QUANTIFYING BABY CRYING RHYTHM ABNORMALITIES USING MULTILAYER PERCEPTRON

H. Dahmani[a, b], CH. Tadj[b]
[a] Mohamed Boudiaf University, M'sila, Algeria.
[b] ETS (Ecole de Technologie Supérieure), Quebec University, Montreal, QC, Canada.
anout_1999@yahoo.fr

ABSTRACT

There are no studies to quantify rhythm of baby crying signals until now. In the present study, we propose that the introduction of the temporal rhythm metrics as used in the languages identification domain may well characterize the rhythm of newborn infant's sounds. Then, it may improve diagnostic accuracy and helps quickly identify, categorize and discriminate between the sick and healthy babies, therefore, provide more reliability to the evaluation of the health of the newborns. The repeated bursts of expiratory sounds were detected and selected from the crying records of 295 full-term babies of 1 to 90 days old. These recordings occur under some stimulus as pain, fever, diaper change and other different states.

The relevance of the proposed rhythm features to distinguish the healthy from the sick babies is approved. The results of various classification experimentations of the correct classification were more than 80%. Indeed, the findings were very promising and confirm the ability of rhythm parameters to characterize the baby crying signal well.

KEYWORDS

Signal Processing, baby crying signal, classification, pathological baby crying signals, rhythm metrics, and neural networks

1. INTRODUCTION

"Normally; infantile sound-making begins when the neonate has to abandon its aquatic adaptation inside the mother and adjust to the air environment outside. There are 'brief inspirations and effortful prolonged expirations signaled by cries, struggles, and coughs". It is a definition of many other expressive definitions describing the first cry sound of newborn infants by Bosma et al. in [1].

Crying is the beginning of human vocal communication. So, we have different crying to express different messages: hunger, boredom, situations of discomfort, colic, discharge at the end of the day, pain, etc. It is crucial that we focus on these messages.

So far, a considerable number of studies have been carried out to investigate the characteristics of the cry for many purposes. The most important of these goals is perhaps the cry category identification and the discrimination between normal and abnormal newborn infants to investigate whether these analyses can be used for medical diagnosis.

The majority of these investigations have focused on the fundamental frequency and harmonics of the cry as acoustic measures [2-22]. While, the non-segmental or supra-segmental or merely prosodic characteristics as intonation, stress, melody, and rhythm, are not smooth to investigate, especially when it comes to cries of newborn infants. Peter Ostwald in [17] reported that the acoustic methods were introduced by Fairbanks (1942) and Lynip (1951). Since then there have been some outstanding contributions to the acoustical structure of infant sounds, especially cry-vocalizations (Truby 1960, Wasz-Hocker*terd.*1963, Ringel and Kluppel 1964, Sedl&kovh1964, Lind *et al.* 1966, Lieberman *et a/.*1971). In the same work, Peter Ostwald reported from Crystal article dated 1970, that Prosodic features-intonation, rhythm and voice-qualities-are surely among the essential acoustical features during the first three months of life. In reality, this study entitled "Non-segmental phonology in first-language acquisition" was published in [23]. Lester and Zeskind et al., too, confirmed that despite the common assumption that the infant crying is rhythmic there is an insufficiency of works which studied rhythm features in infant cry temporal signals. Indeed, Lester wrote in [24]: "It has been suggested that there are rhythmical patterns in the cry, but this has not yet been quantitatively demonstrated. To our knowledge, however, few such studies have been published. Indeed, even attempts to use spectral analysis to examine the rhythmicity in crying and cardiac respiratory activity". While Zeskind et al. reported in [21] that, "Wolff (1967) long ago described the temporal morphology of crying as a rhythmic repetition of an expiratory sound, a brief pause, an inspiratory sound, and a second pause before the beginning of the next cry expiration. Although no test of rhythmicity was conducted, variations in the durations of these components were suggested to reflect different conditions of infant arousal. The primary cry associated with hunger, for example, was depicted as more stable in duration and more rhythmical in the organization than a cry in response to painful stimulation".

There are many works which have been interested in the physiological and temporal aspect of the cries. Many scholars investigated the respiratory patterns associated with the neonate or infant cry and so the time factors of the cry signal [1, 18-21, 24-32]. For example, we cite some findings to demonstrate the general orientation of these scholars. They observed the respiratory patterns activity, and they confirmed that the expiratory phase is longer than the inspiratory phase. They described rhythmic characteristics of particular cry under some stimulus as pain as this cry has a shorter expiration durations. Another statement about the duration of expiratory segments increases during the first eight months as confirmed by Prescott [33]. Zeskind et al. in [22] investigated the effects of the temporal structure of infant crying on mothers' perceptions.

In fact, all these studies established that the patterns of crying activity reveal the temporal rhythmic structure of the infant cry and its changes with the stimulus, time, healthy state and age.

We can expect that being able to quantify rhythmic differences objectively and reliably would facilitate to study the rhythmicity in crying and even facilitate the diagnosis of certain diseases in newborn infants in early stages.

Therefore, the present study aims to investigate the temporal morphology and rhythmic structure underlying the repeated expiratory and inspiratory sounds in a sustained bout of crying of newborn babies through the use of rhythm metrics used primarily in the identification of languages domain.

Indeed, some duration-based metrics designed to capture differences in speech rhythm between and within languages have been developed (Ramus and Grabe parameters) [34-39]. These rhythmic metrics were derived from an acoustically based segmentation of the speech signal into the vocalic and consonantal interval.

1.1. Rhythm Parameters

Ramus et al. [39] based their quantitative approach to speech rhythm on the purely phonetic characteristics of the speech signal. They measured vowel durations and the duration of intervals between vowels. They computed %V, the sum of vocalic interval duration divided by the total duration of the sentence (and multiplied by 100), and ΔV, the standard deviation of vocalic interval duration, and a third parameter ΔC, the standard deviation of the consonantal interval duration. They argued that a combination of %V and either ΔC or ΔV provided the best acoustic correlate of rhythm classes. Ramus in [40] suggested that a normalization procedure might be necessary for both ΔV and ΔC, dividing them by the mean. Grabe and Low [36]do not relate speech rhythm to phonological units such as inter-stress intervals or syllable durations. Instead, they calculate durational variability in successive acoustic-phonetic intervals using Pairwise Variability Indices (PVI). The raw Pairwise Variability Index (rPVI) is given by equation (1):

$$rPVI = \sum_{k=1}^{N-1} \left| d_k - d_{k+1} \right| / (N-1) \qquad (1)$$

Where d_k is the length of the k^{th} vocalic or intervocalic segment and N is the number of segments.

A normalized version of the PVI index (noted nPVI) is defined by:

$$nPVI = \sum_{k=1}^{N-1} \left| \frac{d_k - d_{k+1}}{(d_k + d_{k+1})/2} \right| / (N-1) \qquad (2)$$

Dellow [34] used a new rhythm metric: VarcoC, the standard deviation of consonantal interval duration divided by the mean consonantal duration (multiplied by 100). He found VarcoC more robust than ΔC for discrimination of stress-timed English and German from syllable-timed French. White and Mattys [41] showed that both VarcoV (for vowels) and VarcoC (for consonants) were robust to changes in speech rate and discriminative in language identification. Liss et al. [42] proposed a new normalized rhythm metric based on the duration of successive combined vocalic and consonantal intervals, as an approximation to syllable duration: VarcoVC, the normalized standard deviation, nPVI-VC and rPVI-VC, the normalized and raw pairwise variability indices. Lastly, Dellow and Wagner [43] assume that the distinction between rhythms can be made by using voiced and voiceless segments instead of the vocalic and consonantal segments.
Liss et al. [42]also confirmed the ability of the rhythm metrics to distinguish normal control speech from dysarthria and to discriminate its subtypes. They reported that the most predictive variables were VarcoC, rPVI-VC, and nPVI-VC with an overall rate of correct discrimination of 85%.

Therefore, the purpose of this work is to assess temporal characteristics of baby's signals cry by using these rhythm metrics to distinguish between the healthy and sick babies.

The rest of the paper is organized as follows: in section 2, we describe the subjects, stimuli, and the segmentation procedure with the calculation of the rhythm metrics. The third part, the results of classification and statistical analyses are explored with the discussion. The fourth section is devoted to the conclusion and future perspectives.

2. METHOD AND MATERIAL

2.1. Subjects

The subjects consisted of two groups of newborn infants. Group one is composed of 151 healthy full-term newborn infants (32boys & 119 girls). The group consists of 144 sick full-term newborn infants (97boys & 47 girls). The recordings selected were made in the neonatology departments of Al-Sahel Hospital (Lebanon). The average gestational age of the infants was 39 weeks (range = 38-40 weeks). All of the infants displayed an Apgar score of 8 or more at 5 minutes *(*Apgar >= 8). The age of each baby at the time of audio recording was from to 1 to 90 days postpartum. All the babies are born to Arabic families.

2.2. Cry Stimuli

Cry samples were selected from the newborn cry-based diagnostic system database. A description of the data collection technique was presented in the former works which already investigated these data [44-50]. The cries were recorded with 44.1 kHz sampling rate and a sample resolution of 16 bits. In general, the average duration of the recorded signals is 90 seconds.

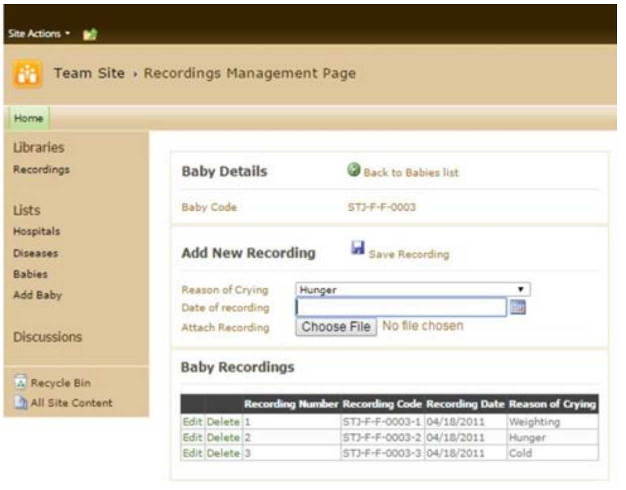


Figure 1: Then interface of the database on the ETS site (ETS: Ecole de Technologies Supérieures)

2.3. Calculation of the Rhythm Metrics

The baby crying signal is different from the speech signal. The first evident question is: if the vowel and the consonant were the based units for the rhythm speech metrics, how about the basic units for the rhythm baby cry metrics?

As a matter of fact, the cry is a term associated with the total duration of the acoustic signal, including all inspirations, expirations, and pauses, from the beginning of the emission until the infant keeps quiet. It is composed of a sequence of cry units, as illustrated in Figure2.

A cry unit or episode of cry shall consist of some utterances. As reported by Prescott [18]: "A cry utterance was defined as one complete respiratory cycle upon which cry sound was overlaid. Commencing with the first expiratory sound and terminating an instant before the onset of the following expiratory sound" (See figure 3).

This repetition of expiratory and inspiratory sounds produced during a sustained bout of crying has long been described as rhythmic [51]. So, we think it is axiomatic to consider the expiration and inspiration sound like the primary rhythmic units underlying the repeated units of crying.

Accordingly, we propose the new rhythmic metric relatively to expiratory and inspiratory segments.

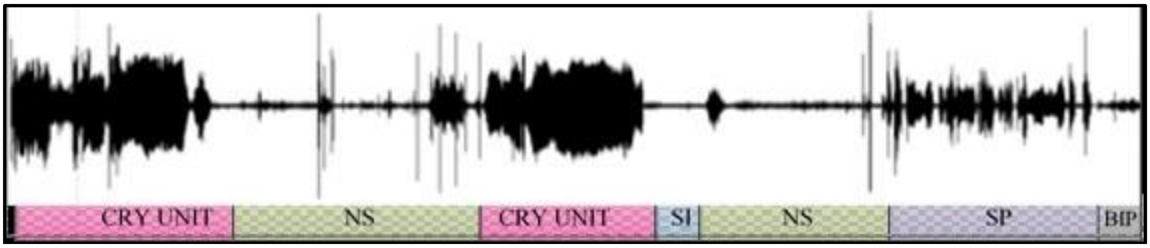


Figure2: An example of the utterance of cry (EXP – Expiration; IN – Inspiration; INSV – Voiced inspiration)

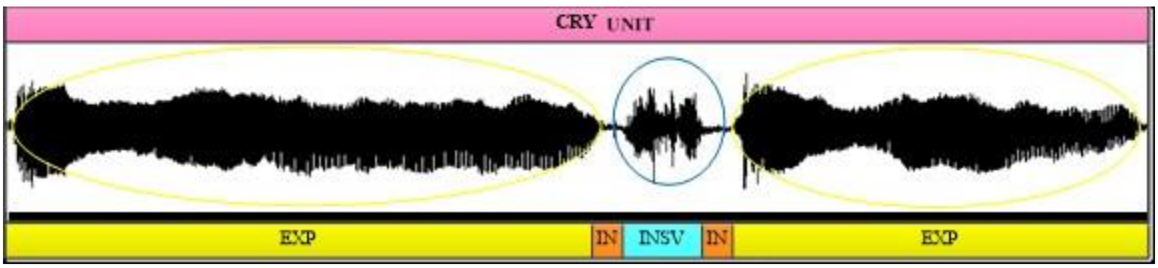
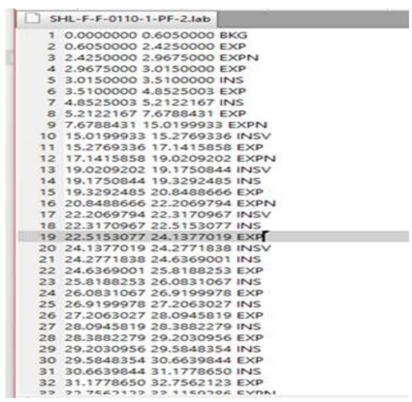


Figure 3: An example of the new segmentation of a unit or episode Cry

All cries samples were analyzed using Praat and Wavesurfer software[52, 53]. For calculation of the rhythm parameters, Expiration and inspiration manual segmentation was performed and labeled by Professor Chakib Tadj research team by visual inspection of crying waveforms and spectrograms according to standard segmentation criteria[44-46]. Figures 2, 3 and 4 illustrate different types of segments that may constitute a cry signal of any baby. We can notice, expiration and inspiration segments labeled by EXP and INS respectively. Other segments for example marked by CRN: a segment of cry with noise, PSN: pseudo cry with noise segment, BKG Symbolizes a background noise segment, and there are other labels of another type of segments.

However, to illustrate the rhythmical structure of the cry, we have to carry out a new segmentation. This new segmentation consists of extracting episodes of cries containing consecutive expiration and inspiration segments by eliminating any segment labeled other than expiration or inspiration. The minimum number required for periods of inspiration-expiration is two for each cry unit or an episode of cry. As a result, the duration of the cry units is variable. So, for each subject, we obtain a varied number of cry units with a varied number of expiration-inspiration cycle. Expiration and inspiration interval durations were extracted using a custom Matlab script on the boundary label files resulting from the manual segmentation already performed. Then, a new automatic segmentation was accomplished under Matlab. We extract the duration of its expiration and inspiration duration segments to calculate every rhythm metrics for each baby.

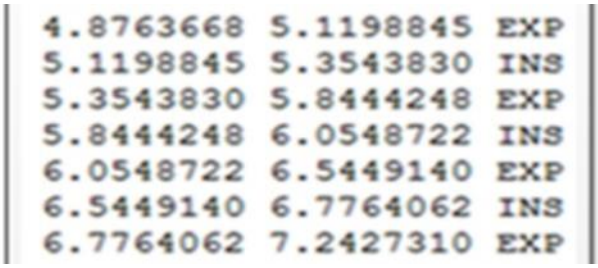


SHL-F-F-0110-1-PF-2.lab

1	0.0000000	0.6050000	BKG
2	0.6050000	2.4250000	EXP
3	2.4250000	2.9675000	EXPN
4	2.9675000	3.0150000	EXP
5	3.0150000	3.5100000	INS
6	3.5100000	4.8525003	EXP
7	4.8525003	5.2122167	INS
8	5.2122167	7.6788431	EXP
9	7.6788431	15.0199933	EXPN
10	15.0199933	15.2769336	INSV
11	15.2769336	17.1415858	EXP
12	17.1415858	19.0209202	EXPN
13	19.0209202	19.1750844	INSV
14	19.1750844	19.3292485	INS
15	19.3292485	20.8488666	EXP
16	20.8488666	22.2069794	EXPN
17	22.2069794	22.3170967	INSV
18	22.3170967	22.5153077	INS
19	22.5153077	24.1377019	EXP
20	24.1377019	24.2771838	INSV
21	24.2771838	24.6369001	INS
22	24.6369001	25.8188253	EXP
23	25.8188253	26.0831067	INS
24	26.0831067	26.9199978	EXP
25	26.9199978	27.2063027	INS
26	27.2063027	28.0945819	EXP
27	28.0945819	28.3882279	INS
28	28.3882279	29.2030956	EXP
29	29.2030956	29.5848354	INS
30	29.5848354	30.6639844	EXP
31	30.6639844	31.1778650	INS
32	31.1778650	32.7562123	EXP

Figure 4: An example of the original lab files for a baby crying signal recording

For each cry unit, we have measured the durations of the inspiration and expiration, segments. Then we computed the standard deviation, the raw and normalized pairwise for both inspiration and expiration segments. Others parameters were calculated too. The rhythm measures are summarized with a brief description in Table 1.



4.8763668	5.1198845	EXP
5.1198845	5.3543830	INS
5.3543830	5.8444248	EXP
5.8444248	6.0548722	INS
6.0548722	6.5449140	EXP
6.5449140	6.7764062	INS
6.7764062	7.2427310	EXP

Figure 5: An example of the new label file resulted for the new segmentation to extract rhythm features

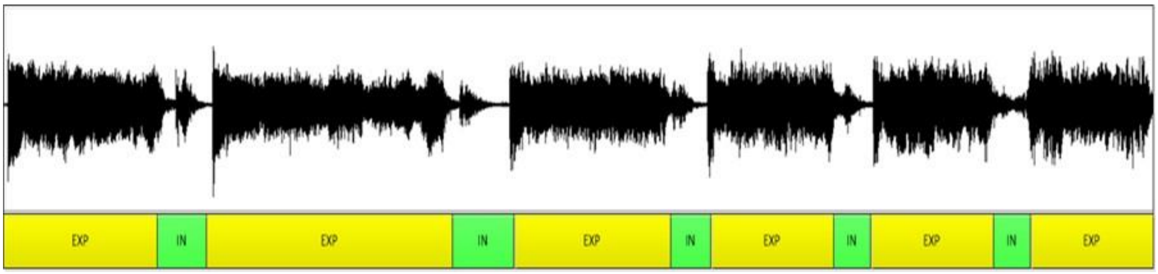


Figure 6: Cry unit resulted from the new automatic segmentation

Table 1: Rhythm features with a brief description

N	Parameter	Parameter Description
1	%EXP	Percent of expiration duration segments of episode cry
2	ΔINS	The standard deviation of inspiration intervals.
3	ΔEXP	The standard deviation of expiration intervals.
4	rPVI- EXP	Raw pairwise variability index for expiration intervals. Mean of the differences between successive expiration intervals.
5	nPVI-EXP	Normalized pairwise variability index for expiration intervals duration. Mean of the differences between successive expiration intervals divided by their sum (× 100).
6	rPVI-INS	Raw pairwise variability index for inspiration intervals duration. Mean of the differences between successive inspiration intervals.
7	nPVI-INS	Normalized pairwise variability index for inspiratory intervals duration. Mean of the differences between successive expiration intervals divided by their sum (× 100).
8	VarcoEXP	The standard deviation of expiratory intervals divided by mean expiratory duration (x 100).
9	VarcoINS	The standard deviation of inspiratory intervals divided by mean inspiratory duration (x 100).
10	ECD	Episode Cry Durations
11	EXPD	Expiration segment Durations
12	INS	Inspiration segment Durations
13	REXP	Range Expiration segment Durations
14	RINS	Range Inspiration segment Durations

Mean values and standard deviation, minimum, maximum and range of episode cry expiration, inspiration interval durations per baby group are reported in Table 2 and figure7.

Table 2: Corpus Statistics

Health state and sex of babies		Sum of Mean Cry episode durations(sec)	Sum of mean expiration interval durations(sec)	Sum of mean inspiration interval durations(sec)
Sick	Male	**1482.58**	159.77	59.42
	Female	845.37	93.27	35.11
	Total	2327.95	253.04	94.53
Healthy	Male	806.80	69.76	22.46
	Female	**1679.12**	191.32	73.05

	Total	**2485.92**	261.08	95.51

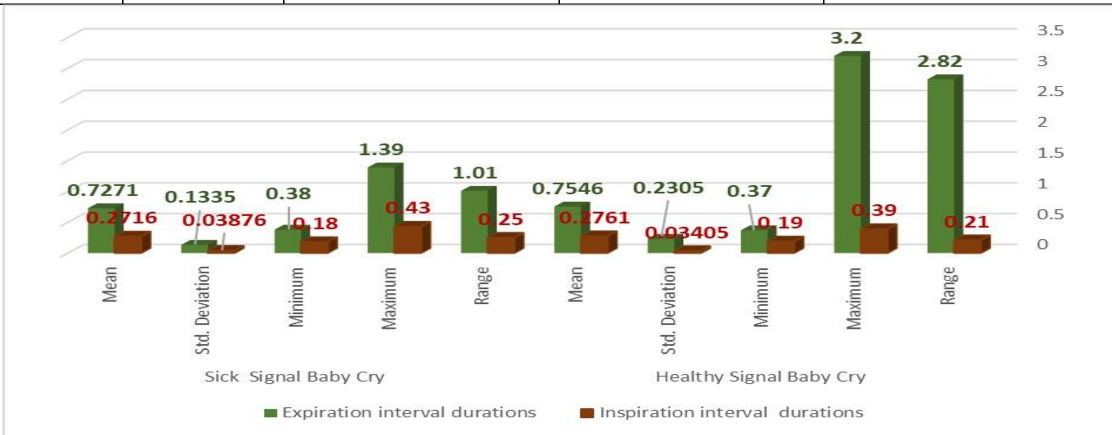


Figure 7: Episode Cry, Expiration, and Inspiration interval durations for both healthy and sick babies

3. CLASSIFICATION AND STATISTICAL ANALYSES

The primary objective of data analysis was to identify whether rhythm metrics could robustly distinguish sick babies from healthy ones. Toward this end, a series of statistical (Analysis of variance: ANOVA and non-parametric test) and neural network classification analysis was undertaken under the SPSS v.23 statistical package (IBM, 2017).

Analysis of variance (ANOVA) and non-parametric test are tow statistical analysis looking for group differences. Anova is a parametric test which compares the variances within the group to the **variance** between the groups to see whether the differences between groups are “big enough” to say that the groups come from different populations. The assumptions of the parametric one-way ANOVA are a normal distribution and homogeneous variances. However, there are very few data that is exactly normally distributed. So, the best way to deal with non-normal distributions is to use robust tests: to use a non-parametric test which does not need any assumption. It is why we use its results for comparison to those of ANOVA test (for more details see [54-57].

Multilayer Perceptron's (MLP) is undoubtedly the most common architecture used today because of their power, flexibility, and ease of use. The multilayer perceptron **(MLP)** employed in predictive applications are supervised in the sense that the model-predicted results can be compared against known values of the target variables. This structure is known as a feed-forward architecture because the connections in the network flow forward from the input layer to the output layer without any feedback loops. The multilayer perceptron (MLP) has one layer of neurons for inputs also called predictors (rhythm parameters in our case), another for outputs or target variables : tow categorical variables (in our case: healthy and sick baby's outputs), with one or more hidden layers between them (see figure 7)[58-61].

For the performance of the analyses, multi-layer perceptron (MLP) network models were developed. We assigned 53.3 %, 16.6 % and 30.1 of the cases to the training, testing, and to holdout groupings respectively. Where the model is built on the training set, the prediction errors are calculated using the holdout or validation set, and the test set is used to assess the generalization error of the final model. This test set should be locked away until the model calibration process is finished to prevent underestimation of the real model error.

3.1. Results and Discussion

3.1.1. Statistics Analysis Results

To determine if the metrics demonstrated significant group differences; a series of one-way analyses (ANOVA) and non-parametric test were conducted. Table 4 reported P values for each parameter [62, 63]. Only seven parameters of the metrics established significant group differences ($p < .05$). By ANOVA method: ΔEXP, ΔINS, VarcoEXP, VarcoINS, INSD, REXP, and RINS. The pertinent parameters selected by non-parametric test were: ΔINS, nPVI-EXP, VarcoEXP, VarcoINS, ECD, REXP, and RINS (See Table 3).

Table 3: ANOVA and non-parametric Analyzes results

	Non-parametric test	ANOVA Test
Parameter	**Sig (p value)**	**Sig (P value)**
%EXP	0.074	0.659
ΔEXP	0.763	**0.012**
ΔINS	**0.000**	**0.000**
rPVI- EXP	0.622	0.065
nPVI-EXP	**0.021**	0.062
rPVI-INS	0.29	0.303
nPVI-INS	0.544	0.338
VarcoEXP	**0.000**	**0.006**
VarcoINS	**0.000**	**0.000**
ECD	0.729	**0.042**
EXPD	0.233	**0.055**
INSD	**0.029**	0.111
REXP	**0.014**	**0.002**
RINS	**0.000**	**0.000**

3.1.2. Neural network classification results

Three analyses were conducted. The first analysis focused on seven relevant rhythm metrics determined by ANOVA analysis ΔEXP, ΔINS, VarcoEXP, VarcoINS, INSD, REXP, and RINS. In the second analysis, were the seven parameters are given by the non-parametric test: ΔINS, nPVI-EXP, VarcoEXP, VarcoINS, ECD, REXP, and RINS were included. Finally, the third analysis included all of the Rhythm metrics.

For the MLP network, we employed the Gradient descent optimization algorithm. Gradient Descent is the essential technique and the foundation of how we train and optimize intelligent systems [64, 65]. The transfer function hyperbolic tangent used for all units in the hidden layer. The activation function softmax utilized for all units in the output layer (see figures 8 and 9).

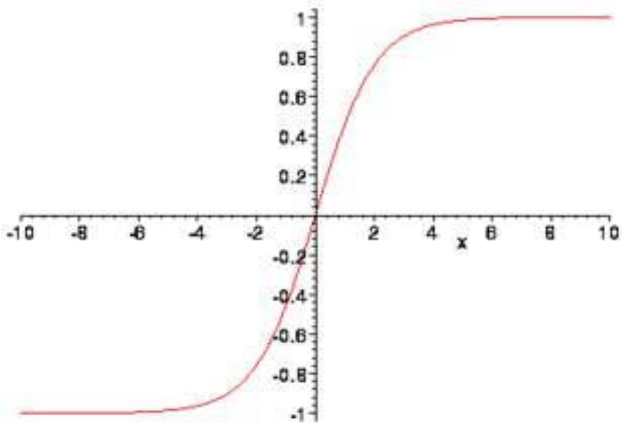


Figure 8: The transfer function hyperbolic tangent

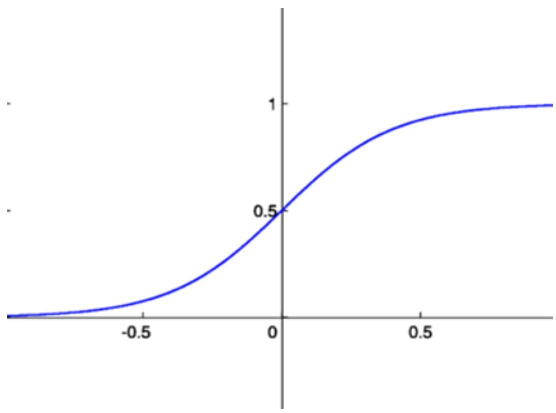


Figure 9: The activation Function Softmax

Before using the input data records to the ANN, a normalization process took place so that the values with wide range do not prevail over the rest. The standardized approach was applied: Subtract the mean and divide by the standard deviation.

These networks were trained in an iterative process. A single hidden sub-layer architecture was followed. Five units were chosen in the hidden layer and two units in the output layer. The schematic representation of the neural network is shown in Figure 10.

Classification results 1: Two categories of babies groups were established: healthy babies versus sick babies. For the first experiments, we have considered only two sets of data: training and holdout sample with 70% and 30 % respectively of the overall data (see table 4).

Table 4: Case Processing Summary given by SPSS neural network analysis.

		N	Percent
Sample	Training	485	69.9%
	Holdout	209	30.1%
Valid		694	100.0%
Excluded		0	
Total		694	

As mentioned in the above, the classification results concern only the seven ANOVA's parameters of the fourteen variables. These parameters: ΔEXP, ΔINS, VarcoEXP, VarcoINS, INSD, REXP, and RINS were entered into the Multilayer Perceptron (MLP) as inputs. The MLP accurately classified 80.6% of the babies into their respective groups (69.9% with the holdout sample; see Table 5). Thus, the results of these only seven predictors demonstrate the ability of established rhythm metrics to distinguish among baby groups.

Table 5: Classification results with Anova rhythm metrics for analysis 1.

Sample	Observed	Predicted: Sick Signal	Healthy signal	Percent correct
Training	Sick Signal Baby Cry	191	48	79.9%
	Healthy Signal Baby Cry	46	200	81.3%
	Overall Percent	48.9%	51.1%	80.6%
Holdout	Sick Signal Baby Cry	73	36	67.0%
	Healthy Signal Baby Cry	27	73	73.0%
	Overall Percent	47.8%	52.2%	69.9%

Classification results 2: As before, the predictor variables identified by non-parametric were entered into MLP. ΔINS, nPVI-EXP, VarcoEXP, VarcoINS, ECD, REXP, and RINS resulted in correct classification of 77.9% of the babies by training sample and by holdout sample an accurate classification of 67.5% of the babies (see Table 6 for a detailed summary of classification results).
These results are not far from the previous ones. We notice a slight decrease from 80.6% to 77.9% (from 69.9% to 67.5% with holdout sample).

Table 6: Classification results with the non-paramedic rhythm parameters

Sample	Observed	Predicted: Sick Signal	Healthy Signal	Percent Correct
Training	Sick Signal Baby Cry	179	60	74.9%
	Healthy Signal Baby Cry	47	199	80.9%
	Overall Percent	46.6%	53.4%	77.9%
Holdout	Sick Signal Baby Cry	65	44	59.6%
	Healthy Signal Baby Cry	24	76	76.0%
	Overall Percent	42.6%	57.4%	67.5%

Classification results 3: The MLP using all metrics. It resulted in a correct classification of 96.1% of newborns. The holdout procedure resulted in correct classification of 82.3% of the babies (see Table 7 for a detailed summary of classification results). We notice that the percentage of 96.1 is remarkable.

Now, we add testing sample as indicated in table 9. From this analysis, 391 cases (56.3%) were assigned to the training sample, 94 (13.5%) to the testing sample, and 209 (30.1%) to the holdout sample (see table 8). The choice of the records was done randomly. The whole effort was targeted to the development of an ANN that would have the ability to generalize as much as possible.

Table 7: Summary classification results with the 14 selected rhythm parameters

Sample	Observed	Predicted Sick Signal Baby Cry	Predicted Healthy Signal Baby Cry	Predicted Percent Correct
Training	Sick Signal Baby Cry	230	9	**96.2%**
	Healthy Signal Baby Cry	10	236	95.9%
	Overall Percent	49.5%	50.5%	96.1%
Holdout	Sick Signal Baby Cry	86	23	78.9%
	Healthy Signal Baby Cry	14	86	**86.0%**
	Overall Percent	47.8%	52.2%	82.3%

Table 8: Case Processing Summary by adding the testing sample

		N	Percent
Sample	Training	391	56.3%
	Testing	94	13.5%
	Holdout	209	30.1%
Valid		694	100.0%
Excluded		0	
Total		694	

The MLP achieved a 90.5% classification accuracy for the training sample, 84.2% with holdout grouping and 88.7% with the testing sample. It is remarkable to have the high levels of training, testing and holdout classification accuracy indicating the robustness of these variables in classification (see table 9).

Table 9: Overall system accuracy obtained for the totality of the 14 selected rhythm parameters

Sample	Observed	Predicted Sick Signal Baby Cry	Predicted Healthy Signal Baby Cry	Predicted Percent Correct
Training	Sick Signal Baby Cry	167	19	89.8%
	Healthy Signal Baby Cry	16	168	**91.3%**
	Overall Percent	49.5%	50.5%	90.5%
Testing	Sick Signal Baby Cry	47	6	88.7%
	Healthy Signal Baby Cry	7	55	88.7%
	Overall Percent	47.0%	53.0%	88.7%
Holdout	Sick Signal Baby Cry	90	19	82.6%
	Healthy Signal Baby Cry	14	86	**86.0%**
	Overall Percent	49.8%	50.2%	84.2%

These results lead to the suggestion that metrics which capture the temporal variation of inspiration and expiration segments are useful in distinguishing healthy from sick babies.

Independent variable importance analysis performs a sensitivity analysis. It calculates the importance of each predictor in constructing the neural network. The study uses the combined training and testing samples. A sensitivity analysis to compute the significance of each predictor is applied. It creates a chart displaying importance and normalized importance for each predictor. The importance graph (Figure 11) shows the domination of VarcoINS and REXP followed distantly by other predictors.

The majority of classification results reported by this study were more than 80% successful in classifying babies into their appropriate group, and for the holdout sample were more than 70% successful. The VarcoINS rhythm metrics is the most reliable parameter (VarcoINS was the most significant predictor variable determined by statistical analysis and by all the experimentations).

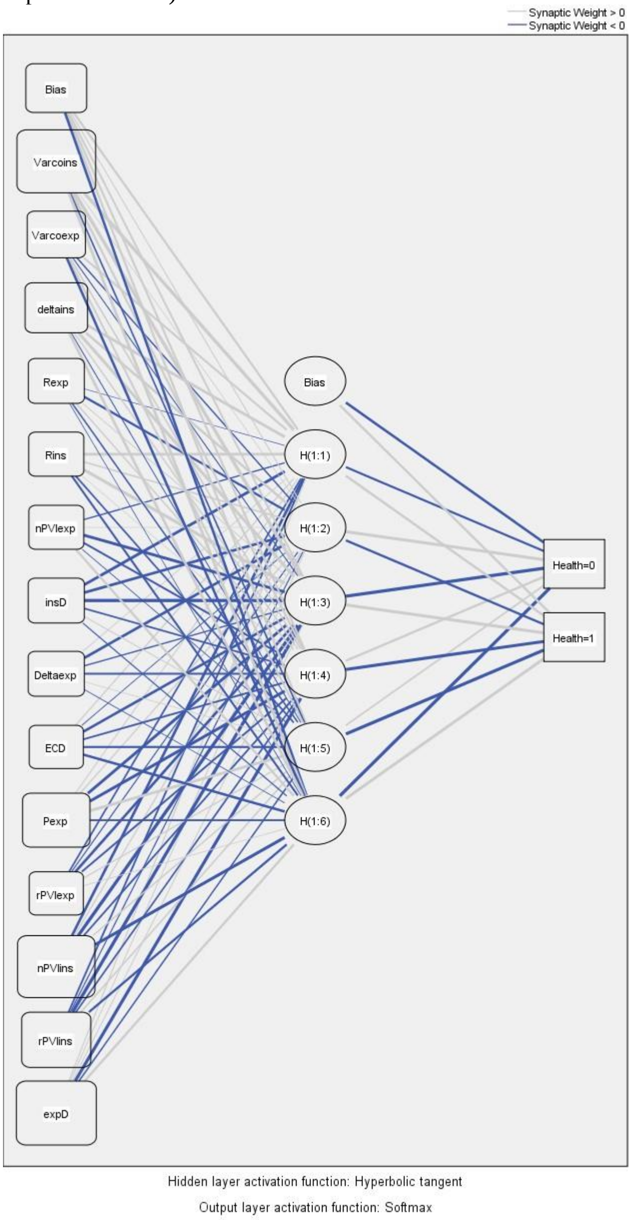


Figure 10: Multi-layer perceptron network architecture used in the last experience classification

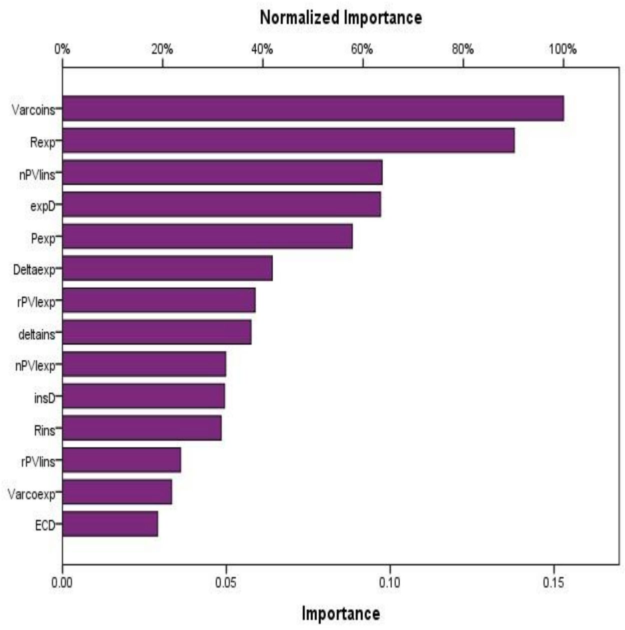


Figure 11: Rhythm metrics importance analysis

4. CONCLUSION AND PERSPECTIVES

In conclusion, this is the first study trying to quantify rhythm in the baby crying signal using the temporal aspects. Indeed, the purpose of this study was to define and quantify the rhythm features of infant cry systematically. We utilized metrics based on both expiratory and inspiratory intervals duration. The results of the MLP classification of this preliminary survey are auspicious. The study demonstrates that these rhythm parameters are useful in distinguishing healthy cries from that produced by unhealthy ones.

Indeed, the reliability of the rhythm metrics proposed in this paper is proved compared to many works which addressed the same issue: discrimination between healthy and sick babies throughout their cries signals. These studies utilized different kinds of acoustic features and many technics of soft computing or stochastic methods and a separate database of baby cries. We can distinguish among these works those using neural networks as method of classification and a different kind of acoustic analysis [3, 4]. The results were from 70 % to 100%.

A long-term goal associated with this research is to determine the usefulness of these parameters in the identification of infants at risk for various health conditions and why not determine the reason or the context of baby cries. We can investigate another rhythm metrics (for example: combined rhythm parameters which include the duration of both expiratory and inspiratory intervals) with larger population.

ACKNOWLEDGMENT

This work is supported by the Bill and Melinda Gates Foundation grant number OPP1067980. Thank the staff of Neonatology departments at Sainte Justine Hospital and Sahel Hospital for their cooperation in the data collection process.

REFERENCES.

[1] J. F. Bosma, H. Truby, and J. Lind, "Cry motions of the newborn infant," *Acta Paediatrica*, vol. 54, pp. 60-92, 1965.

[2] A. Chittora and H. A. Patil, "Data collection and corpus design for analysis of normal and pathological infant cry," in *Oriental COCOSDA held jointly with 2013 Conference on Asian Spoken Language Research and Evaluation (O-COCOSDA/CASLRE), 2013 International Conference*, 2013, pp. 1-6.

[3] M. Hariharan, J. Saraswathy, R. Sindhu, W. Khairunizam, and S. Yaacob, "Infant cry classification to identify asphyxia using time-frequency analysis and radial basis neural networks," *Expert Systems with Applications*, vol. 39, pp. 9515-9523, 2012.

[4] M. Hariharan, R. Sindhu, and S. Yaacob, "Normal and hypoacoustic infant cry signal classification using time–frequency analysis and general regression neural network," *Computer methods and programs in biomedicine*, vol. 108, pp. 559-569, 2012.

[5] M. Hariharan, S. Yaacob, and S. A. Awang, "Pathological infant cry analysis using wavelet packet transform and probabilistic neural network," *Expert Systems with Applications*, vol. 38, pp. 15377-15382, 2011.

[6] M. Kia, S. Kia, N. Davoudi, and R. Biniazan, "A detection system of infant cry using fuzzy classification including dialing alarm calls function," in *Innovative Computing Technology (INTECH), 2012 Second International Conference on*, 2012, pp. 224-229.

[7] M. Kia, S. Kia, and K. M. Mashhadi, "A Detection System of Infant's Cry, Using Fuzzy Classification: from Theory to Practice," *J. Information*, vol. 2, pp. 114-127, 2012.

[8] Y. Mima and K. Arakawa, "Cause Estimation of Younger Babies' Cries from the Frequency Analyses of the Voice-Classification of Hunger, Sleepiness, and Discomfort," in *Intelligent Signal Processing and Communications, 2006. ISPACS'06. International Symposium on*, 2006, pp. 29-32.

[9] S. Ntalampiras, "Audio pattern recognition of baby crying sound events," *Journal of the Audio Engineering Society*, vol. 63, pp. 358-369, 2015.

[10] S. Orlandi, C. A. R. Garcia, A. Bandini, G. Donzelli, and C. Manfredi, "Application of pattern recognition techniques to the classification of full-term and preterm infant cry," *Journal of Voice*, vol. 30, pp. 656-663, 2016.

[11] A. Prochnow, S. Erlandsson, V. Hesse, and K. Wermke, "Does a 'musical'mother tongue influence cry melodies? A comparative study of Swedish and German newborns," *Musicae Scientiae*, p. 1029864917733035, 2017.

[12] O. F. Reyes-Galaviz, S. D. Cano-Ortiz, and C. A. Reyes-García, "Evolutionary-neural system to classify infant cry units for pathologies identification in recently born babies," in *Artificial Intelligence, 2008. MICAI'08. Seventh Mexican International Conference on*, 2008, pp. 330-335.

[13] R. Robu, F. Feier, V. Stoicu-Tivadar, C. Ilie, and I. Enătescu, "The analysis of the new-borns' cry using NEONAT and data mining techniques," in *Intelligent Engineering Systems (INES), 2011 15th IEEE International Conference on*, 2011, pp. 235-238.

[14] J. Saraswathy, M. Hariharan, S. Yaacob, and W. Khairunizam, "Automatic classification of infant cry: A review," in *Biomedical Engineering (ICoBE), 2012 International Conference on*, 2012, pp. 543-548.

[15] K. Wermke, Y. Ruan, Y. Feng, D. Dobnig, S. Stephan, P. Wermke, *et al.*, "Fundamental frequency variation in crying

of Mandarin and German neonates," *Journal of Voice,* vol. 31, pp. 255. e25-255. e30, 2017.
[16] G. Zamzmi, C.-Y. Pai, D. Goldgof, R. Kasturi, Y. Sun, and T. Ashmeade, "Automated Pain Assessment in Neonates," in *Scandinavian Conference on Image Analysis*, 2017, pp. 350-361.
[17] P. Ostwald, "The sounds of infancy," *Developmental Medicine & Child Neurology,* vol. 14, pp. 350-361, 1972.
[18] R. Prescott, "Infant cry sound; developmental features," *The Journal of the Acoustical Society of America,* vol. 57, pp. 1186-1191, 1975.
[19] H. Truby and J. Lind, "CRY SOUNDS OF THE NEWBORN INFANT1, 2," *Acta Paediatrica,* vol. 54, pp. 8-59, 1965.
[20] P. S. Zeskind and R. G. Barr, "Acoustic Characteristics of Naturally Occurring Cries of Infants with" Colic"," *Child development,* pp. 394-403, 1997.
[21] P. S. Zeskind, S. Parker-Price, and R. G. Barr, "Rhythmic organization of the sound of infant crying," *Developmental psychobiology,* vol. 26, pp. 321-333, 1993.
[22] P. S. Zeskind, A. Wilhite, and T. R. Marshall, "Infant cry as a graded signal: Experimental modifications of durations of pauses and expiratory sounds alter mothers' perceptions," *Early Development and Parenting,* vol. 2, pp. 99-105, 1993.
[23] D. Crystal, "Non-segmental phonology in language acquisition: A review of the issues," *Lingua,* vol. 32, pp. 1-45, 1973.
[24] B. M. Lester, "There's more to crying than meets the ear," *Infant Crying: Theoretical and Research Perspectives. New York: Plenum Publishing Corp,* pp. 1-2, 1985.
[25] A. Langlois and R. Baken, "Development of Respiratory Time-factors in Infant Cry," *Developmental Medicine & Child Neurology,* vol. 18, pp. 732-737, 1976.
[26] S. M. Grau, M. P. Robb, and A. T. Cacace, "Acoustic correlates of inspiratory phonation during infant cry," *Journal of Speech, Language, and Hearing Research,* vol. 38, pp. 373-381, 1995.
[27] C. N. WILDER, "Respiratory patterns in infant cry," *Developmental psychology,* vol. 6, pp. 293-301, 1974.
[28] M. P. Robb and A. Goberman, "Application of an acoustic cry template to evaluate at-risk newborns: Preliminary findings," *Neonatology,* vol. 71, pp. 131-136, 1997.
[29] P. Lieberman, "Primate vocalizations and human linguistic ability," *The Journal of the Acoustical Society of America,* vol. 44, pp. 1574-1584, 1968.
[30] P. Lieberman, "On the evolution of language: A unified view," *Cognition,* vol. 2, pp. 59-94, 1973.
[31] P. Lieberman, "The physiology of cry and speech in relation to linguistic behavior," in *Infant Crying*, ed: Springer, 1985, pp. 29-57.
[32] P. Lieberman, K. S. Harris, P. Wolff, and L. H. Russell, "Newborn infant cry and nonhuman primate vocalization," *Journal of Speech, Language, and Hearing Research,* vol. 14, pp. 718-727, 1971.
[33] R. Prescott, "Cry and maturation," *Infant communication: Cry and early speech,* pp. 234-250, 1980.
[34] V. Dellwo, "Rhythm and speech rate: A variation coefficient for∆ C," *Language and language-processing,* pp. 231-241, 2006.
[35] E. Grabe, "The IViE labelling guide, version 3,": http://www.phon.ox.ac.uk/files/apps/IViE//guide.html#.
[36] E. Grabe and E. L. Low, "Durational variability in speech and the rhythm class hypothesis," *Papers in laboratory phonology,* vol. 7, 2002.
[37] E. Grabe and E. L. Low, "Acoustic correlates of rhythm class," *Laboratory phonology,* vol. 7, 2002.
[38] W. Grabe, "Reading research and its implications for reading assessment," *Fairness and validation in language assessment,* pp. 226-262, 2000.
[39] F. Ramus, M. Nespor, and J. Mehler, "Correlates of linguistic rhythm in the speech signal," *Cognition,* vol. 73, pp. 265-292, 1999.
[40] F. Ramus, "Language discrimination by newborns: Teasing apart phonotactic, rhythmic, and intonational cues," *Annual Review of Language Acquisition,* vol. 2, pp. 85-115, 2002.
[41] L. White and S. L. Mattys, "Calibrating rhythm: First language and second language studies," *Journal of Phonetics,* vol. 35, pp. 501-522, 2007.
[42] J. M. Liss, L. White, S. L. Mattys, K. Lansford, A. J. Lotto, S. M. Spitzer, *et al.*, "Quantifying speech rhythm abnormalities in the dysarthrias," *Journal of Speech, Language, and Hearing Research,* vol. 52, pp. 1334-1352, 2009.
[43] V. Dellwo and P. Wagner, "Relationships between rhythm and speech rate," 2003.
[44] L. Abou-Abbas, H. F. Alaei, and C. Tadj, "Segmentation of voiced newborns' cry sounds using wavelet packet based features," in *Electrical and Computer Engineering (CCECE), 2015 IEEE 28th Canadian Conference on*, 2015, pp. 796-800.
[45] L. Abou-Abbas, H. Fersaie Alaie, and C. Tadj, "Automatic detection of the expiratory and inspiratory phases in newborn cry signals," *Biomedical Signal Processing and Control,* vol. 19, pp. 35-43, 5// 2015.
[46] L. Abou-Abbas, L. Montazeri, C. Gargour, and C. Tadj, "On the use of EMD for automatic newborn cry segmentation," in *2015 International Conference on Advances in Biomedical Engineering (ICABME)*, 2015, pp. 262-265.
[47] Y. Kheddache and C. Tadj, "Acoustic measures of the cry characteristics of healthy newborns and newborns with pathologies," *Journal of Biomedical Science and Engineering,* vol. Vol.06No.08, p. 9, 2013.
[48] Y. Kheddache and C. Tadj, "Characterization of Pathologic Cries of Newborns Based on Fundamental Frequency Estimation," *Engineering,* vol. Vol.05No.10, p. 5, 2013.
[49] Y. Kheddache and C. Tadj, "Resonance Frequencies Behavior in Pathologic Cries of Newborns," *Journal of Voice,* vol. 29, pp. 1-12, 1// 2015.
[50] L. Abou-Abbas, C. Tadj, C. Gargour, and L. Montazeri, "Expiratory and inspiratory cries detection using different signals' decomposition techniques," *Journal of Voice,* 2016.
[51] P. H. Wolff, "The role of biological rhythms in early psychological development," *Bulletin of the Menninger Clinic,* vol. 31, p. 197, 1967.
[52] P. Boersma, "Praat: doing phonetics by computer," *http://www. praat. org/,* 2006.
[53] T. KTH, "WaveSurfer," *URL: http 二" www. speech. kth, se/wavcsurfer*.
[54] W. J. Conover and R. L. Iman, "Rank transformations as a bridge between parametric and nonparametric statistics," *The American Statistician,* vol. 35, pp. 124-129, 1981.
[55] C. R. Hicks, "Fundamental concepts in the design of experiments," in *Fundamental concepts in the design of experiments,* ed: Holt, Rinehart and Winston, 1963.
[56] D. E. Hinkle, W. Wiersma, and S. G. Jurs, "Applied statistics for the behavioral sciences," 2003.
[57] R. C. Littell, *SAS*: Wiley Online Library, 1996.
[58] H. B. Demuth, M. H. Beale, O. De Jess, and M. T. Hagan, *Neural network design*: Martin Hagan, 2014.

[59] R. Hecht-Nielsen, "Theory of the backpropagation neural network," *Neural Networks,* vol. 1, pp. 445-448, 1988.
[60] T. Mikolov, M. Karafiát, L. Burget, J. Cernocký, and S. Khudanpur, "Recurrent neural network based language model," in *Interspeech*, 2010, p. 3.
[61] D. George and P. Mallery, *IBM SPSS statistics 23 step by step: A simple guide and reference*: Routledge, 2016.
[62] O. Aalen, "Nonparametric inference for a family of counting processes," *The Annals of Statistics*, pp. 701-726, 1978.
[63] A. Agresti, *An introduction to categorical data analysis*: John Wiley, 2007.
[64] L. Bottou, "Large-scale machine learning with stochastic gradient descent," in *Proceedings of COMPSTAT'2010*, ed: Springer, 2010, pp. 177-186.
[65] A. Govan, "Introduction to optimization," in *North Carolina State University, SAMSI NDHS, Undergraduate workshop*, 2006.

Chakib Tadj: Professor at École de Technologie Supérieure (ETS), University of Quebec, Montreal, Canada. He received a Ph.D. degree in signal and image processing from ENST Paris, Paris, France, in 1995. His primary research interests include signal processing, speech recognition, pervasive computing, and multimodal systems.

Habiba DAHMANI: associate professor at Electrical Engineering Department, Faculty of technology, Mohamed Boudiaf University in Msila, Algeria. She received her Bachelor of electronic Engineering in in 1995, Master in signal processing in 2003 and her Ph.D. in 2015 from Badji Mokhtar University, Algeria. She works as a researcher in INRS in Canada at Institute of Energy, Materials and Telecommunications at Quebec University in Montreal for finishing Her Ph.D. Dissertation. She has a Postdoctoral position as a researcher at ÉTS (École de Technologie Supérieure), Quebec University, Montreal, Canada in 2016. Her primary research interests include signal processing, healthy and pathological speech processing, digital communications and Currently, Babies' Crying signal.