

## Algorithm of Acoustic Analysis of Communication Disorders within Moroccan Students

Brahim Sabir<sup>1\*</sup>, Bouzekri Touri<sup>2</sup> and Mohamed Moussetad<sup>1</sup>

<sup>1</sup>Physics Department, University Hassan II Mohammedia, Faculty of Science Ben M'Sik Casablanca, Morocco

<sup>2</sup>Language and Communication Department, University Hassan II Mohammedia, Faculty of Science Ben M'Sik, Casablanca, Morocco

\*Corresponding author: Brahim Sabir, Physics Department, University Hassan II Mohammedia, Faculty of Science Ben M'Sik Casablanca, Morocco, Tel: +212523314635/36; E-mail: [sabir.brahim@hotmail.com](mailto:sabir.brahim@hotmail.com)

Rec date: Jan 8, 2016, Acc date: Jan 18, 2016, Pub date: Jan 25, 2016

Copyright: © 2016 Sabir B, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

### Abstract

**Objective:** Communication disorders negatively affect the academic curriculum for students in higher education. Acoustic analysis is an objective leading tool to describe these disorders; however the amount of the acoustic parameters makes differentiating pathological voices among healthy ones not an easy task. The purpose of the present paper was to present the relevant acoustic parameters that differentiate objectively pathological voices among healthy ones.

**Methods:** Pathological and normal voices samples of /a/, /i/ and /u/ utterances, of 400 students were recorded and analyzed acoustically with PRAAT software, then a feature of acoustic parameters were extracted. A statistical analysis was performed in order to reduce the extracted parameters to main relevant ones in order to build a model that will be the basis for the objective diagnostic.

**Results:** Mean amplitude, jitter local absolute, second bandwidth of the second formant and Noise-to-Harmonic Ratio; are relevant acoustic parameters that characterize pathological voices among healthy ones, for the utterances of vowels /a/, /i/ and /u/ Thresholds of the acoustic parameters of pathological /a/, /i/, and /u/ were calculated. A training model was built and simulated on Matlab, and a comparison between Hidden Markov Model and K-Nearest Neighbors classification methods were done (Hidden Markov Model had a rate of recognition of 95% and K-Nearest Neighbors within the reduced acoustic parameters reached a recognition rate of 97%).

**Conclusion:** Through the identified parameters, we can objectively detect pathological voices among healthy ones for diagnostic purposes. As a future work, the present approach is an attempt toward identifying acoustic parameters that characterize each voice disorder.

**Keywords:** Communication disorders; Acoustic analysis; PRAAT; Classification methods

### Introduction

9.3% of hard science students in Morocco have a serious problem related to voice disorders [1]. Thus, understanding Acoustic features of speech will mainly discriminate objectively normal voices from pathological voices of these voice disorders [2]. The analysis of speech disorders remains essentially clinical, and the instrumental measures are not widespread in clinical practice. The most used are the acoustic and aerodynamic measures [3]. Techniques often used by doctors for diagnosis and symptoms analysis of vocal pathologies, as are invasive endoscopy. However, it is possible to identify pathologies using certain parameters of speech signal. Various conventional techniques [4] have been used to identify pathological voices (Cepstrum, LPC spectrogram) for extracting certain voice parameters (pitch, formants, jitter, Shimmer...). This paper proposes a model to differentiate normal voices from pathological voices.

A dataset is constructed by recording speech utterances of a set of /a/, /i/ and /u/. The speech signal is then analyzed in order to extract the

acoustic parameters such as amplitude, intensity, formant frequencies, bandwidths of formants, jitter, Harmonic to noise ratio...etc.

We aim to examine the abilities of objective acoustic analysis methods to detect pathological voices from students with speech disorders. In order to reach this objective a model will be built to recognize pathological voices among healthy ones.

### Praat software

Written by Paul Boersma and David Weenink at the University of Amsterdam, Praat is a computer program with which you can analyze, synthesize, and manipulate speech. It is available for many different computer operating systems and can be downloaded for free from <http://www.praat.org/> [5,6].

### The amplitude

Is the size of the vibration, and this determines how loud the sound is. For our experiment, the vowels (/a/, /i/, /u/) were recorded with the levels: high, neutral and low. The magnitude of pressure variation is perceived as volume changes.

### Jitter (local, absolute)

The absolute local jitter (in seconds) is the mean absolute (non-negative) difference of consecutive intervals:

$$Jitter (seconds) = \Sigma i = 2N |T_i - T_{i-1}| / (-1)$$

Where  $T_i$  is the duration of the  $i^{th}$  interval and  $N$  is the number of intervals.

(An interval is the time between two consecutive points) [5].

### Bandwidths (B1, B2, B3 and B4)

Bandwidth is a measure of frequency band of a sound, especially a resonance. Bandwidth is determined at the half-power (3 dB down) points of the frequency response curve [5].

### Proposed Method

The voice recordings consist of utterances from pathological and healthy speech, recorded by 150 students (80 females, 70 males) with the subjects' ages ranged from 19 to 23 years old. The database contains phonation of the vowel /a/, /i/ and /u/, with the levels of loudness: neutral, low, high and low\_high\_low (combined high and low). The recorded files are in wav format, and the sampling frequency is down sampled to 50 KHz (Pratt sampling frequency), within a mono-channel. Acoustic analysis was performed with PRAAT Software program. The following 23 parameters were analyzed:

Mean amplitude, energy, mean power, mean pitch (F0), standard deviation, mean F1, mean F2, B1 first bandwidth of F1, B2, B3, B4, all components of jitter (5 components), all components of shimmer (6 components) and mean noise to harmonic ratio (NHR). In the

proposed method we have used a set of acoustic parameters; which are relevant to describe a pathological voice among healthy ones. After the analysis step, the reduced acoustic parameters are (mean amplitude, the bandwidth B2 of the second formant F2, jitter local absolute and NHR) which are relevant for /a/, /i/ and /u/. A vector of reduced acoustic parameters was constructed as an input of the classification methods. A model was built based on obtained data and simulated on matlab, and a comparison between artificial neural network [7], K-Nearest Neighbors, and Hidden Markov Model classification methods was done.

- Input audio of /a/, /i/ and /u/ utterances.
- Feature extraction: all acoustic parameters acquisition  $A_i$  for pathological ( $P_i$ ) and healthy voices ( $H_i$ ).
- Feature reduction: based on the calculated ratio ( $R_i$ ) <pathological, healthy>, define the relevant parameters, in the proposed method: mean amplitude, bandwidth B2, jitter local absolute, NHR and HNR are the relevant parameters.
- $R_i = P_i / H_i$
- Training stage of the vector constructed by relevant acoustic parameters.
- Classification stage: HMM and KNN will be applied in order to have high accuracy of classification of pathological and healthy voices.
- Calculate the degree of the severity of the pathological voice based on the  $R_i$  ratio.

### Results and Discussion

#### Acoustic parameters of /a/, /i/ and /u/ vowels

| Parameter                              | H (/a/) | P (/a/)  | H (/i/) | P (/i/)  | H (/u/) | P (/u/)  |
|--|---------|----------|---------|----------|---------|----------|
| Mean amplitude (Pa)                    | 0,00026 | -0,00012 | 0,00034 | -0,00003 | 0,00019 | -0,00025 |
| Total energy (Pascal <sup>2</sup> sec) | 0,053   | 0,062    | 0,034   | 0,053    | 0,040   | 0,077    |
| energy in air (Joule/m <sup>2</sup> )  | 0,00013 | 0,00016  | 0,00008 | 0,00013  | 0,00010 | 0,00019  |
| (mean pitch F0) Hz                     | 231     | 146      | 273     | 156      | 277     | 148      |
| (minimum pitch Hz)                     | 204     | 137      | 244     | 148      | 261     | 141      |
| (maximum pitch) Hz                     | 249     | 158      | 298     | 168      | 241     | 159      |
| (mean F2) Hz                           | 1249    | 1361     | 2445    | 2188     | 790     | 1379     |
| (B1) first bandwidth in Hz             | 139     | 360      | 113     | 73       | 134     | 190      |
| (B2) Hz                                | 172     | 1621     | 945     | 178      | 220     | 535      |
| (B3) Hz                                | 486     | 785      | 500     | 986      | 514     | 770      |
| (B4) Hz                                | 640     | 971      | 732     | 1178     | 337     | 945      |
| Jitter (local): %                      | 0,32    | 0,80     | 0,26    | 0,52     | 0,32    | 0,46     |
| Jitter (rap): %                        | 0,14    | 0,42     | 0,10    | 0,23     | 0,16    | 0,19     |
| Jitter (ppq5): %                       | 0,15    | 0,37     | 0,12    | 0,25     | 0,17    | 0,23     |
| Jitter (ddp): %                        | 0,41    | 1,25     | 0,31    | 0,70     | 0,49    | 0,58     |

|                                  |         |         |         |         |         |         |
|----------------------------------|---------|---------|---------|---------|---------|---------|
| Jitter (local, absolute):seconds | 0,00001 | 0,00007 | 0,00001 | 0,00005 | 0,00001 | 0,00004 |
| Shimmer (local): %               | 2,1     | 4,0     | 1,8     | 2,1     | 2,2     | 2,8     |
| Shimmer (apq3): %                | 1,0     | 2,1     | 0,6     | 0,9     | 1,1     | 1,3     |
| Shimmer (apq5)%                  | 1,2     | 2,6     | 1,0     | 1,4     | 1,4     | 1,9     |
| Shimmer(apq11)%                  | 1,7     | 3,2     | 2,1     | 2,1     | 2,0     | 2,6     |
| Shimmer (dda): %                 | 3,0     | 6,4     | 1,9     | 2,6     | 3,3     | 4,0     |
| Shimmer (local)in dB):           | 0,19    | 0,35    | 0,17    | 0,18    | 0,22    | 0,25    |
| Mean noise-to-harmonics ratio:   | 0,006   | 0,040   | 0,254   | 0,009   | 0,003   | 0,007   |

**Table 1:** Acoustic parameters of /a/, /i/, and /u/ vowels for pathological and healthy voices (results extracted from PRAAT software) ( H: Healthy, P:Pathological).

### Ratios “pathological”, “healthy” of the acoustic parameters

Based on the ratios, main parameters that characterize pathological voices among healthy ones were extracted.

| Parameter                        | Ratio((Pathological, Healthy) /a/) | Ratio((Pathological, Healthy) /i/) | Ratio((Pathological, Healthy) /u/) |
|----------------------------------|------------------------------------|------------------------------------|------------------------------------|
| Mean amplitude (pascal)          | -0,5                               | -0,1                               | -1,3                               |
| (B2)Hz                           | 9,4                                | 0,2                                | 2,4                                |
| Jitter (local, absolute):seconds | 7,0                                | 5,0                                | 4,0                                |
| Mean NHR                         | 6,7                                | 0,04                               | 2,3                                |
| HNR(dB)                          | 0,68                               | 1                                  | 0,89                               |

**Table 2:** Ratio pathological, healthy for relevant parameters.

Acoustic measures of: Mean amplitude, second bandwidth of the second formant, jitter (local absolute), mean NHR, and HNR are more relevant acoustic parameters that characterize a pathological voice among healthy one.

### Results of amplitude

We notice, that for all (/a/, /i/, /u/) pathological utterances: Mean amplitude <0. The loudness low, high low has specific result compared to high, normal and neutral loudness [8].

| Amplitudes of /a/, /i/and /u/ utterances |             |                   |            |                  |                |                      |                     |                           |
|--|-------------|-------------------|------------|------------------|----------------|----------------------|---------------------|---------------------------|
|  | high_normal | high_pathological | low_normal | low_pathological | neutral_normal | neutral_pathological | low_high_low_normal | low_high_low_pathological |
| /a/ in 10 <sup>-5</sup> Pa               | 25          | -29               | 20         | -10              | 20             | -20                  | 30                  | 10                        |
| /i/ in 10 <sup>-5</sup> Pa               | 16          | 34                | 10         | -10              | 20             | -10                  | 100                 | -30                       |
| /u/ in 10 <sup>-5</sup> Pa               | 37          | -26               | 10         | -10              | 20             | -30                  | 10                  | -30                       |

**Table 3:** Amplitudes of /a/, /i/and /u/ utterances.

### Noise harmonic ratio

As shown in the Figure 1, for /a/ utterance, the percent of noise present in the pathological signal compared with healthy signal is relevant in high loudness.

And this difference will increase if we do not take into account low high low loudness.

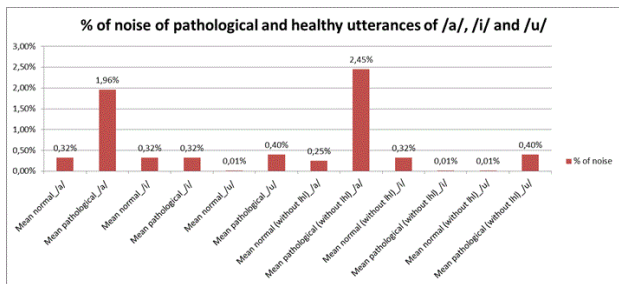


Figure 1: % of noise of pathological and healthy utterances of /a/, /i/ and /u/.

### Results of NHR

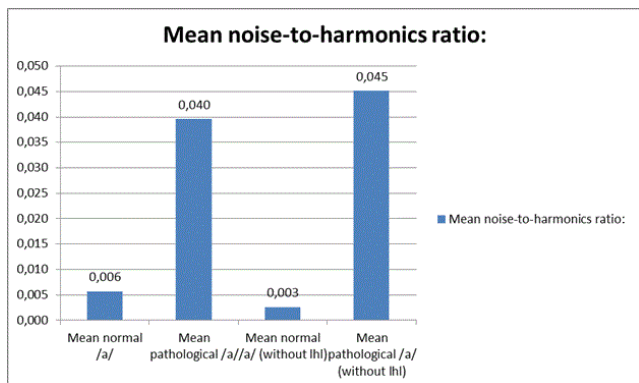


Figure 2: Mean NHR of pathological and healthy /a/.

We notice the effect of low, high, low loudness on the result. However, the difference between pathological and healthy utterance is clear.

### Results of Bandwidths

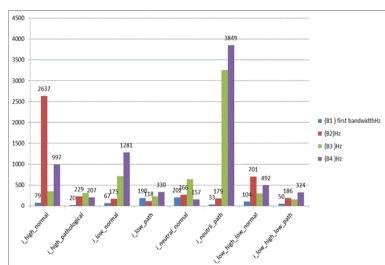


Figure 3: Bandwidths of four first formants of /i/.

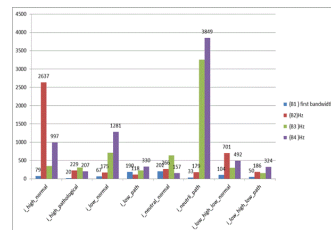


Figure 4: Bandwidths of four first formants of /a/.

For /i/ utterance, the most significant bandwidth is the first bandwidth B<sub>1</sub> for high loudness. Besides, B<sub>4</sub>, B<sub>3</sub> and B<sub>2</sub> respectively for low, neutral and low high low loudness.

For the /a/ utterance B<sub>3</sub> is most significant for high and low high low loudness. B<sub>4</sub> and B<sub>2</sub> are most relevant for normal and neutral loudness (Figure 4).

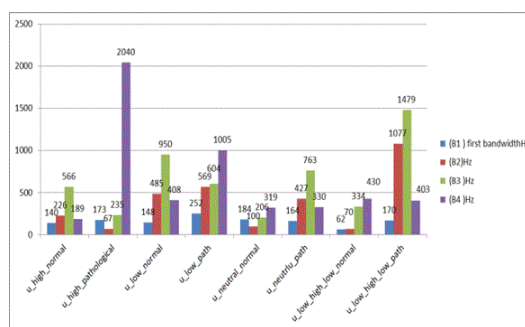


Figure 5: Bandwidths of four first formants of /u/.

We notice that B<sub>4</sub>, B<sub>3</sub>, B<sub>3</sub> and B<sub>2</sub> are most significant respectively for high, low, neutral and low high low loudness. Based on studied bandwidths, the second bandwidth is considered relevant for all utterances (Figure 5).

### Jitter local absolute

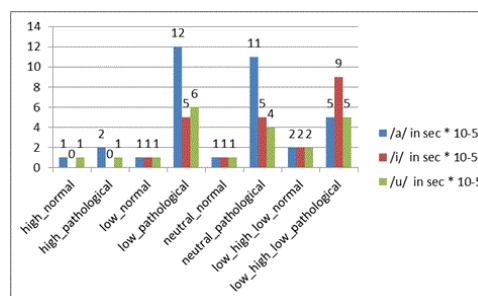


Figure 6: Jitter local absolute for /a/ , /i/ and /u/ utterances of pathological and healthy voices.

Significant values are for /a/ utterance related to low and neutral loudness. Also we notice for /i/ and /u/ utterances that low loudness, neutral and low high low loudness have significant values. However, for high loudness, we notice that there is no significant variation between pathological and healthy values (Figure 6).

## Classification

### Classification with KNN

The 4 relevant acoustic parameters were extracted for the input audio signal, and a classification was performed using K-Nearest neighbors' algorithm (Figure 7):

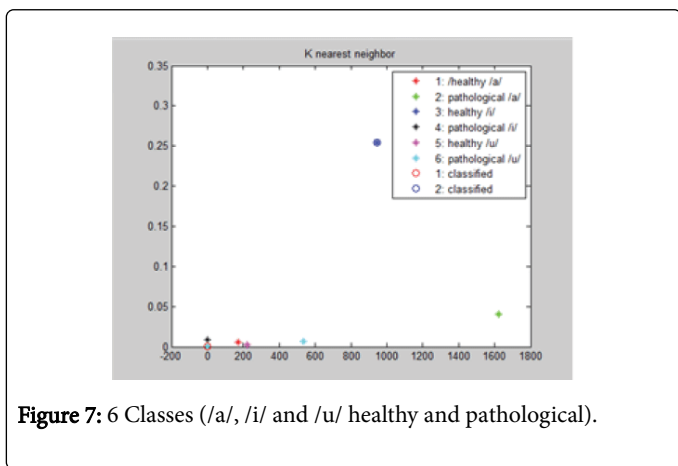


Figure 7: 6 Classes (/a/, /i/ and /u/ healthy and pathological).

A recognition rate of: 97% was obtained.

### Classification with HMM

The audio input will be the set of observation symbol sequences which will be tested against trained HMM classifier. The probability of an observation sequence was calculated using forward algorithm.

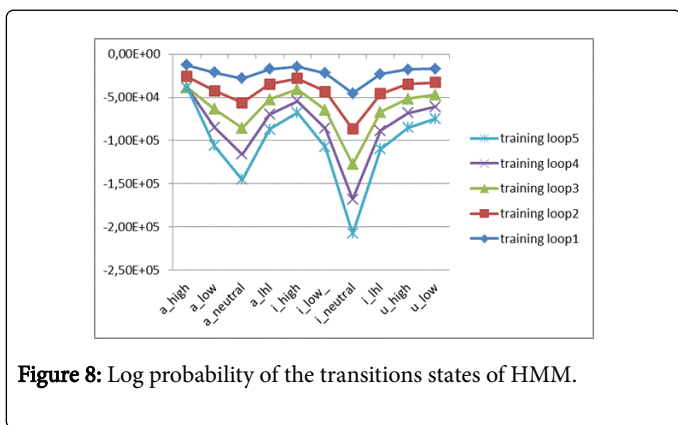


Figure 8: Log probability of the transitions states of HMM.

Obtained recognition rate: 95%.

A set of 4 states were used. Low\_high\_Low vowel affects negatively the HMM recognition (picks on "low\_high\_low utterances": a\_lhl and i\_lhl) (Figure 8) [9].

## Conclusion and Future Work

A set of acoustic parameters (mean amplitude, jitter local absolute, NHR, and the bandwidth of the second formant) are sufficient to characterize objectively pathological voices. As a future work, to extend the proposed method to all phonemes in order to build a database of normal/pathological utterances, which can be an initial point to define objectively the symptoms of each voice disorder.

## References

1. Touri B, Talbi M, Soubhi F, Lima L (2012) Psycho-communication disorders among Moroccan university hard science major students, *Procedia - Social and Behavioral Sciences* 116: 5058-5062.
2. Leong K, Hawkshaw MJ, Dentchev D, Gupta R, Lurie D, et al. (2013) Reliability of objective voice measures of normal speaking voices. *J Voice* 27: 170-176.
3. Zhang Y, Jiang JJ (2008) Acoustic analyses of sustained and running voices from patients with laryngeal pathologies. *J Voice* 22: 1-9.
4. Eadie T, Philip C (2004) Classification of Dysphonic Voice: Acoustic and Auditory-Perceptual Measures, *Journal of Voice* 19: 1-14.
5. Boersma P, Weenick D (2006) Praat Manual. Amsterdam: University of Amsterdam, Phonetic Sciences Department.
6. Boersma P, Weenink D (2009) Praat: doing phonetics by computer (Version 5.1.17) [Computer program].
7. Linder R, Albers A, Hess M, Pöppel S, Schönweiler R (2006) Artificial Neural Network-based Classification to Screen for Dysphonia Using Psychoacoustic Scaling of Acoustic Voice Features, *Journal of Voice* 22: 155-163.
8. GeoGebra software: www.Geogebra.org.
9. Gillespie A, Dastolfo C, Magid N, Gartner-Schmidt J (2014), Acoustic Analysis of Four Common Voice Diagnoses: Moving Toward Disorder-Specific Assessment, *Journal of Voice* 28: 582-588.