Data Collection of Infant Cries for Research and Analysis

Anshu Chittora and Hemant A. Patil, *Gandhinagar, Gujarat, India*

Summary: Analysis of infants cries may help in identifying the needs of infants such as hunger, pain, sickness, etc and thereby develop a tool or possible mobile application that can help the parents in monitoring the needs of their infant. Analysis of cries of infants who are suffering from neurologic disorders and severe diseases, which can later on result in motor and mental handicap, may prove helpful in early diagnosis of pathologies and protect infants from such disorders. The development of an infant cry corpus is necessary for the analysis of infant cries and for the development of infant cry tools. Infant cry database is not available commercially for research, which limits the scope of research in this area. Because the cry characteristics changes with many factors such as reason for crying, infant's health and weight, age, etc, care is required while designing a corpus for a particular research application of infant cry analysis and classification. In this paper, the ideal characteristics of the corpus are proposed along with factors influencing infant cry characteristics, and experiences during data collection are shared. This study may help other researchers to build an infant cry corpus for their specific problem of study. Justification of the proposed characteristics is also given along with suitable examples.

Key Words: Infant cry–Spectrographic analysis–Cry modes–Cepstrum analysis–Database.

INTRODUCTION

Infant cry signal carries several levels of information, for example, reason for crying (hunger/pain/colic pain/discomfort due to wet diaper/loneliness), integration of vocal production system, maturity of respiratory system, language acquisition, and maturity of central nervous system (CNS). Numerous research has been done in the infant cry analysis from the viewpoint of parental perception of infant cry, developmental aspects of infants from their cry analysis, effects of various factors on infant cry patterning, etc.[1–5] In the medical domain, cry is studied to find out the acoustic features underlying the cry, to understand the possible causes of cry, and the effects of causes apparent in the cry. In the analysis of pain cries of newborns, it is observed that newborns modulate the supralaryngeal tract considerably following a painful stimulus than they would in spontaneous cries.[6] In invasive pain cry stimulus, the cry produced is rated as urgent by parents, and acoustic features such as high fundamental frequency (F_0), longer crying bouts, fewer harmonics, and greater variability of the F_0 are found.[7] In newborns, dysphonic cries show anger.[8] Along with it, newborn cries do not show much difference in the highest and lowest F_0, and there are no differences in the cry acoustics according to gender.[9] Gender-specific differences in F_0 and formant frequency patterns appear at the age of 11.[10] Cry has also been studied for the parents' perception of infant cry. It is shown in the past that mothers can recognize their infants from their cries. However, recent studies show that fathers are as good as mothers in recognizing their newborn's cries, and it depends on the time spent with the infant by parents.[11] The perception of cry from parents is reflected in the parenting depending upon how parents perceive the underlying message sent through the cry signal. Cry draws the attention of the parents when it is different. Parents' response to the infants' cry depends on the cry acoustics (F_0, changes in F_0 [ie, intonation pattern], loudness, duration, rhythm, etc.), psychological makeup, and living conditions of the parents. Negative response or no response to infant cries may result in child abuse.[12] Another study[13] reported that the synchrony of arousal between infant and caregiver results in changes in the neurobehavioral mechanisms, and the changes in the intensity of arousal are reflected in graded and dynamic acoustic signal. Deviations in the cry signal are noticed by the parents, and misunderstanding these deviations may compromise infant care and parental effectiveness.[14] Moreover, in the same study, it is reported that the infants with abnormal cries should be referred for full neurologic evaluation. Infant crying in the early 3 months of age is a signal of vigor that is evolved to reestablishment of parental contact.[15]

Research in infant cry analysis using signal processing methods such as spectrograms was started by a team of Scandinavian researchers in 1960 (Lester, et al.). In the initial two decades, the cry analysis used mainly the spectrographic analysis.[16] The infant cry signal has been studied mostly by the researchers from the medical field and those who study the development of infants. In the medical field, infant cry has been analyzed for identifying the features particular to a pathology. Most of the work done in this direction is the classification of normal and deaf infants, asphyxiated infants, and those who suffer from heart-related diseases. Another discipline where infant cries are studied is the behavioral sciences.[17] In this field, the cry pattern is analyzed to see the effect of parents and infants relationship under various conditions such as parents' literacy, response of parents to infants' cries, possible use of nicotine during pregnancy, etc.[18] Linguists also use the cry signal analysis for studying language acquisition by the infants and studying the effects of multilingual atmosphere on child's learning.[19] Interestingly, language acquisition also plays an important role in designing of speech

Accepted for publication July 13, 2016.
From the Dhirubhai Ambani Institute of Information and Communication Technology, Near Indroda Circle, Infocity, Gandhinagar, Gujarat, India.
Address correspondence and reprint requests to Anshu Chittora, Dhirubhai Ambani Institute of Information and Communication Technology, Near Indroda Circle, Infocity, Gandhinagar, Gujarat, India. E-mail: anshu_chittora@daiict.ac.in
Journal of Voice, Vol. 31, No. 2, pp. 252.e15–252.e26
0892-1997

http://dx.doi.org/10.1016/j.jvoice.2016.07.007

robots. Furthermore, contribution is also made by the engineering field in this area.

Nowadays, engineering professionals are showing interest to develop an infant cry analyzer, signal processing algorithms, and classification of pathologic infant cries from normal infant cries. In this direction, work has been done toward classification and analysis of infant cry types (hunger, pain, and pleasure cries), and classification of pathologic cries where different set of pathologies were considered by different researchers.[20–25] However, there is little work performed in the field of infant cry analysis and classification.

Infant cry analysis is also important for the identification of the infants who are susceptible to sudden infant death syndrome (SIDS). SIDS is the condition of infant death where the reason of death remains unanswered even after a thorough medical examination and autopsy. Sudden unexpected death in infancy (SUDI) or sudden unexpected infant deaths (SUID) are also used as synonyms to SIDS. Distinction between SIDS and SUID is generally difficult. Most of the SIDS deaths occur during 2–6 months of age. The chance of deaths due to SIDS reduces after 1 year of age. It has been reported by parents of the SIDS victims that the cries of their infants are strange or different from their siblings and other normal infants. Stark and Nathanson studied the cries of a male infant who died at the age of 6 months. They found that compared with normal infants, the infant cries were shorter and weaker. Colton and Steinschneider reported the cry characteristics of a female infant who died at the age of 63 days and reported that the F_0 was lower, duration was longer, and sound pressure level was higher than the normal infants group and SIDS siblings group.[26] Infants who die of SIDS have abnormalities in the brain stem, which helps to control functions like breathing, blood pressure, and arousal, and abnormalities in serotonin signaling.[27]

The main challenge in the infant cry analysis and classification is the unavailability of a statistically meaningful database (or at least sufficiently large number of infant cry samples). Collecting a database requires permissions from the hospital authorities and parents as well. Getting pathologic cry signals of infants who are suffering from some disease is furthermore difficult. Getting a statistically significant data is a challenging task. Most of the researchers working in this area have their own database with different sets of infant cry types, recording conditions, age groups, different pathologies, and different weights of infants. A standard database for the task is not available, which also restricts the comparison of different research works. Cry signal characteristics change with several factors such as reason for crying, age of the infant, etc. For example, birth cry cannot be mixed with other cry types because it is a response of the infant to exposure to air atmosphere from fluid atmosphere. It is an indicator of proper functioning of the infant's respiratory system. Therefore, the first cry has distinct F_0 which is different from the rate of vibration of vocal folds due to poor neural control of CNS over vocal system. However, there is turbulent noise present due to nonlinear interaction of the vocal excitation source and system. In pathologic cry analysis, cry characteristics change with the severity of the disease. In such cases, long-term follow-up of the infant is required, which is a difficult task. All these effects altogether pose challenge to the researchers to work in this area and contribute toward it. In this paper, the effects of these factors on the infant cry analysis are given, and a guide for data collection is suggested so that the researchers who are interested to work in this area can collect their data accordingly and design their corpus which can be used potentially in their research work. This study will help in possible standardization of infant cry corpus preparation. The paper is organized as follows: Section 2 describes the ethical issues in infant cry data collection and protection of human rights. Section 3 gives details of metadata preparation, and ideal characteristics of infant cry corpus are listed in Section 4. Experiences during data collection and details of database collected are reported in Section 5 and Section 6, respectively. Illustration of several factors influencing the cry characteristics are given in Section 7, and the paper is summarized in Section 8.

ETHICAL ISSUES IN INFANT CRY DATA COLLECTION AND PROTECTION OF HUMAN RIGHTS

Infant cry recording is a very sensitive task that involves dealing with newborn human participants. Researchers have considered pain cry and hunger cry for infant cry analysis. To record the pain cry, how stimulation should be given to the infants has been an issue of debate for a long time. In the medical domain, pain cry is recorded by giving a rubber snap on an infant's foot. As long as this is done by a medical practitioner, parents generally do not object. However, for others, they do not allow it to be performed on their child. Even hospital authorities also do not allow using such practice for data collection purpose. Data collection of pain cry is possible only during *immunization process* or during treatment if injection is given. There are several advantages of data collection of pain cries during the immunization process, such as the following: the amount of stimulation is controlled (which is not the case in other methods), cries collected are from the same age groups as for a particular vaccine, specific age is defined, and generally, in immunization units, silence is maintained. Recording of hunger cry is even more difficult. For data collection of hunger cry, the only option is to wait until the infant gets hungry. Another important issue is to convince the parents for the purpose of data collection and getting their approval. For normal infants, this part is comparatively easy. However, for infants who are sick, especially those who have severe pathologies, this task becomes very difficult. In these cases, emotional status of the parents should be considered, and all their decisions, even if these are not positive for data collection, should be respected without hurting their sentiments.

To look into these sensitive issues such as the method of cry stimulation and protection of human rights, the Institutional Ethics Committee (IEC) is formed by hospital authorities. The objective of the IEC is to ensure a competent and consistent ethical review mechanism for health and biomedical research proposals dealt by the committee as prescribed by the ethical guidelines for biomedical research on human participants (Indian Council of Medical Research[28] in India or Council for International Organizations of Medical Sciences and World Health Organization guidelines[29] or respective country's guidelines). The ethical committee ensures that the procedures used for scientific research on human participants do not harm the participants under study. It looks into the feasibility of the research area and the methods used in the study for research and its application for future use. The composition of IEC is multidisciplinary, and it includes experts from several disciplines such as medical, legal, and social welfare area, lay persons, and clinicians from different institutes to ensure independence in the composition of the committee.

The research proposal for infant cry research and data collection should be reviewed by the IEC, and the consent for data collection needs to be taken before collection of infant cries. During data collection, general principles in biomedical research involving human participants should be followed as mentioned in the "Ethical Guidelines for Biomedical Research on Human Participants" described by the Indian Council of Medical Research, New Delhi, or respective hospitals' guidelines (for other countries). Based on these recommendations, few important guidelines for the infant cry data collection are as follows:

(a) The participation of infants must be voluntary in nature. No one can be forced to be a part of this study.
(b) Before participation in the data collection, parents must be informed about the purpose of the study and the method of data collection. Written consent must be obtained from the parents (because the participants are minors in this study).
(c) The data and associated metadata of the parents and infants who have participated in the study should be kept confidential.
(d) Data can be collected using a handheld recording instrument to ensure minimum risk or non-intrusion to the infant.
(e) Data should be collected by the researcher to avoid chances of mishandling of the data and privacy of the participants.
(f) Data should be collected in the presence of medical practitioners to avoid any possible harm due to accidents and get their feedback in the research work for normal vs pathologic cases.

METADATA PREPARATION

Before collection of the infant cry data, parent consent form and participant information form must be prepared with due care. The parent's consent form must indicate the purpose of study, method of data collection, terms of compensation in case of any injury that occurred during data collection, and privacy conditions. For the preparation of the subject (infant) information form, the purpose of the study is very important. For example, in normal healthy infants, the cry acoustics are dependent on the reason for crying. However, cry acoustics are dependent on the pathology and its severity in pathologic cry analysis. Because it is difficult to collect data from infants, it is advisable to get as much information about the infant in the participant information form as possible, so that the database can be used for several purposes. The details required in the participant information form are name of infant, date of birth, weight at the time of birth, gestation age (GA) at the time of birth, current weight, reason for crying, gender, details of siblings, history of any sibling deaths, parent's name, any genetic disease to either of the parent, educational qualification (in some studies of child development it has been observed that infant's language development is directly related to the mother's educational qualification[30]), and date of recording with *.*wave* file name. In case of sick infants, details of disease and doctors comments about the severity of disease are necessary. A template for metadata preparation (Subject Information form) is shown in Appendix I.

IDEAL CHARACTERISTICS OF INFANT CRY CORPUS

The desirable characteristics of the infant cry database are as follows:

(a) The reason for crying should be the same in a particular class for infant cry classification task.
(b) The age group under consideration should be the same.
(c) The weights of the infant should not vary too much.
(d) In the study of premature infants, GA is important to consider for infant cry analysis.
(e) During data collection of pain cries, pain stimulation must be the same for all infants.
(f) The number of cries as well as the number of infants should be significantly high to give statistical meaning to the findings or results.
(g) The cry utterance should be long for the study of reasons for crying.
(h) Sampling frequency (F_s) should be kept high.
(i) In neonatal infants (0–1 month), cry characteristics change very frequently because of rapid development of the respiratory system and its coordination with CNS. This group can be considered separately.
(j) Infants having less than 3 months of age are obligate nose breathers; this group of infants can be dealt in different class from their older infants. After the age of 3 months, infants have separate paths for food and air (for breathing), and hence, it becomes more like an adult system. Thus, they form a separate group in the cry analysis.
(k) In pathologic cry analysis, the number of participants should be high to give better understanding of the reason for crying. Collecting more cry samples from very few infants (participants) may not give statistically significant or reliable results.
(l) In the study of pathologic cries, the reason for crying should be the same. Reason for crying will give changes in acoustic features in addition to the changes in acoustic features due to presence of pathology.
(m) For infants who have history of sibling's deaths, their cries can be analyzed separately. Furthermore, such infants are considered as high-risk infants in the study of SIDS.

EXPERIENCES DURING DATA COLLECTION

Some of the experiences during data collection are as follows:

(a) On the day of recording, concerned doctors and concerned staff were also informed about the purpose of study, and they were very cooperative during recording of cry samples and in describing the pathology of the infants. They tried to maintain silence as much as possible.

(b) Sometimes, it was difficult to convince the parents to give consent for recording. However, no one was forced to give permission for cry recording; rather, recording was done for only those infants whose parents voluntarily agreed for this task.

(c) Because of the presence of four to six doctors and several infants and their mothers in the outdoor patient department room, the noise-level was sometimes very high. Such cases were ignored in the analysis.

(d) At some time during recording of an infant's cry, cries of other infants present in the same outdoor patient department was dominating. It made the recording task very difficult. Such cases were excluded while recording. In case of too much noise, it was better to stop recording the infant cries. Instrument settings can be adjusted, and it should be placed near the infant's mouth to minimize the surrounding noise effects to a certain extent.

(e) Sometimes as soon as the voice recorder was placed near the infant for recording of cry, the infant kept quiet and was found to start observing the recording instrument. Cries from such a scenario were also discarded from the analysis as transition in infant's behavior changes the prosodic content of the cry.

(f) Almost all the infant cries were spontaneous. In acquiring a cry sample, *none* of the infant was given any external stimulation for data collection purpose.

DATABASE COLLECTED

In this work, to show the results and illustrations, we have used two databases which were collected from various hospitals of India. While cries were recorded, doctor's comments were also recorded. Database collected by Buddha and Patil[31] is referred to as *Corpus I* and the other database collected by Chittora and Patil[32] is referred to as *Corpus II* in this paper. *Corpus I* was collected with a portable Cenix digital recorder (Cenix Degicom Co. Ltd., Gyeonggi-do, Korea) with an external microphone. The sampling frequency of the recording was 12 kHz, and it was quantized at 16-bits pulse code modulation (PCM). *Corpus II* was collected with a portable handheld, battery operated recorder (Zoom H4n, Zoom North America, Hauppauge, NY) at a sampling frequency of 44.1 kHz with 16bits quantization in stereo recording mode. It has adjustable sampling frequency up to 96 kHz and quantization up to 24-bits and USB audio interface. The corpus statistics are shown in Tables 1–4. In this paper, infant cry data refer to the infant cry recordings. However, the database is defined as the collection of all these recordings.

TABLE 1.
Distribution of Infants in Database

S. No.	Type	Corpus I	Corpus II
1.	Normal infants	93	61
2.	Pathologic infants	56	26

TABLE 2.
Distribution of Samples Over Gender

S. No.	Gender	Corpus I	Corpus II
1.	Boy	114	38
2.	Girl	60	49

Note: Sometimes more than one cry samples are collected from a single infant.

FACTORS INFLUENCING CRY PERFORMANCE

The ideal and desirable characteristics of the infant cry corpus are mentioned in Section 4. In this section, justification of these characteristics is given with suitable examples using earlier studies on infant cry analysis and narrowband spectrograms that represent the joint time-frequency energy density of infant cry signal. Spectrographic analysis has been used in infant cry analysis by other researchers.[27,33] Several factors which influence the cry characteristics need to be considered while collecting the infant cry signal for corpus preparation. Some of these factors are discussed in the following sections.

Variability of acoustic features with age (newborn to 1 year)

An important factor to be considered in the analysis of infant cry is variability of acoustic features with age. The effect of acoustic features is shown in Figures 1–3 using *Corpus II* cry samples. In these figures, the spectrograms of the infant cry signals of different ages, ie, 2 months, 6 months, and 12 months old, are shown. All the cry signals are from normal infants and the reason for crying (pain) is the same. From the spectrograms, it is observed that as the infant grows older, the cry duration increases. In younger infants, the pauses between the cries are longer and

TABLE 3.
Distribution of Samples Over Age for *Corpus I*

Age	Number of Samples
1 d	23
<1 wk	38
>1 wk and <1 mo	31
1–6 mo	37
6 mo to 1 y	22
1–1.5 y	13

Abbreviations: d, days; mo, months; y, years.

TABLE 4.
Distribution of Samples Over Age for *Corpus II*

Age	Total
1–7 d	6
<1 mo	7
1–4 mo	37
4–8 mo	20
8 mo to 1 y	14
1–2 y	3

Abbreviations: d, days; mo, months; y, years.

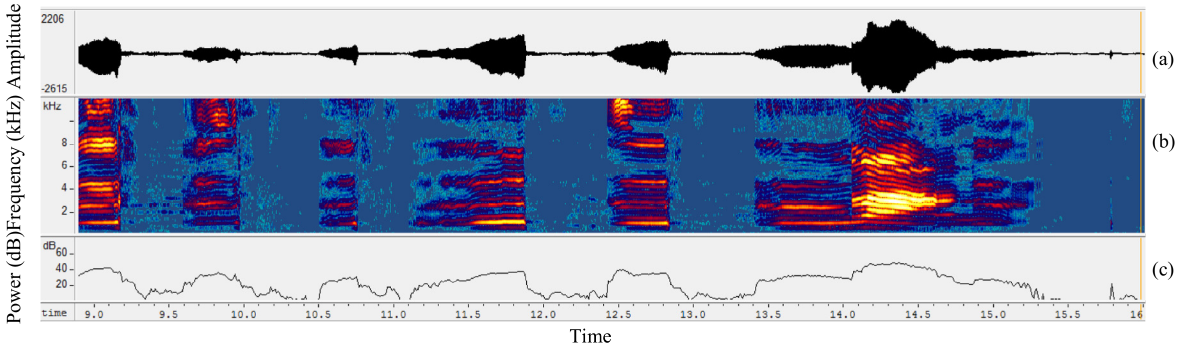

FIGURE 1. Spectrogram of a 2-month-old infant's cry sample. (a) Time-domain signal, (b) Narrowband spectrogram, and (c) Power plot of **A**.

cries are of short duration. As the infant grows, he or she learns to control the respiratory movements and has better neural control. Because of this, the cry duration becomes longer. In neonates, the respiration rate is 40 rpm (respirations per minute), which reduces to 30 rpm in infants aged 12 months, whereas in adults, respiration rate is 20 rpm. Another important observation is that the power in the cry signal increases with growing age. It also shows the improvement in the muscular strength of the voice production system. It has already been reported by researchers that with the growing age of the infant, pitch and formants of the cry signal decrease because of increase in vocal tract length.

From Figure 4, it is observed that with age the respiratory control of the infant becomes better, and it results in longer duration of cries. An increase in age also results in higher energy in the cries. Along with this, instead of having abrupt changes in the energy at the inspiratory durations, the energy transitions become smoother (compare newborn and 20th day cry signal energies). Moreover, we can observe a vibration pattern in F_0 contour compared with only a rising pattern that is visible in the first-day cry.

All these changes in the cry pattern cause difficulties in the analysis of the infant cry. However, in case of adults, the change in the F_0, formants, and duration features for the same pronunciation are comparatively small (because of well-defined rules for the pronunciation of a particular phoneme).

The effect of the development of speech production system has been observed in our analysis of an infant whose cries were recorded for 20 days from the day of birth. The results of analysis are shown in Table 5, where the mean fundamental frequency (mean F_0), maximum fundamental frequency (max F_0), and unvoicing ratio (VUV ratio, which is the ratio of number of

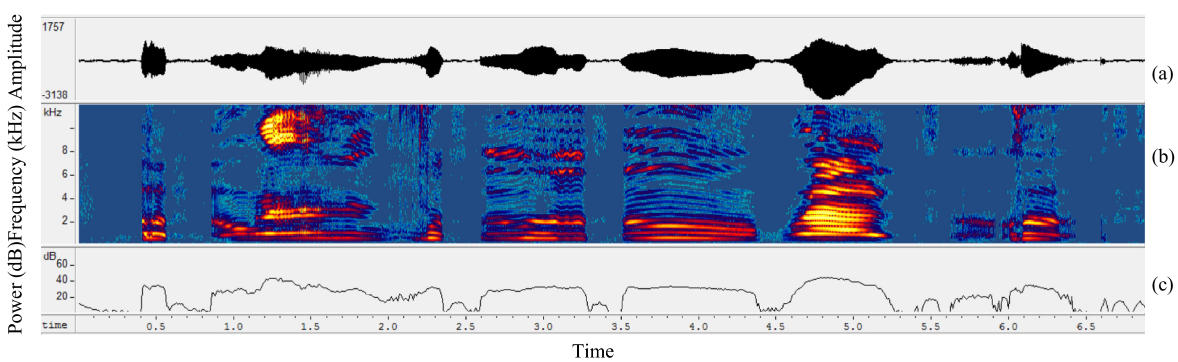

FIGURE 2. Spectrogram of a 6-month-old infant's cry (a) Time-domain signal, (b) Narrowband spectrogram, and (c) Power plot of **A**.

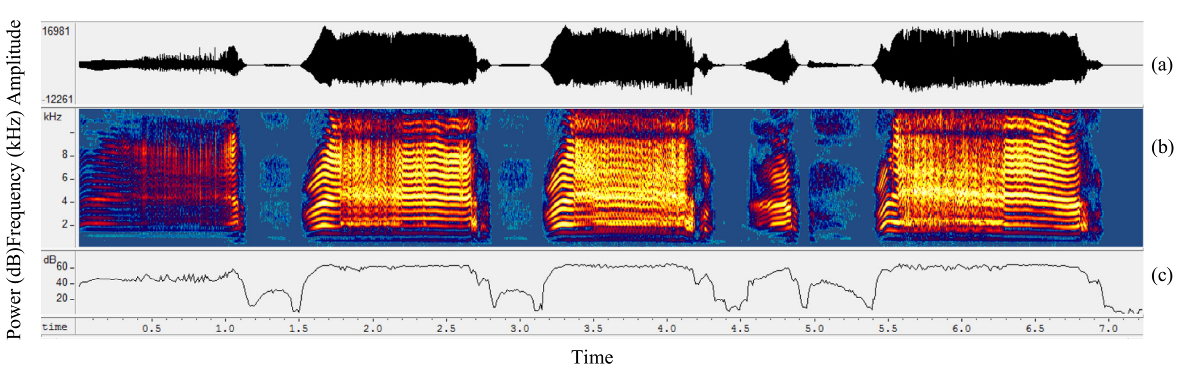

FIGURE 3. Spectrogram of a 12-month-old infant's cry sample (a) Time-domain signal, (b) Narrowband spectrogram, and (c) Power plot of **A**.

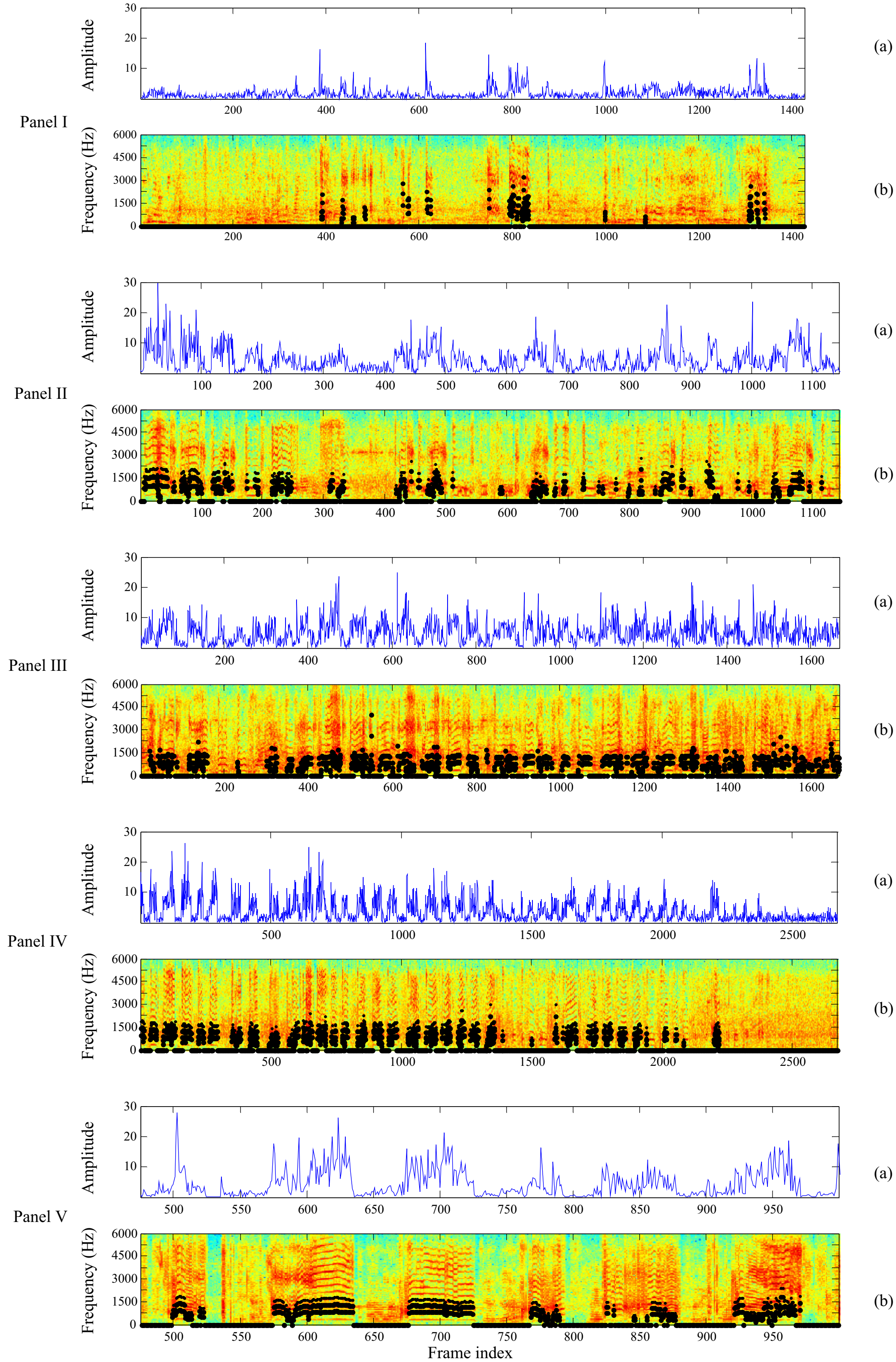

FIGURE 4. Variation of acoustic features in infant's cry with age. Panel I: birth cry, Panel II: day 1 cry, Panel III: day 4 cry, Panel IV: day 11 cry, and Panel V: day 20 cry. In all the subfigures (a) represents short-time energy of the cry signal, where *x*-axis is the frame index and *y*-axis is the signal energy, and (b) shows the spectrogram and fundamental frequency contour of the cry superimposed on the spectrogram of the cry where the *x*-axis is the frame index and the *y*-axis is the frequency in hertz. All the samples are taken from *Corpus I*.

TABLE 5.
Variation of Acoustic Features with Age on *Corpus I*

Age in Days	Mean F_0	Max F_0	VUV Ratio
0	428.57	800	0.9552
1	393.44	857.14	0.679
3	400	1333	0.5913
10	406.6	750	0.71
15	406.78	1090.9	0.64
20	400	857.14	0.62

unvoiced frames to total number of frames in a cry) of the cry is calculated from the cry recording. Ratio of unvoicing in the cry is also calculated to show the development of the vocal folds.

From Table 5, it can be observed that the mean F_0 of the infant cry reduces with the development of the vocal system. However, with the maturity of neural control system and speech production system, the infant learns to modulate the cry and can increase the F_0 or pitch up to 1 kHz. These high-pitched cries are used by the infants to draw the attention of the caretaker in case of emergency. The voicing content in the cry also increases with age, and the cries become rhythmic.

Variability of acoustic features with weight of the infants

Cry characteristics in infants with different weights follow the same trends as that of age. Infants with higher weight in the respective age group show dominant prosodic marks in their cry. From the 10th day after birth, when the post-birth weight loss is usually regained, there is a steady increase in weight so that during the first 3 months, an average baby gains about 2 lb per month, or nearly 1 oz per day. At 5 months, the birth weight is doubled. Beginning at 6 months, there is only a 1-lb gain per month in weight so that the birth weight is tripled at the end of the first year and quadrupled at the end of the second year.[34]

Figures 5 and 6 show the variability in the acoustic features for the infant cries of two infants with almost the same age but with different weights (samples from Corpus II are used). When the spectrograms of the low weight and high weight infants are compared, it can be observed that for low birth weight infant, the duration of cry is small, and it has poor control on the respiratory system. F_0 is also high in the low birth weight infant. In premature low birth weight infants, the pitch is found to be higher than the normal full-term infants. The power content and prosodic variations are high in the high birth weight infant. Improvement in the weight of the infant is correlated with development of neural, muscular, and anatomical structures. It is also an indication of the integrity of the various anatomical structures with the neural system. In adults, because the voice production system is fully developed, the variations in weight of the participants do not change the results of speech analysis or recognition. However, recently, an attempt has been made to estimate the height of a person from the state-of-the-art spectral features, namely, Mel frequency cepstral coefficients.[35]

Cry type differences

The cry characteristics also change with the reason for crying. The researchers working on the infant cry analysis divided the

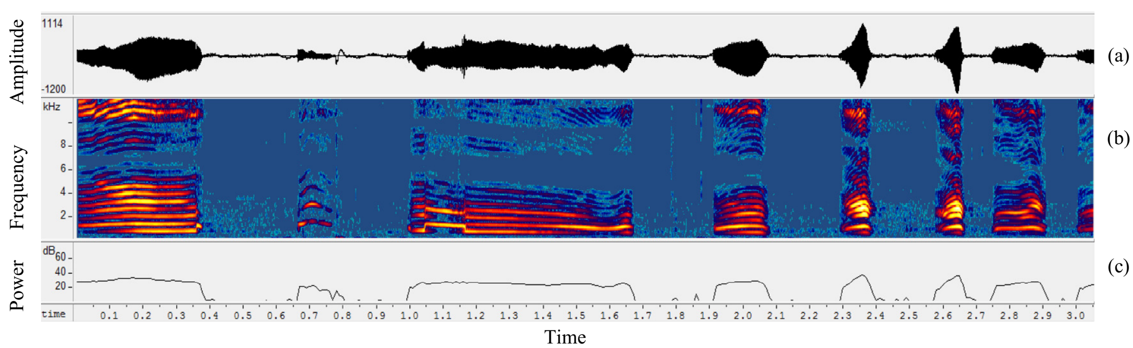

FIGURE 5. Infant of 2.1 kg weight, of 4.5 months of age. (a) Time-domain signal, (b) Narrowband spectrogram, and (c) Power plot of **A**.

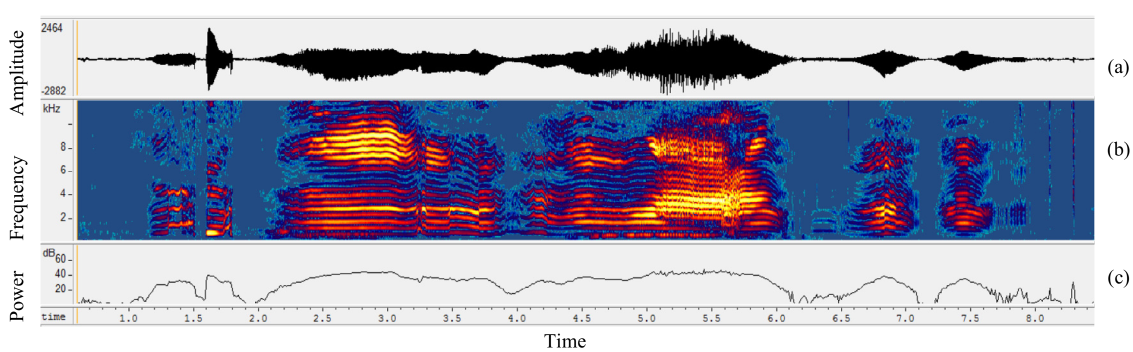

FIGURE 6. Infant of 6.2 kg weight, of 5 months of age. (a) Time-domain signal, (b) Narrowband spectrogram, and (c) Power plot of **A**.

cry types in four types, namely, birth, hunger, pain, and pleasure. The characteristics of the cry such as durational features and F_0-based features vary with the type of cry. For example, the mean maximum and minimum pitch for the hunger cries are 550 Hz and 390 Hz, whereas these parameters are 650 Hz and 360 Hz, respectively, for pleasure cries. On the other hand, in birth cries, these are 550 Hz and 450 Hz, and for pain cries the values of maximum and minimum pitch are 650 Hz and 400 Hz, respectively.[36] Sometimes, infants also cry because of fatigue; however, this cry is very difficult to identify and collect. Hence, infant cry due to fatigue is ignored in our analysis, and most of the literature that deals with signal processing aspect of infant cry does not have an analysis of this type of cry.

Anatomical differences in airways

Infants have a proportionately larger head and tongue, narrow nasal passages, an anterior and cephalad larynx (at a vertebral level of C3–C4), a long epiglottis, and a short trachea and neck. Vocal folds in infants are 3–5 mm long, and the composition of the vocal folds is uniform. The vocal folds in infants are much smaller than the vocal folds of adults, and they lack lamination seen in adults. This lamination plays an important role in the theories of phonation studies. Vocal fold length reaches to 7.5 mm by the age of 5.[37] These anatomic features make neonates and most young infants obligate nasal breathers until about 3–4 months of age. The cricoids cartilage (subglottis) is the narrowest point of airway in children. All these anatomic differences make signal processing of the infant cry difficult than the adult speech signal. The position of the larynx in infants is close to the base of the skull. This high position of the larynx helps in forming a closed passage from the nose to the lungs. The newborn infants can move the larynx upward in the nasophagus. The soft palate and the epiglottis effect a double seal and liquids can flow through the larynx while air flows through the nose through the larynx and through the trachea down to the lungs. There is *no* possibility of choking by having lodged into the larynx as is the case with adults. Moreover, infants are obligate nose breathers, ie, they do not breathe through the mouth even in case of nose blocking.[38]

Shorter length of vocal tract results in high formant frequencies as they have *inverse* relationship.[39] Normally, the first two or three formants are studied in infant cry analysis. The first two formants of infant cry are observed at 1100 Hz and 3000 Hz, respectively. This can also be approximated from the vocal tract length of the infants, which is approximately 7 cm in length. The tongue in the human newborn is long and thin compared with that of the adult human being. The tongue is positioned in the oral cavity and does not have an almost circular shape. This difference in the tongue shape makes it impossible for a newborn to produce the supralaryngeal vocal tract area function that is necessary to produce sounds.[36]

Variability of acoustic features with GA in premature infants

In premature infants, the cry characteristics change with the GA. GA is the age of the infant counted from the date of conception of the fetus and is measured in weeks. A normal pregnancy ranges from 38 to 42 weeks, and infants born before 37 weeks are considered premature. It has been shown that infants born with 31–33 weeks of GA have smaller duration of cries (~1.2 seconds) compared with infants born with 38–41 weeks of GA (~2.6 seconds). The mean maximum pitch and mean minimum pitch values are also higher in infants with low GA. In infants with 31–33 GA, the mean minimum and maximum F_0 values are 990 Hz and 470 Hz, respectively. The values of these parameters are 750 Hz and 370 Hz, respectively, in infants born with 38–41 weeks of GA.[40]

Variability of acoustic features with pathologies

All pathologies effect differently to different organs in the infants as well as in adults. In the spectrographic analysis of the cries, various differences in the cry modes, are observed for different pathologies. Some of the spectrographic analysis of pathologic infant cries is given below.

Cry of an infant suffering from laryngomalacia

Spectrographic analysis. From the spectrogram shown in Figure 7, the following are observed:

(1) Dysphonation and inspiratory phonation are dominating.
(2) Double harmonic break, glottal roll, and glides are totally absent.
(3) Spectral resolution is poor.

Cry of an infant suffering from asthma

Spectrographic analysis. Due to frequent inhalation, inspiratory phonation is observed in the spectrogram shown in Figure 8. Double harmonic break is visible in the spectrogram of the cry.

(1) Because of problem in breathing, inspiratory phonation is frequent in spectrogram.
(2) Rising and falling modes are present. It is similar to normal infant cry.

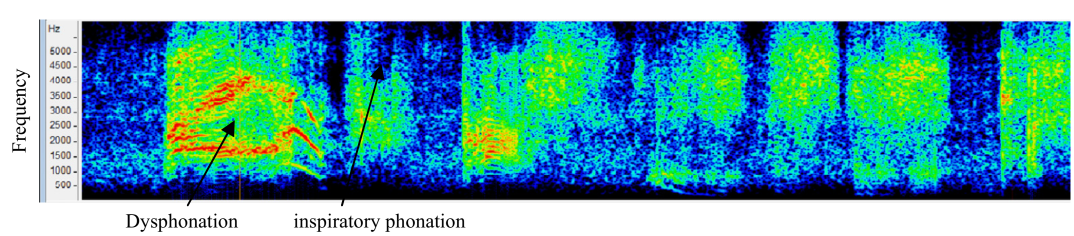

FIGURE 7. Spectrogram of an infant's cry who is suffering from laryngomalacia.

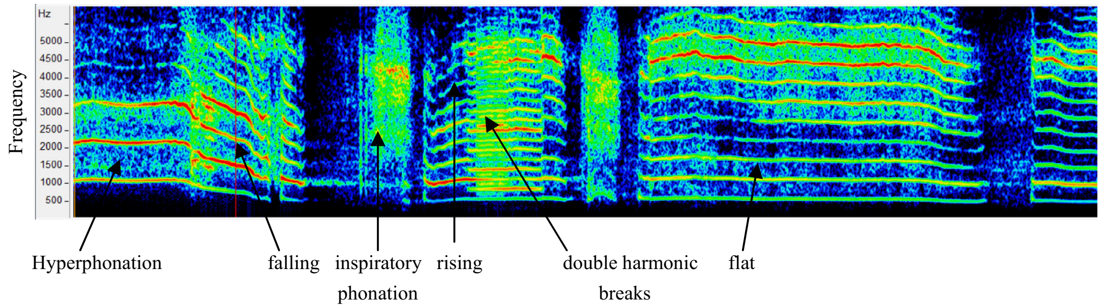

FIGURE 8. Cry modes present in the spectrogram of an infant's cry suffering from asthma sample taken from *Corpus I*.

Cry of an infant suffering from congenital heart disease

Spectrographic analysis. Spectrogram of an infant cry suffering from congenital heart disease is shown in Figure 9. It can be observed that the melody type is rising/falling, the same as in a normal infant. Glottal roll is present in the spectrogram, and dysphonation is dominating in the spectrogram (Figure 10).

Cry of a normal infant

Spectrographic analysis. Spectrographic analysis of the normal infant cry shows that the infant cry has rising cry melody followed by falling melody pattern. Inspiratory phonations are visible in the infant cry. However, the duration of the inspiratory phonations is shorter than the pathologic cases.

From the above analysis of the spectrograms of normal and pathologic infant cries, it can be observed that pathologic cries have vibrations, double harmonic breaks (many times also correlated to muscular pain[41]), and dysphonation (ie, *unstructured* energy distribution) modes in their spectrograms. Higher percentage of dysphonation, presence of double harmonic breaks, and glottal roll may be attributed to the presence of some disorder, although these are not yet found to be specific to any disease or pathology. Thus, which pathologies should be considered for the analysis of infant cries during data collection is an important aspect. In specific cases such as deafness and asphyxia, differences in the cry patterns are observed when compared with normal infants, even through auditory inspection of the infant cries. However, differences in the cry pattern do not give reliable classification or distinction among infant cries. For pathologies which are not due to neurologic disorder, the class-specific features in the cry may not be reflected in cry acoustics. Thus, a detailed study of the pathologies which are to be considered in the analysis is compulsory to avoid wrong selection of pathologies. On the other hand, pathologies that reflect differences in the infant cry acoustics are difficult to find among infants. Thus, creating a statistically significant database is a challenging task. In such cases, suggestions from medical practitioners and nurses can prove to be tremendously helpful before starting data collection.

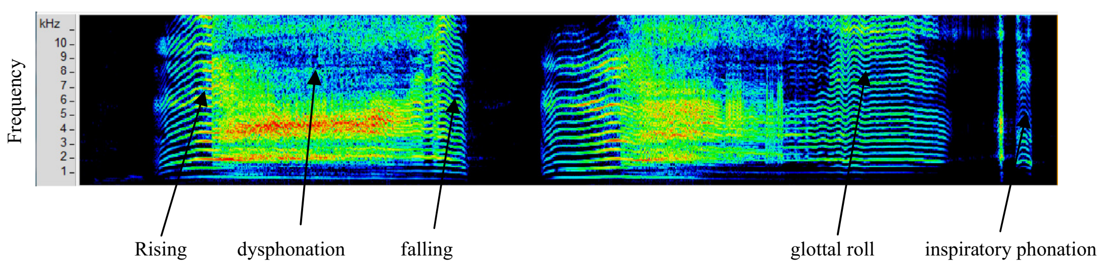

FIGURE 9. Cry modes present in the spectrogram of an infant suffering from congenital heart disease sample taken from *Corpus II*.

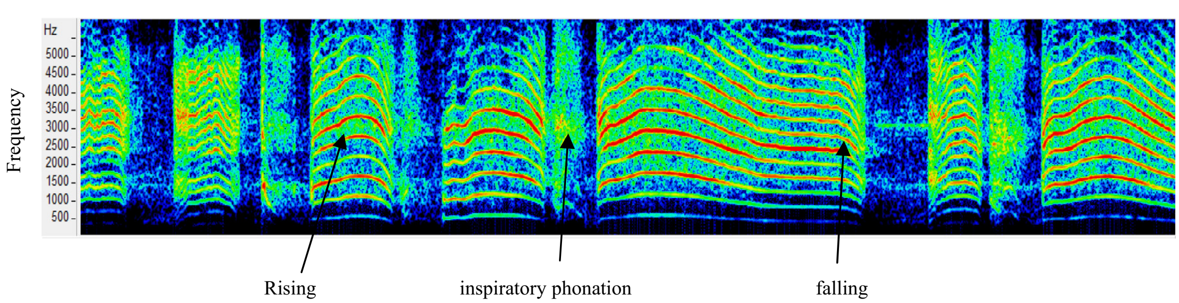

FIGURE 10. Spectrogram and cry modes present in the spectrogram of a healthy normal infant's cry sample taken from *Corpus I*.

Birth cry or first cry analysis

Birth cry is considered as a separate cry type because it conveys important information about the health of the infant. Birth cry is a symbol of the beginning of a healthy life. This is the first response of the newborn to the external world after coming out from the liquid atmosphere of the uterus to the outside air. It ensures proper functioning of the lungs, and this is the first time lungs full up with air and expand to their fullest capacity. It also helps the babies to get rid of any amniotic sac present in the lungs and nasal cavity. After birth, if an infant does not cry, it implies that something may be wrong with the infant and that infant has to be rigorously examined by the pediatrician. Pediatricians use the Apgar score to test the health of the newborn. To evaluate the Apgar score, five parameters (namely, complexion, pulse rate, reflex, activity, and respiration) are rated on a 0–2 scale, and they are summed to get the Apgar count. Infants with a score of 7 or above are normal healthy infants, whereas those with a score of 3 or below are regarded as critical.

Stimulation for cry production

As discussed in Section 2, during collection of the pain cries of infants, the important question is how the stimulation should be given to the infants to elicit pain which can result in crying. The procedure should be such that it follows the human ethics and guidelines, and is acceptable to the parents of the infants as well. The amount of stimulation should be kept the same for all the infants during data collection. The best possible solution to this is to collect infant cry data during vaccination. It has two advantages, namely, pain stimulation is not objectionable to the parents, and it is fully controlled by the experienced practitioner. Secondly, infants of the same age group will be covered in the analysis, which adds accuracy and consistency to the results of the research.

Room acoustics

The environment of data collection should be as silent as possible. However, it is difficult in infant cry analysis because in hospitals, many infants may cry at the same time. Moreover, parents may interfere in the recording to soothe their babies. Hence, efforts must be made to minimize the noise.

Time and duration of infant cries

In case of pain cries, recording should be done as soon as the stimulation is given to elicit the infant cry. In case of pathologic cries (neurologic disorders), it has been noted that the latency (duration between stimulation and cry production) is higher compared with normal health infants. This feature may help in classifying the pathologic infant cries. It is always better to have complete recording of a cry utterance (from start to end when infant stops crying), to study the cry behavior using prosodic features. This analysis may help in identifying the reasons for crying.

Recording device

A small handheld voice recorder (with high sampling frequency and excellent bit depth or resolution) is preferred so that one can move it closer to crying babies as soon as they cry. Recent advances in recording instruments have made available very small size microphones which can be attached to the infant's clothing so that spontaneous cries can be recorded, and there is no need to give stimulation for cry generation. Such recording can be helpful in studying the developmental aspects of the infants and for a long-term follow-up in case of sick infants. Data should be recorded at high sampling frequency (F_s) such as 22 kHz or 44.1 kHz. Reasons for using high sampling frequency are as follows:

(1) The F_0 in infants is higher than the male voice, female voice, and voices of children (due to small mass of the vocal folds). F_0 in infants are found in the range of 350 Hz to 1.2 kHz. High sampling frequency gives better resolution in F_0 estimated using computer-based algorithms.
(2) The vocal tract length in infants is small (about 7.5 cm) which is half of the vocal tract length in adults. Small vocal tract length results in high formant frequencies almost double than adult's formants. The theoretical values of formants in infants are 1.1 kHz, 2.2 kHz, and so on. Thus, using lower values of sampling frequency limits the number of formants to be covered in the available spectrum of the infant cry signal, which is half of the sampling frequency (due to Shannon's sampling theorem).
(3) Choosing a lifter size (ie, window in cepstrum domain) in cepstrum analysis of infants is a difficult task with smaller sampling frequency. Changing the lifter size by a single sample value in low F_s signal results in large variation in the cepstrum source and system response separation (called as deconvolution). This effect is illustrated in Figure 11, using the infant cry sample of 12 kHz (*Corpus I*). It can be observed from Figure 11A–Figure 11D that as the lifter size is changed from 10 samples to 25 samples, the system response captures source information instead of system-related information, ie, formants. Thus, to use cepstrum analysis in an effective way, a large sampling frequency is required.
(4) Higher F_s provides a sufficient number of samples between two glottal closure instances, which are necessary to detect glottal closure instances parameters.

SUMMARY AND CONCLUSIONS

In this paper, important factors affecting the data collection of infant cries are discussed with suitable references and examples. It is shown that the prosodic marks, duration, voicing, power, and short-time energy of the infant cry signal changes significantly with the infant's age, weight, and reason for crying. Analysis of infant cries, without considering these variables (ie, factors), may give misleading experimental results. This study may be helpful to researchers who want to work in the area of infant cry analysis for the purpose of possible medical diagnosis, developmental studies, infant's or parent's behavioral studies, and cry analysis and signal processing research.

Apart from the factors reported in this paper, other factors may also affect the cry characteristics such as change of place, crowd,

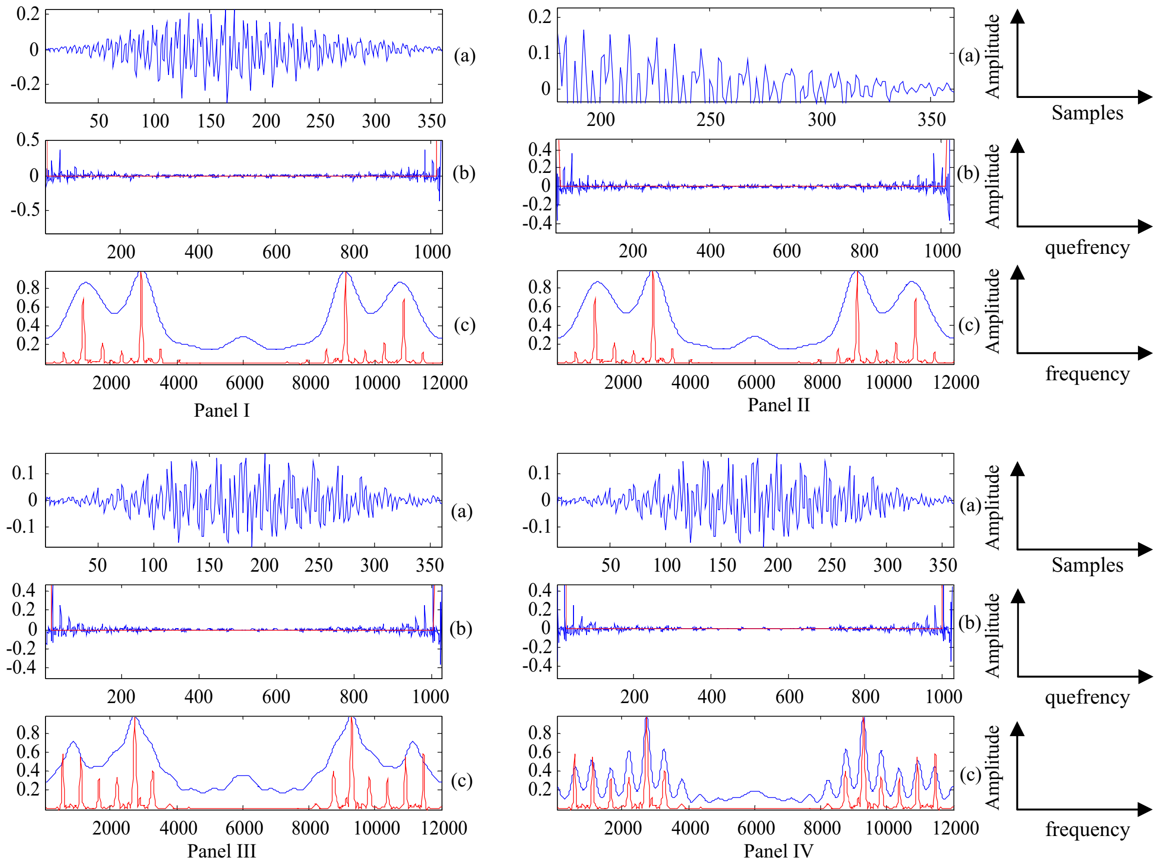

FIGURE 11. Cepstrum analysis with variable lifter size: Panel I: 10 samples, Panel II: 12 samples, Panel III: 20 samples, and Panel IV: 25 samples (for sampling frequency of 12 kHz for infant's cry). In all subfigures are (a) time-domain waveform, (b) cepstrum of **A**, and (c) system response spectrum.

improper handling of the infant, etc, although it is difficult to cover all the factors in an analysis and considering more factors in data collection process make the task even more difficult. Hence, it is up to the researcher to balance the research problem and availability of infant cry signals when analyzing infant cry studies. Similarly, in the participant information form, other factors can also be included, keeping in the mind the possible use of the database for future research. For example, in the study of effects of a drug on infant development during prenatal or postnatal period, one can add a column indicating the use of the drug and its dosage in daily routine or a combination of the drugs to find out their possible consequences. In another example, one can also study the effect of multilingual environment in infant cry patterns and relate it to their language acquisition skills. In all, it is up to the vision and area of specialization of the researcher to decide which questions or details should be asked while collecting infant cry samples so that these can be used in the long term and maximum information can be explored for particular research problem.

Acknowledgements

Authors would like to thank authorities of DA-IICT, Gandhinagar, India, for providing necessary resources for this study. We also like to thank Department of Electronics and Information Technology (DeitY) and Department of Science and Technology (DST), Government of India, New Delhi, India for help in providing partial support for carrying out this research work. We acknowledge the contributions made by the members of Speech Research Lab, DA-IICT, Gandhinagar, Gujarat, India.

SUPPLEMENTARY DATA

Supplementary data related to this article can be found online at doi:10.1016/j.jvoice.2016.07.007.

REFERENCES

1. Juntunen K, Sirvi P, Michelsson K. Cry analysis in infants with severe malnutrition. *Eur J Pediatr*. 1978;128:241–246.
2. Sherman M. The differentiation of emotional responses in infants. II. The ability of observers to judge the emotional characteristics of the crying of infants, and of the voice of an adult. *J Comp Psychol*. 1927;7:335–351.
3. Zeskind P, Sale J, Maio M, et al. Adult perceptions of pain and hunger cries: a synchrony of arousal. *Child Dev*. 1985;56:549–554.
4. Lewis M, Feiring C. Infant, mother, and mother-infant interaction behavior and subsequent attachment. *Child Dev*. 1989;60:831–837.
5. Chess S, Thomas A, Birch H. Characteristics of the individual child's behavioral responses to the environment. *Am J Orthopsychiatry*. 1959;29:791–802.
6. Clarici A, Travan L, Accardo A, et al. Crying of a newborn child: alarm signal or protocommunication. *Percept Mot Skills*. 2002;95:752–754.
7. Porter FL, Miller RH, Marshall RE. Neonatal pain cries: effect of circumcision on acoustic features and perceived urgency. *Child Dev*. 1986;57:790–802.

8. Cecchini M, Lai C, Langher V. Dysphonic newborn cries allow prediction of their perceived meaning. *Infant Behav Dev*. 2010;33:314–320.
9. Michelson K, Eklund K, Leppanen P, et al. Cry characteristics of 172 healthy 1-to-7-day-old infants. *Folia Phoniatr Logop*. 2002;54:190–200.
10. Lee S, Potamianos A, Narayanan S. Acoustics of children's speech: developmental changes of temporal and spectral parameters. *J Acoust Soc Am*. 1999;105:1455–1468.
11. Gustafsson E, Levrero F, Reby D, et al., 2012. Fathers are as good as mothers at recognizing the cries of their baby, Nature Communication, no. 2713, pp. 1–6.
12. Gordon S. 2004. Health day. [Online]. Available at: https://consumer.healthday.com/caregiving-information-6/infant-and-child-care-health-news-410/infant-crying-can-trigger-abuse-in-some-parents-521678.html. Accessed June 15, 2016.
13. Zeskind PS. 2013. Infant crying and synchrony of arousal, Evolution of Emotional Communication.
14. LaGasse LL, Neal AR, Lester BM. Assessment of infant cry: acoustic cry analysis and parental perception. *Ment Retard Dev Disabil Res Rev*. 2005;11:83–93.
15. Soltis J. The signal functions of early infant crying. *Behav Brain Sci*. 2004;27:443–490.
16. Wasz-Hockert O, Michelson K, Lind J. Twenty five years of Scandinavian cry research. In: Lester BM, Zachariah Boukydis CF, eds. *Infant Crying - Theoretical and Research Perspective*. New York and London: Plenum Publishing Corporation; 1985:83–104.
17. Barr RG. 2006. Crying behaviour and its importance for psychosocial development in children, Encyclopedia on Early Childhood Development, vol. 1, pp. 1–8.
18. Fisher J, Rowe H, Hiscock H, et al., 2011. Understanding and responding to unsettled infant behaviour, Australian Research Alliance for Children and Youth, pp.1–60.
19. Beckman ME. 2015. The emergence of compositional structure in language evolution and development, Interspeech, Germany. Available at: http://www.superlectures.com/interspeech2015/isca-medalist-the-emergence-of-compositional-structure-in-language-evolution-and-development. Accessed June 15, 2016.
20. Garcia JO, Reyes Garcia CA. 2003. Mel-frequency cepstrum coefficients extraction from infant cry for classification of normal and pathological cry with feed-forward neural networks, Proceedings of the International Joint Conference on Neural Networks, pp. 3140–3145.
21. Reyes-Galaviz OF, Cano-Ortiz SD, Rayes-Garcia CA. 2008a. Evolutionary-neural system to classify infant cry units for pathologies identification in recently born babies, 7th Mexican International Conference on Artificial Intelligence, pp. 330–335.
22. Reyes-Galaviz OF, Cano-Ortiz SD, Reyes-Garcia CA. 2008b. Validation of the cry unit as primary element for cry analysis using an evolutionary-neural approach, Mexican International Conference on Computer Science, pp. 261–267.
23. Varallyay G Jr, Benyó Z, Illényi A, et al., 2004. Acoustic analysis of the infant cry: classical and new methods, 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEMBS, Vol. 1. pp. 313–316.
24. Hariharan M, Sindhu R, Yaacob S. Normal and hypoacoustic infant cry signal classification using time-frequency analysis and general regression neural networks. *Comput Methods Programs Biomed*. 2012;108:559–569. Elsevier.
25. Kheddache Y, Tadj C. Acoustic measures of the cry characteristics of healthy newborns and newborns with pathologies. *J Biomed Sci Eng*. 2013;6:796–804.
26. Colton RH, Steinschneider A. The cry characteristics of an infant who died of the sudden infant death syndrome. *J Speech Hearing Disord*. 1981;46:359–363.
27. Patil HA. Cry baby: using spectrographic analysis to assess neonatal health from an infant's cry. In: Neustein A, ed. *Advances in Speech Recognition, Mobile Environments, Call Centres and Clinics*. Springer-Verlag; 2010:323–348.
28. ICMR, 2006. Ethical Guidelines for Biomedical Reserach on Human Participants [Report]. New Delhi: Director General, Indian Council for Medical Research, 2006. Available at: http://icmr.nic.in/ethical_guidelines.pdf. Accessed June 15, 2016.
29. Council for International Organizations of Medical Sciences (CIOMS) with WHO, 2002. International Ethical Guidelines for Biomedical Research Involoving Human Subjects, Geneva, 2002. Available at: http://www.recerca.uab.es/ceeah/docs/cioms.pdf. Accessed June 15, 2006.
30. Parent_qualification: Available at: http://www.literacytrust.org.uk/assets/0001/3375/Hamer_NCT_research_overview_Parent_child_communication_p15-20_Mar12.pdf. Accessed June 15, 2016.
31. Buddha N, Patil HA. 2007. Corpora for analysis of infant cry, International Conference on Speech Databases and Assessments, Oriental COCOSDA, Hanoi, Vietnam, pp. 43–48.
32. Chittora A, Patil HA. 2013. Data collection and corpus design for analysis of normal and pathological infant cry, Oriental COCOSDA held jointly with International Conference on Asian Spoken Language Research and Evaluation (O-COCOSDA/CASLRE), New Delhi, India, Vol. 2, pp. 1–6.
33. Xie Q, Ward RK, Laszlo CA. Automatic assessment of infants' level of distress from the cry signals. *IEEE Trans Speech Audio Process*. 1996;4:253–265.
34. Huelke DF. An overview of anatomical considerations of infants and children in the adult world of automobile safety design. *Annu Proc Assoc Adv Automot Med*. 1998;42:93–113.
35. Arsikere H, Lulich SM, Alwan A. Estimating speaker height and subglottal resonances using MFCCs and GMMs. *IEEE Signal Proc Lett*. 2014;21:159–162.
36. Lester BM. Inroduction - there's more to crying than meets the ear, In: Lester BM, Boukydis Zachariah CF, eds. *Infant Crying - Theoretical and Research Perspective*. New York and London: Plenum Press; 1985:1–27.
37. Kent RD. 2003. The MIT Encyclopedia of Communication Disorders [Report]/Massachusetts Institute of Technology, pp. 49.
38. Lieberman P. The physiology of cry and speech in relation to linguistic behaviour. In: Lester BM, Boukydis Zachariah CF, eds. *Infant Crying - Theoretical and Research Perspective*. New York and London: Plenum Press; 1985:29–57 [Ch. 2].
39. Quatieri TF. *Discrete-Time Speech Signal Processing*. ed 2. Pearson Education; 2004:68 [ch. 3].
40. Thoden C-J, et al. Sound spectrographic cry analysis of pain cry in prematures. In: Lester BM, Boukydis Zachariah CF, eds. *Infant Crying - Theoretical and Research Perspective*. New York and London: Plenum Press; 1985:105–117 [Ch. 5].
41. Xie Q, Ward RK, Laszlo CA. *Determing Normal Infant's Level of Distress from Cry Sounds*. 2009:1094–1097 Canadian Conference on Electrical and Computer Engineering, Vancouver, BC, vol. 2.