ORIGINAL ARTICLE

Hearing Screening through Frequency Analysis of Auditory Brainstem Response Using PhysioNet Data

Catherine Manuela Lee-Ramos,[1] Al Francis L. Bontogon,[2] Angelica S. Collanto,[2] Patrick John P. Labra, MD,[3] Luis G. Sison, PhD[4] and Charlotte M. Chiong, MD, PhD[3]

[1]*Manufacturing Engineering and Management, De La Salle University, Manila, Philippines*
[2]*Electronics, Communications and Computer Engineering Department, Ateneo de Manila University, Quezon City, Philippines*
[3]*Philippine National Ear Institute, National Institutes of Health, University of the Philippines Manila, Manila, Philippines*
[4]*Electrical and Electronics Engineering Institute, University of the Philippines Diliman, Quezon City, Philippines*

ABSTRACT

Objectives. Responding to the reality of neonate patients with delayed childhood development due to late diagnosis of and intervention on hearing impairment, this study aims to determine the features based on time-frequency domain of auditory brainstem response (ABR) signals and to test the protocol on ABR signals from PhysioNet.

Methods. This is done by pre-processing, performing time-frequency analysis, and characterizing hearing impairment using the dominant features of the ABR. In this study, normal (N) and hearing impaired (HI) ABR adult human signals were acquired from Physionet.org, a publicly available database. Considering its high signal-to-noise ratio, numerous filters and transformations were applied to extract the ABR. Consequently, the features acquired — dominant frequency and bigrams, were used as data classifiers.

Results. Initial results using only N classifiers, that is features from the Normal dataset, and bandpass Chebyshev filter with a lower cut-off frequency of 60 Hz show that the tests yielded low to middle sensitivity. Further tests were done to improve the sensitivity that incorporated the HI classifiers, used data filtered with a low cut-off frequency of 300 Hz, and data divided per stimulus intensity level.

Conclusion. Conclusions made are 1) data with both N and HI classifiers have higher sensitivity than those using only N classifiers, 2) data with a Chebyshev cut-off frequency of 300 Hz have a higher sensitivity than those with 60 Hz, and 3) data divided per intensity level have a higher sensitivity than data analyzed as a whole, and that features with stimulus intensity in middle ranges have a better distinction between HI and N patients.

Keywords: evoked potentials, auditory, brain stem, delayed diagnosis, hearing loss humans

INTRODUCTION

Early detection of hearing loss is crucial for mental and social development of the child. This is primarily addressed by Republic Act (RA) 9709 known as "Universal Newborn Hearing Screening and Intervention Act" which establishes a universal newborn hearing screening program for the Prevention, Early Diagnosis and Intervention of Hearing Loss.[1]

RA 9709 mandates all newborns in the Philippines to go through hearing screening prior to hospital discharge or within three months after birth for those born outside hospitals. Prior to discharge, the baby's hearing is tested by either otoacoustic emission (OAE) or automated auditory brainstem response (AABR). Depending on the screening modality used, this may be a one-step test (AABR) or a two-step test with a re-screening session.[2]

eISSN 2094-9278 (Online)
Published: September 28, 2023
https://doi.org/10.47895/amp.v57i9.4336

Corresponding author: Catherine Manuela Lee-Ramos
Manufacturing Engineering and Management
De La Salle University
2401 Taft Ave., Malate, Manila, Philippines
Email: catherine.ramos@dlsu.edu.ph
ORCiD: https://orcid.org/0000-0002-0099-3554

Since the implementation of the Universal Newborn Hearing Screening and Intervention Program, only an estimated 10% of the hearing-impaired Filipino babies are being screened. This is due to the high cost and lack of availability of hearing screening devices.[1] One of the current solutions is the Hearing for Life (HeLe) project, a collaborative project between the University of California, Davis, the University of California, Berkeley, and the University of the Philippines that aims to develop a low-cost telehealth device that will enable local health units to implement newborn hearing screening. HeLe, an Auditory Brainstem Response (ABR) machine for neonates is currently being locally designed and fabricated. In conjunction with the hardware, a software algorithm is being developed and tested so that the entire system would have the capability of detecting automatically whether the subject is normal or hearing-impaired. Furthermore, the algorithm should be able to display the results immediately after the test.

For the past two decades, there has been a limited number of ABR research. Traditionally, most Automated Auditory Brainstem Response (AABR) machines rely on time domain analysis of the ABR signals. Current methods of detecting hearing impairments measure the latency of Wave V as a function of auditory stimulus. The latency is then measured against a standard model.[3,4] If the latency falls within the normal range, then the patient is tagged by the machine as normal hearing and the result of the test would be a PASS. However, if the latency falls outside of the normal range of the standard model, then the patient is tagged as hearing-impaired and the result of the test would be a FAIL. In this study, the researchers will try to develop a different algorithm based on the frequencies present in the ABR signals.

The study was done in collaboration with Philippine National Ear Institute (PNEI). Clinicians from PNEI were also consulted with regards methodology such as how individual sweeps for one stimulus are processed and analyzed. The researchers aim to analyze and determine what are the distinguishing features present in the ABR signals of normal hearing patients (N) and those with hearing impairments (HI). Once the distinguishing features have been obtained, further testing could be done to determine if these features could be used to automatically give a PASS (normal hearing) or a REFER (hearing-impaired) result immediately after the test. Unlike previous ABR algorithms, which focus on detection of specific features such as Wave V, the analysis of the entire short-latency auditory-evoked response waveform will be performed to identify any feature that may distinguish a waveform as “PASS” or “REFER”.

METHODS

Data

Using the ABR dataset from PhysioNet, an MIT database, the authors employed methods to analyze the ABR signals using the frequency components present in these signals and determine what are the unique frequencies in the normal and hearing-impaired subjects. The Evoked Auditory Responses in Normal and Hearing-impaired databases were both sampled at 48 kHz.[5] As of 2021, only the PhysioNet data were publicly available. The data from PhysioNet were validated and used in submissions in peer reviewed journals.[6,7] There are two separate databases in PhysioNet, a Normal Database and Hearing-impaired Database.[6,7] As shown in Figure 1, the data used in this study consists of eight normal and seven hearing-impaired adult subjects. There are a total of eight hearing-impaired adult subjects in the PhysioNet database but one was excluded because it used a different stimulus frequency. For the hearing-impaired, they were confirmed with clinical tests. The patients were tested using different audio stimulus intensities: 10-55, 60, 65, 70, 75, 80, 85, 90, 95, and 100 dB.

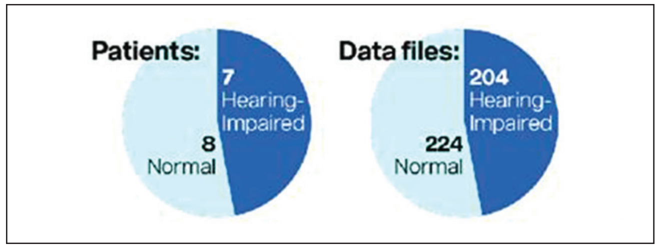

Figure 1. ABR Dataset from PhysioNet.

Each subject would have a number of files per stimulus intensity. This can be seen in Tables 4 and 5. For example, at an intensity level of less than 60 dB, there would be 23 hearing-impaired files and 98 normal hearing files. Overall, there were 224 files from normal subjects and 204 files for hearing impaired. For this study, the files were considered independent of each other. Each file contains about 989 sweeps which are then averaged and analyzed as a single ABR signal. That is at each time point, the arithmetic mean of 989 data points is obtained. Each sweep is initiated by an audio signal of 4-kHz tones with a duration of 1 ms. These are then padded with silence to get a total stimulus length of 41.7 ms.

Preprocessing

Processing and analysis of the raw data from Physionet.org were done using Matlab (Mathworks, Inc., MATLAB). Prior to analyzing the ABR signal, different processing techniques were employed to determine which one would perform best in removing significant noise or unwanted signals from the ABR signals. The noise present was due to electromagnetic interference (60 Hz) as well as the native EEG signals. Two cut-off frequencies were used, 60 Hz and 300 Hz. It was determined that the best cut-off frequency for the filters which yielded higher sensitivity is 300 Hz.

All sets of ABR signals will first undergo pre-processing to remove noise and eliminate unwanted frequencies thus isolating the signals needed. To further specify and explain this step, Figure 2 is shown below. The first step is to obtain the ABR data. Next, 900-1000 sweeps of the ABR signal are averaged. Third, the audio signal is removed so that the response waves may be isolated. Next, a moving average filter is applied which is essentially a low-pass filter to remove the high frequency noise. A high-pass filter with a cut-off frequency of 1 Hz is then applied to remove the DC offset or to center the ABR signal along the x-axis (time). Finally, a Chebyshev bandpass filter is applied to isolate the relevant ABR frequencies. There were several filters tested, Chebyshev, Bessel, and Butterworth filters. However, Chebyshev was used as this was the filter that gave a result closest to the "textbook" ABR signal in appearance.

The study used two types of bandpass filters, one with a cut off of 60 Hz to 1500 Hz and another with a cut off of 300 Hz to 1500 Hz. The frequency spectra of each of the steps are plotted for verification. Figure 3 shows the raw, unprocessed ABR data for 42 ms. The spikes aside from immediately after the stimulus cannot be differentiated because of the various noise present. On the other hand, a post-processed ABR

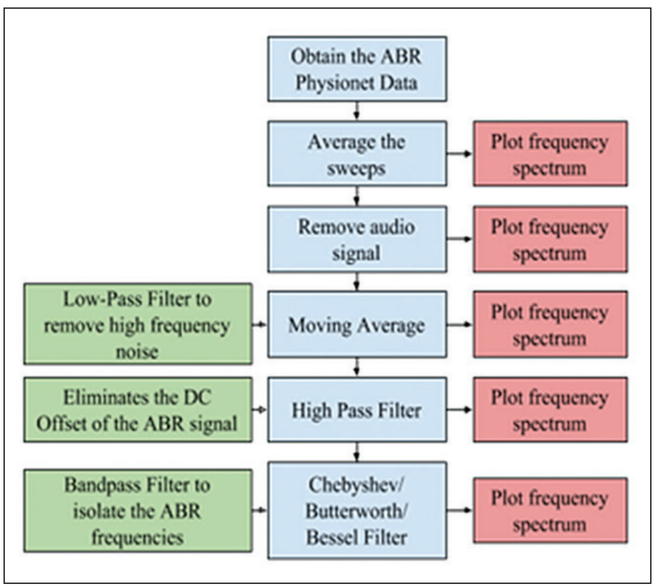

Figure 2. Preprocessing of ABR Data.

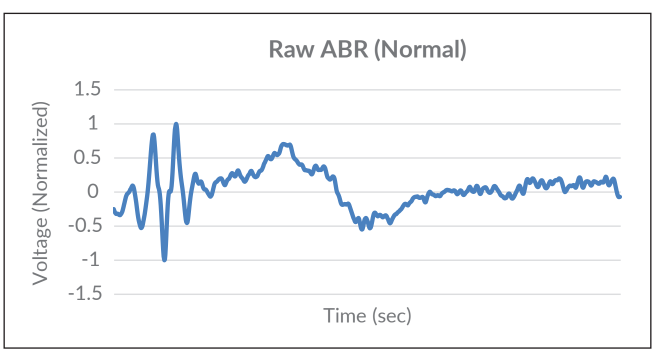

Figure 3. Raw ABR signal of normal hearing subject for 42 ms.

signal from a hearing-impaired subject and the corresponding spectrum after the high pass filter is shown in Figure 4.

Frequency Analysis

To be able to analyze the frequency, the ABR signal's frequency spectrum was retrieved using the Short-Time Fourier Transform (STFT). One sweep would correspond to 2002 datapoints. For 989 sweeps per file, there would be around 2 million datapoints. The averaged ABR signal would be divided into 39 segments. The segmented 2002 points will have 100 points (2ms) window size with 50% (50 points) overlap. However, only the first 10ms (9 segments) was analyzed since this already contains waves I-V of the ABR signal as shown in Figure 6. The top three dominant frequencies per segment are obtained. To obtain the dominant frequencies, a threshold was set so only those frequencies with sufficient power relevant to the analysis would be used. Frequencies outside the threshold would be disregarded.

To determine the frequency components that are present in the ABR signals, the frequency resolution was determined as 480 Hz which is computed accordingly:

$$\text{Frequency resolution} = Fs/Nd \qquad (1)$$

where, *Fs* is the sampling frequency (48 kHz) and *Nd* is the number of datapoints. The STFT in Matlab would show which frequency components are present in each segment. The dominant frequencies per segment were obtained and formed a frequency set. The top 20% of the frequency sets,

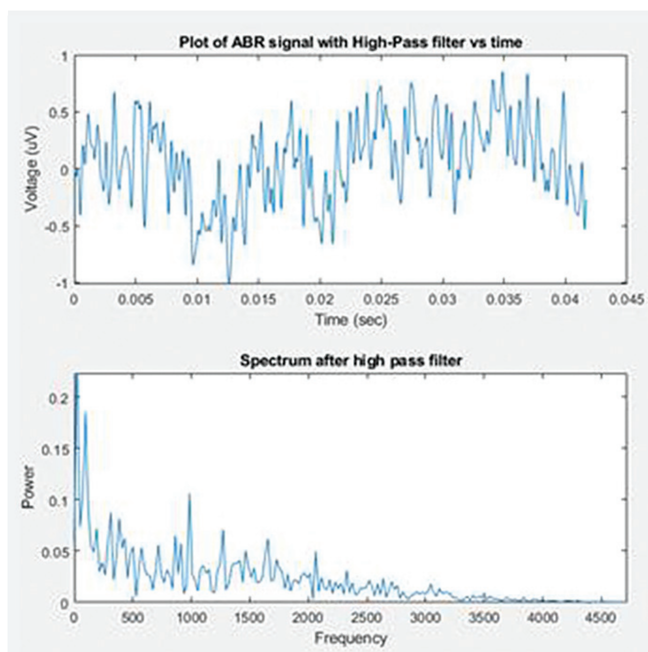

Figure 4. Filtered ABR signal of hearing-impaired subject and its corresponding frequency spectrum in Hz.

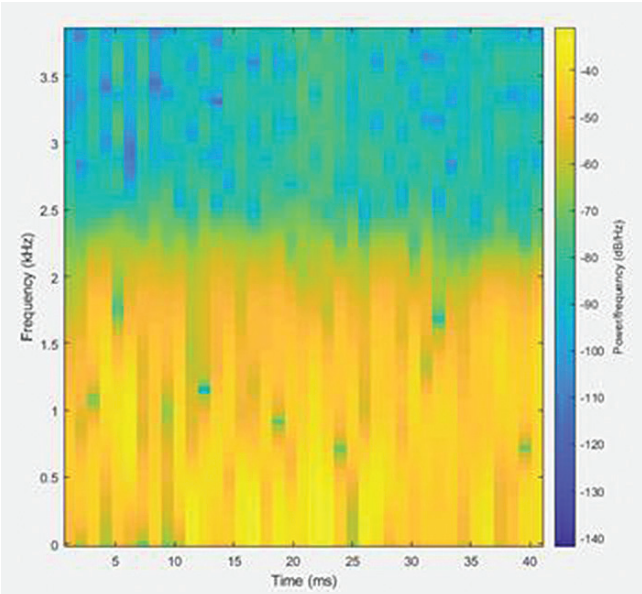

Figure 5. Spectrogram of averaged ABR signal (80dB) of hearing-impaired subject.

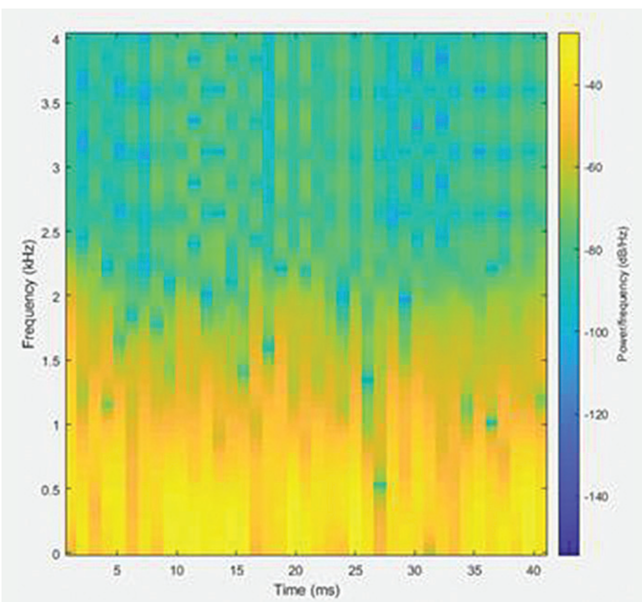

Figure 6. Spectrogram of averaged ABR signal (80dB) of normal hearing subject.

following the Pareto principle, for both normal and hearing-impaired were obtained. The frequency sets can be visualized using the spectrogram shown in Figures 5 and 6. From both figures, most of the high-power frequencies for N and HI subjects were at frequencies below 2.4 kHz.

The bigram method is then used wherein it analyzes the frequency sets or the dominant frequencies from one time segment to the next and is considered a feature.[8] Figures 7 and 8 illustrate the bigram method. The bigram of the

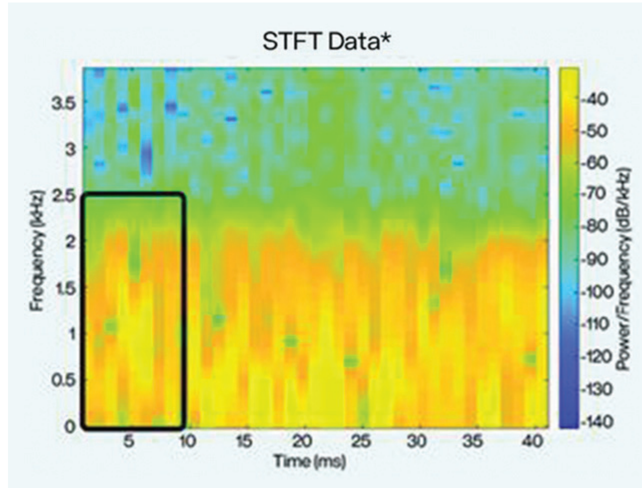

Figure 7. STFT Data.

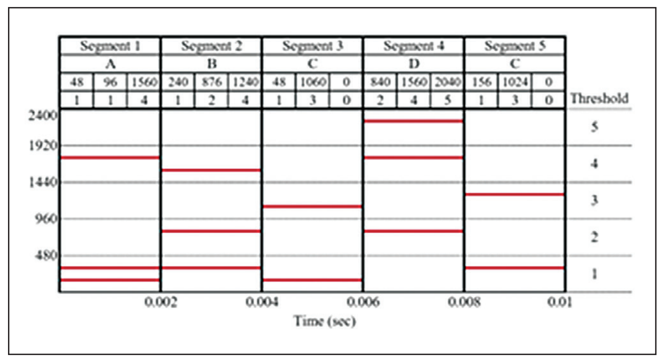

Figure 8. Sample of Bigrams.

marked 10 ms box in Figure 7 is shown in Figure 8. In Figure 8, the left most numbers would be the frequencies used for each threshold. Thus, frequencies up to 480 would correspond to bin 1, frequencies up to 960 would be bin 2, and so on. For segment 1, the three frequency components with the highest contribution during that time duration would be 48, 96, and 1560 Hz. This would correspond to bins 1, 1, and 4, respectively. Similarly, the same is done for segments 2 to 5. Each unique set of frequencies would then be assigned a symbol, in this case, the frequencies 48, 96, and 1560 Hz would be assigned the symbol A and so on. A bigram would be a pair of symbols, so 48-96-1560 Hz followed by 240-876-1240 Hz would have the bigram AB. Thus, the formed bigrams for Figure 8 would be as follows, AB, BC, CD and DC. The bigrams would be tested for uniqueness and checked for occurrence among the files. The top 20% unique bigrams, again using the Pareto principle would be used as classifiers for the files. However, the sensitivity and specificity is increased if bigrams from both normal and hearing-impaired are used, that is the bigrams with the highest difference of occurrence between N and HI are used as classifiers as shown in Figure 9.

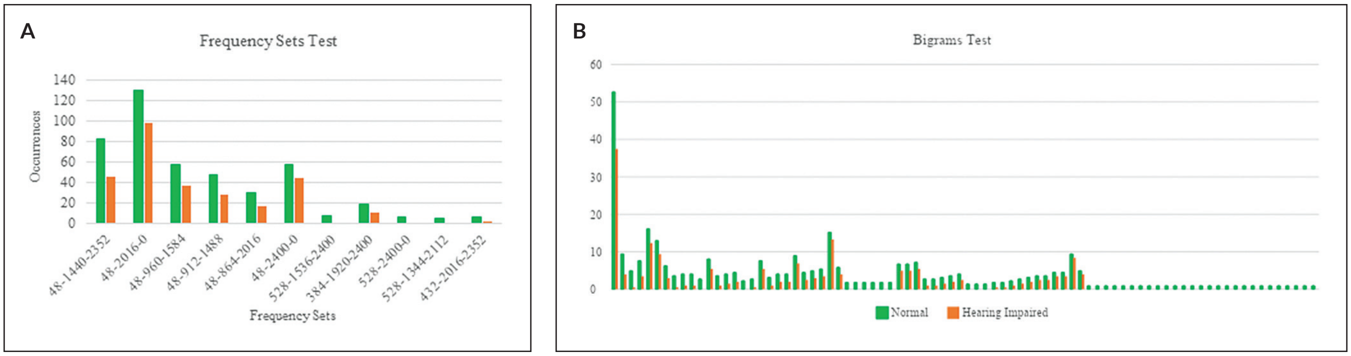

Figure 9. Frequency sets **(A)** and Bigrams **(B)** with highest difference in occurrence between N and HI (frequency is in Hz).

RESULTS

In analyzing the data, the frequency sets and bigram method were used. Instead of analyzing the frequency components of an ensemble ABR signal, the frequency sets of each time segment were obtained. There were 36 unique frequency sets for the N and HI ABR. However, only 11 frequency sets were obtained by getting those with highest difference in occurrence between N and HI as shown in Figure 9. This procedure was replicated for the bigram test. Figure 9 (bottom pane) shows the multiple top bigrams with highest difference in occurrence between N and HI. A total of 377 unique bigrams were extracted from the dataset and Figure 9 which showed eighty-two (82) were used as classifiers.

Five iterations of the Bigram test were performed depending on whether either or both Normal (N) or Hearing-Impaired (HI) classifiers were used, whether or not the bigrams were analyzed per intensity level, and the lower cutoff frequency set for the filter. The results from these iterations are discussed throughout the section.

Table 1 shows the sensitivity (SN) and specificity (SP) using 224 files from normal subjects and 204 files for hearing-impaired. They were obtained from the test conducted with only N classifiers and using a filter with a lower cutoff frequency of 60 Hz. Table 2, on the other hand, shows the SN and SP when both N and HI classifiers were used. It can be observed from the values in the two tables that the second iteration yielded results that are considerably better than the first, which means that using both N and HI classifiers achieves better performance than using only N classifiers.

The third iteration also uses both N and HI features since it was previously determined that doing so yields better performance rates than using only N classifiers. Unlike in the first two, however, the lower cutoff frequency of the filter in this iteration is set to 300 Hz. The results from this test are shown in Table 3. Comparing the values in Tables 2 and 3, it can be seen that the filter allows for significantly better sensitivity and slightly improved specificity if the lower cutoff frequency is set to 300 Hz instead of 60 Hz.

In the fourth and fifth iterations, the unique bigrams for N and HI signals were then classified and analyzed per intensity level. A lower cutoff frequency of 60 Hz was set for the filter in the fourth iteration, the results of which are shown in Table 4. From the table, true positive rates ranging from 60.87 to as high as 100% were achieved. As for the fifth and final iteration, with a cut-off frequency of 300 Hz, the bigram features were able to distinguish HI from 83.33 to 100% of the time based on the values presented in Table 5.

DISCUSSION

The bigram test, an algorithm used in a previous study to classify electrocardiogram signals with normal sinus rhythm and those with atrial fibrillation was also used in this study to classify whether a signal was from a normal or hearing-impaired subject.[8] The bigram method involves pairing up frequency sets of ABR signals from both N and HI patients to produce even more distinct features that may be used to classify the ABR signals. In this study, five iterations of the

Table 1. Bigram Test using only Normal Classifiers (Lower Cutoff – 60 Hz)

Attribute	Percent
Specificity	85.27 (95% CI: 79.94 - 89.64)
Sensitivity	59.61 (95% CI: 52.73 - 66.59)

Table 2. Bigram Analysis Results using both Normal and Hearing-impaired Classifiers (Lower Cutoff – 60 Hz)

Attribute	Percent
Specificity	87.50 (95% CI: 82.44 - 91.53)
Sensitivity	79.31 (95% CI: 73.21 - 84.74)

Table 3. Bigram Analysis Results using both Normal and Hearing-impaired Classifiers (Lower Cutoff – 300 Hz)

Attribute	Percent
Specificity	88.39 (95% CI: 83.46 - 92.28)
Sensitivity	96.55 (95% CI: 93.06 - 98.61)

Table 4. Bigram Analysis Results for Varied Stimulus Intensities using both Normal and Hearing-impaired Classifiers (Lower Cutoff – 60 Hz)

dB	# HI files	# N files	Specificity	Sensitivity
Lower 60	23	98	66.32 (95% CI: 56.07 - 75.56)	60.87 (95% CI: 38.54 - 80.29)
60	12	13	92.31 (95% CI: 63.97 - 99.81)	100.00 (95% CI: 73.54 - 100.00)
65	12	17	94.12 (95% CI: 71.31 - 99.85)	91.67 (95% CI: 61.52 - 99.79)
70	20	13	92.31 (95% CI: 63.97- 99.81)	90.00 (95% CI: 68.30 - 98.77)
75	20	10	90.00 (95% CI: 55.50 - 99.75)	80.00 (95% CI: 56.34 - 94.27)
80	24	13	92.31 (95% CI: 63.97 - 99.81)	87.50 (95% CI: 67.64 - 97.34)
85	20	15	73.33 (95% CI: 44.90 - 92.21)	90.00 (95% CI: 68.30 - 98.77)
90	24	11	90.91 (95% CI: 58.72 - 99.77)	83.33 (95% CI: 62.62 - 95.26)
95	24	18	94.44 (95% CI: 72.71 - 99.86)	83.33 (95% CI: 62.62 - 95.26)
100	24	16	93.75 (95% CI: 69.77 - 99.84)	95.83 (95% CI: 78.88 - 99.89)

Table 5. Bigram Analysis Results for Varied Stimulus Intensities using both Normal and Hearing-impaired Classifiers (Lower Cutoff – 300 Hz)

dB	# HI files	# N files	Specificity	Sensitivity
Lower 60	23	98	95.92 (95% CI: 89.88 - 98.88)	100.00 (95% CI: 85.18 - 100.00)
60	12	13	92.31 (95% CI: 63.97 - 99.81)	83.30 (95% CI: 51.59 - 97.91)
65	12	17	88.23 (95% CI: 63.56 - 98.54)	100.00 (95% CI: 73.54 - 100.00)
70	20	13	92.31 (95% CI: 63.97 - 99.81)	100.00 (95% CI: 83.16 - 100.00)
75	20	10	90.00 (95% CI: 55.50 - 99.75)	100.00 (95% CI: 83.16 - 100.00)
80	24	13	92.31 (95% CI: 63.97 - 99.81)	95.83 (95% CI: 78.88 - 99.89)
85	20	15	93.33 (95% CI: 68.05 - 99.83)	95.00 (95% CI: 75.13 - 99.87)
90	24	11	90.91 (95% CI: 58.72 - 99.77)	100.00 (95% CI: 85.75 - 100.00)
95	24	18	94.44 (95% CI: 72.71 - 99.86)	100.00 (95% CI: 85.75 - 100.00)
100	24	16	93.75 (95% CI: 69.77 - 99.84)	91.67 (95% CI: 73.00 - 98.97)

bigram test were performed. A total of 82 bigrams were used for classifying the signals from the normal and hearing-impaired subjects. However, this may have to be readjusted when a different hardware is used. Raw data from Physionet.org indicates that the noise present is significant. Thus, care should be taken to pre-process the data. One iteration of the bigram test was performed on the different stimulus intensity shown on Tables 4 and 5. This was done to see if the stimulus intensity would affect the test. Eventually, this could be of benefit if the testing time could be reduced by using fewer stimulus intensities.

Three factors that significantly improved the classification performance were identified. First, the overall accuracy was found to be higher when both N and HI classifiers were used than when only the N classifiers were considered. The results have shown a 20% true positive rate improvement. Second, using a Chebyshev filter with a cut-off frequency of 300 Hz demonstrated better rates than using one with a 60-Hz cut-off. Analysis of the results revealed that the percentage of false negatives reduced from around 20% to 3%. Lastly, the results further improved when the data had been partitioned according to their intensity levels. Data divided per intensity level have more accuracy than data analyzed as a whole, and that middle level ranges have a better accuracy than at the extremes. All the true negatives have percentages of 90% and above. All intensities, except 60 Hz, have a false negative rate of less than 10%.

The raw data from PhysioNet were successfully processed to acquire the required ABR signal for analysis. Various signal processing techniques were properly implemented wherein the ABR signal to be analyzed resembles an ABR signal from that of a commercial device. Although the PhysioNet data can still be further processed as there are some parameters that can still be modified to produce new findings.

The entire procedure from the processing of data to the results of frequency analysis were executed in MATLAB. Instead of hours, the whole process was done in 8 - 30 seconds. This test time proves vital to ease up the newborn hearing screening process. Though this time might still vary depending on the capabilities of the HeLe hardware. It is also recommended to use on a hardware with a higher sampling rate to enable finer distinction in the frequencies present in the ABR signals.

Currently, the study is only limited with the usage of data from PhysioNet. For further trials, it would be best to test the algorithm using data of newborn subjects from commercial devices. Moreover, the number of sweeps of a certain data might be a factor in the analysis of frequency. Testing the current algorithm using a lower number of sweeps (600 or less) might result in new findings. Though it demands more time, another possibility is to test the algorithm for each sweep as there might be some features to consider.

With the current study focused on newborn hearing screening, chances are that the frequency analysis algorithm can also be explored for other auditory problems. Furthermore, since frequency is the center feature, the algorithm could be used to model and trace the changes in frequency along the auditory pathway. It would be interesting, for example, to see the frequency components present in the different waveforms (I to V), particularly Wave V, as this might give a clue on the conditions of the pathway. This is the recommendation as well of Paulraj, a closer look at Wave V as this is instrumental in classifying normal and hearing-impaired subjects.[9] This will supplement studies based on the traveling wave delay of our auditory system.

CONCLUSION

This study shows that there is merit in studying the frequency components of the ABR signal. It might lead to new ways of classifying normal or hearing-impaired subjects other than looking at the latencies of the different waveforms. By combining both time and frequency domain in the bigram method, new features can be extracted from the ABR signals.

Statement of Authorship

All authors certified fulfillment of ICMJE authorship criteria.

Author Disclosure

All authors declared no conflicts of interest.

Funding Source

The study was funded by the Philippine-California Advanced Research Institutes (PCARI).

REFERENCES

1. Hele.ph, Hearing for Life [Internet]. 2018 [cited 2018 Oct]. Available from: http://hele.ph/.
2. Republic Act 9288 An Act Promulgating A Comprehensive Policy And A National System For Ensuring Newborn Screening, Vol. 2. 2018 (Phil.).
3. Starr A, Achor J. Auditory brain stem responses in neurological disease. Arch Neurol. 1975 Nov;32(11):761-8. doi: 10.1001/archneur.1975.00490530083009.
4. Hecox K, Galambos R. Brain stem auditory evoked responses in human infants and adults. Arch Otolaryngol. 1974 Jan;99(1):30-3. doi: 10.1001/archotol.1974.00780030034006.
5. Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, et al. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation. 2000 Jun;101(23):e215-20. doi: 10.1161/01.cir.101.23.e215.
6. Silva I, Epstein M. Estimating loudness growth from tone-burst evoked responses. J Acoust Soc Am. 2010 Jun;127(6):3629-42. doi: 10.1121/1.3397457.
7. Silva I, Epstein M. Objective estimation of loudness growth in hearing-impaired listeners. J Acoust Soc Am. 2012 Jan;131(1):353-62. doi:10.1121/1.3666024.
8. Santos M. Electrocardiogram Analysis Using Constellation Maps and Bigrams. [MS Thesis, Ateneo de Manila University]. 2011.
9. Paulraj MP, Subramaniam K, Bin Yaccob S, Bin Adom AH, Hema CR. Auditory evoked potential response and hearing loss: a review. Open Biomed Eng J. 2015 Feb;9:17-24. doi: 10.2174/1874120701509010017.