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]

METHOD

Recording sessions

The recordings were made in the Pediatrics Department at Sainte-Justine Hospital under the supervision of medical personnel. The cries of newborns 1 day to 1 month old were recorded using a small recording device placed 10 cm from the baby's mouth, at a sampling rate of 44.1 kHz. These include full- and pre-term newborn cries (premature infants at under 37 weeks' gestation and full-term infants at 37 to 42 weeks' gestation), healthy infants, and infants suffering from pathologic conditions.

The cries of each infant were recorded two or three times at intervals of at least 1 hour over a period of not more than 10 days. The conditions during which the cries were recorded were: hunger, blood sampling, and diaper changing. The date and time of recording, and the infant's identification, birth date, diagnosis, ethnicity, and so on, were noted for each crying episode. To confirm the health status of the infant or to determine that an undiagnosed health problem was, in fact, present during recording, we noted the results of a follow-up examination conducted 6 months after the cries had been recorded. We further categorized the cry recordings by pathology and by gestational age. A sample from the newborn cry database (CDB) that we created is presented in Table 1. The CDB contains 3000 cry samples of 1-second duration from 65 newborn babies.

Procedure

Noise filtering and segmentation of the recordings into useful and nonuseful categories were performed in the first step of the assessment procedure using *Praat* (Boersma and Weenink, University of Amsterdam, Amsterdam, The Netherlands). For the acoustic analysis of newborn cries, an automated assessment procedure was performed using *MATLAB* (MathWorks, Natick, MA) to measure the characteristics defined in Table 2. The main steps of the assessment procedure are presented in Figure 1.

We applied the following procedure for acoustic characteristic measurement:

- Division of the cry samples of 1-second duration into overlapping frames of n = 1024 samples, resulting in 512 covered frames, and the multiplication of each frame by the Hamming window.
- For each data frame, the first five acoustic measures defined in Table 2 were estimated as follows:
  - Estimation of F0 and its 10 harmonics using the SIFT algorithm.[25]
  - Estimation of RF1, RF2 using the modified covariance method based on autoregressive power spectral density (AR-PSD).[18,19]
  - Calculation of the RF1:RF2 ratio (V_{ratio}).
  - Determination of the tuning between the RFs and the first 10 harmonics.
- For each cry sample of 1-second duration, we estimated:
  - the average F0, RF1, and RF2 (A_{F0}, A_{RF1}, and A_{RF2});
  - the average RF1:RF2 ratio (A_{ratio});
  - the tuning percentage of (P_{TUP});
  - the tuning duration (D_{TUP}) for RF1, RF2 and its average (A_{TUP});
  - the transition duration (D_{TRP}) for RF1, RF2 and its average (A_{TRP});
  - the extent of the transition;
  - the RF2 slope (V_{SLP}) and its average (A_{SLP}; absolute values were taken for the slope and the extent of the transition).

The following formulas were used:

$$A_{F0} = \frac{\sum F0}{N_{Total}},\ A_{RF1} = \frac{\sum RF1}{N_{Total}},\ A_{RF2} = \frac{\sum RF2}{N_{Total}},$$
$$A_{ratio} = \frac{\sum V_{ratio}}{N_{Total}},$$

$$P_{TUP} = \frac{N_{TUP}}{N_{Total}},\ A_{TUP} = \frac{\sum D_{TUP}}{N_{TUP}},\ A_{TRP} = \frac{\sum D_{TRP}}{N_{TRP}},$$
$$A_{SLP} = \frac{\sum V_{SLP}}{N_{TRP}}$$

N_{Total}: Total number of segments in cry samples of 1-second duration; N_{TUP}: Number of frames where tuning occurred

TABLE 1.
Pathologies Studied

Gestational Age	Pathology	Age at Recording (d)	Sample Size
Full-term newborn (*t*)	Healthy	1–3	1010
	Vena cava thrombosis	2–3	77
	Meningitis	14–15	115
	Peritonitis	28	20
	Asphyxia	4–28	190
	Lingual frenum	1–2	141
Preterm newborn (P)	Healthy	3–28	764
	IUGR-microcephaly	1	78
	Tetralogy of Fallot	17	53
	Gastroschisis	6–20	134
	IUGR-asphyxia	30	148
	RDS	28	270

TABLE 2.
Cry Characteristics Measured

Characteristic	Definition
Fundamental frequency (F_0)	Vibration frequency (in Hz) of the vocal folds.
Harmonics	Multiples of fundamental frequency
RF1, RF2 (Hz)	The first and the second vocal tract RFs
RF1:RF2 ratio	Derived using RF1 and RF2
Tuning (TUP)	Blocks of 1024 samples during which an RF remains close to the harmonics of F0 (distance <100 Hz)
Tuning duration (ms)	The time between the start and end of tuning
Transition duration (TRP) (ms)	Duration between two subsequent TUPs, or the time between the start and end of transition
Transition extent (Hz)	Frequency change between the transition start and end
RF2 slope (Hz/ms)	Derived using transition extent and duration measures.

between RF1, RF2, and the F0 harmonics; N_{TRP}: Number of transition durations in cry samples of 1-second duration.

- Calculation of the median and interquartile range of the evaluated characteristics of pathologic cries and healthy cries.

RFs estimation

The cry signal as a speech signal contains both vocal tract and pitch information. In this article, we use the source filter linear model of speech production in which the glottis and the vocal tract are considered independently. So, passing the glottal pulse train through the filter makes it possible to remove the correlation between the vocal tract information and the pitch information.[19]

Assuming that the vocal tract acts as the filter, then, according to the all-pole autoregressive model, the cry signal is the output of an all-pole filter $1/A(z)$, $A(z)$ being called the inverse filter.

Consequently, the cry sample is a weighted linear combination of n previous samples described by the equation:

$$y(k) = \sum_{i=1}^{n} a_i y(k-i) + w(k)$$

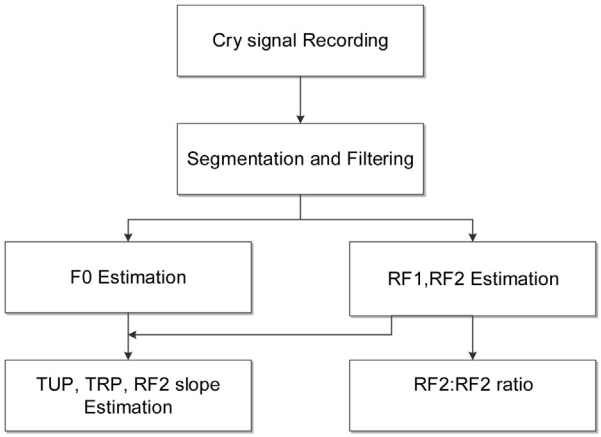

FIGURE 1. Methodology. ("signal"; "filtering," "RF1, RF2," "slope," and "ratio").

where n is the number of samples in a frame; $k = [1 \ldots n]$; $y(k)$ is the cry signal; $w(k)$ is the prediction error or (LP) residual, which contains the excitation contribution to the cry signal; a_i are the $A(z)$ polynomial coefficients; and n is the model order.

However, the formant structures of the model are defined by the zero of the filter $A(z)$, which corresponds to the power spectral density peaks of the cry signal.

The filter equation is as follows:

$$A(z) = 1 + \sum_{i=1}^{n} a_i z^{-i}$$

The power spectral density is given by:

$$\mathrm{PSD}(f) = \frac{1}{\left|1 + \sum_{i=1}^{n} a_i e^{-2jpfiT}\right|^2}$$

The RFs are therefore obtained by peak picking in the AR-PSD, which is estimated using the modified covariance method. This method reduces spectral line splitting and the bias of frequency estimation.[26] It is based on simultaneously minimizing the forward and backward prediction errors w(k) in the least squares sense.

The model order n is chosen by means of the following formula: n = ~0.5 Fs (kHz), for Fs = 44.1 kHz, n = 22. This relation comes from the physical constraint n = 2 LFs/c, where L is the length of the vocal tract = 8.5 cm, and c = 34 cm/ms, which is the speed of sound.[18] Formant splitting may occur if n is overestimated, whereas an underestimation of n smoothes the spectrum and causes the misallocation of spectral peaks.[18]

Fundamental frequency estimation

In this article, the modified SIFT algorithm is used to determine F0. It was demonstrated that this algorithm includes the autocorrelation properties and the cepstral pitch analysis technique.[17] In addition, the performance of this algorithm has been tested on the database of real newborn cries.[18,26]

This method is based on determining the autocorrelation sequence after performing glottal inverse filtering to attenuate the influence of the vocal tract.

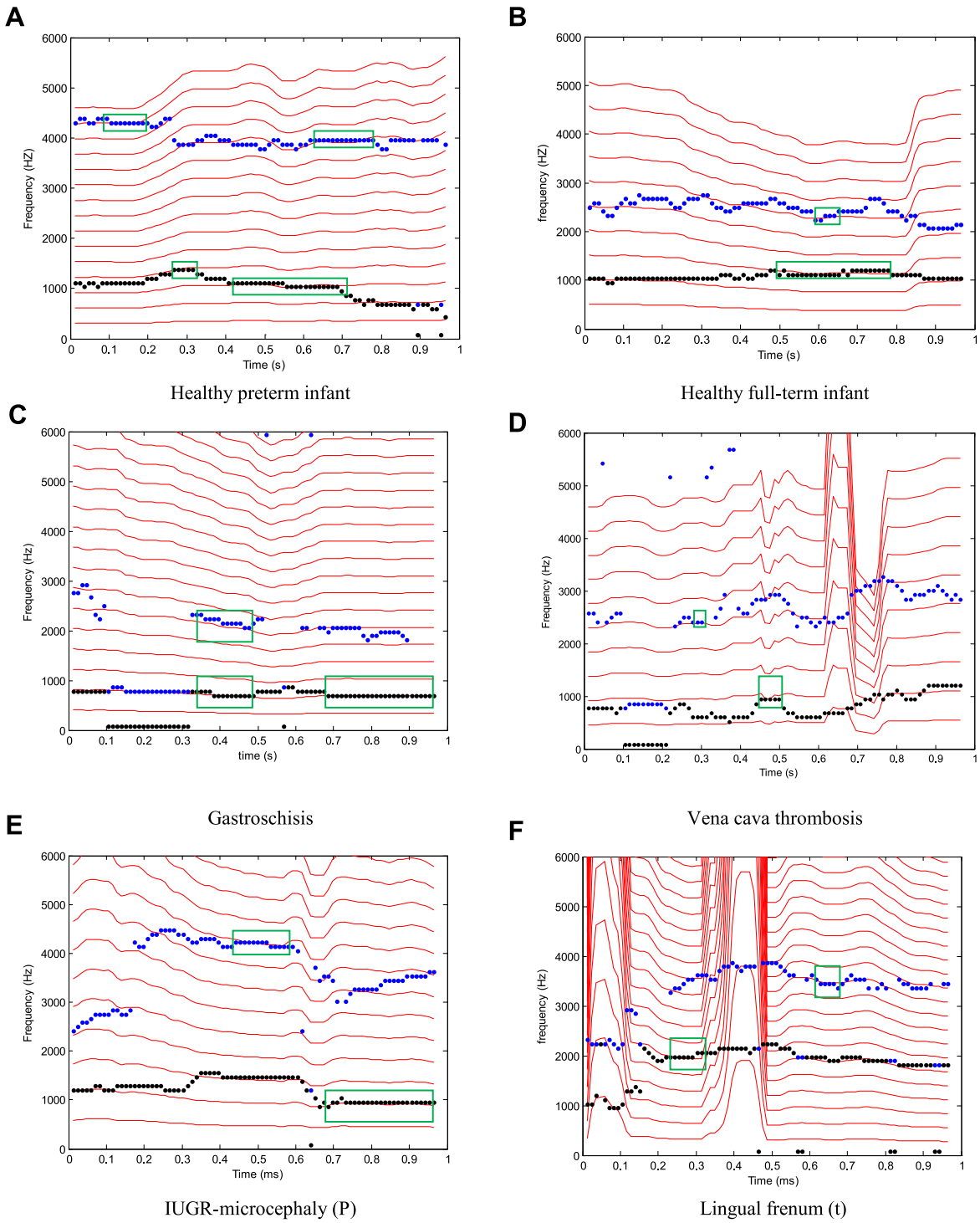

FIGURE 2. A, Healthy preterm infant (P); B, Healthy full-term infant (t); C, Gastroschisis (P); D, Vena cava thrombosis (t); E, IUGR-microcephaly (P); F, Lingual frenum (t). Fundamental frequencies and their harmonics (red lines), resonance frequencies: RF1 (black dots) and RF2 (blue dots), tuning between the RFs, and F0 harmonics (green rectangles).

The coefficients $\{a_i\}$ $i = [1,\ldots,4]$ of the inverse filter $A(z)$ are determined by a forward linear predictor which minimizes the prediction error in the least squares sense.[25]

Knowing $\{a_i\}$, the inverse filter output $\{s_k\}$ can be calculated using the equation:

$$s_k = y_k + \sum_{i=1}^{4} a_k y_{k-i}$$

Then, the output autocorrelation $\{r_i\}$ sequence is calculated as the autocorrelation sequence of $\{s_k\}$.

$$r_j = \sum_{k=0}^{N-1-j} s_k s_{k+j}, j = [1, \ldots, N-1]$$

Then, peak picking is performed and a decision algorithm in which the peak value is compared with the voice threshold is applied.[14,25] F0 is estimated using the following equation:

$$T0 = 1/F0 = \arg\max\{r_j\}$$

The contour of the F0 result is then smoothed by the median filter.

RESULTS AND ANALYSIS

Plots of typical examples of the estimated F0 and its harmonics with RF1, RF2 superimposed are presented on a color scale in Figure 2. This figure shows that early harmonics-resonance interaction exists in both the healthy and pathologic cries of newborns. Figure 2A,B shows that healthy cries contain relatively long periods of adjusted RF1 tuning with a fast transition duration time. In Figure 2C, RF1 is low with a long TUP. Figure 2D shows low RF1, RF2 with fast tuning and long transition duration for RF1 and unstable transition duration for RF2. Figure 2E shows high RF1, RF2 with long transition duration and relatively fast tuning for RF2. In Figure 2F, F0 is more irregular and the RFs are moving to the nearest harmonic to form a tuning event. Sudden pitch changes are noted in Figure 2D,F), which cause changes in the harmonic and RFs structures.

The various results of our investigation into the distribution of the acoustic characteristics of healthy and pathologic cries that we studied are presented by pathology and by gestational age using a box-and-whiskers plot (25th, 50th, and 75th percentile) in Figures 3 and 4. The recordings used for this assessment are those of premature newborn cries (*p*) and those of full-term newborn cries (*t*), and include healthy infants and infants suffering from the pathologies listed in Table 2.

F0, RF1, and RF2 estimation results

Table 3 shows the median and interquartile range of the average F0, RF1, and RF2 for all the pathologies studied. According to these results and those shown in Figure 3A, the median and interquartile range of the average F0 in the case of asphyxia (*t*), vena cava thrombosis (*t*), and intrauterine growth retardation (IUGR)-microcephaly (*P*) are higher than those of healthy full-term newborns. The median and interquartile range of the average F0 for healthy preterm newborns is lower than that of healthy full-term newborns. This difference may be attributable to the difference in age at recording. The smallest interquartile range of the average F0 is found in the tetralogy of Fallot (*P*) and the largest interquartile range of the average F0 is found in lingual frenum (*t*) after asphyxia (*t*).

As for the estimated RF1 average, its highest median and interquartile range are found in Lingual frenum (*t*), as shown in Figure 3B and Table 3. In cases of IUGR-asphyxia (*P*) and the tetralogy of Fallot (*P*), the estimated median is high compared with those of healthy full- and pre-term newborns. But the interquartile range of the average RF1 in the case of the tetralogy of Fallot (*P*) is the smallest. The lowest medians for the estimated RF1 average are found in meningitis (*t*) and peritonitis (*t*).

The estimation results in Table 3 and Figure 3C on the average RF2 indicate that the median of the average RF2 for IUGR-asphyxia (*P*) is higher and the interquartile range larger than those of the other pathologies studied or of healthy newborns. We also found that the median is high and the interquartile range large for lingual frenum (*t*), and that the median of the average RF1 and RF2 in healthy full-term newborns is lower and the interquartile range smaller than those of healthy preterm newborns. The lower medians are evident in both vena cava thrombosis (*t*) and peritonitis (*t*).

RF1:RF2 ratio and RF2 slope evaluation results

The medians and interquartile ranges of the A_{ratio} are shown in Table 3 and Figure 3D. The results indicate a median value that is relatively the same for both healthy full- and pre-term infants and also for newborns suffering from IUGR-microcephaly (*P*) and respiratory distress syndrome (RDS) (*P*), but with different interquartile ranges. The highest median, A_{ratio}, is associated with the cries of newborns with lingual frenum (*t*), and the lowest with those with vena cava thrombosis (*t*). As for asphyxia (*t*), IUGR-asphyxia (*P*), the tetralogy of Fallot (*P*), and gastroschisis (*P*), the A_{ratio} cry signal medians are higher than those in healthy cry signals, and for meningitis (*t*) and peritonitis (*t*), they are lower. The A_{SLP} median and interquartile range results for both the healthy and pathologic newborn cries are given in Table 3 and illustrated in Figure 3E. These results indicate that the median of A_{SLP} in healthy full-term newborns cries is slightly higher than in those of healthy preterm newborns. The newborn cries of infants with IUGR-microcephaly (*P*) or vena cava thrombosis (*t*) are characterized by a high A_{SLP} median and a large interquartile range, unlike the newborn cries of newborns with lingual frenum (*t*), which are characterized by a smaller A_{SLP} median compared with either healthy or pathologic newborn cries. Also, not surprisingly, the A_{SLP} median is lower than those of healthy newborns for asphyxia (*t*), meningitis (*t*), RDS (*P*), IUGR-asphyxia (*P*), the tetralogy of Fallot (*P*), and gastroschisis (*P*).

TUP and TRP evaluation results for RF1 and RF2

The various TUP and TRP measurements of the RFs are shown in Table 4 and Figure 4. The results shown in Figure 4A–C indicate that the median and interquartile range of P_{TUP} between RF1 and the harmonics of F0 are relatively the same in healthy full- and pre-term newborns, and A_{TRP} is the same in both cases. But the A_{TUP} median is the highest in the healthy preterm cries studied. The P_{TUP} median between RF1 and the F0 harmonics is the highest in lingual frenum (*t*), meningitis (*t*), gastroschisis (*P*), the tetralogy of Fallot (*P*), and IUGR-microcephaly (*P*). The lowest P_{TUP} medians are found in IUGR-asphyxia (*P*) and in vena cava thrombosis (*t*), but with a large interquartile range. Higher A_{TUP} medians for RF1 in lingual frenum (*t*), meningitis (*t*), peritonitis (*t*), gastroschisis (*P*), and RDS (*P*) are

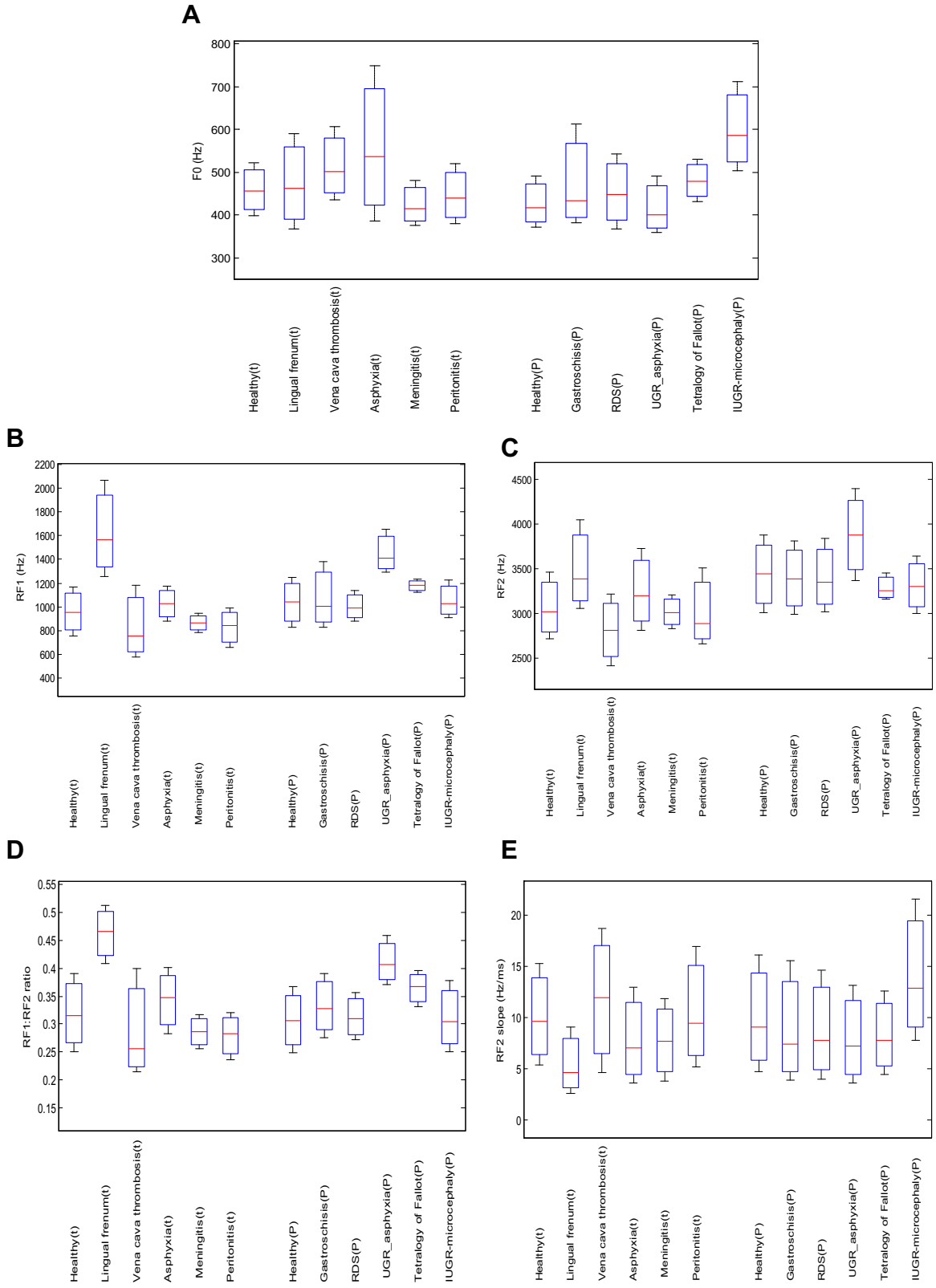

FIGURE 3. Box-and-whiskers plots for the averages of (**A**) F0, (**B**) RF1, (**C**) RF2, (**D**) the RF1:RF2 ratio, and (**E**) the RF2 slope by pathology and gestational age.

noted compared with those of the other pathologies studied and those of healthy newborns. The lowest A_{TUP} median and the highest A_{TRP} median for RF1 with a large interquartile range are noted in vena cava thrombosis (t). The lowest A_{TRP} medians for RF1 are found in lingual frenum (t), meningitis (t), IUGR-asphyxia (P), and the tetralogy of Fallot (P).

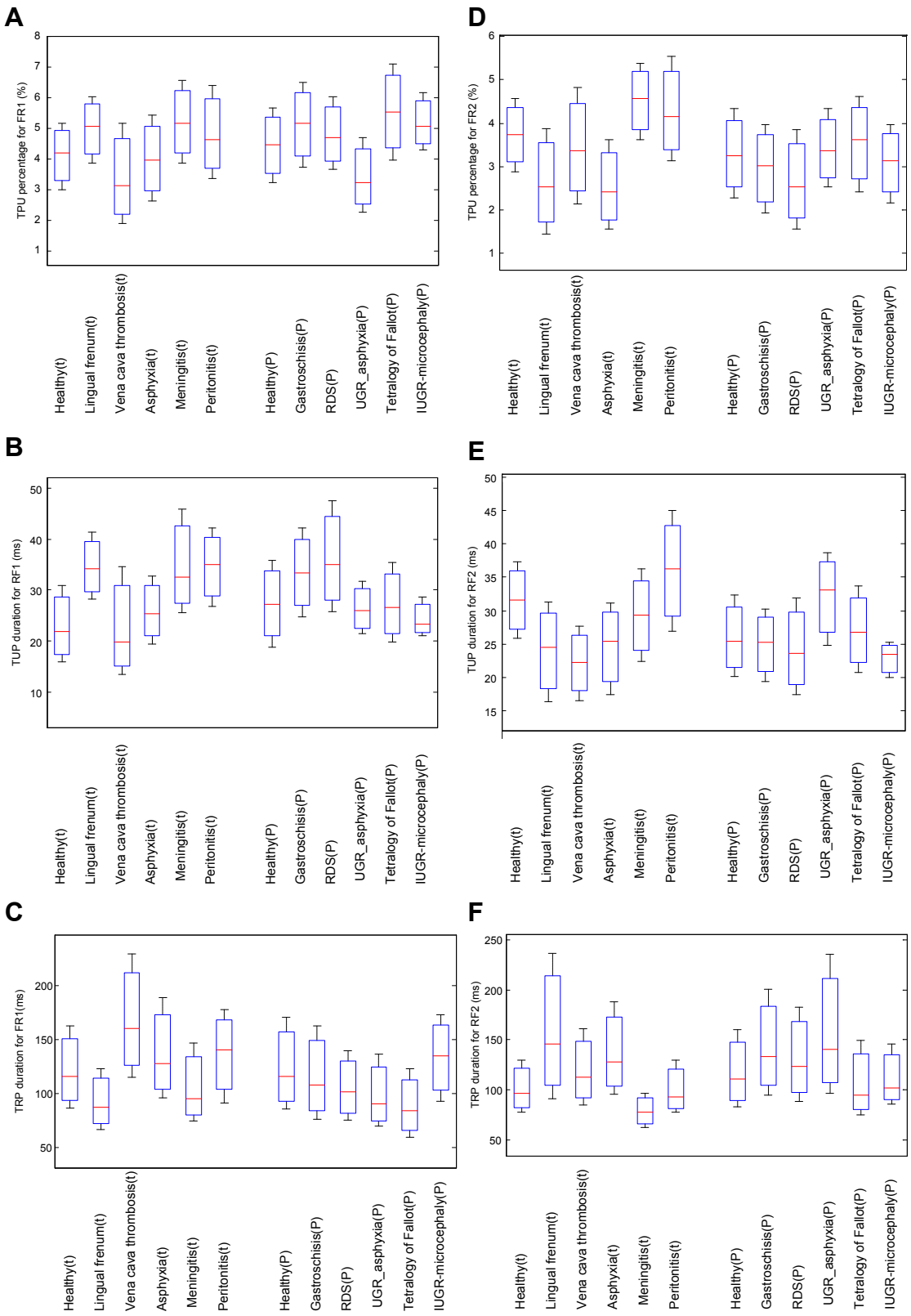

FIGURE 4. Box-and-whiskers plots for the average of (**A**) P_{TUP} for RF1, (**B**) A_{TRP} for RF1, (**C**) A_{TRP} for RF1, (**D**) P_{TUP} for RF2, (**E**) A_{TRP} for RF2, and (**F**) A_{TRP} for RF2, by pathology and gestational age.

As for P_{TUP}, A_{TUP}, and A_{TRP} between RF2 and the F0 harmonics, Table 4 and Figure 4D–F indicate that the estimated median of P_{TUP} and A_{TUP} in healthy full-term newborns are higher than those of preterm newborns. Unlike P_{TUP} and A_{TUP}, the A_{TRP} median is the lowest in healthy full-term cries. We note the same lower P_{TUP} median in lingual frenum (*t*), vena cava thrombosis

TABLE 3.
Measured Characteristics of Cries

	A_{F0} (Hz)		A_{RF1} (Hz)		A_{RF2} (Hz)		A_{ratio}		A_{SLP} (Hz/ms)	
Pathology	Median	Interquartile Range	Median	Interquartile Range	Median	Interquartile Range	Median	Interquartile Range	Median	Interquartile Range
Healthy (*t*)	456	94	955	311	3024	559	0.31	0.11	9.5	7.43
Lingual frenum (*t*)	462	168	1569	609	3393	738	0.46	0.08	4.63	4.82
Vena cava thrombosis (*t*)	501	128	754	456	2817	601	0.25	0.14	11.88	10.58
Asphyxia (*t*)	536	272	997	255	3002	635	0.35	0.1	6.98	7.06
Meningitis (*t*)	415	79	867	118	3009	281	0.28	0.05	7.67	6.06
Peritonitis (*t*)	440	105	846	248	2885	632	0.28	0.06	9.41	8.8
Healthy (*P*)	417	89	1045	314	3447	641	0.3	0.09	9.05	8.52
Gastroschisis (*P*)	433	174	1009	414	3388	620	0.33	0.08	7.35	8.79
RDS (*P*)	449	131	991	192	3347	616	0.31	0.06	7.7	8.01
IUGR-asphyxia (*P*)	400	98	1411	269	3878	773	0.4	0.06	7.14	7.15
Tetralogy of Fallot (*P*)	479	74	1180	79	3256	222	0.37	0.05	7.37	6.07
IUGR-microcephaly (*P*)	587	156	1028	241	3300	482	0.3	0.1	12.85	10.32

TABLE 4.
Measured Characteristics of Cries

	P_{TUP} for RF1 (%)		A_{TUP} for RF1 (ms)		A_{TRP} for RF1 (ms)		P_{TUP} for RF2 (%)		A_{TUP} for RF2 (ms)		A_{TRP} for RF2 (ms)	
Pathology	Median	Interquartile Range	Median	Interquartile Range	Median	Interquartile Range	Median	Interquartile Range	Median	Interquartile Range	Median	Interquartile Range
Healthy (*t*)	4.2	1.6	22	11.2	116.1	57	3.7	1.3	31.6	8.6	96.5	39.1
Lingual frenum (*t*)	5	1.6	34.2	9.8	87.5	42.3	2.5	1.8	24.4	11.3	146.3	109
Vena cava thrombosis (*t*)	3.1	2.5	19.8	15.9	160	85.1	3.4	2	22.3	9.3	112.2	56.9
Asphyxia (*t*)	4	2	25.4	9.9	127.7	69.5	2.5	1.5	25.4	10.3	127.7	69.4
Meningitis (*t*)	5.2	2	32.6	15.3	95.5	53.8	4.6	1.3	29.3	10.4	78	25
Peritonitis (*t*)	4.6	2.3	35	11.7	140.5	64.3	4.2	1.8	36.3	13.6	92.8	39.5
Healthy (*P*)	4.5	1.8	27.2	12.7	116.1	64	3.3	1.5	25.4	9	110.7	58.7
Gastroschisis (*P*)	5.2	2.1	33.3	13	108	64.7	3	1.5	25.2	8.2	133.8	79
RDS (*P*)	4.7	2.3	35.1	16.4	101.5	47.9	2.5	1.7	23.6	10.8	123.2	70.6
IUGR-asphyxia (*P*)	3.3	1.8	25.9	7.6	90.5	50.1	3.4	1.4	33	10.5	140.4	104.3
Tetralogy of Fallot (*P*)	5.5	2.3	26.5	11.7	84	47.3	3.6	1.7	26.7	9.7	94.6	55.9
IUGR-microcephaly (*P*)	5	1.4	23.2	5.64	135	60	3.1	1.4	23.5	4	101.9	44.9

TABLE 5.
Cry Characteristics Associated With the Pathologies Studied

Pathologies	Cry Characteristics
Lingual frenum (*t*)	↑RF1, ↑RF1 tuning, ↓RF1 transition, ↑RF2, ↓RF2 tuning, ↑ RF2 transition, ↑ RF1:RF2 ratio, ↓slope RF2
Vena cava thrombosis (*t*)	↑F0, ↓RF1, ↓RF1 tuning, ↑RF1 transition, ↓RF2, ↓RF2 tuning, ↑RF2 transition, ↓RF1:RF2 ratio, ↑slope RF2
Asphyxia (*t*)	↑F0, ↑RF1 transition, ↑RF2 transition, ↑RF1:RF2 ratio, ↓slope RF2
Meningitis (*t*)	↓RF1, ↑RF1 tuning, ↓RF1 transition, ↑RF2 tuning, ↓RF2 transition, ↓RF1:RF2 ratio, ↓slope RF2
Peritonitis (*t*)	↓RF1, ↓RF2, ↑RF1 tuning, ↑RF1 transition, ↑RF2 tuning, ↓RF2 transition, ↓RF1:RF2 ratio
Healthy (*P*)	↑RF1, ↑RF2, ↓RF2 tuning, ↑RF2 transition
Gastroschisis (*P*)	↓RF1, ↑RF1 tuning, ↓RF1:RF2 ratio,↑RF2 transition, ↓slope RF2
RDS (*P*)	↓RF1, ↑RF2, ↑RF1 tuning, ↓RF2 tuning, ↑RF2 transition, ↓slope RF2
IUGR-asphyxia (*P*)	↑RF1, ↓RF1 tuning, ↓RF1 transition, ↑RF2, ↑RF2 tuning, ↑RF2 transition, ↑RF1:RF2 ratio, ↓slope RF2
Tetralogy of Fallot (*P*)	↑RF1 tuning, ↓RF1 transition, ↑RF2 tuning, ↓RF2 transition, ↑ RF1:RF2 ratio, ↓slope RF2
IUGR-microcephaly (*P*)	↑F0, ↑RF2, ↑RF1 tuning, ↑RF1 transition, ↓RF2 tuning, ↑ slope RF2

(*t*), and RDS (*P*). In meningitis (*t*) cries, we found a higher P_{TUP} median and a lower A_{TRP}. The highest A_{TUP} median is found in peritonitis (*t*) cries, and the highest median and the largest A_{TRP} interquartile range are found in the cries of newborns suffering from lingual frenum (*t*) and IUGR-asphyxia (*P*).

DISCUSSION

An average value is greatly influenced by extreme values, which is not the case for a median. For this reason, we chose to represent our evaluation results in the form of medians and interquartile ranges in this work. This method of display is useful because it makes possible both global and local dispersion of the data results.

In this study, we first focus our attention on the RF1, and RF2, the RF1:RF2 ratio, and the RF2 slope, which notify us of changes in the configuration of the vocal tract. To understand the behavior of F0 and RF1 and RF2 in healthy and pathologic newborn cries, we concentrated our investigation on the percentage and duration of tuning (P_{TUP} and A_{TUP}), as well as on the transition duration (A_{TRP}) for both RF1 and RF2. These two acoustic phenomena can tell us whether or not a true relationship exists between F0 and the RFs at an early age in the newborn and also whether or not the various medical conditions we studied influence the extent of these acoustic phenomena.

The association we have established between the characteristics of newborn cries and the pathologies studied is provided in Table 5. This association is based on our comparison of the cries of pathologic and healthy newborns.

According to the study by Seifert et al[16] in a normal group of children aged 3.8 to 10.2 years, A_{ratio} varies between 0.5 and 0.63. In our study, we found a significant difference in this parameter between healthy newborns aged 1 day to 1 month and pathologic newborns of the same age. This is presumably

TABLE 6.
Analysis of Variance Based on Gestational Age

	Healthy		Pathologic	
Characteristic	*F**	*P*†	*F**	*P*†
A_{F0}	218.064	0.000	0.052	0.824
A_{RF1}	311.061	0.000	0.683	0.428
A_{RF2}	63.762	0.000	11.304	0.007
A_{ratio}	225.770	0.000	0.143	0.713
A_{SLP}	43.570	0.000	0.036	0.853
$P_{TUP\ for}$ RF1	58.332	0.000	0.623	0.448
$A_{TUP\ for}$ RF1	25.119	0.000	0.012	0.913
$A_{TRP\ for}$ RF1	44.890	0.000	1.304	0.280
$P_{TUP\ for}$ RF2	322.175	0.000	0.743	0.409
$A_{TUP\ for}$ RF2	536.565	0.000	1.304	0.280
$A_{TRP\ for}$ RF2	0.780	0.377	0.458	0.514

* One way ANOVA F-test statistic.
† The P-value is the probability of being greater than the F-statistic.

because successful word production requires greater and more rapid changes in vocal tract geometry than cries. In spite of the increase in RF1 and RF2 in healthy preterm newborn cries compared with healthy full-term newborn cries, we found no significant difference in either case for the estimated $A_{\text{ratio.}}$ In the case of pathologic cries, the values were higher than they were for healthy newborns for some pathologies and lower for others. However, lower values of A_{ratio} are attributable to the decrease in RF1 values, and higher values of A_{ratio} are the result of increased RF1 values. A_{ratio} is related to the position of the tongue in the oral cavity. RF1 is associated with tongue height, and RF2 refers to tongue advancement.[10] These relationships are confirmed by the results of our study. Among the congenital malformations, lingual frenum (*t*) cries are characterized by the highest RF1 and A_{ratio} of all the healthy and pathologic cries studied. RF2 was also the highest among full-term newborn cries, and its A_{SLP} was the lowest among the newborn cries studied. We can infer from this that a reduced A_{SLP} can be explained by the relative slowness of tongue movement, which is induced by the shortness of the transition and the long RF2 transition duration (A_{TRP}) (Table 4). However, in this case, we noted a decrease in A_{TRP}, along with an increase in (P_{TUP}) and (A_{TUP}) for RF1, compared with the values of these parameters in healthy full-term newborns.

In another congenital malformation, gastroschisis (*P*), the average RF2 slopes (A_{SLP}) are lower for newborn cries, which contain a high percentage of RF tuning (P_{TUP}) and a long duration (A_{TUP}) of RF1 tuning. In the case of the neurologic disease IUGR-microcephaly (*P*), the highest medians of A_{F0} and A_{SLP} were observed and also a high P_{TUP} median and a low A_{TUP} median for RF1. A long transition duration (A_{TRP}) was also noted for RF1. This behavior is possibly the result of very rapid changes in the configuration of the vocal tract.[9] In the other neurologic diseases, in IUGR-asphyxia (*P*), for example, high A_{RF1}, A_{RF2}, and A_{ratio} values and a slow A_{SLP} and long A_{TRP} were observed. Asphyxia cries are characterized by a high A_{F0} and A_{ratio}, a low A_{SLP} and high A_{TRP} for RF2, and a high A_{TRP} for RF1. In the heart defect diseases, the cries of newborns with vena cava thrombosis (*t*) are characterized by a high A_{F0} and the lowest A_{RF1} and A_{RF2}, and an A_{ratio} with a high (A_{SLP}) and a high (A_{TRP}) for RF1 and RF2. Tetralogy of Fallot (*P*) cries contain the highest percentage of RF1 tuning (AP_{TUP}) and a shorter RF1 and RF2 transition duration (A_{TRP}). In this case, A_{ratio} is high with a low A_{SLP}.

However, in the infectious diseases category, meningitis (*t*) and peritonitis (*t*), A_{RF1} and A_{ratio} are lower. In meningitis (*t*) cries, the A_{SLP} and A_{TRP} for both RF2 and RF1 are lower with a high P_{TUP} and A_{TUP} for RF1. Unlike meningitis (*t*), peritonitis (*t*) cries are characterized by a low A_{RF2} and a high A_{TRP} for RF1. In the respiratory disease category, A_{SLP} is lower in RDS cries, and P_{TUP} for RF1 and A_{TRP} for RF2 are high.

We can surmise from the results of this study that there are three tuning patterns between the RFs and F0 in newborn cries. The first pattern is made up of a high percentage of long TRP times with short transition duration times, as in the case of lingual frenum (*t*) and meningitis (*t*). In the second pattern, we find a high percentage of short TRP times with long transition duration times, as in IUGR-microcephaly. In the third pattern, the tuning percentage is low, but with a long duration time and long transition duration time, as for RF2 in healthy full-term newborns.

We note from our results that a decrease in RF1 and RF2 becomes more apparent in healthy newborns with increasing gestational age, which seems more relevant than the age at recording in this case.

To further investigate the influence of gestational age (full- and pre-term) on the characteristics measured, we performed an analysis of variance for healthy and pathologic cries for each characteristic, to determine whether or not they change with gestational age. Our results, shown in Table 6, indicate significant statistical differences in all the assessed characteristics between healthy premature cries and healthy full-term cries. Only for the average transition duration of RF2 (A_{TRP}) was no significant difference found based on gestational age. In the case of pathologic cries, significant statistical differences ($F = 11.304$, $P < 0.05$) between the average RF2 (A_{RF2}) of premature pathologic cries and full-term pathologic cries was noted. For the remaining characteristics, no significant statistical differences were found based on gestational age. We have determined, therefore, that all the characteristics assessed, except for the transition duration of RF2, depend on gestational age in healthy cries. In pathologic cries, with the exception of the average RF2, the assessed characteristics do not depend on gestational age, but rather on the pathology itself.

CONCLUSIONS

In this article, a new automated method for newborn cry analysis is presented. This measurement method allows us to associate the most relevant characteristics with pathologies of interest and also to extend the method to other pathologies.

The goal of this study has been to examine the possibility of classifying pathologic and healthy cries on the basis of their acoustic attributes. We focused our investigation on the distribution of RFs and F0 in the cries of newborns aged 1 day to 1 month. To identify the early relationship between these two frequencies, the tuning between the RFs and F0 and the transition duration were evaluated. Variation patterns of the RF1:RF2 ratio and the RF2 slope were also examined, based on the healthy and pathologic cries studied.

The calculation of percentile values for each assessed characteristic provides details about its evolution in healthy and pathologic cries. The distribution of these acoustic variables could be used to develop a quantitative metric to differentiate pathologic cries from healthy ones. We conclude from our results that, although laryngeal phonation and vocal tract–based articulation are independently controlled systems in newborns, these two systems interact according to the status of the newborn. Also, we infer that the characteristics of cries in the newborns we studied who were suffering from pathologies do not vary with gestational age, but rather mostly according to the pathology itself, which alters vocal folds and vocal tract resonance.

The results of this study suggest that the acoustic measures that we studied (A_{F0}, A_{RF1}, P_{TUP}, A_{TUP}, A_{TRP}, A_{SLP}, and A_{ratio}) show significant differences among diseases. The automated approach presented in this article for assessing these characteristics can be successfully used to distinguish pathologic cries from healthy cries.

Acknowledgments

This work is supported by the Bill & Melinda Gates Foundation through the Grand Challenges Explorations Initiative. Special thanks to Dr. Barrington and members of the Neonatology group at Saint-Justine Hospital in Montreal (QC) for helping us to collect samples for the Infant Cry Database.

REFERENCES

1. Michelson K, Todd de Barra H, Michelson O. Sound Spectrographic cry analysis and mothers perception of their infant's crying. In: Lewis FR, ed. *Focus on nonverbal communication research*. New York, NY: Nova Science Publishers; 2007:31–64.
2. Wasz-Hockert O, Michelsson K, Lind J. Twenty-five years of scandinavian cry research. In: Lester BM, Boukydis CFZ, eds. *Infant crying*. New York, NY: Plenum; 1985:83–104.
3. Michelson K, Eklund K, Leppanen P, Lyytinen H. Cry characteristics of 172 healthy 1- to 7-day-old infants. *Folia Phoniatr Logop*. 2002;54:190–200.
4. Lester BM, LaGasse LL. Crying. In: Haith MM, Benson JB, eds. *Encyclopedia of infant and early childhood development*. San Diego, CA: Academic Press; 2008:332–343.http://www.sciencedirect.com/science/article/pii/B9780123708779000451.
5. LaGasse LL, Neal AR, Lester BM. Assessment of the infant cry: acoustic cry analysis and parental perception. *Ment Retard Dev Disabil Res Rev*. 2005;11:83–93.
6. Soltis J. The signal functions of early infant crying. *Behav Brain Sci*. 2004; 27:443–490.
7. Buder EH, Chorna LB, Oller DK, Robinson RB. Vibratory regime classification of infant phonation. *J Voice*. 2008;22:553–564.
8. Cecchini M, Lai C, Langher V. Dysphonic newborn cries allow prediction of their perceived meaning. *Infant Behav Dev*. 2010;33:314–320.
9. Kim Y, Weismer G, Kent RD, Duffy JR. Statistical models of f2 slope in relation to dysarthria severity. *Folia Phoniatr Logop*. 2009;61:329–335.
10. Kim Y, Kent RD, Weismer G. An acoustic study of the relationships among dysarthria-type neurological diseases and dysarthria severity. *J Speech Lang Hear Res*. 2011;54:417–429.
11. Wermk K, Mende W, Kempf A, Manfredi C, Bruscaglioni P, Stellzig-Eisenhauer A. Interaction patterns between melodies and resonance frequencies in infants' pre-speech utterances. In Proceedings of the 4th International Workshop. Models and Analysis of Vocal Emissions for Biomedical Applications, 2005; 187–190.
12. Kheddache Y, Tadj C. Frequential characterization of healthy and pathologic newborn cries. *Am J Biomed Eng*. 2013;3:182–193.
13. Kheddache Y, Tadj C. Characterization of pathological cries of newborns based on fundamental frequency estimation. *Engineering*. 2013;5:272–276.
14. Manfredi C, Mende W, Bruscaglioni P, Wermke K. Resonance development and formant tuning phenomena in infant's crying. In: Proc. 3rd Int. Workshop MAVEBA, 2003; 35–38.
15. Poissant SF, Petersa KA, Robb MP. Acoustic and perceptual appraisal of speech production in pediatric cochlear implant users. *Int J Pediatr Otorhinolaryngol*. 2006;70:1195–1203.
16. Seifert E, Oswald M, Bruns U, Vischer M, Komis M, Haeusler R. Change of voice and articulation in children with cochlear implants. *Int J Pediatr Otorhinolaryngol*. 2002;66:115–123.
17. Kasuya H, Ogawa S, Mashima K, Ebihara S. Normalised noise measure to evaluate the pathological voice. *J Acoust Soc Am*. 1986;80:1329–1334.
18. Manfredi C, Bocchi L, Orlandi S, Spaccaterra L, Donzelli GP. High-resolution cry analysis in preterm newborn infants. *Med Eng Phys*. 2009; 31:528–532.
19. Fort A, Manfridi C. Acoustic analysis of newborn infant cry signals. *Med Eng Phys*. 1998;20:432–442.
20. Kheddache Y, Tadj C. Acoustic measures of the cry characteristics of healthy newborns and newborns with pathologies. *J Biomed Sci Eng*. 2013;6:796–804.
21. Sundberg J, Filipa M, La B, Gill BP. Formant tuning strategies in professional male opera singers. *J Voice*. 2013;27:278–288.
22. Varallyay G, Illényi A, Benyo Z, Farkas Z, Katona G. An attempt to detect hearing disorders by acoustic features of the infant cry. In Proceedings of the Forum Acusticum Congress, Budapest, 2005, 1–6.
23. Verduzco-Mendoza A, Arch-Tirado E, Garcia CA, Ibarra JL. Qualitative and quantitative crying analysis of newborn babies delivered under high risk gestation. In: de Esposito A, Hussain A, Marinaro M, Martone R, eds. *Multimodal signals: cognitive and algorithmic issues*. Berlin Heidelberg: Springer; 2009:320–327. 5398.
24. Michael C, Robb P, Cacace AT. Estimation of formant frequencies in the infant cry. *Int J Pediatr Otorhinolaryngol*. 1995;32:57–67.
25. Markel JD. The SIFT algorithm for fundamental frequency estimation. *IEEE Trans Audio Electroacoustic*. 1972;20:367–377.
26. Manfredi C, Tocchioni V, Bocchi L. A robust tool for newborn infant cry analysis. Proceedings of the 28th IEEE, EMBS Annual International Conference, 2006, 509–512.