Development and Validation of a Computerized Screening Tool for Infant Cry

Project funded by AIISH Research Fund (ARF)

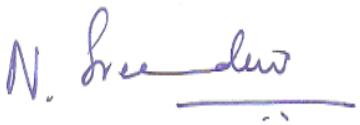
(2014-2015)

Sanction number: SH/CDN/ARF-10/2014-15

Total grants: 4, 43, 000/-

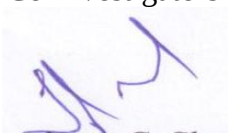
Total duration of the project: 9.10.2014 to 30.10.2015

Principle Investigator



Dr. N. Sreedevi
Reader & Head
Department of Speech- Language Sciences

Co-Investigators



Dr. Jayashree C. Shanbal
Reader & Head
Department of Prevention of Communication Disorder

and

Mr. Arunraj K
Clinical Assistant
Department of Audiology

Research Officer

Ms. Neethu Thoduvayil

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Continued...

Continued…

ACKNOWLEDGEMENTS

We extend our sincere gratitude to our Director, Prof. S .R. Savithri, All India Institute of Speech and Hearing, Mysore, for sanctioning and funding the project.

We extend our gratitude to Dr. B. Krishnamurthy, Dean and Director, Mysore Medical college and Research Institute for permitting us to conduct the study in Cheluvamba hospital, Mysore. We thank all the staff of the neonatal ward especially Ms. Yeshoda, Mr. Chennayya shetty and Ms. Leelavathi, staff nurses for their cooperation during cry recording. We also thank, Dr. Mahadevappa, ENT specialist, AIISH, Mysore, and other POCD staff members for their timely help.

We sincerely thank Mr. Akshay M, for his preparation of the software "ICry". We thank Dr. M. S. Vasanthalakshmi, Reader in Biostatistics and Mr. C. D. Santhosh, Lecture in Biostatistics, AIISH, Mysore, for helping us with statistics. We thank Ms. Irfana, JRF and Ms. Amulya P Rao, JRF for their generous contribution in completing the project report and Ms. Sushma S, Research Officer at Dept of SLS, for her timely help in completing the project.

Finally, we thank all the parents and caretakers for permitting us to record the cry samples of their newborn babies.

CHAPTER 1

Introduction

"Crying is a biological siren, alerting the caregiving environment about the needs and wants of the infant and motivating the listener to respond" (Zeskind & Lester, 2001, p 149). In humans, crying not only refer to lacrimation (or tears) but several loosely coordinated behaviors, involving changes in respiration (gasping), and vocalization (wailing) (Vingerhoets, Bylsma, & Rottenberg, 2009) and emotional tears are only shed by humans (Frey, 1985; Murube, 2009c). Compared to infants of other species, human infants stand out for their high degree and long duration of helplessness. Baby crying has therefore been referred to as the ''acoustical umbilical cord'' (Ostwald, 1972), serving to establish and maintain a close connection between the infant and the caregiver. A cry is infant's first verbal communication and it is a message of urgency or distress. Cry sound is nature's way of ensuring that adults attend to the baby as quickly as possible and it is very hard for most people to listen to a crying baby. Almost everyone recognizes that infants cry for many reasons and crying is a normal response.

To the pediatrician and obstetrician, the cry at birth is a welcome sign that the infant is breathing normally. Delay in the establishment of respiration may signify serious brain damage, or at least the risk of it, from deprivation of oxygen. The first cry also removes foreign material to improve pulmonary capacity in the first days of life which is a defense mechanism to increase body temperature.

The cries of infants have stimulated curiosity among physicians, parents, scientists, speech pathologists long before sound spectrography or even acoustic recording devices were known. Early scientists working in the 19th century believed that the first sounds, contained

information about the child, its physical and emotional well being and about the cry provoking situation.

Infant cry is a product of the respiratory and phonatory systems. The respiratory system accounts for variability due to rapid maturational changes. Fisichelli, Haber, Davis, and Karelitz (1966) pointed out about an inspiratory voicing in neonatal cry and hence it is difficult to find the expiratory cycle. Once infant matures, virtually cry becomes expiratory. Cry behavior varies according to the cry evoking events or stimuli. Birth cry made within minutes after birth is partially pain, hunger and environmental related. Hence, birth cry is primarily a response to the infant's external stimulus. Further, it serves to assist in the cardio respiratory organization in the new environment. At two weeks of age hunger cries were characterized by a higher F0 of phonation, greater F0 range, longer mean cry durations, a larger portion of phonation to silence and more complex melodic contour than discomfort cry. Pain cries are elicited through pin prick, base of foot stung with rubber band and in few studies blood samples were taken. Hunger cries were elicited withholding food from infant at the normal feeding time. Startle cries elicited by clapping loudly near infants head or suddenly dropping the child to a table top. Sedlackova (1964) noted birth cry to be high frequency signal. In the cry samples, pain cry resulted with highest F0, followed by hunger then startle at lowest F0.

Since early 1960s researchers have used spectrum to correlate acoustical data with a variety of pathological conditions and they have reported normal ranges for several spectrographically measurable features in a number of pathological conditions. There has been considerable interest in utilizing infant cry analysis as a measure of developmental status in babies with pre- and perinatal risk factors, such as prenatal substance exposure (e.g., Lester et al., 2002) or premature birth (e.g., Goberman & Robb, 1999). There has also been

interest in utilizing infant cry to identify babies at risk for developmental or medical conditions including hearing impairment (Várallyay, Benyó, Illényi, Farkas,& Kovács, 2004) and, recently, autism spectrum disorders (Esposito & Venuti, 2010; Sheinkopf, Iverson, Rinaldi, & Lester, 2012).

Manual inspection of spectrograms of infants cries have been a gold-standard method for detecting acoustic features in cry sounds, the timing and onset of cry vocalizations and the fundamental frequency (F0) of cry. Such visual inspection has also been used to describe melodic variations in F0 across an utterance, and voicing or periodicity in the cry utterance. However, although these manual approaches have the advantage of detecting F0, the drawback is that the process is tedious and slow, constraining the amount of time for analysis of the data.

Over time, advances in computing have allowed for increased power and flexibility in acoustic analysis, including the ability to utilize robust techniques for the accurate estimation of fundamental frequency and other acoustic features of cry vocalizations. Findings of several studies reveal the importance of cry as a useful window for early detection of diseases (infants with profound hearing loss, various syndromes such as Down syndrome, Cri- Du-Chat syndrome etc, infant voice pathologies, various metabolic pathologies, cerebral palsy etc).

More recent approaches have utilized computer-assisted methods or commercially available speech analysis software packages to code acoustic aspects of infant cry. For example, one approach uses a computer cursor that is moved along a digitally displayed sound spectrogram to quantify aspects of cry duration and selects locations on the spectrogram for acoustic analysis (Zeskind & Barr, 1997; Goberman & Robb, 1999; Wermke & Robb, 2010). In this way, resulting portions of the cry were subjected to a Fast Fourier

transform (FFT) to yield information from the power spectrum in the cry (e.g., maximum/minimum F0). As mentioned previously, abstracting quantitative data from spectrograms by this method is time consuming (LaGasse, Neal, & Lester, 2005). In addition, these approaches utilized speech analysis tools that were designed to extract acoustic information from adult speech.

Given the anatomical differences in the vocal tract of infants, there is a need for tools designed to track and extract the F0 and other acoustic features from infant cries specifically. Advances in computing capacity in the recent years allow researchers to utilize methods for quantifying aspects of the sound spectrum that can be expected to yield more accurate estimates of F0 and related parameters automatically. Automated approaches have the advantage of fast analysis of very large data sets, objective assessment, the ability to quantify multiple data points, and the flexibility to yield derivative measures (e.g., jitter, pitch contours, etc.), which have the potential to increase the applied value and clinical utility of cry assessment.

1.1 Need for the Study

Analysis of cry in infants is vital as it aids in the knowledge of adequate functioning of the neuro-respiratory and the phonatory systems and also to derive at diagnostic and prognostic indicators if the analysis proves to be deviant from normal. Attempts at analyzing infant cry are still in the preliminary stages and have not yet been recognized as a widespread tool in assessing neuro-respiratory and phonatory functions in infants. In the international arena of infant cry analysis, previous work has shown a relationship between acoustic characteristics of the cry and diagnoses related to neurological damage, SIDS, prematurity, medical conditions, and substance exposure during pregnancy. In the Indian context also there have been a number of attempts on acoustic analysis of infant cries that focus on

detection of conditions such as heart diseases or pulmonary diseases (Daga & Panditrao, 2011; Radhika Rani, Chandralingam, Anjaneyulu & Satyanarayana, 2012) in young infants.

Also, infant cry analysis could prove as an efficient method of identification of infants with communication disorders especially in cases when it is not feasible to carry out objective tests on infants such as high risk babies and babies in the Neo-natal Care units (NICU). Developing an automated robust acoustic tool for the assessment of infant cry can be thought of as another effortless means of verifying the activation and functioning of major physiological systems of the infant in the early days itself. Hence it is essential to pursue research with larger sample sizes in order to obtain normative data and standardize the acoustic and other vocal parameters to assist in early detection of disorders in high-risk infants.

In this context, the present study scrutinizing the relevance and vitality of the cry behavior of infants aims to develop a non-invasive computerized screening tool to identify infants who may be at risk for typical development. In addition with enhancement of the database of infant cries, the tool could serve as one of the means of early follow up protocol for infants with potential communication disorders.

1.2 Objectives of the Study

1. To study and compare the acoustic parameters of the cries in healthy and high risk newborns
2. To develop a computerized screening tool for screening the newborns cries based on the acoustic pattern of the newborns cries.
3. To validate the sensitivity and specificity of the screening tool using the data base obtained on a pilot basis.

1.3 Brief Method of the Study

Participants

Two groups of newborns were selected for the study. Group I was a control group which consisted of 220 healthy newborns within 7 days of birth with average mean age of 4 days. These newborns were recruited for the study on the basis of no significant pre or natal history as checked by a high risk register, screening audiometric evaluation, medical records, and a formal interview with the parents. And group II consisted of 92 newborns within 30 days of birth with one or more risk factors. The risk factors considered for the study included low birth weight of not more than 2000 grams, birth asphyxia, meconium aspiration syndrome and infantile jaundice.

The study was conducted in 4 phases, which included

Phase I, Recording of cry samples from newborns

Stimulated pain cry samples of both healthy and high risk newborns were recorded using an audio recorder. Intramuscular immunization (Hepatitis B) was used as the pain stimulus for the healthy newborns and toe tapping for high risk newborns.

Phase II, Acoustic analysis of newborn cries

Acoustic analysis of infant cry was performed by extracting the parameters from the cry samples of newborns from both groups I and II. The cry samples were analyzed and the acoustic parameters such as frequency, intensity, formants and other parameters were extracted using PRAAT software version 5.4.01.

Phase III, Development of Computerized screening tool software

The overall statistical value of the analyzed parameters of infant cry in healthy newborns was provided to a software professional for the development of computerization tool. Software was developed specifically for recording, analyzing and the interpretation of the acoustic signal.

Phase IV: Pilot Validation of the developed screening tool

The developed computerized software was administered on the total cry recording of newborns in both groups. The sensitivity and the specificity of the tool was calculated based on this using appropriate formulae.

CHAPTER 2

Review of Literature

Crying is behaviour; in fact it is a sequence of behaviour pattern that is part of a large behaviour repertoire of infants. Compared to infants of other species, human infants are notable for their high degree and long duration of defencelessness. Baby crying has therefore been referred to as the ‘‘acoustical umbilical cord’’ (Ostwald, 1972), serving to establish and maintain a close connection between the infant and the caregiver. Babies cry from hunger, pain, or to bring the caregiver nearer. Crying is often viewed as a distress call and solicits care giving (Acebo & Thomam, 1992; Zeifman, 2001).Crying is a perseverative behaviour as crying once initiated, is a self sustaining behaviour that continues regardless of its cause (Zeifman, 2001). Vocal crying is present in newborns, while tearing develops in the third or fourth month. Firmly, as a form of communication, crying is the beginning of vocalization and may have implications for the development of speech and language.

The function of the birth cry is said to be entirely physiological, having to do with establishment of normal respiration and the oxygenation of the blood. To the pediatrician and obstetrician, the cry at birth is a welcome sign that the infant is breathing normally. Delay in the establishment of respiration may signify serious brain damage, or at least the risk of it, from deprivation of oxygen. The first cry is said to have other physiological purposes, such as to remove foreign material, to improve pulmonary capacity in the first days of life and a defense mechanism to increase body temperature (Long & Hull, 1961),

There has been extensive research in the field of infant cry analysis for the past several years. The ground work was laid in the 1960s by a Scandinavian team including Wasz-Hockert, Lind, Vuorenski, Partanen, and Valanne. They defined crying as a distinct behavioral state through which basic sensations such as hunger and pain are expressed; and

cry as a normal developmental phenomenon. Their sound spectrographic study results revealed acoustical properties of the cry, which opened the door to using the cry for medical diagnosis. Later with the advance in the research aided by high-speed computer technology and interdisciplinary impact, the signal processing and analysis of the cry is greatly improved. That led to the understanding of the physiological and anatomical basis of cry production.

Despite all those researches, clinicians were unaware about the potential applications of infant cry analysis. Factors like physical limitations of the spectrogram analysis and the lack of a conceptual model of cry production where one can correlate the cry output to specific pathologic conditions were the issues. In an attempt to address all these issues, Golub and Corwin (1985) developed an approach for cry analysis that utilizes a conceptual model of infant cry production which is called as the physio-acoustic model.

2.1 Physioacoustic Model

Golub's (1980) physioacoustic model of crying (Figure 1) was designed to select acoustic features that would identify infants at medical risk. Model assumes the source filter theory where the production of any acoustic waveform is explained based on its source, and filter function. In infant cry production, source is the air from the lungs which is being pushed throughout the vocal folds of the larynx and filter function is carried out by rest of the vocal tract. Model assumes three levels of central processing of the muscles contributing to the source filters of crying. First level is the upper processor which is involved in choosing and modulating the state of action of the child (eg: fussiness). Middle processor involves in the vegetative states such as swallowing, coughing, digestion and crying. The third lower processor controls the relevant muscle groups including the sub glottal, supra glottal, glottal and facial muscles which are coordinated for cry production.

Major assumption of the model is the independent control of sub glottal (respiratory), supraglottal (filter) and glottal (laryngeal) muscles. Based on the muscle control assumption if we can isolate the differences in cry caused by glottal, sub glottal and supraglottal malfunctions, then we can locate its physiological or anatomical abnormalities. To elaborate on the cry production, this model divides cry production into 4 parts. The first part is the sub glottal (respiratory) system that is responsible for developing the pressure below the glottis necessary for driving the vocal fold source. The second part is the sound source located at the larynx. It has two subparts as a periodic source and a turbulence noise source. The third source of cry production is the vocal and nasal tracts located above the larynx, which act as acoustic filters that have a transfer function. The fourth part is the radiation characteristics that describes the filtering of the sound between the mouth of the infant and the microphone placed at some distance.

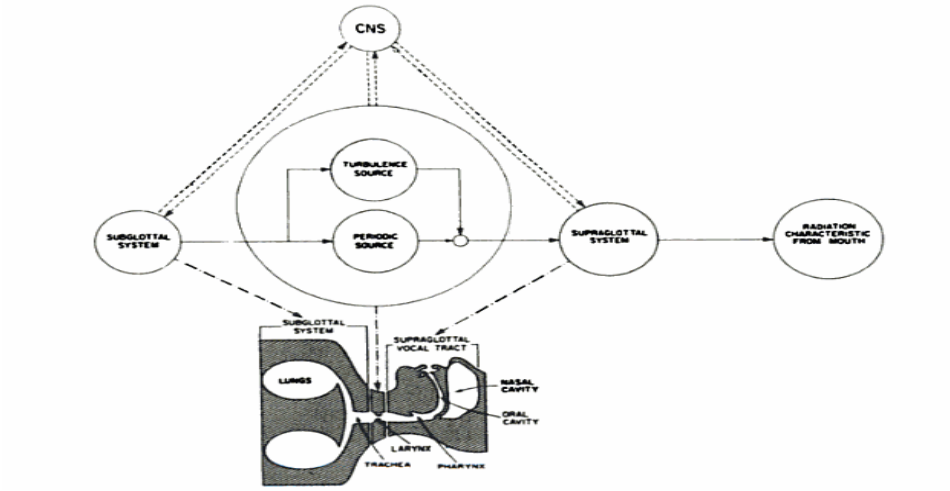


Figure 2.1., Schematic representation of Physioacoustic Model (Golub 1980)

Golub also suggested the crucial role of cranial nerve X (the vagus) in influencing the glottal muscle process which affects the modes of phonation of cry and its fundamental frequency. There have been elaborations of the Golub's model, most importantly by Lester (1987). He added cranial nerves IXth, XIth, XIIth along with vagus nerve which are important to cry acoustics. These nerves influence the muscles of larynx, pharynx, chest and neck. He

also suggested the effect of phrenic and thoracic nerves and the nuclei of these pathways in cry acoustics. Lester's elaboration on the model emphasize the fundamental frequency (both its average and variability) and the formant frequencies associated with the cry as critical measures reflecting central nervous system (CNS) functioning. Lester and Boukydis (1990) described how the central control of these characteristics of crying involves the brain stem, mid brain and limbic system with later involvement by the cortex.

Variation in tension on the laryngeal muscles, cricothyroid and vocalis, and the abdominal respiratory muscle are thought to be responsible for fundamental frequency (F0). There are three identifiable cry modes of vocal fold vibration: basic cry or phonation or fundamental frequency (F0), high pitch cry or hyperphonation (1000–2000 Hz), and noisy or turbulent cry (dysphonation) (Lester, 1984; Golub, 1989). Damage to vagal cranial nerve complex is related to atypical patterns of F0. Rapid shifts or variability in F0 suggests instability in the neural control system. The brainstem also controls the contour and cross-sectional airway of the supraglottal system, which shapes formant frequencies. Only the first two formants are usually measured: The first formant (F_1) occurs at approximately 1100 Hz and the second formant (F_2) at approximately 3300 Hz (Lederman et al., 2004). Other regions of the brain participate in cry initiation (limbic system, hypothalamus), configuration of the cry pattern (midbrain), and motor coordination of respiration, larynx, and articulation (reticular activating system) (Zeskind & Lester, 2001). The extensive vagal innervation and CNS coordination underlying infant cry result from modulation by the autonomic nervous system. Cry initiation, including latency to cry or threshold (amount of stimulation required), has been associated with sympathetic arousal. Respiratory modulation and temporal patterning of the cry (long or short duration of cry utterance, number of utterances, duration of inspiratory period, and number of short unvoiced sounds or utterances) reflect parasympathetic mechanisms of homeostasis. Modulation of the overall contour of F0 as

well as the amplitude or intensity of the cry reflects both faciliatory and inhibitory autonomic mechanisms (Lester, 1984).

Physiological regulatory model of Porges and Maita (1994) is another physiological model that bears on cry production and acoustics. Model discusses the importance of vagal tone in the expression and regulation of emotion. Vagus system has two branches and each branch has two source nuclei as dorsal motor nucleus and nucleus ambiguous. The source nuclei of right vagus, nucleus ambiguous, control the larynx and sino – atrial node of the heart and therefore it is important to both vocal intonation and vagal tone. Thus vagal control of the right side of the larynx produces changes in vocal intonation as well as changes in heart rate, creating a cardiovascular state associated with specific emotions.

Lester and Zeskind (1982) proposed bio behavioral model, another model of autonomic control of cry features. They distinguish between the act of crying and the sound of crying. Both the aspects tap different behavioural systems. Sympathetic nervous system is implicated in excitatory responses which control the cry production and it influences characteristics such as the threshold to cry and the duration of sustained crying. Parasympathetic nervous system is involved in inhibitory responses and spectral components of the cry or acoustic characteristics of the sound (e.g. fundamental frequency, dysphonations) are regulated by the system through regulation of the vagus. According to this model, when the infant is aroused, sympathetic and parasympathetic system acts in opposition and this lead to acoustic outcomes such as frequent pitch shifts (eg, from phonation to hyperphonation), which reflect poor organization/modulation of these systems in infant. Thus departures from the mean on either cry production (sympathetic system) or spectral characteristics (parasympathetic system) could reflect poor bio behavioral functioning.

In summary all three models of cry production proposes vagal system as the primary source of variation in cry acoustics. Deficits in either brainstem functioning or higher brain functioning may affect vagal control of the cry and this may lead to abnormalities in cry acoustics especially fundamental frequency. Models of cry production enable us to correlate medical abnormalities with acoustical measurements. These cry models may allow us to identify disease processes reliably that affect cry production system in a consistent manner. These disorders include respiratory problems, structural deficits of the vocal cords or vocal tract, muscular abnormalities, and abnormalities of the peripheral and central nervous systems. Tables 2.1 (a) and (b), list the commonly used cry characteristics in infant assessment and their associated biological mechanism (Corwin, Lester, & Golub, 1996; Zeskind & Lester, 2001; Lester et al., 2002). The first four characteristics evaluate the entire cry (Cry Latency to Short Utterances) and the remaining characteristics are calculated for each cry utterance and may be averaged across utterances.

Table 2.1

Cry Characteristics and Associated Mechanisms

Characteristic	**Definition**	**Biological Mechanism**
Cry Latency	Time from known stimulus to onset of the first utterance (cry sound)	Arousal from limbic–hypothalamic system
Threshold	Number of applications of stimulation to elicit a cry	Arousal from limbic–hypothalamic system
Utterances	Number of cry sounds across cry	Neural control of respiratory system
Short Utterances	Number of unvoiced sounds across cry	Unstable respiratory control

Continued……

Continued……..

Phonation	Cry mode resulting from vocal fold vibration between 350 and 750 vibrations or cycles per second (Hz)	Neural control of muscular tension in vocal folds and air flow through the glottis
Hyperphonation	Cry mode caused by a sudden upward shift in F0 to _1000 Hz	Neural constriction of the vocal tract
Dysphonation	Cry mode caused by noisy or inharmonic vibration of the vocal folds	Unstable respiratory control
Cry Mode Changes	Number of times cry modes change during an utterance	Instability in neural control of the vocal tract
Fundamental frequency (F0)	Base frequency during vocal fold vibration, heard as the pitch of the cry	Vagal input to larynx and lower vocal tract
First Formant	Frequencies centered at first resonance of F0, approximately 1100 Hz	Neural control of size and shape of upper vocal tract
Second Formant	Frequencies centered at second resonance of F0, approximately 3300 Hz	Neural control of size and shape of upper vocal tract
Duration.	Time (ms) from onset to offset of cry utterance	Neural control of respiratory system
Duration of Inspiration (Interutterance Interval)	Time (ms) between cry utterances, which is a measure of breath holding is evaluated by the 2nd inspiratory period	Neural control of respiratory system
Amplitude.	Intensity of the cry (dB). Heard as loudness	Neural control of respiratory system and capacity
Variability in F0	Changes in F0	Instability in neural control of the larynx and lower vocal tract
Variability in F_1, F_2	Changes in formant	Instability in neural control of upper vocal tract
Variability in Amplitude	Changes in intensity within an utterance or averaged across utterances	Instability in neural control of the respiratory system

Truby and Lind (1965) proposed that an infant's crying episode involves distinct phases of phonatory activity. In turn, these phases of phonatory activity were assumed to

correspond to changes in an infant's behavioral status across the crying episode. Attack phase is the initial phase and is displayed acoustically as a period of high amplitude expiratory cry phonation. This phase is thought to reflect an automatic or "reflexive" distress response by the infant after the presentation of the pain stimulus. The middle or *cruising* phase is displayed acoustically as a decrease in the cry amplitude and the introduction of voicing during the inspiratory phase of the crying/respiratory cycle. This phase is thought to represent a levelling off of the infant's behavioral response to the cry-eliciting stimulus. The latter *subdual* phase is depicted as a further decrease in the overall amplitude of the infant's cry, while voicing continues during inspiration. Figure 2 shows the representation of a pain cry.

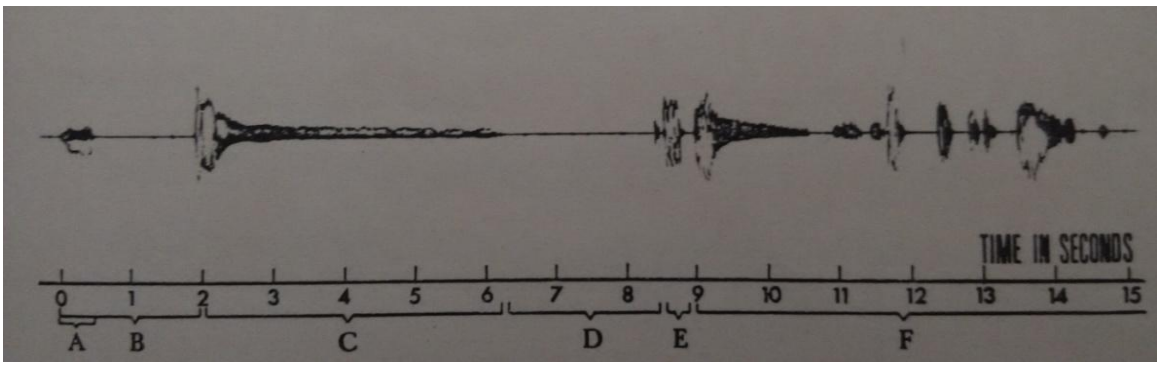


Figure 2.2., Oscillogram of a pain cry. (A) the baby is pinched and the examiner says the word "now';(B) represents the first latency period; (C) is the first expiratory cry; (D) is the second latency period; (E) a phonated inspiration; (F) second cry signals and phonated inspirations; (G) the signals gradually get smaller and smaller. The cry signal (C) is the one which we invariably analyse. (Wasz-Hockert , Lind, Vuorenkoski, Partanen & Valanne 1968)

2.2 Development of Anatomical and Physiological Systems in Infants

The normal development and production of cry is best understood by considering the development of anatomical and physiological systems in a new born. There are three component systems for infant cry production which includes the respiratory system (lungs and trachea), the vocal cords (larynx), and the vocal tract (pharynx, oral and nasal cavities). Following section discusses the anatomical development of each system separately.

2.2.1 Respiration

In the newborn, each cry cycle displays a consistent displacement of respiratory volume, which suggests a rather precise coordination between crying and respiration (Bosma, Truby & Lind., 1965). However, the newborn cry contains evidence of inspiratory voicing, but sometime between 1 and 3 months nearly all cries are produced in the expiratory phase of tidal breathing (Fischelli, Haber, Davis, & Karelitz., 1966). Development of prolonged expiratory phase was explained in terms of newborn's purported lack of a 'hold – back' mechanism (Wolff, 1969; Lieberman, 1975). This mechanism is deemed to be necessary for regulating subglottal air so that long expirations can be produced. In newborns, the rib cage is aligned almost perpendicularly to the spine (Langlois, Baken, & Wfilder, 1980). As such, it imposes a mechanical constraint on the maintenance of a steady sub glottal air pressure required for prolonging the expiratory phase.

Due to the elastic coil of the lungs being relatively low in the newborn, sufficient air pressure can be generated during short expirations for voicing to occur. Increase in lung volume during crying lead to an initial pressure that is too high for a long expiration to be sustained (Lieberman, 1985). At about 3 months, the rib cage begins to assume an adult like configuration in which there is a downward and outward angle between it and the spine (Hixon, 1973). As a consequence the actions of intercostal muscles will compliment those of abdomen. The external intercostals perform the task of lifting and enlarging the rib cage. The internal intercostals can now act more fully to pull it down and decrease size. This major restructuring of the rib cage enables greater control to be exercised over sub glottal air pressure and thereby not only deeper and longer breathing, but also greater variations in the vocal output of crying.

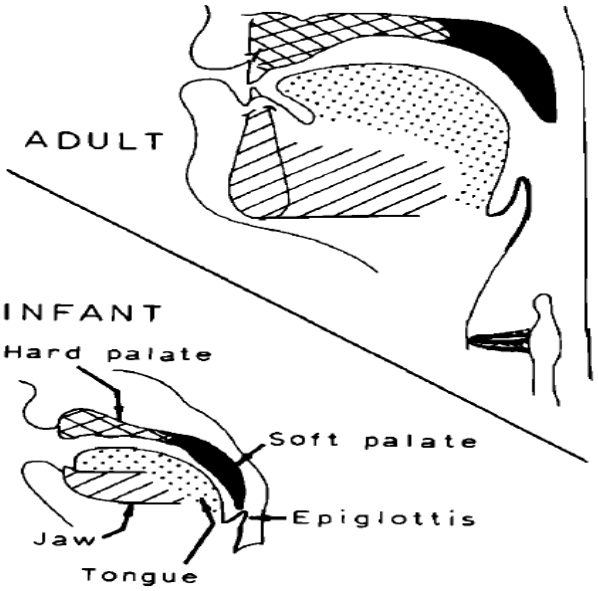


Figure 2.3., Vocal tracts of adult and infant. Compared to the adult, the infant has a gradually sloping oropharyngeal channel, a short broad oral cavity, a short pharynx, and an elevated larynx (Kent & Murray, 1982).

2.2.2 Larynx and pharynx in newborn infants

The anatomy and pathology of the infant larynx is significantly different to that of the adult (Figure 3). The larynx not only is a sound producer in that it contains the glottis and vocal folds, but also serves as a valve to regulate air flow from the respiratory system. According to the myoelastic – aerodynamic theory (Van den Berg, 1958), it shifts from the valve like function to being an organ of phonation with every closure of the glottis. But this sort of functional switching is not well developed in the newborn. A possible constraint in this respect resides in the fact that newborn larynx is situated close to the base of the skull and has a high position relative to the adults of all primate species (Crelin, 1969). Its high position together with velum and the epiglottis forming a seal, allows the human newborn to breathe and ingest liquids at the same time (Laitman et al., 1977). However this arrangement means that newborns are obligate nose breathers, which imposes severe constraints on the respiratory modulation of sound production. Newborn infants do not execute any manoeuvres of their supra laryngeal vocal tracts during vocalizations except for gross laryngeal

manoeuvres. The shape of their supra laryngeal vocal tract appears to approximate a uniform cross section, schwa like configuration.

As the position of the larynx in the human neonate is quite high, the epiglottis is at the level of the first cervical vertebra while the inferior border of the cricoids is at the level of fourth cervical vertebra (Noback, 1928). In adult, these cartilages are respectively at the level of third and sixth cervical vertebrae. The thyroid cartilage of the neonate lies continues to the hyoid bone (Eckenhoff, 1951) placing the epiglottis in fairly close proximity to the velum and keeping the root of the tongue within the oral cavity. The infant tongue is large, much closer to its adult size than is in the oral cavity. The mandible will undergo dramatic downward, forward growth which, together with the downward growth of the upper alveolar process, the upward growth of the lower alveolar process, and the decent of the root of the tongue in conjunction with the descent of the larynx, will eventually enclose the tongue in an oral cavity such as known in the adult (Brodie, 1949). However in the neonate, the short, broad tongue fills the entire length of the palate, laterally with the buccinators, and anteriorly with parts of the jaws so that the mouth is closed by the action of the lips (Scammon, 1923; Brodie, 1950). During the cry, the mouth is wide open and the tongue separated from its apposition with the palate and lower lip (Bosma, 1967).

Earlier discussed cry models helps researchers to understand the relationship between specific acoustical features of the cry sound and the anatomical and physiological characteristics of the infant. These understanding are important to investigate the medical applications of infant cry analysis. However, it is not enough to be able to identify the important cry features; one must be able to extract the features in an accurate and efficient way. For the same, several techniques have been developed for the extraction of acoustical data.

2.3 Cry Analysis Techniques

Different cry analysis across year includes auditory analysis, Time domain analysis, Frequency domain analysis, spectrographic analysis and the recent, computer assisted analysis. Following section summarizes each analysis techniques separately.

2.3.1 Auditory analysis

Human ear is the most readily available means for cry analysis. Hippocrates was the one who described about the diagnostic listening in ancient times. Flatau and Gutzmann (1906) used gramophone to the recordings of 30 neonates and noted 3 infants with higher pitched phonation. Lewis (1936) used International Phonetic Alphabet (IPA) for the first time in an attempt to describe infant cry vocalizations. Wasz-Hockert, Partanen, Vuorenkoski and Valanne (1964) identified hunger, pain, pleasure, and birth cries auditorily using tape recordings. Partanen et al. (1967) demonstrated that the pain cries of healthy infants could be differentiated from the cries of sick babies with one of the following diagnoses: neonatal asphyxia, neonatal brain damage, neonatal hyper bilirubinemia and Down syndrome. It was shown that after a training period of approximately 2 hours 82 paediatricians could diagnose normal versus pathological cries very accurately and differentially diagnose the specific pathology somewhat less accurately.

These studies show that medical information can be obtained from listening to cries. This readily available method of analysis can be improved with experience and training. However it provides only minimal information contained in the cry signal and for significant diagnostic information, more sophisticated techniques are required.

2.3.2 Time domain analysis

For the analysis, time domain information is obtained from devices that graph sound magnitude versus time on a paper strip chart, (Fisichelli & Karelitz, 1963). Wasz- Hockert,

Lind, Vuorenkoski, Partnen, and Valanne (1963) used a direct writing oscillograph to study the time course of the durations and latencies of different kinds of crying. They found irregular phonations in the initial cry record than those phonations which appeared once the infant is fully aroused. After the arousal gradual reduction in both time and intensity of the cry units occurred until the baby stopped crying. Using the measurement of inspiratory as well as expiratory phonations, Wolff (1967, 1969) explained the duration difference between hungry, mad, pain produced and teased crying in 4 day old infants. He also indicated that in pain induced cries, the cry units (one expiratory phonation) are longer in the beginning of the cry recorded than at the end. Time domain instruments had the advantage of being relatively easy to operate and it is inexpensive, reliable and easy to inspect visually. However, it had limitations with this method as presence of signal distortion due to pen inertia and paper speed variation that resulted in poor frequency response.

2.3.3 Frequency domain analysis

Devices performing frequency domain analysis allow one to obtain a representation of the frequency spectrum characteristics of a sound. They utilize a bank of band pass filters. These filters only allow input of a specified frequency range, measure the average magnitude in that range, and give a visual display of the relative magnitude. One can then compare the relative magnitude of a series of frequency ranges. Ostwald, Freedman, and Kurtz (1962) examined the cries from 32 twins using half octave band analyzer. They determined that the variability in pitch measurements and temporal characteristics between the cries of twins could be explained by differences in "weight, size, physical development and vigor of the children recorded". Ostwald (1963) used half- octave analysis to analyze the cry of a normal neonate and found the fundamental frequency to be between 425 and 600 Hz.

Frequency domain analysis devices give information only about the relative magnitude of various frequency ranges and do not give information about timing. Only

limited value frequency information is obtained because of the large and inflexible bandwidth used in the band pass filter.

2.3.3 Spectrographic analysis

Spectrographic analysis integrates both time and frequency domain analysis to produce a permanent visual record which shows the distribution of energy in both frequency and time. It has served as an important and useful device in many areas of signal processing. It was originally developed by Bell Laboratories in the late 1940's. Most studies of the infant cry have utilized the sound spectrogram. Figure 2.4 is the picture representing the Spectrogram -700 series. Several nomenclatures are used in spectrographic cry analysis to explain the cry characteristics. Sirvio and Michelsson (1976) and Michelsson (1980) have made detailed reviews of the various cry parameters measured. Figure 2.5 shows the schematic representation of cry features.

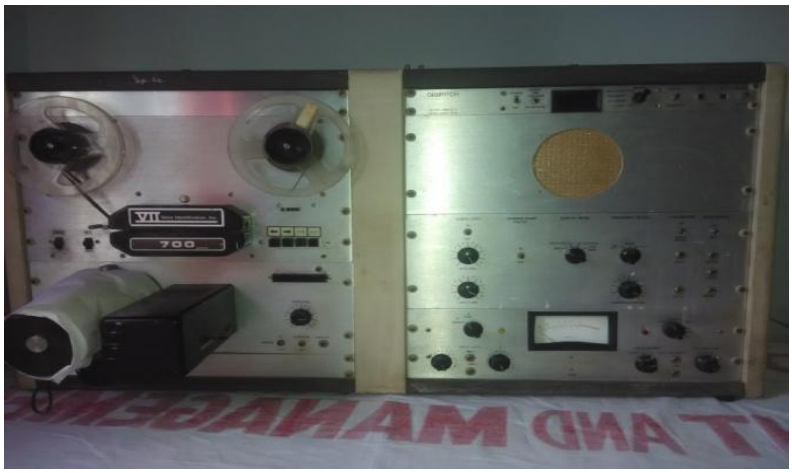


Figure 2.4., Photo representing Spectrograph -700 series

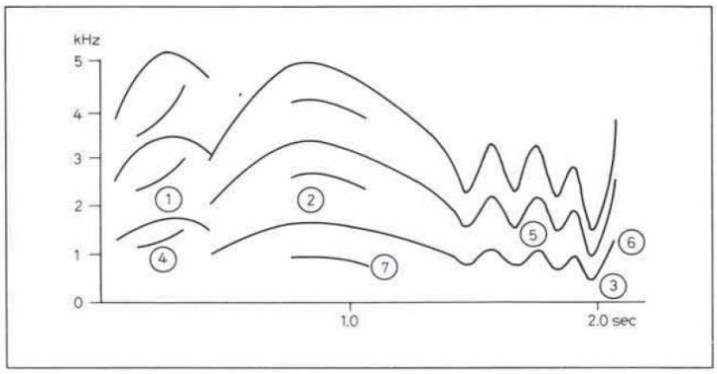


Figure 2.5., 1 = Shift; 2= maximum pitch;3= minimum pitch;4=biphonation;5=vibrato;6=glide; 7= double harmonic break. (Michelsson & Sirvio, 1967)

Scandavian research headed by Wasz- Hockert et al. (1968) defined and classified spectrographically based cry parameters in to two general categories as durational features and fundamental frequency features. Durational features: latency, duration, continuity of the cries and fundamental frequency related measures: F0, shift, glottal roll, vibrato, double harmonic break, bi- phonation, gliding, furcation, glottal plosives. Short description and picture representation of these features are given in the following section

2.3.3.a *Durational Measures*

i. Latency period

Defined as the time between the application of pain stimulus and the on set of crying, latency has been studied using spectrography and direct writing oscillograph. Stimuli for eliciting pain cry were vaccination, pinching, stimulating the child's foot etc. It is a good indicator to evaluate CNS abnormalities and it can be influenced by states of arousal, hunger or time during respiratory cycle (Wasz-Hockert et al., 1968).

ii. Total Cry Duration

Total cry duration is the time from the onset of crying to the end of last phonation before inspiration, independently of whether the signal is continuous or consists of several short phonations. It is one of the easiest part of the cry signal to measure. Usually the onset of the phonation is marked by the high intensity and hence easy to identify on a spectrogram. Decreasing intensity is marked as fade- out on the spectrogram. Inspiration cannot be visualized with ease on a spectrogram and hence auditory perception should always aid the acoustic analysis.

iii*. The Second Pause*

Vuorenskoski et al. (1966) and Michelson and Sirvio (1976) defined the second pause as the time between first and second vocalization. Sirvio and Michelsson (1976) define the end of the vocalization to the next inspiration. The second pause, if measured to be greater than it was considered as abnormal in studies by Vuorenskoski et al. (1966) and Michelsson and Sirvio (1976).

2.3.3.b *Frequency Measures*

i. *Fundamental Frequency*

The maximum pitch and minimum pitch have been measured in all studies. Maximum pitch is the highest measurable point of the fundamental frequency seen on the spectrogram and Minimum pitch is the lowest measurable point in the F0 contour seen on the spectrogram. (Wasz- Hockert et al., 1968) Figure 2.6 depicts the spectrograph of maximum pitch.

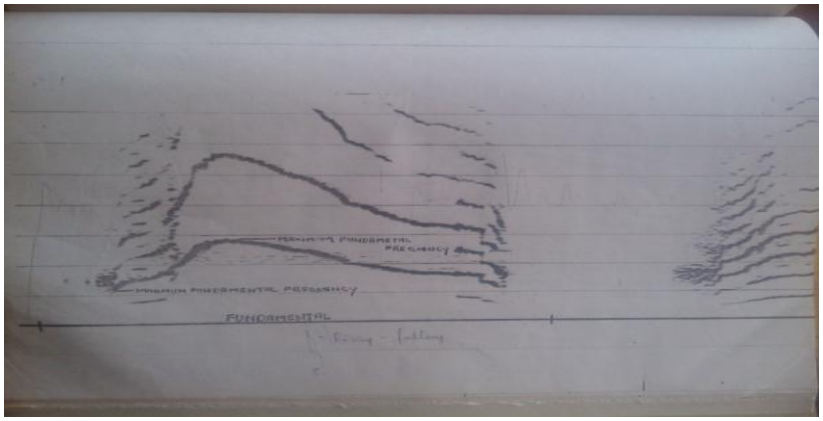


Figure 2.6., Spectrogram depicting the maximum fundamental frequency

ii. *Continuity of signal*

Any signal containing two or more vocalisations of 0.4 sec or more in length is classified as an interrupted signal (Wasz- Hokert et al., 1968). Figure 2.7 shows an interrupted pain cry.

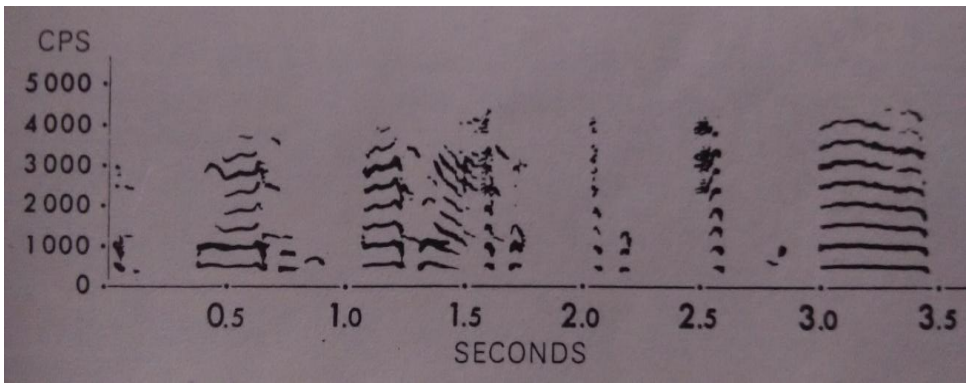


Figure 2.7., Spectrogram depicting an interrupted signal (pain cry) with many vocalizations more than .4 of a second

iii. Shift

Pitch shift is a sudden upward or downward change in pitch level (Wasz-Hockert et al., 1968). Shifts occur in almost every third pain cry of healthy infants, mostly at the beginning of the signals. It can also occur at the end of phonations or in the middle. Shift is usually high pitched than the fundamental and the maximum pitch values can be up to 1000 to 2000 Hz (Theoden & Koivisto, 1980). Figure 2.8 depicts pitch shift in spectrogram.

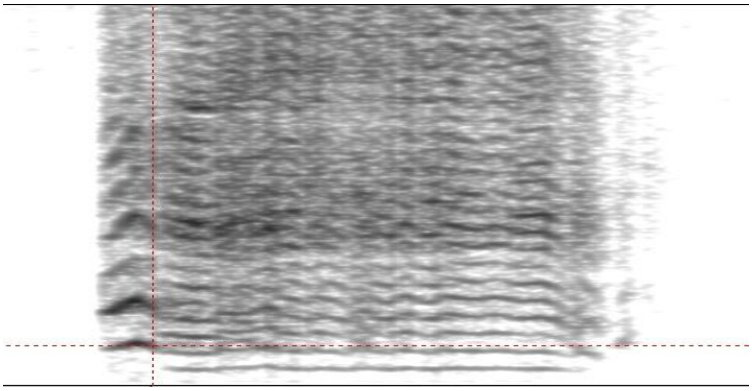
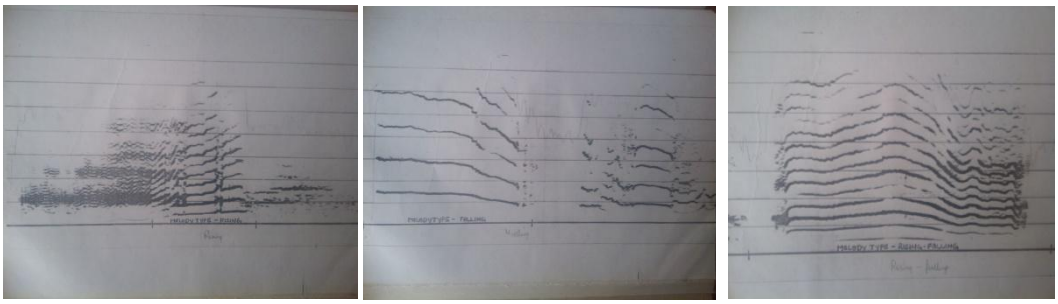


Figure 2.8., Spectrogram demonstrates frequency shift (Abbs 2015)

iv. Melody Type

The different melody types have been defined as a change in the pitch level, when exceeding 10 percent of the pitch during more than 10 percent of the length of the cry (Wasz-Hockert et al., 1968). Melody type has been classified as Falling, rising- falling, rising, and falling-rising and flat. Stark, Rose, and McLagen (1975) described melody as a flat, rising or falling pitch contour, or a combination of these. Figure 2.9 depicts different melody types.



(a) (b) (c)

Figure 2.9., Spectrogram depicting the melody type (a) rising type (b**)** Falling pattern (c) Rising-Falling type

v. Glottal roll

Vocal fry/glottal roll is a periodical phonation of the vocal folds in a lower frequency range, that is below the normal pitch register. It appears quite often at the end of a phonation, falling in pitch and intensity. The vocal fry is seen in the spectrograms as trembling, narrow

harmonics of a low intensity (Wasz-Hockert et al., 1968). Figure 2.10 depicts the glottal roll in a spectrogram.

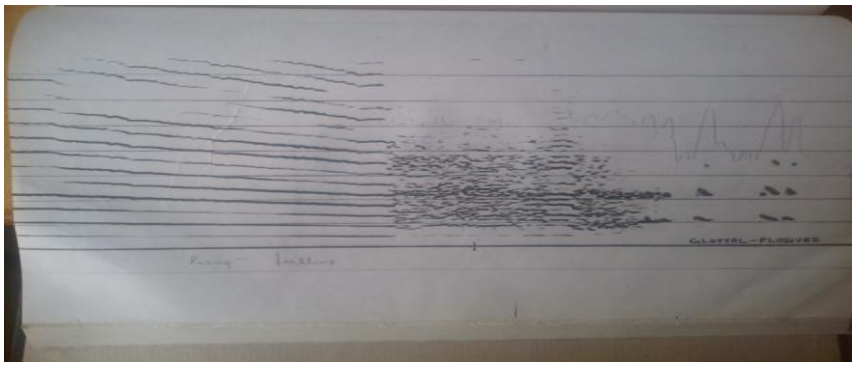


Figure 2.10., Spectrogram depicting the glottal role

vi. Vibrato

It is characterized by frequency variations that appear more clearly in the upper harmonics. It is noted in the signals when at least four successive, rapid up and down vibrations have been noticed (Wasz-Hockert et. al 1968).Figure 2.11 depicts the vibrato in spectrogram.

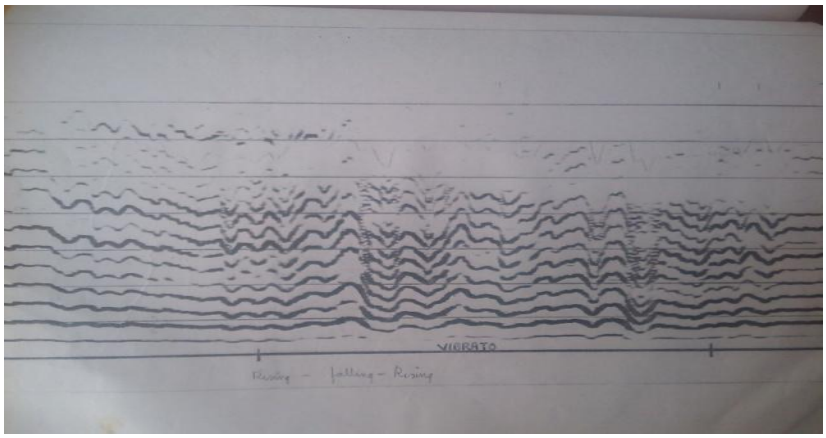


Figure2.11., Spectrogram depicting the vibrato.

vii.Double- Harmonic Break

Also called sub-harmonic break, is a parallel series of harmonics which have the same melody pattern as the F0 which gives an impression of roughness of the voice. (Wasz-Hockert et al., 1968). Figure 2.12 depicts the sub harmonic peak.

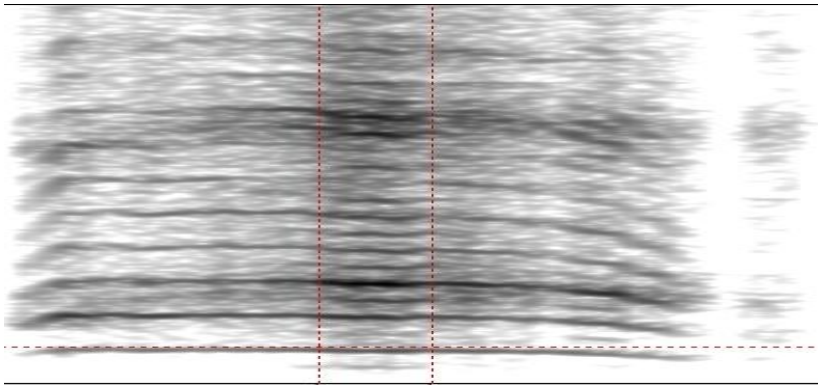


Figure 2.12., Spectrogram depicting the Double–Harmonic (Abbs, 2015)

viii. Bi-Phonation

Also called diplophonation, is a double series of F0 but without a parallel melody form. One series can be falling while the other could be simultaneously rising. It is included in studies when exceeding 0.1s (Michelsson, Sirvio, & Wasz-Ho¨ckert, 1977b) or 0.2 s (Wasz- Hockert et al., 1968). Figure 2.13 depicts the biphonaion in spectrogram.

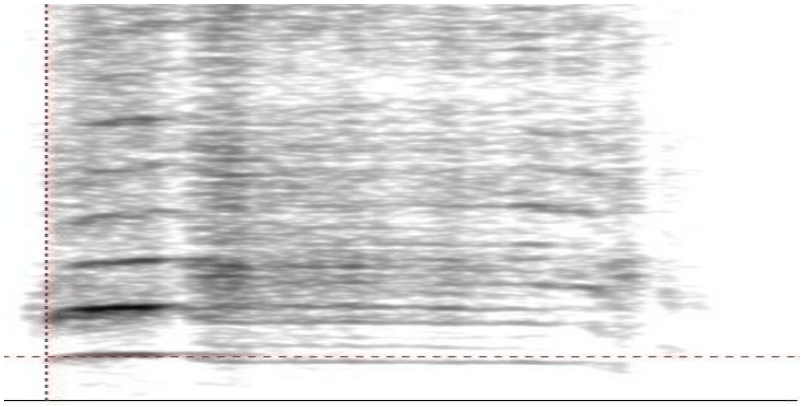


Figure 2.13., Spectrogram depicting the Bi-Phonation (Abbs, 2015).

ix. Gliding

It is the upward and/or downward movement of the fundamental frequency with generally a short duration. Change in frequency must at least be 600Hz in 0.1 seconds (Wasz-Hockert et al., 1968). Stark et al in 1975 have defined it to be a rapid pitch change of 100 Hz in 0.1 s. Figure 2.14 depicts the gliding.

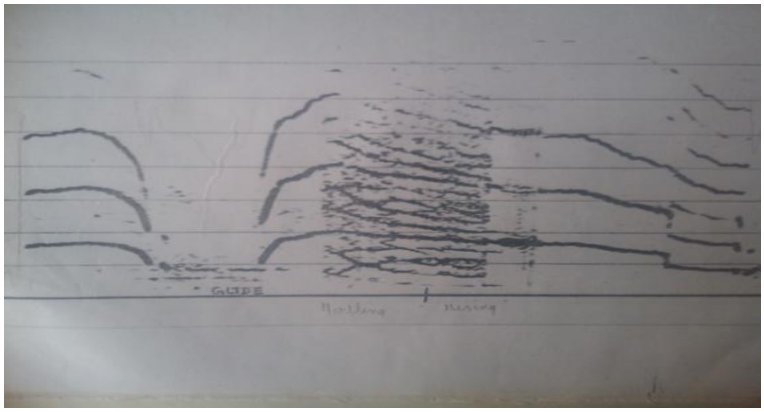


Figure 2.14., Spectrogram depicting Glide.

x. Furcation

This term denotes a "split" in the F0 where a relatively strong cry signal breaks into a series of weaker ones with each of them having its own F0. This is a feature mainly in pathological cries (Wasz- Hockert et al., 1968). Figure 2.15 depicts the furcation.

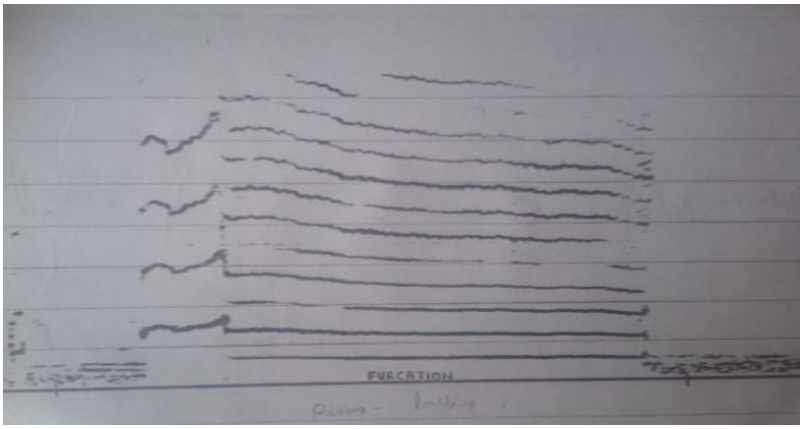


Figure2.15., Spectrogram depicting Furcation.

xi. Tonal Pit

This refers to a rapid downward and upward movement in the F0. When the fall in pitch exceeds 30% and occurs in less than 0.4s, tonal pit measurements are included. Figure 2.16 depicts the tonal pit

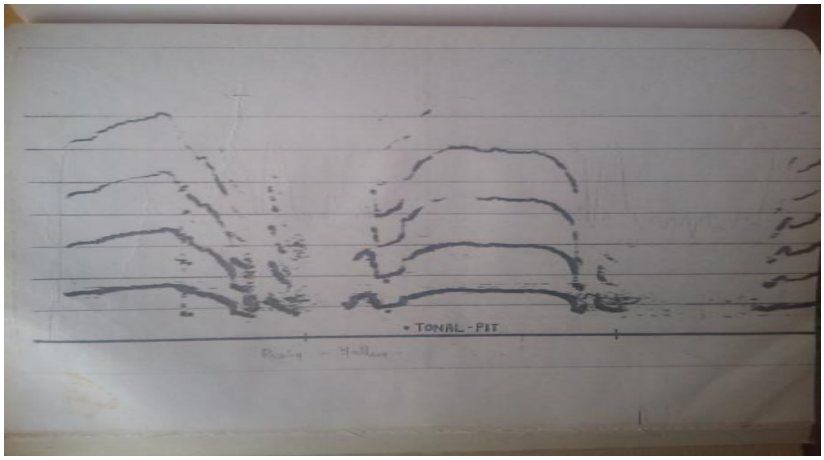


Figure 2.16., Spectrogram depicting Tonal Pit.

Numerous studies have been carried out using spectrographic analysis and authors have mentioned the cry characteristics based on the above mentioned features. And the major contribution was from Michelsson et al. (1977). Table 2.2 shows the brief summary of the studies and respective results.

Table 2.2

Summary of the spectrographic studies

Characteristics	**Authors**	**Type of Cry**	**No.of infnts**	**Age**	**Latency (Sec)**
Latency	Michelsson (1971)	Pain	50	1-10 Days	1.8
	Lester (1977)	Pain	12	Newborn	1.56±0.46
Duration	Michelsson (1971)	Pain	50	1-10 Days	2.0
	Lester (1977)	Pain	12	Newborn	2.54±2.23
Fundamental Frequency	Michelsson (1971)	Pain	50	Newborn	620Hz
	Lester (1977)	Pain	12	Newborn	466±83

Spectrogram had been a useful tool for the advancement of our understanding of infant cry analysis. It is relatively inexpensive and is a good way to visualize acoustic signals. However, it has several limitations which prevent its widespread use in medicine. Physical limitations of spectrogram include the poor dynamic range and inadequate frequency

resolution. It requires visual inspection of the output for interpretation. And expertise is needed to carry out the long and tedious analysis process thus it is not suitable for large scale database (Golub & Corwin, 1982).

Earlier discussed analysis systems give useful acoustical information. However they all have a limitation that the extraction of the acoustical information is a difficult and tedious process. Computer analysis allows more accurate determination of the acoustical information and allows extraction of information that would otherwise be unobtainable.

2.3.4 Computer- based signal processing

The analysis procedure consists of five major steps. (1) recording of the cry; (2) obtaining the parameters of fundamental frequency, formants and amplitude v/s time; (3) sampling the complex F0 contours in order to facilitate the development of F0 features; (4) developing a number of features from the parameters and samples by procedures that include averaging within cry modes and calculation of probability of being in any mode at any point of cry; and (5) conglomerating relevant features into a set of "diagnostic tests". Figure 2.17 shows the computerized spectrogram.

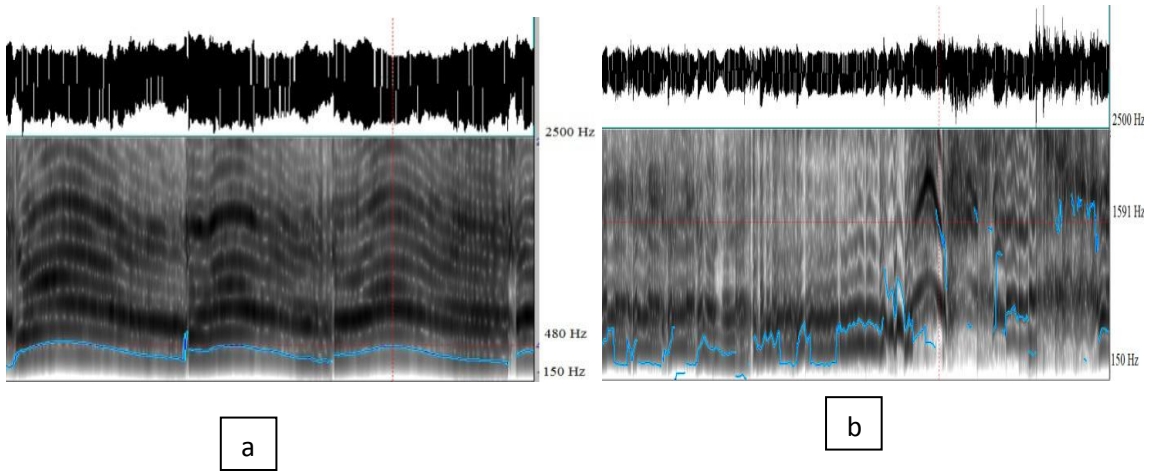


Figure 2.17., Waveform and spectrogram of (a) healthy and (b) pathologic cries (Premature newborn suffering from microcephaly*)* (Kheddache & Tadj, 2013).

Along with the features which are obtained from spectrogram, several acoustic features and quite a lot of feature extraction procedures are used for infant cry analysis. Various features such as LPC-based features (LPC, LPCC, etc.), MFCC and filters bank coefficients are used for feature extraction. There are several feature extraction steps which are generally used in the recent studies. Lederman (2002), in an attempt to classify the infant cries have listed out the common feature extracted procedures of infant cry. Following section briefs those feature extraction procedures.

2.3.4a Pre-Emphasis

The first step of the feature extraction procedure is filtering the signal with a pre-emphasis filter with a transfer frequency; and the main purpose of this step is to effectively eliminate the spectral contributions of the larynx and lips so that the analysis can be asserted to be seeking parameters corresponding to the vocal tract only (Rabiner & Schafer, 1978).

2.3.4b Blocking into Frames and Frame Windowing

The cry signals are non-stationary. Hence, short term analysis must be used in which the signal is divided into quasi-stationary overlapping frames and each frame is multiplied by a suitable Hamming window. The frame length is pre-chosen in order to ensure quasi-stationary of the framed signal on one hand, and on the other hand, to ensure that each frame will include at least one period of the fundamental frequency and to decrease the amount of computations as much as possible. In cry signals the fundamental frequency is known to have a wide range of possible values (about 200- 2500 Hz). Consequently, the appropriate practical range for frames duration should be around 5-25 msec.

2.3.4c Mel-Frequency Cepstrum Coefficient

Many studies have used Mel-Frequency cepstrum coefficient (MFCC) as input feature for infant cry classification analysis (Orozco-García & Reyes-García, 2003; Lederman, et al., 2004; Reyes-Galaviz, Tirado, & Reyes-Garcia,2004; Galaviz & García, 2005).

2.3.4d LPC Analysis

Contrary to other techniques, linear prediction (LP) analysis analyse the acoustic signal in time domain. Linear prediction coefficients (LPC) are a small number that can be used to represent the spectral envelope of the cry signal. The signal power spectrum can be predicted in more detail as the order of LP increase. Autocorrelation LP technique is one of the methods used to segment the cry signal. There are few researches that have used LPC as a selected feature to differentiate between different types of cry (Orozco-García & Reyes-García, 2003; Reyes Galaviz & Reyes Garcia, 2004; Galaviz & García, 2005; Lederman, Zmora, Hauschildt, Stellzig-Eisenhauer, & Wermke, 2008; Santiago-Sánchez, Reyes-Gracia, & Gomez-Gil, 2009; Hariharan, Chee, & Yaacob, 2012). LP analysis technique was used to extract the fundamental component of the cry signal for the pathological cry from healthy baby and baby with hearing impairment (Orozco-García & Reyes García, 2003).

Recently, researchers are widely making use of computer assisted methods or commercially available speech analysis software packages to analyse acoustic aspects of infant cry. Goberman and Robb (1999) made use of computerised method of LTAS (long term average spectrum) display for the acoustic analysis of cry features in preterm infants. Fundamental frequency extraction of inspiratory cries were analyzed using an integrated computer hardware/software package (KAY Elemetrics, CSL-4300) and the software permits digitization, editing, and storage of cry samples (Grau, Robb & Cacace, 1995). Fuamenya, Robb & Wermke (2014) also made use of the similar software package (CSL-4500, Kay-

PENTAX; Montvale, NJ) for acoustic analysis on their study of quantification of the frequency of occurrence of sub harmonic and noise phenomena in healthy infant cries.

Advances in computing capacity allow researchers to utilize methods for quantifying aspects of the sound spectrum that can be expected to yield more accurate estimates of F0 and related parameters. Recently, automated approaches also came into picture. It has been reported that the automated approaches have the advantage of fast analysis of very large data sets, objective assessments, the ability to quantify multiple data points, and the flexibility to yield derivative measures (e,g., jitter, pitch contours, etc), which have the potential to increase the applied value and clinical utility of cry assessment. Automated approaches also have disadvantages such as weakness due to limits in computing power which resulted in difficulties in signal detection and not able to pinpoint important cues when multiple analyses are required to do so. Researchers have developed automated analysis systems for the quantification of cry acoustics with minimized need of visual inspection and manual processing of data in a large sample of infants. Branco, Fekete, Rugulo, and Rehder (2007) utilized automated application of FFTs to detect F0 in cry utterances across whole cry episodes. In this approach, digitized and filtered samples were subjected to an FFT in order to compute the log magnitude spectrum for analysis blocks of specific lengths (eg., 25 ms).Summary variables for each analysis block can then be aggregated to yield summary statistics for a cry episode and for the individual cry utterances (single expiratory period) within cry episodes (a series of expiratory periods).

Other than FFT, alternative automated approaches have also been described. Várallyay, Benyó, Illényi, Farkas, and Kovács (2004) explained the classical and new methods of acoustic analysis of the infant cry. Final goal was reported to detect hearing disorders according to the crying at the earliest possible moment. Thirty seven infants, out of

which 14 of them with severe hearing loss (40-80 dB) and 23 infants with normal hearing in the age range of 1-15 months old were analyzed in the study. Routine otological examination served as the stimuli for pain cry elicitation. Microphone was kept in 1 meter distance from infant's mouth and digital camera (SONY DCR-TRV25) was used for the recording. For the fundamental frequency detection authors, used Smoothed Spectrum Method (SSM) which is based on the spectral analysis and combined with noise filtering and statistical processing. Other methods for fundamental frequency detection including Cepstrum analysis, and local maximum value detection were compared with SSM. The three methods had 95.50%, 96.86% and 97.99% efficiency for Local maximum value detection, Cepstrum analysis and SSM respectively for the detection of fundamental frequency. Lederman, Cohen, Zmora , Wermke, Hauschildt, and Stellzig-Eisenhauer (2002) introduced an automatic classification system based on Continuous Density Hidden Markov Models (CD-HMM).

Galaviz, Tirado and Garcia (2004) have developed an automatic infant cry recognizer for the early identification of pathologies to classify three classes into normal, hypo acoustics and asphyxia. They used Mel frequency cepstral coefficients (MFCCs) for feature extraction and for the classification a Feed Forward Input Delay Neural Network with training based on gradient descent with adaptive back-propagation was used. The accuracy of their proposed system varies from 96.08% to 97.39%.

Orozco-Garcia and Reyez -Garcia (2003) have developed a method based on linear prediction technique and scaled conjugate gradient neural networks for the detection of pathologies from infant cry. The classification accuracy of their proposed method was 91.08% for 314 samples and 86.20% for 1036 samples. Same authors have used MFCC and linear prediction coding techniques for characterizing the infant cry signal and used a feed-

forward neural network for classification with several learning methods. The accuracy of their proposed system was up to 97.43%.

Reggianni, Sheinkopf, Silverman, Xiaoxue Li, and Lester (2013) described and validated the performance of a modern acoustic analyzer specifically designed for infant cry analysis. The authors developed a method to extract acoustic parameters describing infant cries from standard digital audio files. Frame rate of 25 ms and 12.5 ms frame advance was used. There were 2 phases in the cepstral based acoustic analysis as computing frame- level data and then organizing and summarizing this information within cry utterances. They also evaluated the accuracy of the automated system to determine the voicing and F0 as compared to the manually coded voiced segments and pitch period from the spectrogram displays. Results revealed that system detected F0 with 88% to 95% accuracy, depending on tolerances set at 10 to 20 Hz. Receiver operating characteristic analyses demonstrated very high accuracy at detecting voicing characteristics in the cry sample.

All these automated approaches have the potential to speed scientific inquiry, allowing for the study of large number of infants with efficient and rapid analyses. They do not require manual inspection of spectrograms and nor the time consuming task of cursor placement and frame selection used in few computer assisted methods. In this way, automated approaches allow for more rapid analysis that is less prone to observer bias. Since these systems are specifically for the study of infant cry, there is an assumption that these algorithms accurately track F0.

2.4 Cry Analysis Studies in Healthy and Pathological Population

Several attempts were carried out by researchers to show that cry analysis is an additional aid in making diagnosis in clinical paediatrics, especially in newborn period and in

diseases that have affected the central nervous system. Neonatal cries were frequently targeted because they reflect initial neurological state prior to possible shaping by experience. As discussed, earlier studies in the area relied on visual inspection of spectrograms to describe the acoustic features of cry in specific clinical populations. Lately, computer-assisted methods were used for the same. Following section summarises the numerous studies which have been carried out in the infant cry analysis using different analysis techniques.

2.4.1 Cry in healthy full- term infants

Various researchers have carried out cry acoustic analysis with the use of spectrogram. Michelsson, 1971 and Wasz- Hockert et al. (1963, 1964, 1968) analyzed pain cries of more than 300 healthy newborn infants. Study on normal infant crying by Thoden and Koivisto (1980) dealt with prospective analysis of cries of 38 infants from birth to 6 months of age and analyzes the first, second and third phonation after the pain stimulus. In all these studies on the pain cry of healthy newborn infants, the mean maximum pitch of the fundamental frequency without shift has been about 650 Hz, and the mean minimum pitch was 400 Hz. In 80% of the samples, the pain cry had a falling or rising/melody type with a stable pitch and duration of approximately 2.5 secs. Shifts with higher pitch occurred roughly in every 3rd cry. Two third of the cries showed voiced and continues signals. There was common occurrence of glottal roll at the end of phonations preceded by vibrato. Biphonation, glide, furcation and noise concentration were extremely rare in the cry (Wasz- Hockert et al., 1968; Michelsson, 1971; Thoden & Koivisto, 1980).

2.4.2 Low birth weight

Michelsson's (1971) study on low birth weight infants showed that full term infants who were small for gestational age did not differ in their cry characteristics from full term infants with normal birth weights. Cry results of the truly premature infants showed that the more premature the infant, the higher the fundamental frequency. Wasz- Hockert et al. (1971) reported that the more premature the infant was the more the cry differed from the crying of healthy full term infants with shorter duration and higher pitch.

Long time average spectrum (LTAS) characteristics of 10 full term and 10 preterm infants revealed significant difference in the parameters; higher first spectral peak values were observed in preterm babies which indicates higher vocal F0 (Goberman & Robb, 1999). Sangeetha (1999) compared 29 acoustic parameters using Multidimensional Voice Programme (MDVP) in normal children as well as in high-risk babies including premature babies. The study indicated that most of the acoustic parameters including frequency related and intensity related parameters can differentiate normal and abnormal cries of newborns. Minimum fundamental frequency of Very Low Birth Weight Babies (VLBWI) were found to be higher than normal infants (Rautava, Ojala, Parkkola, Rikalainen, Lapinleimu, Haataja, & Lehtonen, 2007). Shinya, Kawai, Niwa, and Yamakoshi (2014) found increased *F0* of spontaneous cries in preterm babies due to their different intrauterine and extra uterine experiences.

2.4.3 Sudden infant death syndrome

Attempts to understand Sudden Infant Death Syndrome (SIDS) have led to considerable interest in the infant's cry. The cries of these children were strange or different in some way when compared to their siblings or other infants (Naeye, Messemet, Specht, & Merritt, 1976). Stark and Nathanson (1975) obtained cry samples from a 4-day old infant

who died of SIDS at 6 months of age. The infant's cries were found to be of shorter duration and having a higher pitch than healthy infants. In contrast, Colton and Steinschneider (1981) evaluated the cries in one SIDS infant and found the cries to be of long duration with a low F0 and low vocal tract resonance (e.g. formant frequencies) compared to healthy term infant. Colton and Steinschneider (1981) study on comparison of healthy term and preterm infants during the first and fourth weeks of life stated the difference between two groups during the first week of life. The SIDS siblings produced cries with significantly greater loudness in the high spectral regions (4–8 kHz) compared to the other groups and their F0 was reported to be 40 Hz lower (though statistically not significant).

Robb, Crowell, Dunn-Rankin, and Tinsley (2007) compared cry physiology of SIDS siblings from healthy term infants. Subjects included 23 HT (Healthy term) and 6 siblings of SIDS infants. History of previous full or half sibling succumbed to SIDS (documented by autopsy) was the primary inclusion criteria for group of SIDS siblings. Each infant was audio recorded in the laboratory within 2 weeks following birth. Cries were elicited via administration of a pain stimulus to the sole of the infant's right foot and audio recorded using a condenser microphone placed at a constant 15 cm from the infant's mouth. For measurement of cry latency; an audible tone was time-locked to administration of the pain stimulus. And that served as an acoustic reference for the latency. The entire crying episodes were analyzed using integrated computer hardware software package (KAY CSL-4300) to obtain the specific measurements: overall cry duration, number of expiratory cries, cry ratio, cry latency, first spectral peak (FSP), spectral tilt (ST), and high-frequency energy (HFE). Among the various features significant differences in the values were identified between the two groups for FSP and ST. The FSP of SIDS siblings was significantly higher compared to the HT group, probably a result of stressful response to the pain stimuli. ST of the cries produced by SIDS siblings was significantly lower than the HT group. It indicates that the

amplitude of the harmonic components comprising the cry remained elevated across the acoustic spectrum (Robb et al., 2007). Lofqvist and Mandersson (1987) stated that a shallow ST in combination with considerable HFE is representative of hyper adductional vocal fold behaviour. Although HFE was not significantly higher among the SIDS siblings compared to the HT group, the overall mean HFE was noticeably larger among the SIDS siblings.

2.4.4 Asphyxia

Michelsson (1971) compared cry signals of 250 infants with asphyxia with 50 healthy full-term infants and 75 healthy preterm infants. The infants were separated into 2 groups: peripheral asphyxia- infants who suffered from apnea and central asphyxia- infants with neuorological problems. The comparison was performed separately for full-term and preterm infants. In both cases major differences in the acoustic signals between the two groups were found, and they differed significantly from normal cries in terms of maximum pitch, biphonation and melody type.

2.4.5 Chromosomal abnormalities

Veurenkoski et al. (1966) analyzed 44 cry signals of infants suffering from cri-du-chat syndrome. Cri-du-chat ("Cry of the cat") syndrome is related to chromosome No 5 which is absent in infants suffering from this disorder. The average pitch of 860 Hz was noted in 44 cries of 8 children and the melody was flat in 36% of the signals and rising in 23%. Michelsson Tupperainen, and Aula, (1980) reported similar results with higher occurrence of flat melody in two infants with cri-du-chat syndrome. Bauer (1968) also reported 600 to 1000 Hz pitch in children with cri-du- chat syndrome.

Lind, Vuorenkoski, Rosberg, Partanen, and Wasz- Hockert (1970) analyzed 120 samples of 30 infants in the age range of 0 to 8 months old with Down syndrome. The

average duration of the cry segments was found to be higher than usual about 4.5 seconds. Minimum and maximum fundamental frequencies were 270 Hz and 510 Hz with flat melody type in 63% of the signals. In a cry analysis of 14 infants with various chromosomal abnormalities, infants with 13 and 18 trisomy showed low pitched and monotonous cries. On the other hand, high pitched cry was noted in infants with chromosome numbers 4 and 5 abnormality.

2.4.6 Diseases of the central nervous system

Michelsson et al. (1977) study on 14 infants with meningitis reported a maximum fundamental frequency of 750 Hz along with a minimum fundamental frequency of 560 Hz, and a rising or falling-rising melody in 24% of the cases. Bi-phonation was found in 49% of the infants and gliding in 11%. The authors concluded that meningitis could be detected early based on the cry signals. They also studied cries of infants with herpes simplex virus encephalitis and found that noise concentration occurred in a frequency range of 2000-3000 Hz. Cries were higher pitched and frequent occurrence of bi-phonation and gliding were more common than healthy controls. In another study (Michelsson, Karskinen, Aulanko, & Rinne, 1984) they examined 248 cry signals of 62 infants suffering from hydrocephalus. Findings showed a maximum and a minimum fundamental frequency of 750 Hz and 430 Hz respectively, a flat melody, bi-phonation in 14% of the cases and gliding in 8%.

2.4.7 In utero drug exposure

Corwin et al. (1992) investigated cries of 404 cocaine-exposed infants and compared to 364 non-exposed infants. Significant differences between the two groups were found including fewer cry utterances, more short cries, and less hyper phonated cry for the cocaine-exposed neonates. The acoustic cry characteristics of drug-exposed infants were investigated

by Lester et al., (2002). Both the groups were compared adjusting for the following covariates: alcohol, marijuana, tobacco, birth weight, social class, and recording site. Cry of cocaine exposed infants reported to have more energy, higher fundamental frequency and a lower second formant than the cry of infants not exposed to cocaine. In regard to opiate exposure, opiate-exposed infants had fewer short utterances and hyper phonation than infants not exposed to opiate.

Kheddache and Tadj (2013) analyzed cries of healthy and newborn with different categories of diseases in order to evaluate a fundamental frequency of these cries. Data base included 2800 cry samples of 1s duration from 48 of newborn babies aged 1 to 30 days. 1774 cries were from healthy newborn (among them 764 are premature) and 1010 from newborn who present some diseases (among them 628 are premature).The conditions were hunger, sampling blood and change of diapers. For the F0 estimation, initially the recorded samples were noise filtered and segmented into useful and non-useful portions. Further the SIFT algorithm (Simple Inverse Filtering Tracking) were used for the estimation of F0 in short intervals typically 20 ms interlaced frame with 10 ms recovering. The main steps of SIFT algorithm includes the a) division of signal in to overlapping frames of 20 ms with 10 ms recovering, b) multiplication of each frame by Hamming window c) performing glottal inverse filtering d) autocorrelation sequence estimation e) peak picking and decision algorithm with voiced threshold f) F0 estimation, T0 = 1/F0 g) smoothing using median filter. The range of F0 for pathologic cries and healthy cry were in the range of 150 Hz-1600 Hz and 450 Hz respectively. Among pathological conditions, neurological problems (asphyxia, microcephali- IUGR (Intra-Uterine Growth Retardation) showed highest average percentage of hyper-phonic cries and the highest average percentage of high pitched cries with 750 < F0 < 1000 Hz was found in the category of neurological problems (micro-cephaly-IUGR).

A collation of literature reports on different medical conditions were reviewed and tabulated below. Table 2.3 (a &b) lists the studies of medical conditions and associated cry measures indexed by acoustic features. These studies support the relationship between neurological status and cry.

Table 2.3 (a)

Studies on different severe medical conditions and the respective cry measure results

Medical Condition	**Authors**	**Cry Measure**
Asphyxia	Michelsson, (1971); Michelsson et al., (1977a); Partanen et al., (1967)	Increased F0, increased F0 instability (biphonation), Increased sub harmonic break, increased or decreased duration
Brain damage	Fisichelli and Karelitz, (1966); Karelitz and Fisichelli, (1962); Sirvio¨and Michelsson, (1976); Wasz-Ho¨ckert et al., (1968)	Increased F0, increased F0 instability (biphonation), decreased duration, increased threshold, increased latency, Increased short utterances
Cri-du-chat	Vuorenkoski et al., (1966); Wasz-Ho¨ckert et al., (1968)	Increased F0
Down syndrome	Fisichelli and Karelitz, (1966); Wasz-Ho¨ckert et al., (1968)	Decreased F0, increased F0 variability, decreased intensity (amplitude)
Hydrocephalus	Michelsson et al., (1984)	Increased F0, increased F0 instability, increased latency
Hypothyroidism	Michelsson and Sirvio (1976)	Decreased F0
Krabbe's Disease	Thode´n and Michelsson, (1979)	Increased F0
Meningitis (Bacterial)	Michelsson et al., (1977b)	Increased F0, increased F0 instability (biphonation), decreased duration
Trisomy 13, 18, 21	Michelsson et al., (1980); Lind et al., (1970); Ostwald et al., (1970)	Decreased F0
SIDS	Stark and Nathanson (1975) Colton and Steinschneider, (1981)	Shorter cry utterances, lower amplitude, extremely high F0 events with rapid shifts.

Continued…..

Continued…..

	Golub and Corwin, (1982); Corwin et al., (1995)	Decreased F0, F_1, F_2, decreased utterance duration, decreased sound pressure, episodes of F_1 _2000 Hz and F_2 at 4000 Hz
	Corwin et al., (1995);	Increased F_1
	Hoffman, Damus, Hillman, & Krongrad (1988)	Increased F_1, increased mode changes

Table 2.3 (b)

Studies on Potential Central Nervous System Insults and the respective cry measure results

Medical Condition	**Authors**	**Cry Measure**
Low birth weight (< 2500 g), small for gestation	Michelsson, (1971)	Increased duration, decreased F0, Increased F0, increased F0 instability (biphonation)
Prematurity	Michelsson et al., (1983)	Increased F0, decreased duration
Preterm infants	Lester, (1987])	Increased F0, increased F0 variability, decreased F_1 variability, decreased amplitude
	Corwin et al., (1992)	Increased short cry utterances with decreased developmental outcome (30 mo)
Hyperbilirubinemia	Wasz-Hockert et al., (1971)	Decreased latency, decreased duration, increased F0, increased F0 variability
	Golub and Corwin, (1982)	Unstable glottic function [mode changes]
	Vohr et al., (1989)	Increased F_1 variability, increased phonation
Lead exposure	Rothenberg et al., (1995)	Low % nasalization, decreased number of cries, increased F0

Continued……

Continued…..

Prenatal opiate exposure	Blinick, Tavolga, & Antopol (1971)	Increased F0, increased likelihood of abnormal cries
	Corwin, Golub, & Potter (1987)	Increased hyperphonation, increased short utterances, increased F0, increased duration of first cry utterance associated with withdrawal symptoms
	Lester et al., (2002)	Increased short utterances, increased hyperphonation
Prenatal marijuana exposure	Lester and Dreher, (1989)	Shorter cries, increased dysphonation, increased F0, increased F0 variability, decreased F_1
	Nugent et al., (1996)	Increased dysphonation, increased F_1
	Lester et al., (2002)	Increased mode changes, increased F_2
Prenatal cocaine exposure	Corwin et al., (1992)	Decreased cry utterances, increased short cry utterances, decreased hyperphonation
	Lester et al., (1991)	Direct effects (excitation) increased duration, increased F0, increased F_1, increased F_1 variability; Indirect effects via growth retardation (depression) increased latency, decreased utterances, decreased amplitude, increased dysphonation
	Lester et al., (2002)	Increased F0, decreased F_2, increased dysphonation, increased amplitude, increased duration 2nd utterance

Plentiful studies on infant cry analysis on both normal and clinical population attempt to develop automated cry analysis tools that highlight the value of cry analysis for determining the medical status of infants. Literature evidences that various acoustic features of the cries of healthy infants are clearly different from the cries of unhealthy infants. A reliable acoustic analysis of cry can serve as a promising newborn screening tool that is a non-invasive means of detecting irregularities in an infant's neurobehavioral integrity. Hence, the present study aimed to develop a quick computerized screening tool for infant cry analysis.

CHAPTER 3

Method

The primary aim of the study was to develop a computerised screening tool for newborns through cry analysis.

3.1 Study Design

The study followed a Standard Group Comparison design including comparison of parameters of infant cry between control group and clinical group.

3.2 Participants

Two groups of newborns were selected for the study. Group I was a control group which consisted of 220 healthy newborns within 7 days of birth with average mean age of 4 days (during the immunisation procedure, before the discharge of the mother and the baby from the hospital).

The inclusion criteria for Group I were,

- No maternal complications during the pregnancy period such as infections, exposure to toxicated drugs, maternal chronic illness, maternal nutritional deficiencies and any other reported by the mother and family members.
- Normal gestation period of within 37 to 42 weeks.
- Normal birth weight of >2500 to 3500 gm
- No other complications/ risk factors during natal or post-natal period such as birth asphyxia, viral/bacterial infections, convulsions, Rh isoimmunisation, delayed birth cry, premature delivery, aspiration of amniotic fluid, neonatal jaundice and also the baby should not be isolated in the NICU for >2 days.

- Passed the screening tests for hearing on Behavioral Observation Audiometry, TEOAE, and showed normal pattern on examination of Primitive reflexes

Group II consisted of 92 newborns within 30 days of birth with one or more risk factors. This was because these babies were kept in the NICU and the admission to the NICU was not permitted for the cry recording within a span of 7 days and it was possible only after the initial medical care. The risk factors considered for the study included low birth weight (not more than 2000 grams) birth asphyxia, meconium aspiration syndrome, infantile jaundice and other disorders. Majority of the high risk newborns were placed in neonatal intensive care unit (NICU) for more than 2 days. The information required was collected through the medical records of the newborn in the hospital and an informal interview with parents/care takers.

3.3 Procedure

The study was conducted in 4 phases, which included

Phase I: Recording of cry samples from newborns

Phase II: Acoustic analysis of newborn cries

Phase III: Development of Computerized screening tool software

Phase IV: Pilot Validation of the developed screening tool

After obtaining an informed consent from the parents/caretakers, demographic information of their infants was obtained and the purpose of the study was subsequently explained to them before proceeding with the recording of the cries.

3.3.1. Phase I- Recording of cry samples from newborns

Cry samples of both healthy and high risk newborns were recorded using Olympus LS-100 Multi track linear PCM audio recorder with 48KHz sampling frequency and 24 bits

quantization, positioned in front of the newborns with an external directional microphone placed 10 cm away from newborn's mouth. Care was taken to minimize the environmental noise with minimum disturbances. The pain stimulus was used to elicit the cry during the immunization procedure (Hepatitis B injection) for healthy newborns by a staff nurse in a hospital set-up. Audible finger snap (around 50 dB) by the researcher served as an indicator for vaccination of newborns by the staff nurse. This served as an acoustic reference for measurement of latency of the infant cry. The pain cry was recorded after the vaccination to get as standardised a cry sample as possible for all infants. This way it was possible to avoid causing unnecessary discomfort to the child, as vaccination is mandatory in the routine schedule for newborns. Furthermore, pain cry has been the most commonly studied cry type used to compare group differences in cry acoustics (Rautava et al., 2007).

However, the recording of cry for at-risk newborns was done using toe tapping as stimuli to elicit pain cry instead of pain cry elicited during the immunization protocol as followed for healthy newborns. The same procedure was maintained for all the infant cry recorded.

The stimuli to elicit cry for at risk newborn was decided based on a pilot study to determine if there was any difference in the acoustic characteristic of infants cry elicited using vaccination or toe tapping procedure. Fifteen healthy newborns were selected randomly for the study. The cry was recorded twice from the same healthy newborns using the two different procedures. PRAAT software version 5.4.01 was used to extract acoustic parameters of duration, frequency, intensity and formants. The non-parametric Mann- Whitney U-test results indicated that there was no significant difference ($p \geq 0.05$) for the acoustic characteristics between the two procedures except for perturbation related parameters (Jitter & Shimmer). Earlier studies also have reported that variations are possible in jitter and shimmer related

parameters depending on the method used. Oguz, Kilic, and Safak (2011) reported there was significant variation in jitter and shimmer values for the same voice sample in two different softwares (Praat and MDVP). Since, there was no significant difference in the parameters extracted with either of the two methods, toe tapping was used to elicit cry from the at-risk population infants. As the pilot study showed variability in jitter and shimmer values, they were excluded from the acoustic analysis of cry signals. The parameters that were extracted using the PRAAT software version 5.4.01are listed in Table 3.1.

Table 3.1

Extracted acoustic parameters

Frequency (Hz)	**Intensity (dB)**	**Duration (s)**	**Noise**	**Formants (Hz)**	**Other parameters**
Mean pitch Minimum pitch Maximum pitch Median pitch Standard deviation of pitch	Mean intensity Minimum intensity Maximum intensity	Latency Voicing duration	Mean Noise to Harmonic Ratio (NHR) Mean Harmonic to Noise ratio (HNR)	First Formant (F_1) Second Formant (F_2) Third Formant (F_3) Fourth Formant (F_4)	Number of pulses Number of periods Number of voice breaks Degree of voice breaks

3.3.2 Phase II -Acoustic analysis of cry samples

Acoustic analysis of infant cry was performed by extracting the parameters from the cry samples of newborns from both Groups I and II. The recorded and collected infant cry samples in Mp3 format were transferred and stored in the 64 bit window 7 operating system computer. PRAAT software version 5.4.01 had been used as a software tool for analyzing the cry samples. Figure 3.1 shows the sample diagram of a cry signal. The cry sample were analyzed and the acoustic parameters such as frequency, intensity, formants and other parameters as mentioned in Table 3.1 were extracted.

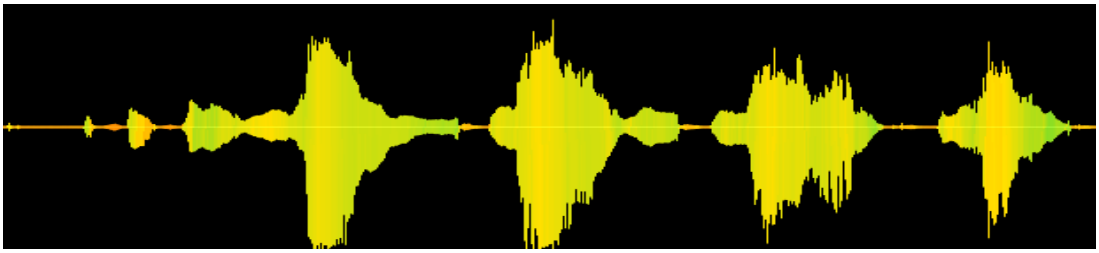


Figure 3.1., A sample of Infant cry waveform

For the analysis of infant cry sample, a total duration of 15s was considered for measurement of various acoustic parameters. This was performed to trim down the total analysis duration in the computerized module. The duration of the infant cry to be considered for analysis was based on results of a pilot study that compared cries of same infants having durations of 10s, 15s and 20s respectively. For this, 15 healthy newborns were selected randomly for the study. The recorded infant cry was analyzed using three different duration (10 sec, 15 sec & 20 sec) in the same sample. PRAAT software version 5.4.01 was used to extract the acoustic parameters of duration, frequency, intensity and formants. The Non-parametric Mann- Whitney U-test results indicated that there was no significant difference ($p \geq 0.05$) for the acoustic characteristics between the analysis of three different durations of the infant cry. However on observation, the first 10s duration of the infant cry had very few cry bouts in majority of the cry samples. Hence, first 15s duration was considered for analysis which contained minimum of 2-3 bouts of infant cry.

Acoustic analysis was performed for first 15 seconds duration from the onset which was selected and segmented from the cry samples to extract the acoustics parameters. Fundamental frequency, intensity, formant frequencies and other parameters were extracted from those samples using wide band spectrogram as shown in figure 3.2 and 3.3a and b.

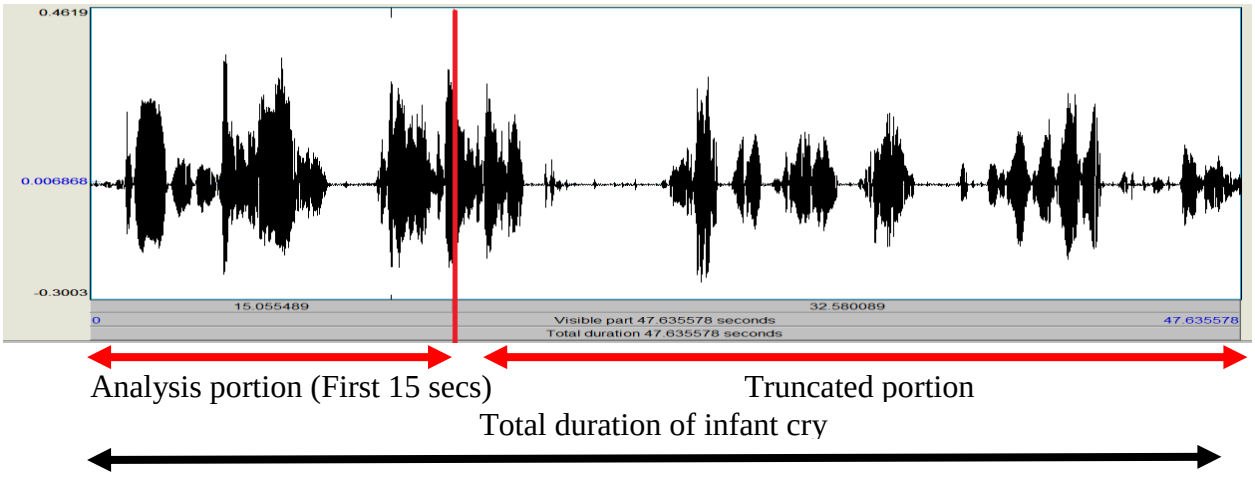
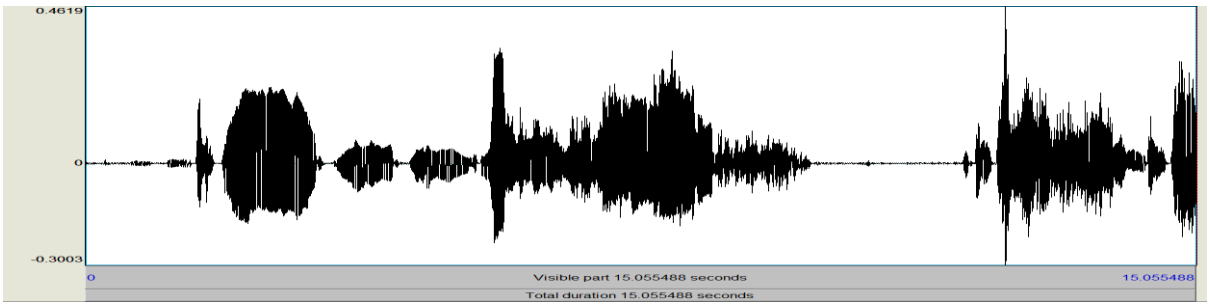


Figure 3.2., Segmentation of first 15 sec duration of the infant cry from total duration

(a)



(b)

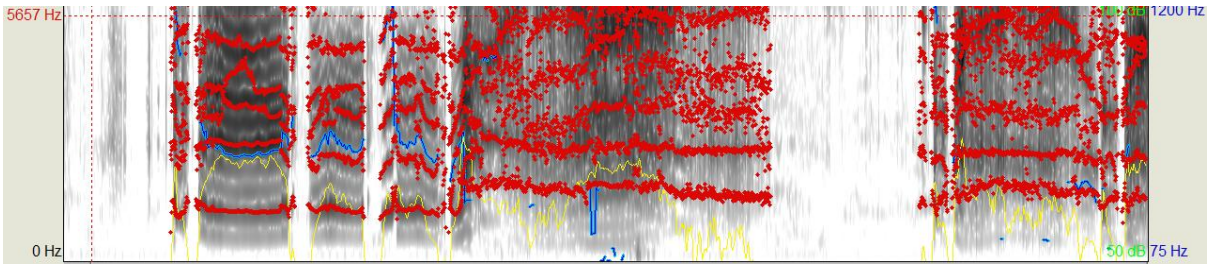


*Figure 3.3: (*a) Truncation of total duration of infant cry to first 15 sec for analysis; (b) Spectrogram of the first 15 sec duration of the infant cry. The red lines indicate formants (F_1

to F_4), yellow lines indicate intensity and the blue line indicates pitch contour of the cry sample.

The following features are extracted from each infant cry sample:

3.3.2a Duration Related Parameters

Latency period:

It is the time lapse between the initiation of audible finger snap and or vaccination procedure and the onset of crying**.** In simple terms, it is the time duration between the finger snap and the first cry sound by the child as identified in the spectrogram.

Voicing duration

It is the duration of voicing in the 15 secs cry signal. For the extraction of voicing duration, both silence and unvoiced components were deleted from the 15secs cry sample and only the voiced signal duration was noted.

3.3.2b Frequency related parameters

Fundamental frequency (F0)

Fundamental frequency represents the rate at which vocal cords vibrate and corresponds to the perceivable pitch of the sound. The mean F0 is the average value extracted from period-to-period fundamental frequency values. Minimum F0 is the lowest fundamental frequency extracted from the period-to-period pitch data and the Maximum F0 is the highest fundamental frequency extracted from the period-to-period pitch data.

PRAAT uses autocorrelation method to extract frequency related measures as it is reported to be more accurate, noise-resistant, and robust than methods based on cepstrum or combs, or the original correlation methods. The analysis window for frequency (Ceiling value point) was set between 75 and 1200 Hz to extract Fundamental frequency (F0), Minimum and Maximum pitch and range of the infant cry sample.

3.3.2c Intensity related parameters

Intensity refers to the period averaged power in the signal. The *mean intensity* value refers to the average intensity values within the specified time domain. Intensity values were calculated using automated method by selecting the required duration of the acoustic sample in the PRAAT software.

3.3.2d Noise related parameters

Harmonic to Noise ratio (HNR) represents the average ratio of harmonic spectral energy in the frequency range which is expressed in dB. HNR of 0 dB means that there is equal energy in the harmonics and in the noise.

3.3.2e Formants

Formants are frequency ranges that characteristically contain a concentration of the acoustic energy. In PRAAT, a formant represents spectral structure as a function of time i.e. formant contour. For extraction of formants using PRAAT, a linear interpolation algorithm method was used by the software. During the acoustic analysis of the infant cry, the maximum formant value was set to 8000 Hz which is appropriate for child population.

3.3.2f Other parameters

Number of pulses (NP)

Glottal flow or glottal pulse is the shape of the glottal source in a single period, corresponding to a puff of air at the glottis. The pitch of the voice is usually set by the frequency of glottal pulses during vowels or voiced consonants. The duration of each cycle/ periods in the acoustic waveform is called the (duration of the) glottal pulse or pitch period length. Glottal pulses are represented as blue vertical lines in the acoustic waveform of PRAAT which is considered as voiced component.

Number of Voice breaks (NVB)

It represents the interruptions in the phonation. The distances between consecutive pulses that are longer than 1.25 divided by the pitch floor provide NVB measurement.

Degree of voice breaks (DVB)

This is the total duration of the breaks between the voiced parts of the signal, divided by the total duration of the analyzed part of the signal.

All the newborn cry data were classified and tabulated. The data included 220 healthy newborn's cry and 93 high-risk newborn's cry. The number of high-risk newborn infant cries considered were Low birth weight (43), birth asphyxia (28) and with other risk factors (20). All the extracted acoustic parameters were subjected to statistical analysis using Statistical Package for the Social Sciences (SPSS) version 21.0 for analysis.

3.3.3 Phase III: Development of Computerized screening tool software

A MATLAB based computerized screening tool was developed in collaboration with software professionals. Codes were developed specifically for recording, analyzing and the

interpretation of the acoustic signal. The installation procedure, menu, navigation and working of the software are given in the Appendix 1 and troubleshooting procedure is given in Appendix 2. The statistical output of the analyzed parameters of both healthy and high risk infant cries were provided for the preparation of the software screening tool, which was the main objective of the present study.

3.3.4 Phase IV: Validation of the developed screening tool

The developed computerized screening software was administered on all the recorded samples of the healthy and at risk infant cries. Hence, the sensitivity and specificity evaluation involved 220 healthy newborns and 92 high-risk newborns. Sensitivity and specificity were calculated using the following formulae:

$$\frac{TP}{TP + FN} = \text{Sensitivity}$$

Where TP is true positive (whether the software identifies the abnormal cries as abnormal) and FN is false negative (where the software identifies the abnormal cry as normal).

$$\frac{TN}{TN + FP} = \text{Specificity}$$

Where TN is true negative (where the software identifies the normal cries as normal) and FP is (where the software identifies abnormal as normal).

Chapter 4

Results

The present study was conducted with the following objectives:

- To study and compare the acoustic parameters of the cries in healthy and high risk newborns.
- To develop a computerized screening tool for screening newborns using infant cries
- To validate the sensitivity and specificity of the developed tool using the infant cry samples collected for the study.

Participants in the present study included two groups- Group I included 220 healthy newborns and Group II included 93 newborns with high risk factors. The group II was subdivided into three subgroups which included babies with Low birth weight (43), Birth Asphyxia (28) and other conditions (22). The other conditions include the infants with Hyperbilirubinia (09), Meconium Aspiration (07), Pneumonia (04) and Cleft lip and palate (02). Acoustic analysis was carried out by extracting the parameters from the cry samples of newborns from both the groups. The derived acoustical parameters were entered in SPSS version 21.0 for statistical analysis. The following statistical procedures were used to analyze the data,

- Descriptive statistics was carried out to derive the mean and standard deviation (SD) values for the acoustic parameters of infant cry in the groups. The acoustic parameters included duration related (latency, duration of voicing), frequency related (mean pitch, minimum pitch, maximum pitch), intensity related (mean intensity, minimum intensity, maximum intensity), Noise related (mean NHR, mean HNR), Formant related (F_1, F_2, F_3, F_4), Other parameters included (number of pulses, number of periods, mean period, NVB and DVB). Independent sample t-test was run to obtain

the significant differences between the healthy newborns and the high risk newborns across the various acoustic parameters for the infant cry. The dependent factors were the two groups and the independent factors included the acoustic parameters studied.

- Multi-variate Analysis of Variance (MANOVA) was carried out with the acoustic parameters as the independent factor and the four groups (i.e., healthy newborns, LBW, Birth asphyxia and others) as dependent factors. Further Post-hoc Bonferroni analysis was run to check for within-subject differences in terms of the acoustic parameters at $p<0.05$ level of significance.

- Discriminant function analysis was used to extract the possible predictors of acoustic parameters to differentiate the healthy newborns and the high risk babies.

Phase II Results: The results of the study are delineated in Table 4.1 which shows the mean and SD values for the acoustic parameters in the two groups (healthy newborns and high risk babies).

Table 4.1

Mean, Standard deviation (SD) and Range [R] for acoustic parameters of cry in healthy and high risk newborns

Acoustic parameters		**Group I (Healthy Newborns)**	**Group II (High risk Newborns)**
Duration	Latency	1.143 (0.65) [0.02-3.2]	1.82 (0.639) [1.07 – 2.8]
	Voicing duration*	8.1 (7.4) [1.22-15.20]	6.67 (2.5) [0.90-75.22]
Frequency	Mean F0*	512 (93.1) [286 - 794]	474 (84) [293 – 773]
	Minimum F0	166 (69.79) [70 -429]	176 (74.81) [75 – 416]

Continued…..

Continued…..

	Maximum F0*	1088 (207.4) [498 – 1273]	981 (264.2) [492 – 1269]
Intensity	Mean Intensity	73 (84.4) [50 - 76]	67 (5.0) [54 – 80]
	Minimum Intensity	47.17(6.5) [24 – 64]	48 (6) [36 – 61]
	Maximum Intensity	75 (4.2) [59 -84]	74 (5) [48 – 89]
Noise related	Mean NHR*	0.1808 (0.080) [.03- .44]	0.2186 (0.091) [.05 - .57]
	Mean HNR*	12.4120 (2.9) [5.34-20.34]	11.2669 (2.9) [4.04-19.21]
Formants	Formant 1*	1125 (139.8) [808- 1667]	1123 (164.9) [830 – 1624]
	Formant 2*	2056 (209.7) [1279 – 2607]	1900 (291) [1328 – 2590]
	Formant 3	3129 (205.9) [1121 -3913]	3172 (203.1) [2506-3655]
	Formant 4*	4084 (200.8) [3558- 4718]	4162 (244.7) [3446 -4783]
Other parameters	No. of Pulses	3691 (1246.4) [364-6431]	3415 (1308.2) [509-7588]
	No. of Periods	3609 (1204.8) [359-6420]	3353 (1289.3) [494-7504]
	Mean Period*	2.1701 (.407) [0-3.47]	2.02 (0.371) [0.84-3.50]
	No. of Voice breaks (NVB)	25 (9.81) [0-15]	27 (9.20) [0-8]
	Degree of Voice breaks (DVB)	44.40 (13.67) [0-46.84]	47.67 (15.42) [0-10.61]

Analysis of results from Table 4.1 reveals that the mean values for latency, Minimum F0, Minimum intensity, mean NHR, Formant 3, Number of voice breaks and degree voice breaks were higher in the high risk groups compared to healthy newborn group.

Further, Table 4.2 shows the mean and SD values within the high risk group. Analysis of results revealed that high risk babies with birth asphyxia showed longer voicing duration and higher mean intensity, minimum and maximum intensity compared to other sub groups such as low birth weight and other conditions. Whereas Mean F0 and minimum F0 was higher in low birth weight newborns. In other conditions, the maximum F0 was the only parameter which was highest compared to the other sub groups.

Table 4.2

Mean and Standard deviation (SD) of acoustic parameters in subgroups of high risk newborns

Acoustic parameters		High risk group		
		Low birth weight	Birth Asphyxia	Other conditions
	Latency	0.71 (0.25)	0.602 (0.23)	0.508 (0.16)
Duration	Voicing duration	6.05 (0.84)	8.31 (0.89)	5.95 (0.82)
	Mean F0	501 (13.68)	444 (16.9)	459.22 (19.13)
Frequency	Minimum F0	194 (10.8)	159 (13.3)	163 (15.1)
	Maximum F0*	960 (34.1)	933 (42.3)	1085 (47.7)
	Mean Intensity	66 (10.8)	69 (13.4)	67 (15.1)
Intensity	Minimum Intensity	47 (0.43)	51 (1.2)	48 (1.36)
	Maximum Int	74 (0.69)	76 (0.85)	74 (.96)
Noise	Mean NHR	0.222 (.08)	0.234 (.091)	0.201 (.10)
related	Mean HNR	12.1 (.21)	11.3 (.18)	10.4 (.21)
	Formant 1	1096 (22.4)	1157 (27.8)	1133 (31.4)
	Formant 2	1830 (35.8)	1965 (44.3)	1957 (50.06)
Formants	Formant 3	3207 (30.9)	3083 (38.4)	3219 (43.3)

Continued…..

Continued…..

	Formant 4	4159 (32.6)	4115 (40.4)	4232 (45.6)
	No. Pulses	3861 (196.9)	3574 (244.06)	3508 (275.3)
Other	No. Periods	3772 (193.08)	3508 (239.2)	3417 (269.9)
parameters	Mean Period	2.049 (.05)	2.318 (.07)	2.219 (0.81)
	NVB	27 (1.4)	28 (1.8)	24 (2.05)
	DVB	47.89(1.02)	47.8(.982)	47.83(1.028)

The results of the independent t-test revealed a list of parameters that showed significant difference between healthy newborns and high risk babies. The findings of the study are broadly presented under the following headings:

1. Cry duration
2. Cry F0
3. Noise related parameters
4. Formant Frequencies
5. Other related parameters such as
    - Number of pulses
    - Number of periods
    - Mean period
    - Number of voice breaks &
    - Degree of voice breaks.

4.1 Cry Duration

Results of the cry duration using independent t-test indicated significant difference in voicing duration [t (311) = -2.6, p <0.05] and no statistical significance was found for latency. The mean value for the voicing duration was higher in healthy newborns (8.1s)

compared to high risk groups (6.67s) whereas the mean value of latency was higher for high risk group (1.82s) compared to healthy group (1.14s) as shown in Table 4. 1.

The analysis of results on MANOVA revealed that there was a significant main effect observed only for voicing duration [F (3, 309) = 5.36, $p<0.05$] and was not seen for latency. Post–hoc Bonferroni test results within the subgroups of high-risk newborns indicated that there was a significant difference for voicing duration between healthy newborns and low birth weight ($p<0.05$). The mean values indicated that babies with low birth weight showed significantly shorter (6.05 sec) voicing duration compared to the healthy newborns (8.1sec). A significant difference was also found between babies of low birth weight and birth asphyxia where in LBW babies showed shorter voicing duration than birth asphyxia (8.31sec) group.

4.2 Cry F0

Analysis of results for Cry F0 as shown in Table 4.1 showed that the average F0 in healthy newborns was 512 Hz with a range of 286 - 794 Hz, while high risk newborns showed lower F0 of 474 Hz with a range of 293 – 773 Hz. The minimum and maximum mean value of healthy newborns were 166 Hz and 1088 Hz respectively and the same for high risk newborns were 176 Hz and 981 Hz. Independent t test revealed that there was a significant difference between healthy newborns and high risk newborns for mean F0 [t = 1.83, $p<0.05$] and maximum F0 [t = 3.84, $p<0.05$].

Results on MANOVA indicated that there was a significant main effect observed for mean F0 [F (3, 309) = 6.446, $p<0.05$] and maximum F0 [F (3,309) = 7.11, $p<0.05$]. Post–hoc Bonferroni test results within the subgroups of high-risk newborns and healthy babies indicated that there was a significant difference between healthy newborns and LBW

($p<0.05$) for both mean F0 and maximum F0 at $p<0.05$ level of significance. Low birth weight newborns had lower mean F0 (501 Hz) and maximum F0 (960 Hz) compared to the healthy newborns**.** A significant difference was also observed between healthy newborns and birth asphyxia for mean F0 and maximum F0 at $p<0.05$ level of significance with the latter showing a lower mean F0 (444 Hz) and maximum F0 (933Hz). There was no significant difference found between healthy newborns and other conditions.

4.3 Cry Intensity

Results as shown in Table 4.1 for mean intensity, minimum and maximum cry intensity for healthy newborns and high risk newborns indicated that the high risk newborns showed lower mean intensity (67 dB) compared to healthy newborns (73 dB). The results also showed a lower mean and maximum intensity for the high risk groups than the healthy newborns. Results of independent t- test indicated that there was no significant difference found for mean, minimum, and maximum intensity values between the healthy newborns and the high risk groups.

Analysis of results on MANOVA within the high risk groups showed that there was a significant main effect seen only for minimum intensity [$F(3, 309) = 3.148$, $p<0.05$]. Results on Post-hoc Bonferroni test showed a significant difference between LBW and birth asphyxia for minimum intensity at $p<0.05$ level of significance. Birth asphyxia newborns showed higher minimum intensity of 51 dB compared to the low birth weight babies (47 dB).

4. 4 Noise Related Values

Analysis of results as indicated in Table 4.1 shows that the mean value of Noise to harmonic ratio (NHR) was found to be higher for high risk newborns (0.218) compared to the

healthy newborns (0.181). The mean Harmonic to noise ratio (HNR) was lower for high risk newborns (11.266) than the healthy newborns (12.412). Independent t-test results revealed that there was a significant difference between healthy and high risk newborns for both mean NHR and mean HNR at $p<0.05$ level of significance.

Results on MANOVA within the groups indicated a significant main effect for mean NHR [$F(3, 309) = 4.182$, $p<0.05$] and mean HNR [$F(3, 309) = 3.77$, $p<0.05$]. Further, Post-hoc analysis using Bonferroni test showed that there was a significant difference in mean NHR and mean HNR between healthy newborns and low birth weight at $p<0.05$ level of significance. Further, Low birth weight newborns showed higher NHR value (0.222) compared to the healthy group (0.18). The HNR was found to be lower for LBW (12.1) compared to healthy newborns (12.41)

4.5 Formant Frequencies

Table 4.1 shows the mean F_1, F_2, F_3 and F_4 values and SD in both healthy and high risk newborns. The results indicated that mean F_1 was similar in healthy newborns (1125 Hz) and high risk newborns (1123 Hz). F_2 and F_4 were found to be higher for healthy newborns compared to high risk newborns. However, healthy newborns were found to have lower F_3 (3129 Hz) mean values compared to the high risk group (3172 Hz). Results on independent t-test indicated a significant difference between the two groups in terms of their F_2 [$t(311) = 5.32$, $p<0.05$] and F_4 [$t(311) = 2.96$, $p<0.05$] mean values whereas no significant difference was found for F_1 and F_3 mean values.

MANOVA test results showed a significant main effect for all the formants, mean F_1 [$F(3, 309) = 11.487$, $p<0.05$]; mean F_2 [$F(3, 309) = 12.008$, $p<0.05$]; mean F_3 [$F(3,309) = 3.581$, $p<0.05$] and mean F_4 [$F(3, 309) = 4.177$, $p<0.05$]. Post-hoc Bonferroni test results indicated a significant difference for F_1 and F_2 formants between healthy newborns and low

birth weight at $p<0.05$ level of significance. The mean values indicated that F_1 (1125 Hz) and F_2 (2056 Hz) in healthy newborn babies which was greater than the F_1 (1096 Hz) and F_2 (1830 Hz) in LBW. Results also indicated that for F_3 a significant difference was found between healthy newborns and birth asphyxia. Significant difference was also found between healthy newborns and other high risk conditions for F_1 and F_4 at $p< 0.05$ level of significance. The mean values of F_1 and F_4 was found to be lower in healthy newborns compared to other conditions.

4.5 Other Parameters

The other extracted acoustic parameters included number of pulses, number of periods, mean periods, number of voice breaks (NVB) and degree of voice breaks (DVB). Table 4.1 displays the mean values of other extracted acoustic parameters in the healthy newborns and high risk babies in the present study.

The analysis of the results indicated that healthy newborns showed higher mean values on number of pulses (3691), number of periods (3609), mean period (2.17) and lower mean values for NVB (25) and DVB (44.40) compared to the high risk newborn group. Independent t-test revealed statistical difference between both the groups only in the mean period [$t (311) = 3.12$, $p<0.05$] and not in any other parameters.

The MANOVA test results within groups indicated a significant main effect for mean period [$F (3, 309) = 6.354$, $p<0.05$]. Post-hoc Bonferroni test results revealed a significant difference between healthy newborns and birth asphyxia ($p<0.05$), birth asphyxia and low birth weight ($p<0.05$) in terms of mean period. Babies with birth asphyxia showed higher mean period value (2.32) than the healthy newborns (2.02) and LBW (2.0).

Discriminant function analysis

Discriminant function analysis (DFA) was used in the present study to predict the group membership of a set of predictors such as acoustic parameters considered in the present study for analysis of infant cry. The characteristics of predictors are related to form groups for the healthy newborns and the high risk groups such as LBW, birth asphyxia, etc. The data in both healthy newborns and high risk newborns was subjected to Discriminant function analysis separately in order to derive functions in both groups. While analysis in healthy group revealed that first Discriminant function (DFI) accounted for 94.6 % of the total among groups variability and second Discriminant function (DF2) for the remaining 6.5 %. In the high risk DF1 accounted for 93.5% of the total among groups variability and DF2 for the remaining 5.5 %.

In the healthy group, for DF1, Wilk's Lambda, **λ** showed significance in the functions of the data analyzed across the acoustic parameters at 0.538, $\chi^2(18) = 190.09$, $p<0.001$. Hence DF1 is significant. DF2 was also found to be statistically significant for the data across the acoustic parameters, **λ** is 0.860, $\chi^2(10) = 46.46$, $p<0.001$. Parameters as function of DF3 was found to be statistically significant with **λ** of 0.944, $\chi^2(4) = 17.60$, $p<0.05$. Eigenvalues and canonical correlation coefficient was more for function 1 and indicated majority of variance. Table 4.3a depicts Wilk's Lambda values and 4.3b indicates Eigenvalues.

Table 4.3a

Wilk's Lambda values for both healthy and high risk infants

Wilks' Lambda				
Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1 through 3	.538	190.085	18	.000
2 through 3	.860	46.461	10	.000
3	.944	17.605	4	.001

Table 4.3b

Eigenvalues for both healthy and high risk infant

Eigenvalues				
Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	.597	79.1	79.1	.611
2	.099	13.1	92.2	.300
3	.059	7.8	100.0	.236

Further to interpret the first Discriminant function DF1, standardized Discriminant function coefficients were considered (see Table 4.4 (a), (b)). Analysis of results indicated that DF1 was found to be most heavily weighted on duration, pitch and intensity related parameters for the subgroups. Interpretation of the results of DF1 and DF2 in terms of group centroids as indicated in Figure 4.1 and results of MANOVA indicate that the parameters on DF1 that can differentiate between healthy newborns and babies with LBW include voicing duration, Mean F0, Maximum F0, Minimum intensity, Mean NHR, Mean HNR, Formant 1, Formant 2. The parameters that differentiate healthy newborns from birth asphyxia include Mean F0, Maximum F0, Formant 3 and Mean period. F_1 and F_4 parameters were found to differentiate healthy newborns and babies with other conditions. The Figure 4.1 and results on MANOVA also indicate the parameters on DF2 that can differentiate LBW and birth asphyxia which includes Voicing duration, Minimum intensity and Mean period.

Table 4.4a

Standardized Canonical Discriminant Function Coefficients for the acoustical parameters

Parameters	**Function**		
	1	**2**	**3**
Voicing duration	-.276	.467	-.586
Maximum pitch	.482	.370	.735
Maximum intensity	-.172	-.271	-.083
Mean Period	.089	.886	.244
F_2	.895	.325	-.152
F_4	-.943	.150	.459

Table 4.4b

Discriminant Functions at group Centroids for the Subgroups

Subgroups	**Function**		
	1	**2**	**3**
Normal	.487	-.044	.005
Low birth weight	-1.475	-.499	-.036
Birth asphyxia	-.803	.650	-.526
Other conditions	-.965	.584	.689

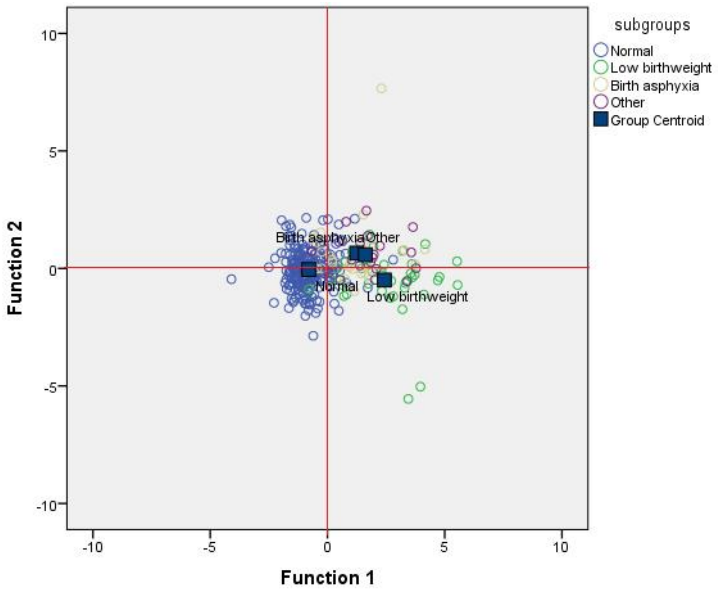


Figure 4.1., Centroid of each group based on Function 1 and Function 2

Classification results based on discriminant functions revealed that, healthy newborn were overall correctly classified 75.1% of infants whereas 74.1% of high risk infants were categorized correctly.

Phase III: Development of the software tool

Development of the software tool was outsourced through a software development firm. The acoustic parameters which were found to be significantly different across the healthy and the high risk infants were provided to the software developer. The screening tool, *ICry* was developed on a MATLAB platform and a standalone application was made for a Windows Operating system. The designed software comes as a package which includes

necessary Matlab Compiler Runtime (MCR) files and other instructional files, which will guide through the installation procedure without any requirement of internet connection. This can be run in any computer with the matlab code which is given in a compact disk format along with the software. Once the installation procedure has been completed, software can be used for infant cry analysis.

Phase IV: Pilot Validation of the developed screening tool: Software *ICry* was developed based on the statistical output of the analyzed parameters of both healthy and high risk infant cries. Pilot Validation of the developed screening tool was carried out by checking the sensitivity and specificity of the entire recorded samples of healthy and at risk new born cries. The calculations revealed 51 % of sensitivity and 61.8 % of specificity.

$$\text{Sensitivity} = \frac{TP}{TP + FN} = \frac{47}{47+ 45} = 51\%$$

$$\text{Specificity} = \frac{TN}{TN + FP} = \frac{136}{84+136} = 61.8\%$$

[*Where TP is True positive (whether the software identifies the abnormal cries as abnormal) and FN is false negative (where the software identifies the abnormal cry as normal). And TN is True Negative (the software identifies the normal cry as normal) where FP is False positive (the software identifies the normal as abnormal)*].

Chapter 5

Discussion

Crying is the infant's most powerful and most often the only means of communication. These cries seem to be uniform, but there are a lot of differences between two infant's cries. An infant cry contains a lot of information about the baby, as hunger, pain, sleepiness or discomfort. It has long been proven that the acoustic characteristics of the cry sound are directly influenced by the infant's physical and psychological state or various external stimuli. Special characteristics of newborn pain cries have been identified in various pathological states such as brain damage (Lester & Boukydis, 1985]; cleft palate (Michelsson et al., 1975]; hydrocephalus [Michelsson et al., 1984]; Sudden Infant Death Syndrome (SIDS) [Colton & Steinschneider, 1981] among others. Using cry as a tool in studies of infants with known medical problems are important in that they have helped to advance our understanding of the relation of cry characteristics to medical status.

The present study shows that there are differences in the acoustic quality of cry between healthy full-term children and high risk babies at the age of 3 to 7 days of birth for the former and within a span of one month for the latter. Acoustic characteristics of infant cry are produced by the larynx and the vocal tract function. These functions are regulated by the central nervous system (CNS) and vagal tone (Golub & Corwin, 1985). Thus, any CNS pathology can affect the acoustic characteristics of cry including the duration characteristics. Cry latency measure is assumed to reflect the time for the CNS to interpret the pain stimulus (Robb, Crowell, Dunn- Rankin & Tinsley, 2007). In the present study, even though, the latency measurements did not show any statistical difference between both groups, high risk population showed higher latency compared to healthy groups. These results are in support with the earlier studies. Infants with diffuse brain damage have been found to have higher threshold and a longer latency between a painful stimulation and a cry response compared to

healthy infants (Karelitz & Fisichelli, 1962; Fisichelli & Karelitz, 1963). Longer cry latency of SIDs siblings showed the generalized delay in the brain circuitry responsible for central arousal (Gislason et al., 2002).

The present study results showed longer voicing duration in healthy newborns compared to the high risk group. Thus, outcome of the analysis also indicated that the high risk newborns had lesser amount of voicing or phonation in their total cry duration. This is in consonance with the earlier studies, which indicated lesser amount of phonation and duration of cries produced by preterm infants (Michelsson et al., 1982; Cacace et al., 1995). Goberman and Robb (1999) measured the percent phonation value i.e. the amount of phonation that remained after the removal of non-voiced portions of the crying episode. The results indicated lower percent phonation in preterm babies in the cry signal. High risk newborns are frequently exposed to intravenous injection as a part of their routine care in NICU. Associated pain, discomfort or distress could have made the infant fatigued, causing reduced sub glottal pressure resulting in cessation of vocal fold vibrations represented as less voicing duration.

In the field of infant cry research, fundamental frequency (F0) has been the most commonly examined measure (Truby & Lind, 1965; Zeskind & Lester, 1978;Wermke & Robb, 2010). Frequencies of the cries of infants affected with various diseases were found to be abnormal when compared with the cries of healthy infants. The present study results revealed significantly higher minimum F0 in high risk babies (specifically in low birth weight newborns) compared to healthy babies. These results are in consonance with the earlier studies; Rautava et. al (2007) reported higher minimum fundamental frequency values of the very low birth weight infants (VLBWI) compared to control group. Several authors have reported higher F0 in pathological population, infants with heart disorders (Radhika Rani et

al., 2012) ; SIDS (Stark & Nathanson, 1975; Corwin et al., 1995; Goberman & Robb, 1999; Robb, Crowell, Dunn- Rankin & Tinsley, 2007) Cleft lip and palate (Vuorenkoski, Lind, Partanen 1966;Goberman & Robb 2005b) preterm infants (Johnston, Stevens, Craig & Grunau 1992; Goberman & Robb 1999). Presences of higher F0 in clinical populations are attributed to several reasons. Higher cry F0 in preterm infants may simply be related to smaller vocal folds, resulting from physical size differences at birth. In contradiction to these justifications, Shinya, Kawai, Niwa and Yamakoshi (2014) suggest that increased *F0* of spontaneous cries in preterm babies were not caused by their smaller body size, but instead might be caused by more complicated neurophysiological states owing to their different intrauterine and extra uterine experiences. Another possibility of high F0 in pain cries are attributed to the additional stress response for the stimuli. During the stress of pain, the laryngeal musculature is tightened, which might have effects of raising F0 (Wasz- Hockert, Lind, Vuorenkoski, Partanen, & Valenne, 1968; Zeskind, 1983; Johnston et al., 1992; Goberman & Robb, 1999).

Contrary to the mean minimum F0, the present study results revealed lower mean and maximum F0 in high risk infants compared to the healthy group. The lower F0 in our study could be attributed other reasons like fatigue and exhaustion of the babies due to their continuous crying in the NICU as the high risk newborns were recipients of intensive care during the recording. Few newborns were crying even before the toe tapping stimulation and hence their baseline arousal level could not be controlled in the present study. This continuous cry of high risk newborns lead to fatigue and also the intensity of the cry signal was observed to be reduced in the high risk newborns which could possibly result in lower fundamental frequencies.

Golub and Corwin (1985) states that although acoustic theory suggests increasing size decreases fundamental frequency [1], F0 has been shown to rise from the age of 0 to 12 months in longitudinal studies (Gilbert & Robb, 1996). This rise has been interpreted as an increasing control of cry production as the child grows older. In the present study, the mean and maximum fundamental frequency values were lower in high risk infants than in controls and this may also be attributed to the reduced control of cry production as stated above.

Recently, Chen ,Yangi and Lin (2014)'s study on preterm infants using LTAS is in support with the current study results. They obtained lower fundamental frequency in preterm babies compared to the full term babies and results were justified based on the external factors which are presence of back ground noise during recording or interaction between the infant and caretaker. Colton and Steinschneider (1981) found the cries of SIDS infant to be longer with a low F0 and low formant frequency compared to healthy term infants.

Both NHR and HNR values represents the noise component of an infant cry. Researchers have used several measurements to account the noise component in a cry. Higher NHR and lower HNR values of high-risk infants in the present study are in consonance with the earlier studies like Fuamenya, Robb and Wermke (2015) who considered noise as visible elements of phonatory noise within cries that was indicated in spectrogram as no visible F0 and harmonics. Compared to healthy infants, those with CLP had lower mean HNR (Fuhr, Reetz, & Wegener, 2012). Preterm infants differ in their acoustic characteristics and show more variability in their amplitude and more harmonic doubling and noise concentration (Robb, 2000). This is supported by the view that phonatory noise (N) is thought to reflect nonlinearities in vocal fold mechanics (Mende, Herzel &Wermke, 1990; Titze, Baken & Herzel, 1993). LTAS display is used to calculated to estimate High Frequency Energy (HFE) which was noticeably larger among the SIDS siblings (Robb et. al 2007). The amount of HFE

is related to the presence of noise elements during cry phonation (Colton & Steinschneider, 1981; Corwin, Lester Sepkoski , Peucker, & Golub, 1995). Presence of higher HFE is a representation of hyper adductional vocal fold behavior.

In the present study, the high risk group showed significant difference for F_2 and F_4 values compared to healthy group. Studies on different clinical population have shown similar results. F_4 values were reported to be significantly higher in Very Low Birth Weight Infants compared to their control match (Rautava et.al 2007). Similarly, Tanja et. al, (2012) reported higher values for F_4 in infants with hearing impairment.

In the present study, high risk newborns showed higher number of voice breaks (NVB) and degree of voice breaks (DVB) compared to the healthy group. DVB is the ratio of the total length of the areas representing voice breaks in the complete voice sample. Higher degree of voice breaks indicates more of silence and distribution of unvoiced components in the infant cry. Analysis in the present study indicated higher degree of voice breaks in preterm LBW infants. Various earlier authors also supported the findings of the current study indicating higher degree of voice breaks in high-risk babies (Sangeetha, 1999; Kheddache & Tadj, 2013). Sangeetha's (1999) study of acoustic parameters using MDVP on normal and high-risk babies showed higher mean value (44.7) of DVB for premature babies compared to normal (30.16) infants. Kheddache and Tadj (2013) reported higher average percentage of no-voice segments for premature cries compared to full-term healthy cries. The authors concluded that the percentage of no-voice segments is not dependent on gestational age, but on the pathology itself and the difference in voice breaks were attributed to the immature innervations of the larynx in preterm newborns.

Chapter 6

Summary and Conclusions

Cry represents a combination of respiratory, laryngeal and vocal tract functions; any unusual or deviant cry patterns are likely to be a reflection of poor organization in either para sympathetic or sympathetic strands of the nervous system. Thus cry behavior is considered as a unique and important prognostic indicator in several disorders. Physicians and researchers have long been aware about the distinctive cry features of infants with certain diseases. Infant cries, both of healthy and impaired infants, have been widely investigated in the past, Besides the scientific interest in the cry production process, the aim of most of the studies was to use the infant cry as a diagnostic tool for certain diseases. Hereby it is assumed that diseases may have an influence on the cry production process. The efforts of researchers in this sense are to show infant cry not only as an acoustic–linguistic event, but also as an indicator of the neurophysiologic status of the infant.

Manual inspection of spectrograms of infants cries have been a gold-standard method for detecting acoustic features in cry sounds, the timing and onset of cry vocalizations and the fundamental frequency (F0) of cry. Such visual inspection has also been used to describe cry patterns since decades. Over the years, computer-assisted methods have been reported to be of more advantageous than manual spectral analysis, especially quantifying aspects of the sound spectrum that can be expected to yield more accurate estimates of F0 and related parameters automatically.

Automated approaches have the advantage of fast analysis of very large data sets, objective assessment and the ability to quantify multiple data points, which have the potential to increase the applied value and clinical utility of the cry assessment. There are multiple challenges to the design of an automated acoustic analysis of infant cry. These automated

approaches have the potential to speed scientific inquiry, allowing for the study of large numbers of infants with efficient and rapid analyses. In addition, they bypass the need for manual inspection of spectrograms and do not require the time-consuming task of cursor placement and frame selection used as in some computer-assisted methods. In this way, automated approaches allow for more rapid analysis that is less prone to observer bias or coding errors.

In this context, the present study scrutinizing the relevance and vitality of the cry behavior of infants aimed to develop a non-invasive computerized screening tool to identify infants at risk for typical development. Participants in the study included two groups of newborns. Group I was a control group which consisted of 220 healthy newborns with no significant pre or natal history as checked by a high risk register, screening audiometric evaluation, medical records, and a formal interview with the parents. And group II consisted of 92 newborns with one or more risk factors such as low birth weight, birth asphyxia, meconium aspiration syndrome and infantile jaundice.

The study was conducted in 4 phases. Phase I involved recording of pain cry samples from newborns using an audio recorder. Intramuscular immunization (Hepatitis B) was used as the pain stimulus for the healthy newborns and toe tapping for high risk newborns. During Phase II, acoustic analysis of newborn cries was performed for extracting various durational and spectral parameters from the cry samples of newborns from both groups I and II. The cry samples were analyzed and the acoustic parameters such as frequency, intensity, formants and other parameters were extracted using PRAAT software version 5.4.01. In Phase III, the overall statistical values of the analyzed parameters of infant cry in healthy and at risk newborns were provided to a software professional for the development of a computerized screening tool. Software was developed specifically for recording, analysis and interpretation of the cry signal. Phase IV involved the pilot validation of the developed screening tool.

The entire recorded sample of the newborns in both groups were run on the newly developed cry analysis software named *ICry*.

The Phase II results of the acoustic analysis showed that the mean voicing duration was significantly higher in healthy newborns compared to the high risk groups, particularly the LBW babies whereas the mean latency value was higher for the high risk group. It was also seen that there was a significant difference between healthy and high risk newborns for the F0 values with the risk babies depicting a lower F0. The intensity related parameters did not vary significantly between the healthy newborns and the high risk groups. There was significant difference between healthy and high risk newborns for both mean NHR and mean HNR at $p<0.05$ level of significance. Mean NHR was higher in the at risk infants and HNR scored higher in the healthy babies.

The findings on formants indicated that mean F_1 was similar in healthy newborns and high risk newborns. F_2 and F_4 were higher and F_3 was lower for healthy newborns compared to high risk newborns. Results on independent t- test indicated a significant difference between the two groups in terms of their F_2 and F_4 mean values whereas no significant difference was found for F_1 and F_3 mean values.

Using Discriminate analysis, DF1 could differentiate between healthy newborns and babies with LBW based on the acoustic parameters voicing duration, Mean F0, Maximum F0, Minimum intensity, Mean NHR, Mean HNR, Formant 1, Formant 2. The parameters that differentiated healthy newborns from birth asphyxia were Mean F0, Maximum F0, Formant 3 and Mean period. F_1 and F_4 parameters were found to differentiate healthy newborns and babies with other conditions. The parameters that differentiated LBW and birth asphyxia included Voicing duration, Minimum intensity and Mean period.

Software *ICry* was developed based on the statistical output of the analyzed parameters of both healthy and high risk infant cries. All the recorded healthy and at risk

infant cries were run in the newly developed tool to validate its sensitivity and specificity. The calculation showed 62 % of sensitivity and 51% of specificity.

In conclusion, the present study developed a software tool for the automatic cry analysis which can be used as an early diagnostic tool to identify infants at risk for poor developmental outcomes. The future studies can aim on building up a more detailed and precise validated cry analyser to screen for infants at risk for various developmental disorders or specific developmental conditions such as autism spectrum disorders with a constant follow up of the newborns identified as at risk.

More basic research may also utilize this system to study normative aspects of infant cry production with larger samples than has been possible in the past. Future attempts in this direction can also intend to make the developed software tool available online so that the general community will have access to a high-quality infant cry analyzer. Upcoming efforts can also involve refining the analysis and output of this system, as well as developing a more user friendly interface to enhance its accessibility for a variety of clinicians and researchers.

Implications of the Results of the Study

1. The study provides an acoustic profile of the crying behaviour of healthy and high risk infants.
2. The findings of the study pave way for the emergence of a non-invasive tool for the screening of new born infant cries.
3. The computerized screening tool aids in early detection of any pathological conditions the infant is susceptible to and facilitates further investigations.
4. Tool developed is an additional valid component of the HRR in screening new born babies for any discrepancies in their neuro-respiratory and phonatory systems.

Limitations of the Study

1) The high risk group in the present study included only low birth weight and asphyxia babies primarily. It is reasonable to assume that both these conditions can be identified at the time of birth itself. However, they were selected as a starting point to examine the applicability of the approach.

2) The screening tool developed has been validated only on the available samples recorded for the study

3) The sensitivity and specificity obtained for the developed tool is relatively low at this stage. Further validation is warranted with the tool on high-risk babies.

4) The installation procedure of the software takes longer duration. Can be improved in future versions

References

Abbs, K. J. (2015). Dysphonations in infant cry: A potential marker for health status. *Masters Abstracts International, 54*(4), 48. ProQuest dissertations and theses: Thesis (M.S.), Bowling Green State University.

Acebo, C., & Thoman, E. B. (1992). Crying as social behavior. *Infant Mental Health Journal. 13*, 67–82.

Aram, D., Hack, M., Hawkins, S., Weissman, B., & Borawski-Clark, E. (1991). Very-low-birth weight children and speech and language development. *Journal of Speech and Hearing Research, 34,* 1169–1179.

Bauer, H. (1968). Phonatrischer Beitrag zum Cri-du-Chat syndrome. *HNO: Wegweisser fur die Faharzlliche Praxis, 6,* 185-187.

Blinick, G., Tavolga, W. N., & Antopol, W. (1971). Variations in birth cries of newborn infants from narcotic-addicted and normal mothers. *American Journal of Obstetrics and Gynecology, 110,* 948 –958.

Bosma, J. F. (1967). Human infant oral function. In J. F. Bosma (Ed.), *Symposium on oral sensation and perception* (pp. 98-110). Illinois: Springfield.

Bosma, J. F., Truby, H., & Lind, J. (1965). Cry motions of the newborn infant. *Acta Paediatrica Scandinavica Supplement, 163,* 61-92.

Branco, A., Fekete, S. M., Rugolo, L. M., & Rehder, M. I. (2007). The newborn pain cry: Descriptive acoustic spectrographic analysis. *International Journal of Paediatric Otorhinolaryngology, 71,* 539–546.

Brodie, A. G. (1949). On the growth pattern of the human head from three months to eight years of life. *American Journal of Anathesia, LXVIII*(2), 209-259.

Brodie, A. G. Anatomy and physiology of the head and neck, musculature. *American Journal of Orthopedics, XXVI,* 831-837.

Cacace, A. T., Robb, M. P., Saxman, J. H., Risemberg, H., & Koltai, P. (1995). Acoustic features of normal-hearing pre-term infant cry. *International Journal of Pediatric Otorhinolaryngology, 33*(3), 213-224.

Chen, L., M., Yangi, Y. H., Lin, C. H., & Lin, Y. C. (2014). Spectrum analysis of cry sounds in preterm and full-term infants. *Proceedings of Rocling, Conference on Computational Linguistics and Speech Processing,* 193-203.

Colton, R., & Steinschneider, A. (1981). The cry characteristics of an infant who died of the sudden infant death syndrome. *Journal of Speech and Hearing Disorders, 46,* 359–63.

Corwin, M. J., Golub, H. L., Potter, M. (1987). Cry analysis in infants of narcotic addicted mothers. *Pediatric Research, 21,* 190A.

Corwin, M. J., Lester, B. M., Sepkoski, C., McLauglin, S., Kayne, H., & Golub, H. (1992). Effects of in-utero cocaine exposure on newborn acoustical cry characteristics. *Pediatrics, 89*(6), 1199 – 203.

Corwin, M. J., Lester, B. M., & Golub, H. L. (1996). The infant cry: What can it tell us? *Current Problems in Pediatrics, 26,* 325–333.

Corwin, M., Lester, B., Sepkoski, C., Peucker, M., Kayne, H., & Golub, H. (1995). Newborn acoustic cry characteristics of infants subsequently dying of sudden infant death syndrome. *Pediatrics, 96,* 73–77.

Crelin, E. S. (1969). *Anatomy of the Newborn: An atlas.* Philadelphia: Lea & Febiger.

Daga, R. P., & Panditrao, A. M. (2011). Acoustical analysis of pain cries' in neonates: Fundamental frequency. *Special Issue of International Journal of Computer Applications (0975-8887) on Electronics, Information and Communication Engineering – ICEICE, 3,* 18-21.

Eckenhoff, J. E. (1951). Some anatomic considerations of the infant larynx influencing endotracheal anesthesia. *Anesthesiology, XII,* 401-410.

Esposito, G., & Venuti, P. (2010). Understanding early communication signals in autism: A study of the perception of infants' cry. *Journal of Intellectual Disability Research, 54,* 219-223.

Fisichelli, V. R., & Karelitz, S. (1963). The cry latencies of normal infants and those with brain damage. *The Journal of Pediatrics, 62,* 724-34.

Fisichelli, V. R., & Karelitz, S. (1966). Frequency spectra of the cries of normal infants and those with Down's syndrome. *Psychonomic Science, 6,* 195–196.

Fischelli, V. R., Haber, R. A., Davis, J., & Karelitz, S. (1966). Audible characteristics of the cries of normal infants and those with Down's syndrome. *Perceptual and Motor Skills, 23,* 774-746.

Flatau, T. S., & Gutzmann, H. (1906). Die Stimme des Sauglings. *Archiv fur Laryngologie und Rhinologie, 18,* 139-151.

Frey, W. H. (1985). *Crying: The mystery of tears.* Minneapolis: Winston Press.

Fuamenya, N. A., Robb, M. P., & Wermke, K. (2014). Noisy but effective: Crying across the first 3 months of life. *Journal of Voice, 29*(3), 281-286. doi: 10.1016/j.jvoice..07.014. Epub 2014 Dec 4.

Fuhr, T., Reetz, H., & Wegener, C. (2015). Comparison of supervised-learning models for infant cry classification. *International Journal of Health Professions, 2*(1), 4-15.

Galaviz, O. F. R., & Garcia, C. A. R. (2005). Infant cry classification to identify hypo acoustics and asphyxia comparing an evolutionary-neural system with a neural network system. In A. Gelbukh, A. de Albornoz, & H. Terashima-Marin (Eds.), *MICAI 2005: Advancecs in artificial intelligence* (pp. 949-958). Mexico: Springer Berlin Heidelberg.

Gilbert, H., & Robb, M. (1996). Vocal fundamental frequency characteristics of infant hunger cries: Birth to 12 months. *International Journal of Pediatric Otorhinolaryngology, 34,* 237-243.

Gislason, T., Johannsson, J. H., Haraldsson, A., Olafsdottir, B. R., Jonsdottir, H., Kong, A.,…..Steffansson, K. (2002). Familial predisposition and cosegregation analysis of adult obstructive sleep apnea and the sudden infant death syndrome. *American Journal of Respiratory and Critical Care Medicine, 166,* 833–838.

Goberman, A. M., & Robb, M. P. (1999). Acoustic examination of preterm and full-term infant cries: The long-time average spectrum. *Journal of Speech, Language, and Hearing Research, 42,* 850–861.

Goberman, A. M., & Robb, M. P. (2005). Acoustic characteristics of crying in infantile laryngomalacia. *Logopedics, Phoniatrics, Vocology*, *30,* 79–84.

Golub, H. L. (1980). *A psychoacoustic model of infant cry production*. Unpublished doctoral dissertation, Massachusetts Institute of Technology, Cambridge.

Golub, H. L. (1989). A physioacoustic model of the infant cry. In B. M. Lester & C. F. Z. Boukydis (Eds.), *Infant crying: Theoretical and research perspectives* (pp. 59-82). New York: Plenum Press.

Golub, H., & Corwin, M. J. (1982). Infant cry a clue to diagnosis. *Pediatrics*, *69,* 197-201.

Golub, H., & Corwin, M. J. (1985). A physio-acoustic model of the infant cry. In B. Lester & C. Boukydi (Eds.), *Infant crying: Theoretical and research perspectives* (pp. 59–80). New York: Plenum Press.

Gracia, J. O., & Reyes-Gracia, C. A. (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, 4,* 3140-3145.

Hariharan, M., Chee, L. S., & Yaacob, S. (2012). Analysis of infant cry through weighted linear prediction cepstral coefficients and probabilistic neural network. *Journal of Medical Systems*, *36*(3), 1309-1315.

Hixon, T. J. (1973). Respiratory function in speech. In F. Minifie, T. J. Hixon, & F. Williams (Eds.), *Normal Aspects of Hearing, Speech and Language* (pp. 73-125). Englewood Cliffs, NJ: Prentice- Hall.

Hoffman, H. J., Damus, K., Hillman, L., & Krongrad, E. (1988). Risk factors for SIDS. *Annals of the New York Academy of Sciences, 533*(1), 13-30.

Johnston, C., Stevens, B., Craig, K., & Grunau, R. (1992). Developmental changes in pain expression in premature, full-term, two- and four-month-old infants. *Pain, 52,* 201–208.

Karelitz, S., & Fisichelli, V. R. (1962). The cry thresholds of normal infants and those with brain damage. *The Journal of Pediatrics, 61,* 679– 85.

Kent, R.D., & Murray, D.A. (1982). Acoustic features of infant vocalic utterances at 3, 6, and 9 months. *The Journal of the Acoustical Society of America, 72*(2), 353-365.

Kheddache, Y., & Tadj, C. (2013). Acoustic measures of the cry characteristics of healthy newborns and newborns with pathologies. *Journal of Biomedical Science and Engineering, 6,* 796-804.

Koivisto, M., Wasz-Ho¨ckert, O., Vuorenski, V., (1970). Cry studies in neonatal hyperbilirubinemia. *Acta Paediatrica Scandinavia, Supplement, 206,* 26-27.

LaGasse, L. L., Neal, A. R., & Lester, B. M. (2005). Assessment of infant cry: Acoustic cry analysis and parental perception. *Mental Retardation and Developmental Disabilities Research Reviews, 11,* 83–93.

Laitman, J. A., Crelin, E., Conologue, G. J. (1977). The function of the epiglottis in monkey and man. *Yale Journal of Biology and Medicine, 50,* 43-54.

Langlois, A., Baken, R. J., & Wilder, C. N. (1980). Pre-speech respiratory behavior during the first year of life. In T. Murry & J. Murry (Eds.), *Infant communication, cry and early speech* (pp. 526-84). Houston: College Hill Press.

Lederman, D. (2002). *Automatic classification of infants' cry*. Unpublished thesis submitted as a partial fulfillment of the requirements for the M.Sc degree. Ben-Gurion University, Negev.

Lederman, D., Cohen, A., Zmora, E., Wermke, K., Hauschildt, S., & Stellzig-Eisenhauer, A. (2002). *Automatic classification of the cry of infants with cleft palate.* Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.331.2976&rep=rep1&type=pdf

Lederman, D., Cohen, A., Zmora, E., Wermke, K., Hauschildt, S., & Stellzig-Eisenhauer, A. (2002). *On the use of hidden Markov models in infants' cry classification.* Retrieved from http://www.ee.bgu.ac.il/~drorle/is_ieee_1202.pdf.

Lederman, D., Cohen, A., Zmora, E., et al. (2004). HMM-based classification of infants' cries. Biomed Eng (submitted for publication).

Lederman, D., Zmora, E., Hauschildt, S., Stellzig-Eisenhauer, A., & Wermke, K. (2008). Classification of cries of infants with cleft-palate using parallel hidden Markov models. *Medical and Biological Engineering and Computing, 46*(10), 965-975.

Lester, B. M. (1984). A biosocial model of infant crying. In L. Lipsitt & C. Rovee-Collier (Eds.), *Advances in infant research* (pp. 167-212). Norwood, NY: Ablex.

Lester, B. M. (1987). Developmental outcome prediction from acoustic cry analysis in term and preterm infants. *Pediatrics, 80,* 529–534.

Lester, B.M., & Boukydis, C. F. Z. (1985). *Infant crying: Theoretical and research perspectives.* New York: Plenum Press.

Lester, B. M., & Boukydis, C. F. (1990). No language but a cry. In H. Papousek, J. Jurgens, & M. Papousek (Eds.). *Nonverbal vocal communication: Comparative and developmental approaches* (pp. 41-69). New York: Cambridge University Press.

Lester, B. M., & Dreher, M. (1989). Effects of marijuana use during pregnancy on newborn cry. *Child Development, 60,* 694 –705.

Lester, B. M., & Zeskind, P. S. (1982). A behavioural perspective on crying in early infancy. In H. E. Fitzgerald, B. M. Lester, & M. W. Yogman (Eds.), *Theory and research in behavioural paediatrics* (pp. 133-180). New York: Plenum Press.

Lester, B. M., Corwin, M. J., Sepkoski, C., Seifer, R., Peucker, M., McLaughlin, S., & Golub, H. L. (1991). Neurobehavioral syndromes in cocaine-exposed newborn infants. *Child Development, 62,* 694-705.

Lester, B. M., Tronick, E. Z., LaGasse, L., Seifer, R., Bauer, C.R., Shankaran, S.,……. 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.

Lewis, M. M. (1936). *Infant speech: A study of beginnings of language.* New York: Harcourt Brace and World, Inc.

Lieberman, P. (1975). *On the origins of language: An introduction to the evolution of speech.* New York: MacMillan.

Lieberman, P. (1985). The physiology of cry and speech in relation to linguistic behavior. In B. M. Lester & C. F. Boukddyis (Eds.), *Infant crying: Theoretical and research perspectives* (pp. 29-57). New York: Plenum Press.

Lind, J., Vuorenkoski, V., Rosberg, G., Partenen, T. J. & Wasz-Höckert, O. (1970). Spectographic analysis of vocal response to pain stimuli in infants with Down's syndrome. *Developmental Medicine and Child Neurology, 12,* 478 –486.

Lofqvist, A., & Mandersson, B. (1987). Long-time average spectrum of speech and voice analysis. *Folia Phoniatrica, 39,* 221-229.

Long, E., & Hull, W. (1961). Respiratory volume-flow in the crying newborn infant. *Pediatrics, 27,* 373.

Michelsson, K. (1971). Cry analysis of symptomless low birth weight neonates and of asphyxiated newborn infants. *Acta Paediatrica Scandinavica, 216,* 1-45.

Michelsson, K. (1980). Cry characteristics in sound spectrographic cry analysis. In T. Murry & J. Murry (Eds.), *Infant communication: Cry and early speech* (pp. 85-105). Houston: College Hill Press.

Michelsson, K., & Sirvio, P. (1976). Cry analysis in congenital hypothyroidism. *Folia Phoniatrica, 26,* 40-47.

Michelsson, K., Tupperainen, N., & Aula, P. (1980). Cry analysis of infants with karyotype abnormality. *Neuropediatrics,11,* 365–376.

Michelsson, K., Ja¨rvenpa¨a¨, A. L., & Rinne, A. (1983). Sound spectrographic analysis of pain cry in preterm infants. *Early Human Development, 8,* 141–149.

Michelsson, K., Karskinen, H., Aulanko, R., Rinne, A. (1984). Sound spectrographic cry analysis of infants with hydrocephalus. *Acta Paediatrica Scandinavia, 73,* 65–68.

Michelsson, K., Sirvio,¨ P., & Wasz-Ho¨ckert, O. (1977a). Pain cry in full-termasphyxiated newborn infants correlated with late findings. *Acta Pediatrica Scandinavia, 66,* 611-616.

Michelsson, K., Sirvio,¨ P., & Wasz-Ho¨ckert, O. (1977b). Sound spectrographic cry analysis of infants with bacterial meningitis. *Developmental Medicine and Child Neurology, 19,* 309–315.

Michelsson, K., Raes., J., Thoden, C. J., & Wasz-Hockert, O. (1982). Sound spectrographic cry diagnostic: An evaluative study. *Journal of Phonetics, 10,* 79-88.

Michelsson, K., Sirvio, P., Koivisto, M., Sovijarvi, A., & Wasz-Hockert, O. (1975). Spectrographic analysis of pain cry in neonates with cleft palate. *Biology Neonate, 26,* 353-358.

Mende, W., Herzel, H., & Wermke, K. (1990). Bifurcations and chaos in newborn infant cries. *Physics Letters A, 145,* 418–424.

Murube, J. (2009c). Tear apparatus of animals. Do they weep? *The Ocular Surface, 7,* 3–9.

Naeye, R. L., Messemet, III, J., Specht, T., & Merritt, T. A. (1976). Sudden infant death syndrome temperament before death. *Journal of Pediatrics, 88,* 511-515.

Noback, G. J. (1928). The developmental topography of the larynx, trachea, and lungs in fetus, new-born, infant, and child. *American Journal of Disordered Child, XXVI,* 515-533.

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

Oguz, H., Kilic, M. A., & Safak, M. A. (2011). Comparison of results in two acoustic analysis programs: Praat and MDVP. *Turkish Journal of Medical Sciences, 41*(5), 835-841.

Orozco, J., Carlos, A., & Garcia, R. (2003). "Detecting Pathologies from Infant Cry Applying Scaled Conjugate Gradient Neural Networks", *Proceedings of European Symposium on Artificial Neural Networks,* 349-354.

Ostwald, P.F. (1972). The sounds of infancy. *Developmental Medicine and Child Neurology, 14,* 350 -361.

Ostwald, P., Peltzman, P., Greenberg, M., & Meyer, J. (1970). Cries of a trisomy 13–15 infant. *Developmental Medicine and Child Neurolog, 12,* 472–477.

Partanen, T. J., Wasz-Hockert, O., Vuorenkoski, V., Theorell, K., Valanne, E., & Lind, J. (1967). Auditory identification pain cry signals of young infants in pathological conditions and in sound spectrographic basis. *Annales Pediatriae Fenniae, 13,* 56 -63.

Porges, S.W., & Maita, A. K., (1994). Vagal tone and the physiological patterns and later behaviour. In H. W. Reese & M. D. Franzen (Eds.), *Biological and*

neurophysiological mechanisms: Life-span Developmental Psychology (pp. 163-179). Mahwah, NJ: Erlbaum.

Rabiner, L. R., & Schafer, R. W. (1978). *Digital processing of speech signals.* Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Radhika Rani, L., Chandralingam, S., Anjaneyulu, T., Satyanarayana, K. (2012). Estimation of fundamental and formant frequencies of infant cries: A study of infants with congenital heart disorder. *Indian Journal of Computer Science and Engineering, 3*(4), 572-584.

Rautava, L., Ojala, S., Parkkola, R., Rikalainen, H., Lapinleimu, H., Haataja, L., & Lehtonen, L. (2007). Acoustic quality of cry in very low birth weight infants at the age of 1 ½ years. *Early Human Development, 83,* 5-12.

Reggiannini, B., Sheinkopf, S. J., Silverman, H. F., Li, X., & Lester, B. M. (2013). A flexible analysis tool for the quantitative acoustic assessment of infant cry. *Journal of Speech, Language, and Hearing Research, 56,* 1416-1428.

Reyes-Galaviz, O., & Reyes-Garcia, C. A. (2004). A system for the processing of infant cry to recognize pathologies in recently born babies with neural networks. *Proceedings of SPECOM 2004: 9th Conference Speech and Computer,* 552-557.

Reyes-Galaviz, O. F., Tirado, E. A., & Reyes-Garcia, C. A. (2004). Classification of infant crying to identify pathologies in recently born babies with ANFIS. *Proceedings of 9th International Conference, ICCHP 2004,* 408-415.

Robb, L. (2000). Emotional Musicality in mother-infant vocal affect, and an acoustic study of postnatal depression. *Musicae Scientiae, 3,* 123-154.

Robb, M. P, Crowell, D. H., Dunn-Rankin, P., & Tinsley, C. (2007). Cry features in siblings of SIDS. *Acta Paediatrica, 96 (10),* 1404-1408.

Rothenberg, S. J., Cansino, S., Sepkoski, C., Torres, L. M., Medina, S., Schnaas, L., Poblano, A., & Karchmer, S. (1995). Prenatal and perinatal lead exposures alter acoustic cry parameters of neonate. *Neurotoxicology and Teratology, 17*(2), 151–160.

Sangeetha, N. (1999). *Acoustic analysis of infant cry.* Unpublished dissertation, University of Mysore.

Santiago-Sánchez, K., Reyes-Gracia, C. A., & Gomez-Gil, P. (2009). Type-2 fuzzy sets applied to pattern matching for the classification of cries of infants under neurological risk. In D. S. Huang, K. H. Jo, H. H. Lee, H. J. Kang, & V. Bevilacqua (Eds.), *Emerging intelligent computing technology and applications* (pp. 201-210). Mexico: Springer Berlin Heidelberg.

Scammon, R. (1923). A summary of the anatomy of the infant and child. In Issac Apt *Pediatrics* (pp. 296-297). Philadelphia: W.B. Saunders.

Sedlackova, E. (1964). Analyse acoustique de la voix de nouveau-nes. *Folia Phoniatrica, 16,* 48-58.

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 Research, 5*(5), 331-339.

Shinya, Y., Kawai, M., Niwa, F., & Myowa-Yamakoshi, M. (2014). Preterm birth is associated with an increased fundamental frequency of spontaneous crying in human

infants at term-equivalent age. *Biology letters*, doi: 10: 20140350 http://dx.doi.org/10.1098/rsbl.

Sirvio, 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 Phoniatrica, 28,* 161–173.

Stark, R. E., & Nathanson, S. (1975). Unusual features of cry in an infant dying suddenly and unexpectedly. In J. Bosma & J. Showacre (Eds.), *Development of upper respiratory anatomy and function: Implications for the sudden infant death syndrome*. Washington, DC: US Printing Office.

Stark, R. E., Rose, S. N., & McLagen, M. (1975). Features of infant sounds: The first eight weeks of life. *Journal of Child Language, 2,* 205-221.

Thoden, C. J., & Michelsson, K. (1979). Sound spectrographic cry analysis in Krabbe's Disease. *Developmental Medicine and Child Neurology, 21,* 400–401.

Thoden, C. J., & Koivisto, M. (1980). Acosutic analysis of normal pain cry. In T. Murry & J. Murry (Eds.), *Infant communication: Cry and early speech* (pp. 124-151). Houston: College Hill Press.

Titze, I. R, Baken, R., & Herzel, H. (1993). Evidence of chaos in vocal fold vibration. In I. Titze (Ed.), *Vocal fold physiology: Frontiers in basic science* (pp. 143-148). San Diego, CA: Singular Publishing Group.

Truby, H., & Lind, J. (1965). Cry sounds of the newborn infant. *Acta Paediatrica Scandinavia, 163,* 39-59.

Van der Berg, J. (1958). Myoelastic-aerodynamic theory of voice production. *Journal of Hearing and Speech, 1,* 227 -244.

Varallyay Jr, G., Benyo, Z., Illenyi, A., Farkas, Z., & Kovacs, L. (2004). Acoustic analysis of the infant cry: Classical and new methods. *Proceedings of the 26th Annual International Conference of the IEEE EMBS,* 313-316.

Vingerhoets, A. J. M., Bylsma, L., & Rottenberg, J. (2009). Crying: A biopsychosocial phenomenon. In T. Fogen (Ed.), *Tears in the Graeco-Roman World* (pp. 439–475). Berlin & New York: de Gruyter.

Vohr, B. R., Lester, B., Rapisardi, G., O'Dea, C., Brown, R. N. C., Peucker, M., Cashore, M. D. W., & William Oh, M. D. (1989). Abnormal brain-stem function (brain-stem auditory evoked response) correlates with acoustic cry features in term infants with hyperbilirubinemia. *Pediatrics, 115,* 303–308.

Vuorenkoski, V., Lind, J., Partanen, T. J., Jejeune, J., Lafourcade, J., & Wasz-Hockert, O. (1966). Spectrographic analysis of cries from children with maladie du cri du chat. *Annales Paediatriae Fenniae, 12,* 174–180.

Wasz-Hockert, O., Partanen, T. J., Vuorenkoski, V., & Valenne, E. (1964). The identification of some specific meanings in infant vocalization. *Experientia, 20,*154.

Wasz-Hockert, O., Koivisto, M., Vuorenkoski, V., Partanen, T., & Lzind, J. (1971). Spectrographic analysis of cry pain in hyperbilirubinemia. *Biology of the Neonate, 17,* 260-271.

Wasz-Hockert, O., Valanne, E., Vuorenkoski, V., Michelsson, K., & Sovijarvi, A. (1963). A analysis of some types of vocalization in the newborn and in early infancy. *Annales Paediariae Fenniae, 9,* 1-10.

Wasz-Hockert, O., Lind, J., Partanen, T., Valenne, E., & Vuorenkoski, V. (1968). The infant cry: A spectrographic and auditory analysis. *Clinics in Developmental Medicine, 2*9-42.

Wasz-Höckert, O., Lind, J., Vuorenkoski, V., Partanen, T., & Valanne, E. (1968). *Clinics in developmental medicine*. The infant cry: A spectrographic and auditory analysis. London: Spastics International Medical.

Wermke, K., & Robb, M. P. (2010). Fundamental frequency of neonatal crying: Does body size matter? *Journal of Voice, 24*(4), 388–394.

Wolff, P. H. (1967). The role of biological rhythms in early psychological development. *Bulletin of the Menninger Clinic, 31,* 197-218.

Wolff, P. H. (1969). The natural history of crying and other vocalization in early infancy. In B. M. Foss (Ed.), *Determinants of Infant Behaviour,* 4 (pp. 81-109). London: Metheun.

Zeifman, D. M. (2001). An ethological analysis of human infant crying: Answering Tinberger's four questions. *Developmental Psychobiology, 39*(4), 265-285.

Zeskind, P. S. (1983). Cross-cultural differences in maternal perceptions of cries of low- and high-risk infants. *Child Development, 54*, 1119 –1128.

Zeskind, P. S., & Barr, R. G. (1997). Acoustic characteristics of naturally occurring cries of infants with "colic". *Child Development, 68*(3), 394-403.

Zeskind, P. S., & Lester, B. M. (1978). Acoustic features and auditory perceptions of the cries ofnewborns with prenatal and perinatal complications. *Child Development, 49,* 580–589.

Zeskind, P. S., & Lester, B.M. (2001). Analysis of infant crying. In L. T. Singer & P. S. Zeskind (Eds.), *Biobehavioral assessment of the infant* (pp. 149-166). New York: Guilford Publications Inc.

==========

Appendix 1

Guide to install and use the software *ICry*

Introduction

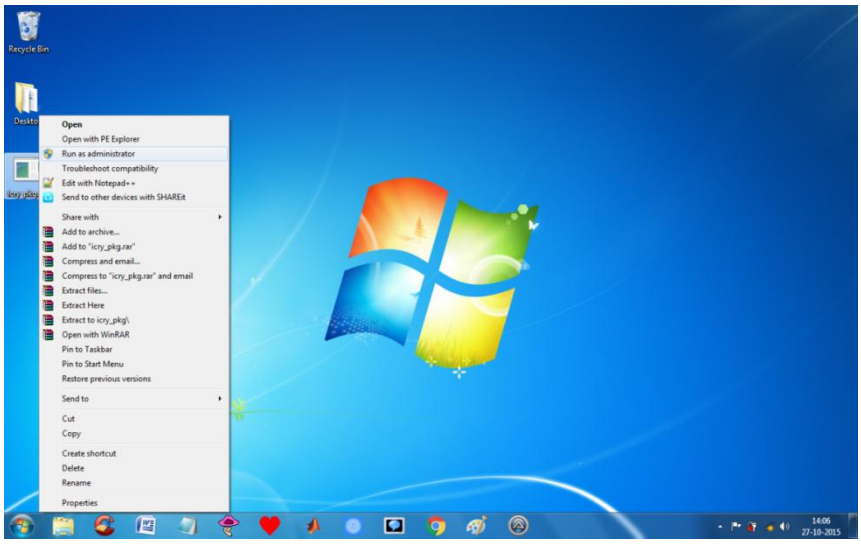
ICry software is an application developed as a screening tool, which takes infant cry samples as the input and provides output in form of Thumbs up (pass) / Thumbs down (At risk cry). This software can be used along with standard screening procedures to provide a more comprehensive idea about the neurophysiological integrity of the infant.

The software was developed on a MATLAB R2009b platform and a standalone application was deployed out of it. *ICry* application requires few prerequisites to be installed before using it. This file will guide you through the installation procedure. A simple troubleshooting file is also provided with the package. Please read it, if you have any problems while installing.

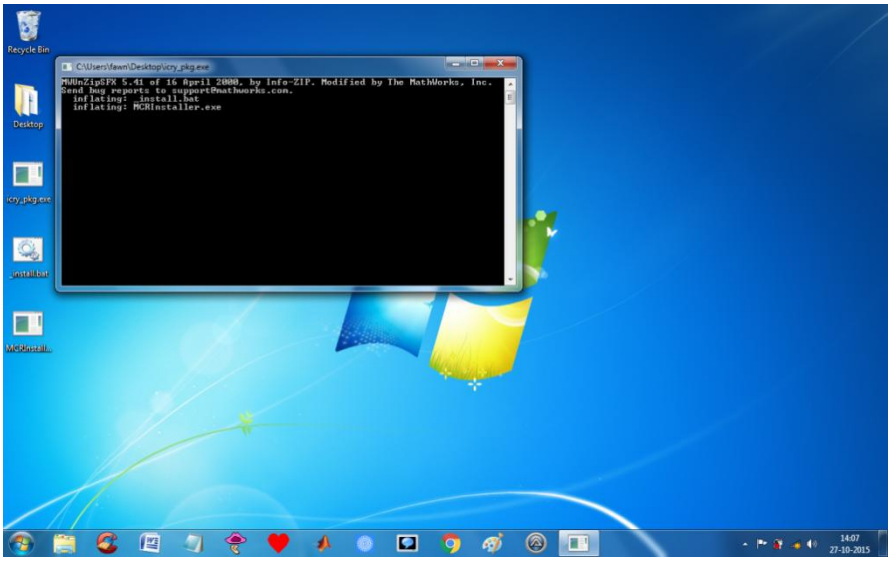
Basically when the *ICry*_pkg executive file is clicked, it automatically installs the Matlab Compiler Runtime (MCR) application, followed by the complete installation of the *ICry* software. Once the installation is complete, 3 additional files will be added to the directory which includes a readme file, a MCR install file and ICry.exe. Based on the data that was provided after data collection, the software was developed by software professionals.

Step 1: Installing the necessary Framework: A readme.txt file has been provided with the software package which describes the minimum requirements that is necessary for the software to run.

Step 2: The software can be installed by right clicking on the ICry_pkg.exe and then clicking on run as administrator. Then click “Yes” to proceed.



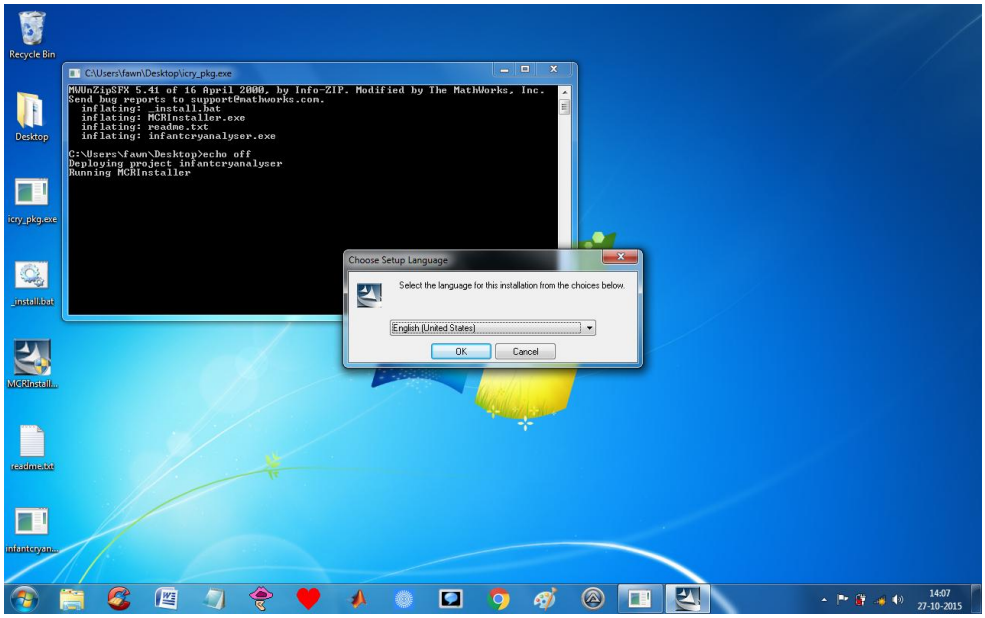
As soon as the package is run, a self guiding extraction procedure takes place where it will guide you through an installation procedure which will ask you to install Matlab Compiler Runtime (MCR). Additional necessary files will be installed along with MCR.



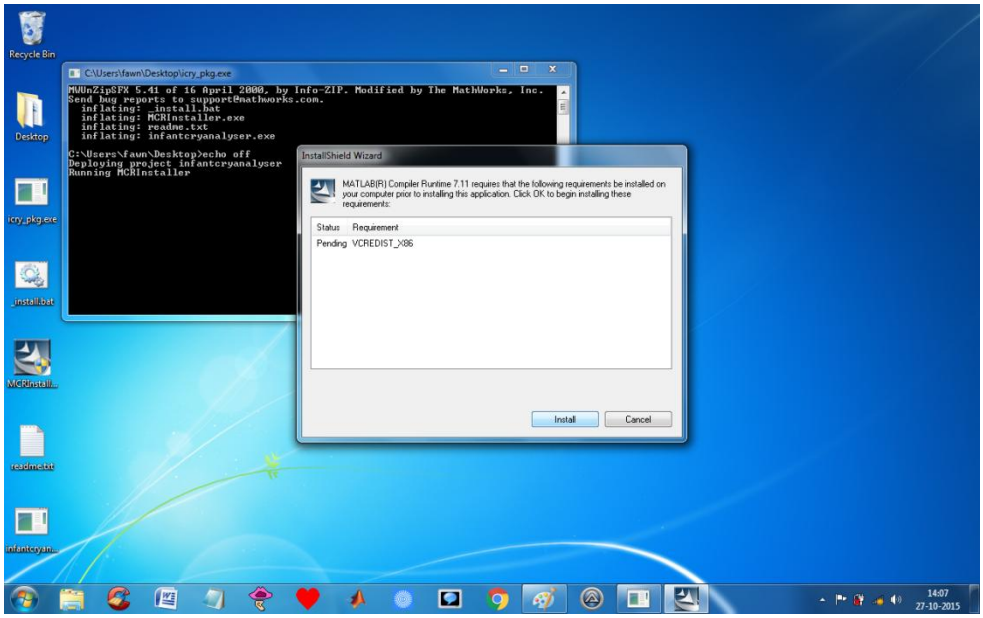
In the black window press “y” at the cursor and then press “enter”. Repeat the same till the language option window is pops up.

Step 3: Please follow through the installation procedure by selecting the desired language.

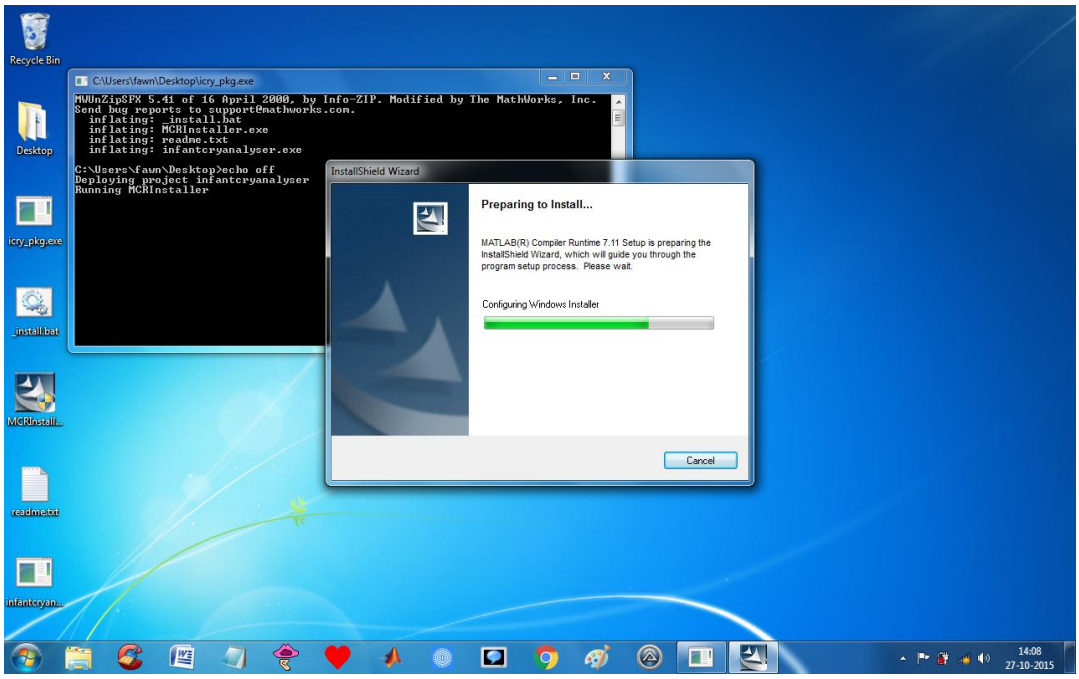
And then click ‘OK’ to proceed.

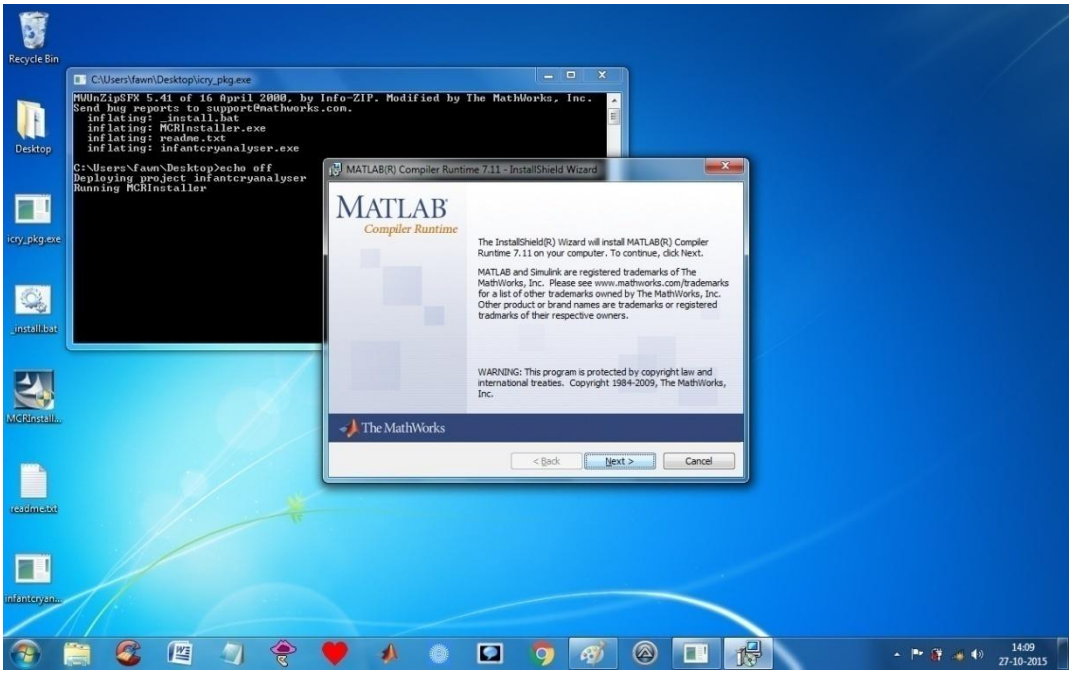


Step 4: Click on ‘Install’ to start the installation

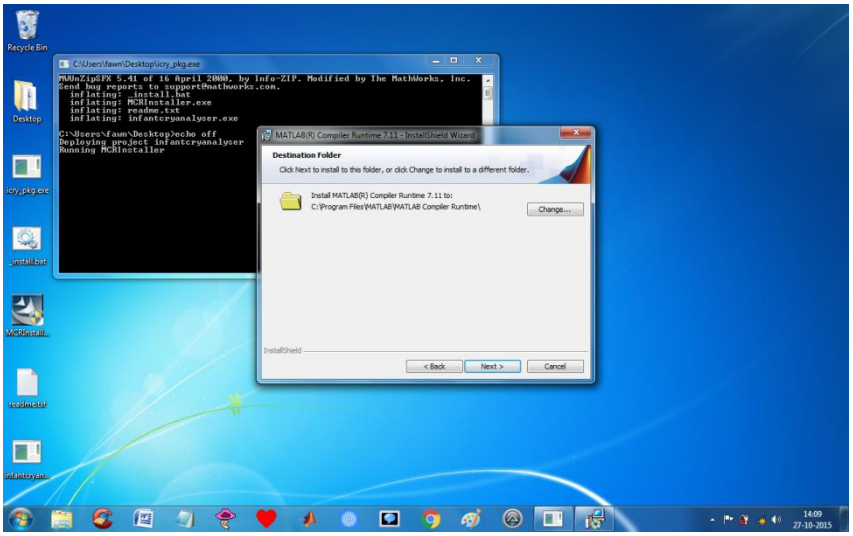


Step 5: This will open a series of dialogue boxes which will display the configuration and the status of the installation procedure.

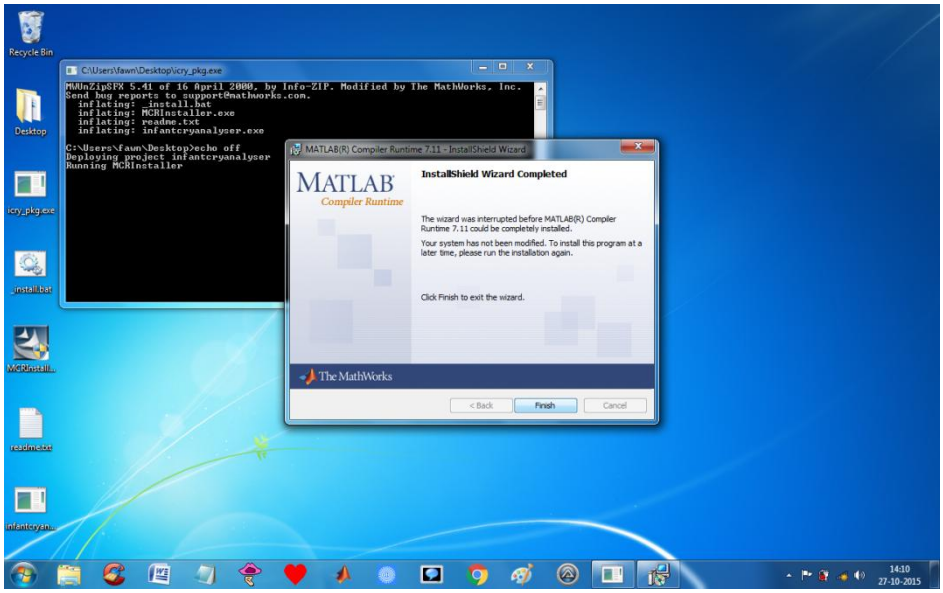




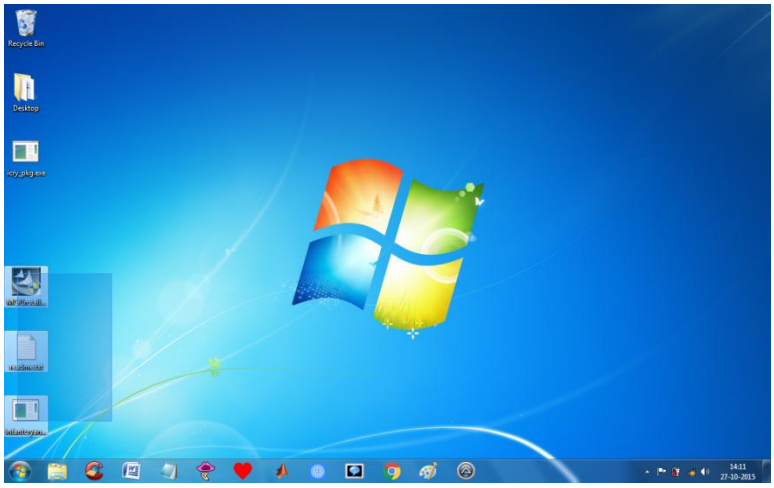
Step 6: Once the necessary packages are installed the path for the Matlab compiler runtime will be specified. By default it shows the path to be the ‘C’ drive, if necessary one can change the default path to other drives as required and continue the procedure.



Step 7: Once the installation procedure completes, the installation wizard is displayed depicting the status of installation. If successfully installed, use the ‘Finish’ button to complete the installation.

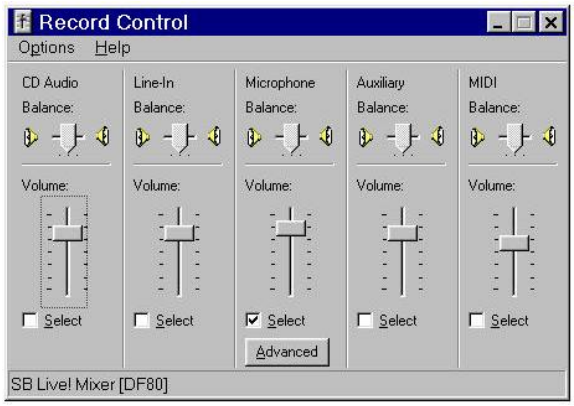


Step 8: Once the installation is complete, 3 additional files will be added to the directory which includes a readme file, a MCR install file, and icry.exe software.



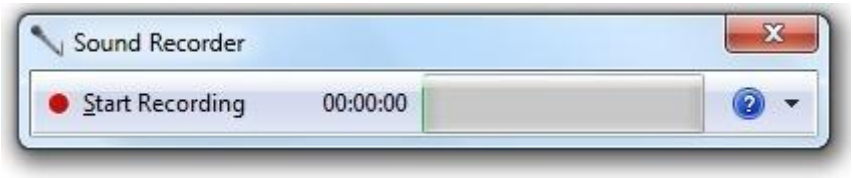
Note: After installation, please run ICry.exe, if you find any errors or issues please refer the troubleshooting file or contact the authors.

Step 9: Before starting the recording, need to connect a microphone or their sound source to the microphone input on your sound card. The next step is to ensure that your computer is set up to record from the microphone. On a Windows machine, you must select the microphone as the source in the Record Control window (see illustration below). The Record Control window can usually be accessed from a speaker icon in the system tray.



Microphone selected as the recording source

Windows *Sound Recorder* program can be used to verify that the microphone is configured correctly. If sounds can be recorded using this program, they can also be recorded in Matlab. If you *can't* record sounds, there is some problem with the configuration.



Windows Sound Recorder in action

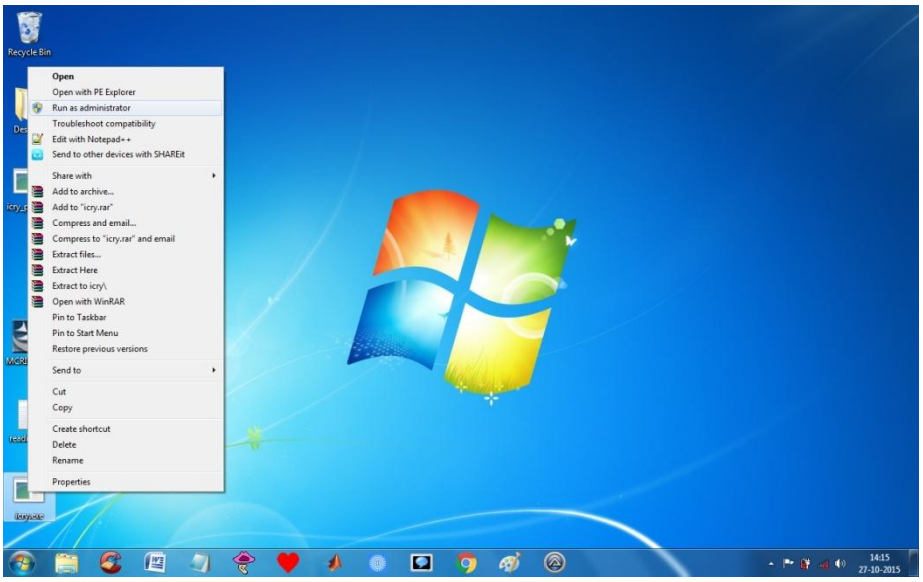
The remainder of this manual will describe the Matlab *Record* program—its inner working and functionality

Guide to use the software i_cry.exe

Step 10: The system is now loaded with the infant cry analyser (i_cry.exe) which can be used to screen infant using their cry signals.

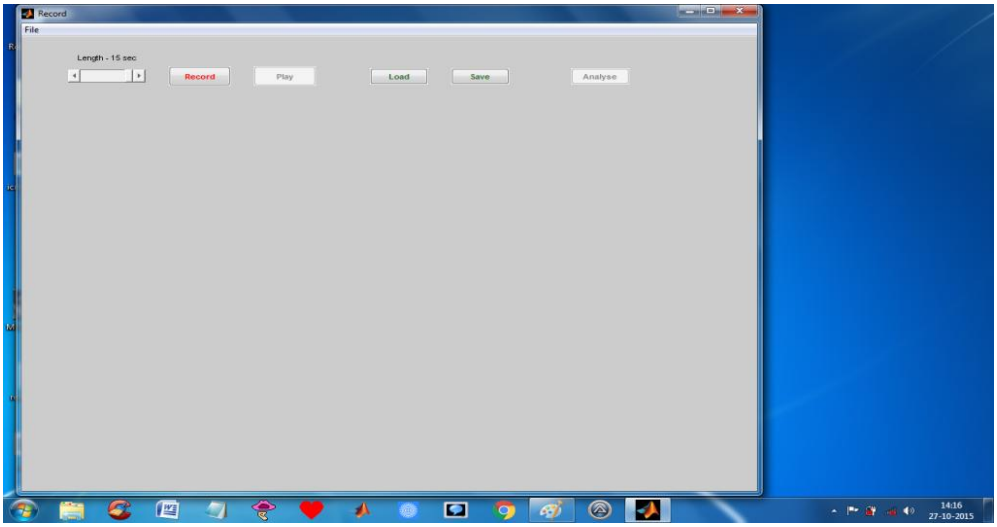
To run the file double click on it or right click on it and open it as the admin.

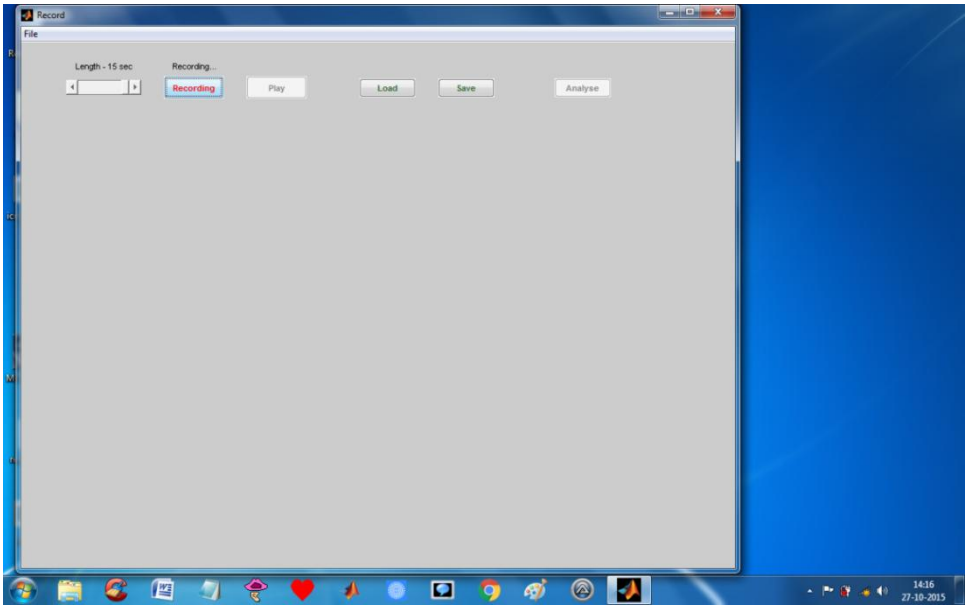
It will take few seconds to load. Please wait till the recording screen pops up.



Step 11: The below picture depicts the recording screen of the software. The time window scale represented by ’Length’ can be modified accordingly to set the time window for analysis.

Once the external microphone and the other recording protocols are met, the record button can be clicked to start the recording.

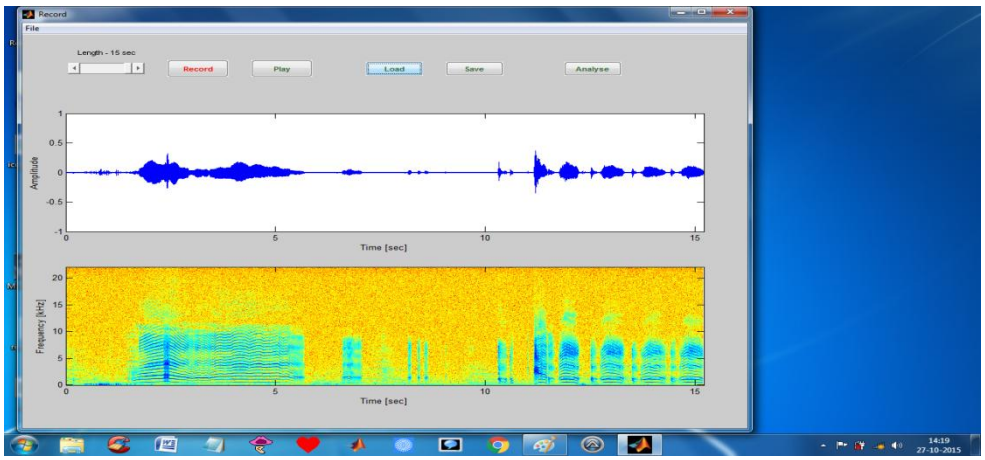




Note: A display above the record button (seen in red) indicates the status of recording.

If any issues with recording is seen, no further screens will be displayed and the software displays an error message saying “10 seconds has elapsed Try again!.” If you obtain this error or any winsound errors (system errors) please check the external microphone connection or sound card settings.

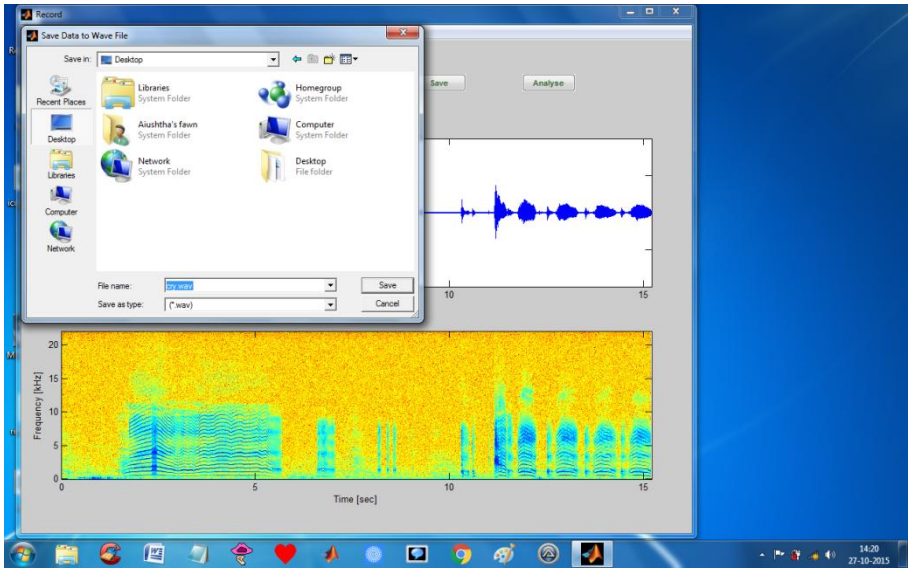
Note: If the recording is successful, the waveform and a narrow/ a broad band spectrogram will be displayed based on the length of the recording.



Step 12: Once the sample has been recorded, the ‘play’ and the ‘save’ button gets highlighted and can be used to replay or save the signal respectively.

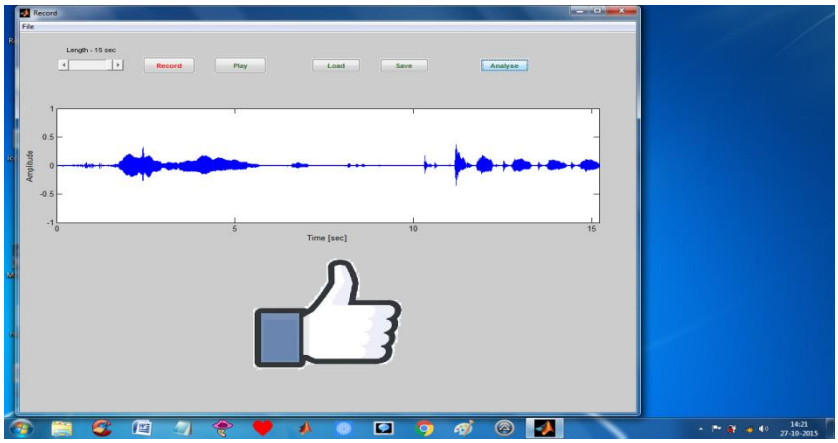
If you click on load button a window will pop up opening the directory containing the file to be analysed**.** It should be noted that only ‘MONO’ sounds can be loaded and analysed.

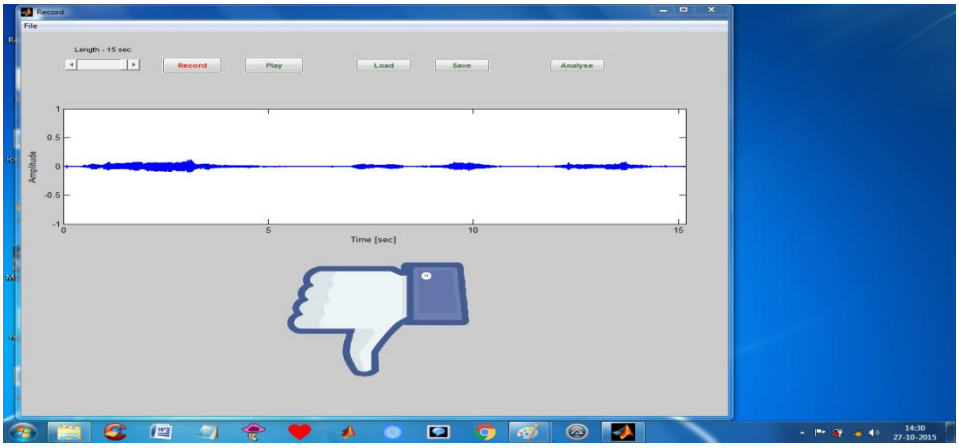
Stereo files should be converted into mono files externally before loading the samples onto the software. If stereo files are loaded, an error message will pop up displaying “cannot load stereo data.”



Step 13: After saving or loading the sample to be analysed, it can be started by clicking on 'analysis' button. As soon as this button is clicked, a pop up window will be displayed depicting the progress of analysis.

Note: Typically the analysis take around 10-15 seconds, please wait for the analysis to complete





Once the analysis process completes thumbs up (PASS) or thumbs down (AT RISK CRY) will be displayed accordingly.

For more information or any queries

Contact: Dr N. Sreedevi
Department of Speech Language Sciences
All INDIA INSTITUTE OF SPEECH & HEARING
MYSORE – 570 006

sreedeviaiish@gmail.com
09449953666

Appendix 2

Troubleshooting Guide

This document provides simple solutions for basic troubleshooting. Please follow the same and Contact the authors if you have any queries.

Troubleshooting Step 01: For Missing dll Errors

It is recommended to reinstall the software if you obtain any .dll errors. Kindly follow the instructions on the screen to install successfully (In case you are reinstalling press y for yes and n for no)

Troubleshooting Step 02: For Side-by-Side Errors

Dear user,

Most Side by Side configuration errors are caused by missing C++ runtime components or corruption.

Explanation:

All Microsoft runtime libraries belonging to different Visual Studio versions such as 6.0, 2003 (.NET 1.x), 2003 (.NET 2.x), 2005 and 2008 are independent and incompatible with each other. See the Related Solution listed below for more information on which version of the Visual C++ Redistributable Package is needed for each MATLAB and/or MCR version. Since MATLAB 7.6 (R2008a) requires the Visual C++ 2005 SP1 Redistributable Package, installing the 2008 version or any other version will not be enough and will lead to errors such as "Side-by-Side configuration error".

Solution:

It is recommended to install all of the following:

Visual C++ 2005 Redistributable Package for (X86)

http://www.microsoft.com/en-us/download/details.aspx?id=3387

Microsoft Visual C++ 2008 Redistributable Package (x86)

http://www.microsoft.com/en-us/download/details.aspx?id=29

Microsoft Visual C++ 2010 Redistributable Package (x86)

http://www.microsoft.com/en-us/download/details.aspx?id=5555

Microsoft Visual C++ Redistributable Packages for Visual Studio 2013

https://www.microsoft.com/en-us/download/details.aspx?id=40784

If you are running a 64-bit system, then you would need to install the (x64) version of the above files. You can download from

http://www.microsoft.com/en-us/download

Hope this helps.

Troubleshooting Step 03: To Check Side by Side Error in the Event Log in Windows 7

1. Open Event Viewer by clicking the **Start** button, clicking **Control Panel**, clicking **System and Security**, clicking **Administrative Tools**, and then double-clicking **Event Viewer**. if you are prompted for an administrator password or confirmation, type the password or provide confirmation.

2. Click an event log in the left pane.

3. Double-click an event to view the details of the event

Troubleshooting Step 04:10 Seconds Elapsed Try Again!

If you find this error check your microphone connection (refer readme getting started part for description)

Troubleshooting Step 05: Can’t Load Stereo Data

ICry software is unable to read and analyse stereo data. If is requested to convert your stereo file into mono file externally and use it in ICry.exe

==========