Resonance Frequencies Behavior in Pathologic Cries of Newborns

Yasmina Kheddache and Chakib Tadj, *Montréal, Quebec, Canada*

Summary: Objective. A new approach to the automatic quantification of the acoustic parameters of the cries of healthy newborns and newborns with pathologies is presented. The purpose of the present study was to examine the relationship between acoustic parameters and pathologies of interest to characterize healthy and pathologic cries of newborns.

Methods. Using *MATLAB*, this study included automatic estimation of F0, RF1, RF2, percentage and tuning duration, transition duration, RF2 slope, and RF1:RF2 ratio. The database used includes full- and pre-term newborns, healthy, and pathologic cries. It contains 3000 cry samples of 1-second duration from 65 newborn babies aged from 1 day to 1 month old.

Results. Statistical analysis results reveal that the distributions of these acoustic cry parameters depend on the pathology of newborn. In this work, we successfully identify the quantitative relationship between the acoustic cry characteristics we examined and the diseases we studied.

Conclusions. Our deduction is that quantification of the variability of these parameters is useful for differentiating the cries of a healthy newborn from those of a newborn with a pathology, and that these data can be used for the early diagnosis of newborn diseases.

Key Words: Newborn cry–Resonance frequencies–Tuning–Transition duration–RF2 slope–RF1:RF2 ratio.

INTRODUCTION

In this article, the newborn cry is assessed acoustically, to correlate its acoustic properties to specific pathologic conditions and to show how acoustic variations in the cry reflect the degree of distress of the newborn. In most crying research, the spectrogram has been used in cry analysis.[1–3] This method of analysis requires manual selection, which in turn requires some technical knowledge. In our work, we automate the evaluation of acoustic cry characteristics to improve infant monitoring in the first days of life.

To further investigate the relationship between infant crying and medical conditions from which infants may suffer, a set of the most important acoustic measures must be selected containing the most relevant information available based on physiological models of cry production. It is also important to understand the physioacoustic structure and levels of the nervous system responsible for muscle control and cry production.

The various acoustic characteristics are formed in accordance with changes in the status of the vocal tract and may be affected by activity along the entire vocal production pathway, including respiration, vocal fold behavior, and vocal tract shape.[4,5]

Greater control of vocalization in infants aged around 1–2 months leads to greater differentiation in their cries, as a result of the physiological and anatomic changes occurring during this period. According to the study by Soltis,[6] at around 3 months of age, crying is contextdependent, intentional, and communicative. At around 7 months, “babbling” occurs.[7] Consequently, because from 1 month of age, infants begin to acquire voluntary control of the vocal tracts,[8] looking at the spontaneous cries of infants during the first weeks of life is crucial if we are to apply this information to the early diagnosis of various newborn pathologies. Our study concentrates on the characterization of the cries of newborns aged 1 day to 1 month.

In this article, the acoustic characteristics of healthy and pathologic cries of pre- and full-term newborns are measured by automated means. The median and interquartile range of these characteristics are used to assess the degree of variation of the following:

- the average resonance frequencies (RF1 and RF2) and the RF1:RF2 ratio;
- the fundamental frequency (F0);
- the percentage and duration of tuning periods (TUPs) for RF1 and RF2 with the first 10 harmonics of F0;
- the duration of the transition periods (TRP) between two consecutive TUPs;
- the RF2 slope.

These acoustic parameters were selected based on previous studies in which their usefulness, either for the development of phonatory and articulatory capabilities or for characterizing the production deficit in dysarthric speech, was reported. In the study by Kim et al[9,10] and Wermk et al,[11] for example, the authors conclude that the irregular behavior of any of these parameters could provide suggest possible neurologic dysfunction.

However, F0 and its variations over time are viewed as an essential component of reliable information on the health status of newborns,[12,13] and RFs reflect important acoustical characteristics of the vocal tract of the infant.[9] TRP and tuning between RFs and the F0 harmonics have not been determined in infants younger than 4 months.[11] These characteristics are interpreted as early articulatory activity in the infant cry[14] and

Accepted for publication April 9, 2014.
From the Department of Electrical Engineering, École de Technologie Supérieure, Montréal Quebec, Canada.

Address correspondence and reprint requests to Yasmina Kheddache, Department of Electrical Engineering, École de technologie supérieure, Montréal, Canada H3C 1K3. E-mail: yasmina.kheddache.1@etsmtl.net
Journal of Voice, Vol. 29, No. 1, pp. 1-12
0892-1997/$36.00

http://dx.doi.org/10.1016/j.jvoice.2014.04.007

are related to the development of phonatory and articulatory capabilities.[11]

The RF1:RF2 ratio is an important marker of the relationship between the first and the second formants and helps us to understand the RFs behavior in healthy and pathologic newborn cries. This parameter is used to evaluate changes in voice and articulation in children with cochlear implants, for example, a study by Poissant et al[15] and Seifert et al.[16]

Also, the RF2 slope has been reported in previous studies as one of the most sensitive indices of vocal tract function for speech production. It has been considered to be an indicator of the severity of dysarthria and of neurodegenerative diseases.[11] This parameter has not yet been measured in newborns or younger infants, however.

The method typically applied for RFs estimation is the parametric autoregressive approach. This method is particularly well suited to the assessment of the newborn cry, which is characterized by higher RFs than those of adults.[9,17] For F0 estimation, we use the simple inverse filtering tracking (SIFT) algorithm, described in a study by Kheddache and Tadj[12] and Kheddache and Tadj.[13]

This article is organized as follows. We first present a physioacoustic model of the newborn cry in Section 2, and characteristics of newborns cries in a medical environment in Section 3. We then explain our methodology for measuring the characteristics of pathologic and healthy cries in Section 4. In Sections 5 and 6, we provide a follow-up analysis of the results and present our conclusions in Section 7.

PHYSIOACOUSTIC MODEL AND NEUROLOGIC BASIS OF THE NEWBORN CRY

The cry signal is the result of coordination among several areas of the brain, which control respiration and the vocal fold vibration from which the cry sounds are produced.[1,5] These sounds are created by the respiratory system (lungs and trachea), the vocal folds (larynx), and the vocal tract (pharynx and oral and nasal cavities).[6]

Cry sounds are generated in the larynx, which contains the vocal folds and glottis, by air being forced through a constricted tube, which causes vibration of the vocal folds at F0. This sound is then filtered as it proceeds through the vocal tract and the lips. The vocal tract acts as a resonance cavity, and its instantaneous shape determines the RFs. So, the cry sound is affected by formant frequency, the size, and contours of upper vocal tract resulting in the audible cry.[8,18]

The shape of the newborn's vocal tract is more like that of a chimpanzee than that of a human adult, where the position of the larynx is higher in the vocal tract. The vocal tract of the newborn is also shorter ($L \approx 8.5$ cm)[18] and has a different structure than that of an adult. This means that it is associated with higher resonances and a higher fundamental frequency than that in adults,[19] because the formant frequencies decrease as the length of the resonance tract increases.

F0 varies from 250 to 450 Hz, with an initial formant at a frequency RF1 of 1100 Hz and a second formant at a frequency RF2 of 3100 Hz.[1] Vocal fold vibrations produce three identifiable modes of cry, defined as follows: (a) a basic cry or phonation with F0 = 350–750 Hz; (b) a cry with a high F0 (750–1000 Hz) or hyperphonation F0 (1000–1800 Hz); and (c) noisy, turbulent, or dysphonic cries.[12,13,20]

Interaction between laryngeal and pharyngeal activity is interpreted by tuning processes between the cry melody and the RFs.[14] The TUP is defined as the time (>20 ms) during which an RF remains close to the harmonic distance (<100 Hz). TRP is time between two consecutive TUPs.[9,18] TRP depends on neuro-physiological maturity and the integrity of the underlying control systems.[11] According to the study by Sundberg et al,[21] RFs tuning at around F0 increases the sound level of the vocal output.

The nervous system innervates the muscles that control the respiratory system, the vocal folds, and the vocal tract.[6] Consequently, the characteristics of cries are influenced by the cranial nerves that innervate the larynx, pharynx, and chest.[6,8]

NEWBORN CRY CHARACTERISTICS IN A MEDICAL ENVIRONMENT

Deficits in brain functioning can affect the vagal control of the cry.[1] Many of the early researchers examined healthy and pathologic cries by observing spectrograms and the spectra of the audio signals of cries. They essentially assert that infant cries are dependent on physical and psychological status, as well as both internal and external stimuli.[5,18,22,23] In spite of differences in measurement procedures, all cry studies have shown that a high F0 is an indicator of a neurologic problem.[5] Hyperphonic cries and very high pitched cries are associated with neurologic problems.[5] Other markers associated with neonatal disease include noisy or dysphonic cries, as well as changes in phonation mode, and variability in F0 and F1.[12,17,20] When a central nervous system disorder is involved, the cry exhibits auditory abnormalities with a high F0 and an irregular melodic contour.[8]

The work conducted in the study by Kheddache and Tadj[12] shows that the average percentages of hyperphonic segments in cry samples are similar for both healthy premature infants and healthy full-term newborns. It also shows that the average percentages of F0 irregularity are slightly higher for premature infants than for healthy full-term newborns. In addition, according to the study by Kheddache and Tadj,[20] the average percentages of dysphonic segments in the cry samples of healthy preterm newborns are higher than those of healthy full-term newborns. This difference may be a result of the immature innervation of the larynx in preterm newborns.

RF1 and RF2 have been used to identify differences between children who are profoundly deaf and those with normal hearing.[15] According to the study by LaGasse et al[5] and Cecchini et al,[8] high variability in RF1 and RF2 (dysregulation) indicates poor or unstable neural control of vocal track and respiration. This characteristic has been associated with hyperbilirubinemia, as well as prenatal tobacco and cocaine exposure. A higher than normal RF1 has been considered to be representative of an excitable neurobehavioral syndrome resulting from the direct effects of cocaine.[24]