.

Comparing the spontaneous cries (*SP*) to its subgroup, the non-distressed cries (*ND*), the non-distressed cries reached better alpha values in 18 out of 19 parameters by trend. This finding might reflect the clean acoustic structure of non-distressed cries in contrast to spontaneous cries in general, which often include for example non-harmonic parts. Compared to all other groups in this study, the non-distressed cry provided the highest alpha values by trend for 16 of 19 acoustic parameters stating the non-distressed cry as the preferable type of crying in infant cry analysis.

For non-distressed cries, acoustic parameters describing the fundamental frequency (F0 P10, F0 median, F0 P90) and the intensity (Intensity P10, Intensity median, Intensity P90) provide the best reliability by trend. Most of the formants (F1 to F5), the jitter, shimmer and the HNR mean have still acceptable reliability. Acoustic parameters like the cry duration, the interquartile ranges of F0 and intensity, the sixth formant and the HNR mean SD have poor reliability values and should therefore used with caution in infant cry analysis.

When including only the acoustic parameters with the highest differences between groups (F0 P10, F0 median, F0 P90), the ND group compared to the PI and PI w/o 1st group had significantly higher Krippendorff's Alpha values. Therefore, fundamental frequencies of infants seem to be more stable for non-distressed cries than for pain cries, supporting our stated preference for the non-distressed group.

Considering Krippendorff's Alpha and ICC, we interpret the results of the comparison as following: in Fig. 2, the ICC line follows the shape of the Krippendorff's Alpha line quite fairly. In addition, Spearman's coefficient revealed significant, moderate correlation between Krippendorff's Alpha and ICC. Therefore, results of the intraclass correlation are comparable to those computed by Krippendorff's Alpha. This supports the validity of our findings. However, ICC values are consistently lower than Krippendorff's Alpha. Krippendorff's Alpha can deal with missing values naturally, whereas ICC needs to replace them by group means. Therefore, Krippendorff's Alpha might be the more valid coefficient, here.

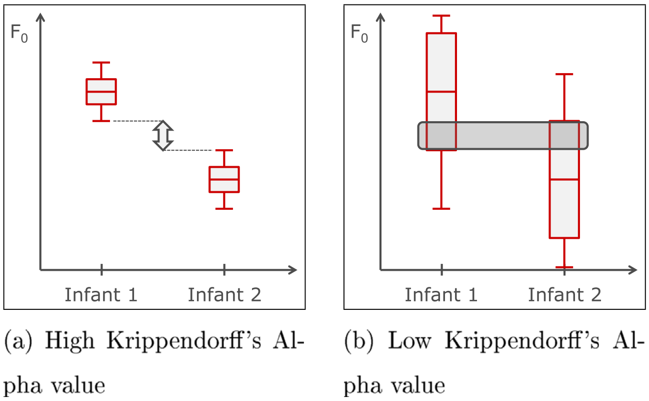
For the results of this study there are two main threats to validity:

(a) only healthy infants were included and
(b) group sizes were not equal.

Regarding (a): For exploring differences between healthy infants and infants with any kind of disorder, it is important to know which type of cry promises the most consistent acoustic parameters and therefore is better suited for infant cry analysis. Based on analyzing the healthy infant cry, we provided answers regarding which type to prefer; the non-distressed cry should be used. For analyzing differences between healthy infants and infants with disorders, the non-distressed cry is a good choice, too; at least the healthy group is known to be as consistent as possible, then. When looking for differences between groups, the similarity of cries in the other groups should be explored, too, as this provides a better understanding about how difficult it will be to find differences between the groups.

Regarding (b): Group sizes of the cry types were considered important as Krippendorff's Alpha might be influenced by group size. For smaller groups it might be easier to achieve better alpha values as less data items have to be similar. For this reason, we verified our results by drawing a randomized sample from each group in the size of the smallest group. By this procedure, we got four groups equal in size. Krippendorff's Alpha computation and comparison of groups were repeated for those randomized groups. The results were not very different from the results of the whole groups and did not lead to any other conclusions.

Summarizing, the reliability of infant cries is best for non-distressed cries. The results of this work indicate, that using the non-distressed cry may be the best choice when homogeneity of cries is required. However, non-distressed cries are still ranging from very low similarity values to acceptable ones.



(a) High Krippendorff's Alpha value (b) Low Krippendorff's Alpha value

Fig. 3. Krippendorff's Alpha influencing the variability of acoustic parameters.

When using acoustic parameters with low consistency values ($\alpha < 0.4$) for determining differences between groups, the impact of the low alpha values should be regarded. Low Krippendorff's Alpha values correlate with a greater variance within a group (Fig. 3), making it more difficult to identify differences between groups. Especially small differences between groups may get lost in high variances within groups.

As a consequence, we propose moving away from univariate statistical methods exploring differences between only one acoustic parameter (e.g. analysis of variances or their non-parametric equivalences). Instead, we recommend using multivariate techniques that explore differences between multiple parameters at the same time. Those techniques might prove more robust against such high variability within groups and might allow reliable classification of cries, though.

References

Andersen, N., 1974. On the calculation of filter coefficients for maximum entropy spectral analysis. Geophysics 39, 69–72.

Apgar, V., 1953. A proposal for a new method of evaluation of the newborn infant. Curr. Res. Anesth. Analg. 32, 260–267.

Arch-Tirado, E., Mandujano, M., Garcia-Torices, L., Martinez-Cruz, C.F., Reyes-García, C.A., Taboada-Picazo, V., 2004. Cry analysis of hypoacoustic children and normal hearing children. Cir. Cir. 72, 271–276.

Artstein, R., Poesio, M., 2008. Inter-coder agreement for computational linguistics. Comput. Linguist. 34, 555–596.

Barr, R.G., Hopkins, B., Green, J.A., 2000. Crying as a sign, a symptom, & a signal: clinical, emotional, and developmental aspects of infant and toddler crying, first ed. In: Clinics in Developmental Medicine, vol. 152 Mac Keith Press and Cambridge University Press, London.

Blinick, G., Tavolga, W.N., Antopol, W., 1971. Variations in birth cries of newborn infants from narcotic-addicted and normal mothers. Am. J. Obstet. Gynecol. 110, 948–958.

Boersma, P., 1993. Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound, in: Proceedings of the Institute of Phonetic Sciences, Amsterdam, pp. 97–110.

Boersma, P., 2009. Should Jitter be measured by peak picking or by waveform matching?. Folia Phoniatr. Logop. 61 305–308.

Boersma, P., Weenink, D., 2013a. Praat: doing phonetics by computer. Manual.

Boersma, P., Weenink, D., 2013b. Praat: doing phonetics by computer. Version 5.3.39.

Branco, A., Fekete, S.M., Rugolo, L.M., Rehder, M.I., 2007. The newborn pain cry: descriptive acoustic spectrographic analysis. Int. J. Pediatr. Otorhinolaryngol. 71, 539–546.

Childers, D.G. (Ed.), 1978. Modern Spectrum Analysis. IEEE Press, New York.

Corwin, M.J., Lester, B.M., Sepkoski, C., McLaughlin, S., Kayne, H., Golub, H.L., 1992. Effects of in utero cocaine exposure on newborn acoustical cry characteristics. Pediatrics 89, 1199–1203.

Crowe, H.P., Zeskind, P.S., 1992. Psychophysiological and perceptual responses to infant cries varying in pitch: comparison of adults with low and high scores on the child abuse potential inventory. Child Abuse Negl. 16, 19–29.

Esposito, G., Nakazawa, J., Venuti, P., Bornstein, M.H., 2013. Componential deconstruction of infant distress vocalizations via tree-based models: a study of cry in autism spectrum disorder and typical development. Res. Dev. Disabil. 34, 2717–2724.

Etz, T., Reetz, H., Wegener, C., 2012. A classification model for infant cries with hearing impairment and unilateral cleft lip and palate. Folia Phoniatr. Logop. 64, 254–261.

Fisichelli, V.R., Karelitz, S., 1966. Frequency spectra of the cries of normal infants and those with Down's syndrome. Psychon. Sci. 6, 195–196.

Fort, A., Manfredi, C., 1998. Acoustic analysis of newborn infant cry signals. Med. Eng. Phys. 20, 432–442.

Furlow, F.B., 1997. Human neonatal cry quality as an honest signal of fitness. Evol. Human Behav. 18, 175–193.

Golub, H.L., Corwin, M.J., 1982. Infant cry: a clue to diagnosis. Pediatrics 69, 197–201.

Grau, S.M., Robb, M.P., Cacace, A.T., 1995. Acoustic correlates of inspiratory phonation during infant cry. J. Speech Hear Res. 38, 373–381.

Green, J.A., Gustafson, G.E., McGhie, A.C., 1998. Changes in infants' cries as a function of time in a cry bout. Child Dev. 69, 271–279.

Hayes, A.F., Krippendorff, K., 2007. Answering the call for a standard reliability measure for coding data. Commun. Methods Meas. 1, 77–89.

IBM, 2011. SPSS Statistics Software. Version 19.0.0.1.

Karelitz, S., Fisichelli, V.R., 1962. The cry thresholds of normal infants and those with brain damage. Disabil. Rehabil. 61, 679–684.

Krippendorff, K., 2003. Content Analysis: An Introduction to Its Methodology, second ed. Sage Publications, Thousand Oaks.

Landis, J.R., Koch, G.G., 1977. The measurement of observer agreement for categorical data. Biometrics 33, 159–174.

Lester, B.M., 1976. Spectrum analysis of the cry sounds of well-nourished and malnourished infants. Child Dev. 47, 237–241.

Lester, B.M., Tronick, E.Z., LaGasse, L.L., Seifer, R., Bauer, C.R., Shankaran, S., Bada, H.S., Wright, L.L., Smeriglio, V.L., Lu, J., Finnegan, L.P., Maza, P.L., 2002. The maternal lifestyle study: effects of substance exposure during pregnancy on neurodevelopmental outcome in 1-month-old infants. Pediatrics 110, 1182–1192.

Lind, K., Wermke, K., 2002. Development of the vocal fundamental frequency of spontaneous cries during the first 3 months. Int. J. Pediatr. Otorhinolaryngol. 64, 97–104.

Lind, J., Wasz-Höckert, O., Rosberg, G., Theorell, K., Valanne, E.H., Partanen, T.J., Vuorenkoski, V., 1967. Sound spectrography in pediatric diagnosis. Acta Paediatr. Scand. 177, 113–119.

Lind, J., Vuorenkoski, V., Rosberg, G., 1970. Spectographic analysis of vocal response to pain stimuli in infants with Down's syndrome. Dev. Med. Child Neurol. 12, 478–486.

Manfredi, C., Bocchi, L., Orlandi, S., Spaccaterra, L., Donzelli, G.P., 2009. High-resolution cry analysis in preterm newborn infants. Med. Eng. Phys. 31, 528–532.

Michelsson, K., Sirviö, P., Wasz-Höckert, O., 1977. Sound spectrographic cry analysis of infants with bacterial meningitis. Dev. Med. Child Neurol. 19, 309–315.

Michelsson, K., Eklund, K., Leppänen, P., Lyytinen, H., 2002. Cry Characteristics of 172 healthy 1- to 7-day-old infants. Folia Phoniatr. Logop. 54, 190–200.

Möller, S., Schönweiler, R., 1999. Analysis of infant cries for the early detection of hearing impairment. Speech Commun. 28, 175–193.

Nugent, J.K., Lester, B.M., Greene, S.M., Wieczorek-Deering, D., O'Mahony, P., 1996. The effects of maternal alcohol consumption and cigarette smoking during pregnancy on acoustic cry analysis. Child Dev. 67, 1806–1815.

Porter, F.L., Miller, R.H., Marshall, R.E., 1986. Neonatal pain cries: effect of circumcision on acoustic features and perceived urgency. Child Dev. 57, 790–802.

Press, W.H., Flannery, B.P., Teukolsky, S.A., Vetterling, W.T., 2002. Numerical Recipes in C: The Art of Scientific Computing, 2 ed. Cambridge University Press, Cambridge.

Rietveld, T., Hout, R.v., 1993. Statistical Techniques for the Study of Language and Language Behaviour. de Gruyter, Berlin.

Robb, M.P., Goberman, A.M., Cacace, A.T., 1997. An acoustic template of newborn infant crying. Folia Phoniatr. Logop. 49, 35–41.

Runefors, P., Arnbjörnsson, E., 2005. A sound spectrogram analysis of children's crying after painful stimuli during the first year of life. Folia Phoniatr. Logop. 57, 90–95.

Runefors, P., Arnbjörnsson, E., Elander, G., Michelsson, K., 2000. Newborn infants' cry after heel-prick: analysis with sound spectrogram. Acta Paediatr. 89, 68–72.

Sheinkopf, S.J., Iverson, J.M., Rinaldi, M.L., Lester, B.M., 2012. Atypical cry acoustics in 6-month-old infants at risk for autism spectrum disorder. Autism Res. 5, 331–339.

Shrout, P.E., Fleiss, J.L., 1979. Intraclass correlations: uses in assessing rater reliability. Psychol. Bull. 86, 420–428.

Sirviö, P., Michelsson, K., 1976. Sound-spectrographic cry analysis of normal and abnormal newborn infants: a review and a recommendation for standardization of the cry characteristics. Folia Phoniatr. 28, 161–173.

Thoden, C.J., Koivisto, M., 1980. Acoustic analysis of the normal pain cry. In: Murry, T., Murry, J. (Eds.), Infant Communication: Cry and Early Speech. College-Hill Press, Houston, pp. 124–151.

Thoden, C.J., Michelsson, K., 1979. Sound spectrographic cry analysis in Krabbe's disease. Dev. Med. Child Neurol. 21, 400–401.

Truby, H.M., Lind, J., 1965. Cry sounds of the newborn infant. Acta Paediatr. 54, 8–59.

Várallyay, G.J., 2007. The melody of crying. Int. J. Pediatr. Otorhinolaryngol. 71, 1699–1708.

Verduzco-Mendoza, A., Arch-Tirado, E., Reyes-Garcia, C.A., Leybon-Ibarra, J., Licona-Bonilla, J., 2012. Spectrographic cry analysis in newborns with profound hearing loss and perinatal high-risk newborns. Cir. Cir. 80, 3–10.

Vuorenkoski, V., Lind, J., Partanen, T.J., 1966. Spectrographic analysis of cries from children with maladie du cri du chat. Ann. Paediatr. Fenn. 12, 174–180.

Wasz-Höckert, O., Lind, J., Vuorenkoski, V., Partanen, T.J., Valanne, E.H., 1968. The infant cry: a spectrographic and auditory analysis, first ed. In: Clinics in Developmental Medicine, vol. 29 Cambridge University Press.

Wermke, K., Mende, W., Manfredi, C., Bruscaglioni, P., 2002. Developmental aspects of infant's cry melody and formants. Med. Eng. Phys. 24, 501–514.

Wermke, K., Birr, M., Voelter, C., Shehata-Dieler, W., Jurkutat, A., Wermke, P., Stellzig-Eisenhauer, A., 2011. Cry melody in 2-month-old infants with and without C lefts. Cleft Palate Craniofac. J. 48, 321–330.

Wolff, P.H., 1969. The natural history of crying and other vocalizations in early infancy. In: Foss, B.M. (Ed.), Determinants of Infant Behavior. Methuen, pp. 81–109.