HMM/GMM Classification for Articulation Disorder Correction among Algerian Children

Abed Ahcène[1, 2] and Guerti Mhania[1]
[1]Laboratoire Signal et Communications, Ecole Nationale Polytechnique, Algeria
[2]Scientific and Technical Research Center for the Development of the Arabic Language, Algeria

Abstract: *In this paper, we propose an automatic classification for Arabic phonemic substitution using a Hidden Markov Model/Gaussian Mixture Model (HMM/GMM) systems. The main objective is to help Algerian children in the correction of articulation problems. Five cases are analyzed in the experiments, 20 Arabic words are recorded by a 20 Algerian children, with age range between 4 and 6 years old. Signals are recorded and stored as wave format with 16kHz as sampling rate, 12 Mel Frequency Cepstral Coefficients (MFCC), with their first and second derivates, respectively Δ and ΔΔ are extracted from each signal and used to the training and recognition phases. The proposed system achieved its best accuracy recognition 85.73%, with 5-stats HMM when the output function is modelled by a GMM with 8Gaussian components.*

Keywords: *Phonemic substitution, HMM/GMM, algerian dialectal, speech recognition, MFCC.*

Received August 29, 2014; accepted October 26, 2014; published online August 22, 2015

1. Introduction

The aim of this paper is to assist Algerian children for correcting their pronunciation and to provide some of the necessary resources. Also, another important point is the correction of functional articulation errors, especially the Arabic phonemic substitution disorder. Three years old, the child should be able to speak clearly. If not, this is one sign of speech pathology. Parents should seek the assistance of speech-language pathologist; who is trained in the diagnosis and remediation of a board rang of speech language disorder among adult and children.

The pathologist follows an evaluation procedure, included perceptual speech evaluation, and guided by a theoretical model. This can take a few science of remediation with a great effort by children and their parents. However, if the remediation procedure takes a long time and the number of children to be re-educated gets greater, the pathologist can not follow the perceptual evaluation. Moreover, in Algeria, there is a huge absence of pathologists and the majority of them settle in the big cities. To take an appointment, parents must make long distances and spend a lot of money.

Speech is the most used mean of interpersonal communication, performing many functions in relation with children. Firstly, it enables them to live in the community and to share their felling and needs in different situations. A great importance has the communication function of speech, which makes possible to transfer content from one person to another and cognitive function of integration, design and conceptualization of thinking. The oral language development is an active process, dependent on the child neuro-cognitive natural ability and his human environment meeting. However, there are several categories of factors which can affect oral languagedevelopment and can cause various dysfunctions.

In this paper, we evaluate the performance of phonemic substitution disorder detection system for Algerian Arabic speaking children. Firstly, we analyse different cases into two main aspects: The phoneme's place of articulation and its voicing property. For this, we use PRAAT software [14]; this latest allows a sound's analysis for different phonemes. The automatic classification procedure is based in Hidden Markov Model/Gaussian Mixture Model (HMM/GMM) Automatic Speech Recognition (ASR) system [16]. The MFCC are used for feature extraction [7].

2. Arabic Language Background

Arabic is the largest Semitic language still in existence and one of the six official languages of the United Nations. It is the official one in Algeria. We can find two major forms of Arabic: Modern Standard Arabic (MSA) and Dialectal Arabic (DA). When MSA is used in writing, news broadcast and movies subtitling, DA is the daily spoken language of people [8].

DA varies from one country to another and sometimes more than one dialect can be found within a country. MSA consists of 40 phonemes where 28 consonants, 6 vowels (three short and three long vowels) and 6 vocalic variant in emphatic context. It is characterized by the presence of emphatic and pharyngeal phonemes with a total of five pharyngeals, three uvular pharyngeal's and four emphatics Table 1 [3].

Six types of syllables are allowed in MSA, which are: [CV], [CV:], [CVC], [CV:C], [CVVC] and [CVCC], this latest appears in some cases as a link between two successive words, where [V] denotes short vowel, [V:] a long vowel and [C] a consonant. All these syllables must start with a consonant, and must contain at least one vowel. In addition, while vowels do not occur word-initially, they occur word-finally or between two consonants [5].

Table 1. Arabic consonants articulation with their International Alphabet Phonetic (IAP) transcription.

Arabic	IAP	Mode and Manner of Articulation
ء	[ʔ]	Voiceless Glottal Plosive
ب	[b]	Voiced Bilabial Plosive
ت	[t]	Voiceless Alveodental Plosive
ث	[θ]	Voiceless Interdental Fricative
ج	[ʒ]	Voiced Palatal Fricative
ح	[ħ]	Voiceless Pharyngal Fricative
خ	[x]	Voiceless Uvular Pharyngeal Fricative
د	[d]	Voiced Alveodental Plosive
ذ	[ð]	Voiced Interdental Fricative
ر	[r]	Voiced Alveolar Liquid
ز	[z]	Voiced Alveodental Fricative
س	[s]	Voiceless Alveodental Fricative
ش	[ʃ]	Voiced Palatal Fricative
ص	[ṣ]	Voiceless Alveodental Fricative
ض	[ḍ]	Voiced Alveodental Plosive
ط	[ṭ]	Voiceless Alveodental Plosive
ظ	[ð̣]	Voiced Interdental Fricative
ع	[ʕ]	Voiced Pharyngeal Fricative
غ	[ɣ)	Voiced Uvular Fricative
ف	[f]	Voiceless Labiodental Fricative
ق	[q]	Voiceless Uvular Plosive
ك	[k]	Voiceless Velar Plosive
ل	[l]	Voiced Lateral
م	[m]	Voiced Bilabial Nasal
ن	[n]	Voiced Alveodental Nasal
ه	[h]	Voiceless Glottal Fricative
و	[w]	Voiced Bilabiale semivowel
ي	[y]	Voiced Palatal semivowel

3. Arabic Speech Disorder

There are a few of researches in the area of speech disorder therapy. An automatic voice disorder classification system was introduced using Arabic digit speech [12]. The development of a novel feature extraction method was proposed for ASR that incorporates distributions of voiced and unvoiced parts and voice onset and offset characteristics in a time-frequency domain to detect voice pathology [13]. The training data influence in ASR was studied for Arabic phonemic substitution problem [1]. A new method for voice disorders classification based on multilayer Neural Network (NN) was developed. The processing algorithm is based on a hybrid technique which uses the wavelets energy coefficients as input of the multilayer NN [15].

A lot of causes that determine speech disorder may act during speech evaluation period. It can fall into many categories: From organic to functional, neurological or psycho-social causes. It can be divided into: Problems of articulation disorders of rhythm of speech, voice or development of language disorders [6]. Articulation disorders are the most widely known in children. It includes distortion, omission, addition and phonemic substitution disorders. In this latest, the child substitutes a phoneme by another during pronunciation. It is produced when the contact level between articulators forward or backward to another place of articulation.

4. Methods and Materials

The proposed system is an HMM/GMM application. Five cases of Arabic phonemic substitution are studied as shown in Table 2 (Appendix 1). The choice of these sounds is proved by a subjective test performed during our work; these cases are the most common pathologies in this rang of children. Each sound has its specific place and manner of articulation [2].

4.1. Feature Extraction

Feature extraction involves analysis of speech signal to reduce the data size before classification. The most widely used speech features are MFCC. The Mel-Cepstral makes use of the auditory system principle; it has high discriminating power at lower frequencies compared to higher frequencies. The frequency bands are equally spaced on the Mel-scale. It is linear below 1khz and logarithmic above 1khz [11].

$$f_{mel} = 2595 log\left(1+\frac{f}{700}\right) \quad (1)$$

Where f is the actual frequency in Hz.

At the beginning, the speech signals are pre-emphasized with a first order high pass filter. The z-transform of the filter is shown in Equation 2:

$$H(z) = 1 - az^{-1} \quad (2)$$

With $0.9 \leq a \leq 1$.

The resulted signal is divided into short frame segment using 32ms Hamming window with a 16ms overlap between two adjacent frames. The main concept of using this window is to minimize the spectral distortion and the signal discontinuities.

$$w(n) = 0.54 - 0.46 cos\left(\frac{2\pi n}{N-1}\right) \quad (3)$$

Where N is the window size and $0 \leq n \leq N\text{-}1$.

The windowed signal is analysed by a Fast Fourier Transform (FFT). The spectrum of each frame is filtered by a set of filters and the power of each band is calculated. The log Mel spectral is converted into time by applying the Discrete Cosine Transform (DCT). Finally, we get:

$$C(i)=\sqrt{\frac{2}{K}}\sum_{j=1}^{K}S_j\cos\left[(j-0.5)\frac{i\pi}{K}\right] \quad (4)$$

Where $i \leq L$, S_j the log filter bank amplitudes, K the number of filter bank channels, L the number of Cepstral coefficients.

The Δ and the $\Delta\Delta$ coefficients are also used. Both can be approximated by [4]:

$$\Delta C_1(i)=0.375\sum_{j=-K}^{K} j(\Delta C_{l-j}(i)) \quad (5)$$

$$\Delta\Delta C_l(i)=\left[\Delta C_{l+1}(i)-\Delta C_{l-1}(i)\right] \quad (6)$$

Where k and l are frames indexes, i is the MFCC component.

4.2. Hidden Markov Model

A HMM can be viewed as a double-embedded stochastic process with an underlying one that is hidden (not directly observable). An HMM is usually represented as $\lambda=\{\pi, A, B\}$. It is characterized by the following parameters [16]:

- N: Is the number of hidden states q_i in the model.
- π: Is the initial state distribution.
- $A=\{a_{ij}\}$ is the state transition probability matrix. where a_{ij} is the probability of taking a transition from state i to state j.
- B: Is the emission probability distribution.

In ASR system the basic problem is expressed as follows: Given the observation sequence as the set of speech parameters $X=\{x_1, x_2, ..., x_T\}$ and a set of state sequence, $S=\{s_1, s_2, ..., s_N\}$, which the word $\tilde{\lambda}$ that maximize the probability $P(\lambda|X)$.

$$\tilde{\lambda}=\underset{\lambda}{argmax}\,P(\lambda/X) \quad (7)$$

Equation 7 can be rewritten by applying bays assumption as:

$$\tilde{\lambda}=\underset{\lambda}{argmax}\,\frac{P(X/\lambda)P(\lambda)}{P(X)} \quad (8)$$

$$\tilde{\lambda}=\underset{\lambda}{argmax}\,P(X/\lambda)P(\lambda) \quad (9)$$

Where $P(X/\lambda)$ represent the posterior probability of observation emitting for the word λ, $P(\lambda)$ represents the prior probability of models.

Assuming statistical independence of observations, it follows that:

$$P(X|\lambda)=\sum_S P(X, S|\lambda) \quad (10)$$

$$P(X/\lambda)=\sum_S \pi_{S_1}b_{S_1}(x_1)\prod_{t=2}^{T}a_{s_{t-1}s_t}b_{S_t}(x_t) \quad (11)$$

To find the best state sequence we use a formal technique called forward-backward algorithm. The forward variable $a_t(i)$ is expressed as:

$$\alpha_t(i)=P(x_1, x_2, \cdots, x_t, q_t=S_i/\lambda) \quad (12)$$

In the same manner we can define the backward variable $\beta_t(i)$, expressed as:

$$\beta_t(i)=P(x_{t+1}, x_{t+2}, \cdots, x_T/q_t=S_i, \lambda) \quad (13)$$

4.3. Gaussian Mixture Model

In speech recognition task the observation does not come from a finite set, but from a continuous space. For continuous output probability density functions, the major candidate is multivariate Gaussian mixture density functions, known as GMM [8]. This is motivated that they can approximate any continuous probability density function. It may express as follows:

$$p(x/\lambda)=\sum_{m=1}^{M}w_{jm}\frac{exp(\frac{-1}{2}(x-\mu_{jm})^t\Sigma_{jm}^{-1}(x-\mu_{jm}))}{(2\pi)^{\frac{D}{2}}\left|\Sigma_{jm}\right|^{\frac{1}{2}}} \quad (14)$$

Where x is a d-dimensional vector; μ_{jm} the mean vector, Σ_{jm} the covariance matrix and w_{jm} is the mixture weight satisfy the stochastic constraint that:

$$\sum_{m=1}^{M}w_{jm}=1 \quad (15)$$

The parameters to be estimated are the mean vector, covariance matrix and the weight vector. These can be calculated by using Baum-welch algorithm.

$$\bar{\mu}_{jm}=\frac{\sum_{t=1}^{T}\xi_t(j,m)x_t}{\sum_{t=1}^{T}\xi_t(j,m)} \quad (16)$$

$$\bar{\Sigma}_{jm}=\frac{\sum_{t=1}^{T}\xi_t(j,m)(x_t-\mu_{jm})(x_t-\mu_{jm})^t}{\sum_{t=1}^{T}\xi_t(j,m)} \quad (17)$$

$$\bar{w}_{jm}=\frac{\sum_{t=1}^{T}\xi_t(j,m)}{\sum_{k=1}^{M}\sum_{t=1}^{T}\xi_t(j,k)} \quad (18)$$

$$\xi_t(j,m)=\frac{\sum i\,\alpha_{t-1}(i)a_{ij}w_{jm}b_{jm}(x_t)\beta_t(j)}{\sum_{i=1}^{N}\alpha_T(i)} \quad (19)$$

5. Experiments and Results

The system is an HMM/GMM speech recognition based. We used it to classify the sound as healthy or pathology. The different steps may be represented by the bloc diagram below as shown in Figure 1.

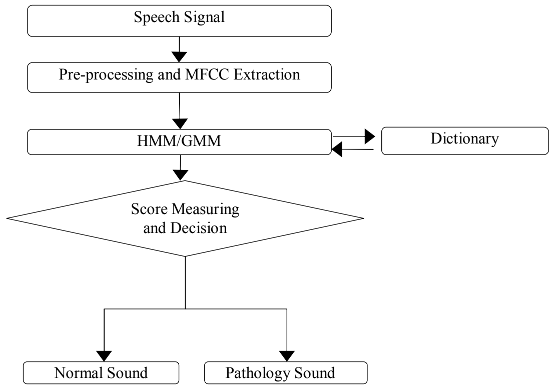


Figure 1. Phonemic substitution classification system.

Two sets of data are used: One for optimal parameters estimation and other for performances evaluations and system validation. The first set is extracted from TIMIT corpora [10]. It was recorded at 16kHz rate with 16bits sample resolution. TIMIT contains a total of 6300 sentences, spoken by 630 speakers from 8 major dialect regions of the United States.

The second corpus is created from a set of 20 Algerian children gradate in-primary school. Each child utters a number of words porters the studied phonemes. A 70% from signals are used in training phase, while the rest of 30% are reserved for testing phase. All children are Arabic native with age range from 4 to 6 years with and without voicing disorder. Test data is recorded in different sessions at primary schools in Algeria. The signals are wave format encoded in 16bits where sampling rate is 16kHz.

5.1. Optimal Parameters Estimation

The proposed system is isolated word recognition so we extract two words [mi] and [ʃi], respectively from SA1 and SA2 utterances. We use data from two regions one for training phase and the other for the recognition phase. Each word is modeled by an HMM with 3, 4, 5 and 6 states. The state transition is left-to-right. Observation probability density functions are modelled using GMM. The number of mixtures of each state is varied to 2, 4, 8 and 16.

We measure the recognition accuracy (%) versus the number of mixture used to model the output density function for different number of states Figure 2. It shows that the recognition accuracy achieved its best value with 5-states HMM when the output is modeled by a GMM with 8 Gaussian components. This best recognition accuracy achieves 97.85 %.

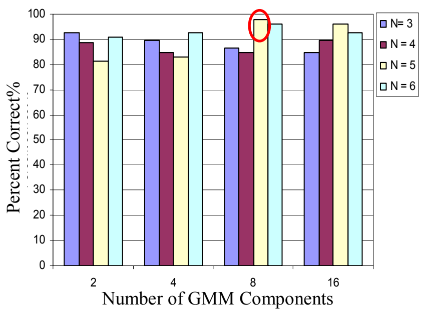


Figure 2. Optimal parameters estimation.

5.2. Phonemic Substitution Classification

In order to improve the use of this system for our application, the phonemic substitution of pathological phonemes is analyzed at different situation. Each phoneme is modeled by 3, 4, 5 and 6-states HMM with GMM of 8 Gaussian components. For each phoneme, we measure the recognition accuracy (%) versus states number as shown in Figure 3.

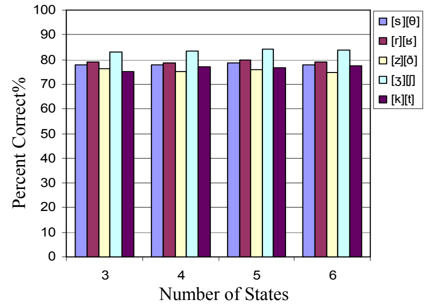


Figure 3. Recognition accuracy versus number of states.

For all phonemes, this later is greater than 74.82% and less than 85.58%. Generally, 5-states HMM achieved a little better recognition accuracy when compared to the others. Moreover, the phonemic substitution classification of [ʒ] by [ʃ] proved the best performance of this system with 85.58%. It is found that 5-states HMM obtained the best phoneme recognition rate of 78.64% and 79.86% for the sounds [s] and [r], respectively. This rate is 76.15% for [z] and 77.35% for [k] using 3 and 6-states HMM, respectively.

For another experiment, we study the number of Gaussian components influence at the system performances. Each phoneme is modelled by 5-states HMM with 2, 4, 8 and 16 GMM. The recognition accuracy (%) versus number of Gaussian components is measured for each phoneme as shown in Figure 4.

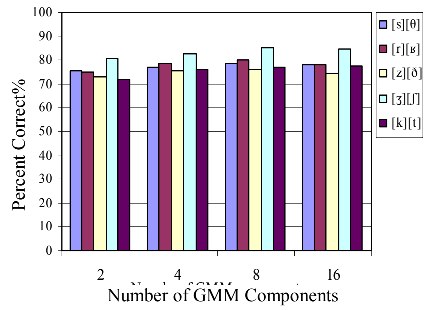


Figure 4. Recognition accuracy versus number of GMM components.

Generally, the system achieves his best accuracy with 8 GMM. It obtained the phoneme recognition rate of 85.36% for the sound [ʒ]. Moreover, with 8 GMM this rate is 78.77%, 80.02%, 75.81% and 77.19%, respectively, for the phonemes [s], [r], [z] and [k].

5.3. Comparative Analysis

In order to compare these results, the proposed system has been trained with other ASR systems: HMM, GMM and HMM/GMM isolated word recognition system. The performance of these three systems is obtained by the recognition accuracy as shown in Table 3. These results improve that HMM/GMM system achieves a better recognition accuracy compared to the other two systems.

Table 3. HMM, GMM and HMM/GMM recognition rate %.

	Configuration N(States):G(Components)			
	5 : 2	**5 : 4**	**5 : 8**	**5 : 16**
GMM	71,06 %	73,36 %	75,56 %	76,92 %
HMM/GMM	76,20 %	78,81 %	80,17 %	79,69 %
	3 : 8	4 : 8	5 : 8	6 : 8
HMM	69,74 %	70,68 %	72,3 %	73,01 %
HMM/GMM	**79,38 %**	**79,62 %**	**80,17 %**	**79,68 %**

5.4. Performances Evaluation

For improving and simplifying its manipulation, the proposed system is built as a graphical computer application. In development process C++ builder was used. Created objects can be used by child. These objects are presented by pictures and audio signals. It may work with a simple computer and a microphone; this latest captures the produced speech signals that respond to changes in air pressure.

The application is tested with two 5-years-old children with phonemic substitution disorder. They got the control of the proposed application with 20 minutes for each therapy session.

First child was a boy who often substitutes the phoneme [ʒ] by [ʃ]. The level of pronunciation correction completed by the boy is measured by the number of phonemes correctly classified versus the total trials. He got his best pronunciation after 3 months later as shown in Figure 5.

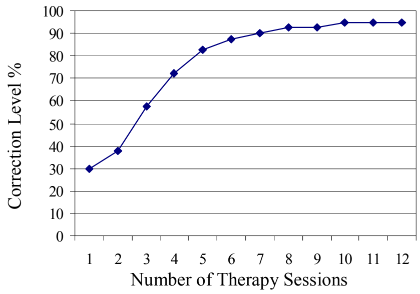


Figure 5. Pronunciation correction completed by first boy [ʒ].

The second trial was performed with another 5-year-old boy. He often substitutes the phoneme [r] by [ʁ]. After six therapy sessions, he demonstrates a better development in their articulation correction as shown in Figure 6.

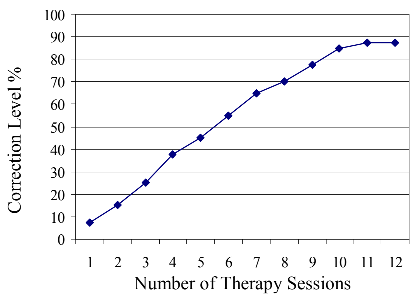


Figure 6. Pronunciation correction completed by second boy [r].

6. Discussion

The obtained results proved the possibility to use this proposed system at automatic classification for Arabic phonemic substitution. The system achieves its best recognition accuracy 85.73%, with a 5 states HMM and 8 GMM to model each state. The palatal affricate [ʒ], which is often substituted by the palatal fricative [ʃ], is classified with a better recognition accuracy compared to the other phonemic substitutions. This may be due to the difference between these phonemes in the voicing propriety.

Moreover, the recognition accuracies of the alveolar [s] and [z] are, respectively, 78.77% and 75.81%. This low rate may be expressed by the articulation places of these different phonemes which are very close with its substituted sounds [θ] and [ð]. For the velar [k], the recognition accuracy is 77.39%, there is a similarity between the two sounds, [k] and [t], in the manner of articulation, the two are plosive. Finally the system achieved a little better recognition accuracy 80.02% for the alveolar [r], this can be explained by the difference between the two sounds [r] and [ʁ], in several levels.

During this work we remarked that child can pronounce the phoneme in initial position better than a final position. Moreover, the phoneme pronunciation can be varied versus the vowel that follows this phoneme, [sa], [su] and [si]. Finally, the choice of words for each phoneme added another problem for this system, this can be seen for the word [ʒima:l] /camels/ and the word [muharriʒ] /clown/. For children, the first word is very easy in pronunciation compared to the second word.

7. Conclusions

An automatic classification for Arabic phonemic substitution system is proposed. Five cases of this problem are analysed: Phonemic substitution of [s], [r], [z], [ʒ] and [k] by [θ], [ʁ], [ð], [ʃ] and [t], respectively. The proposed method is an HMM/GMM based. 12 MFCC are extracted from each signal with their first and second derivate. The system achieves high recognition accuracy, which is 85.73%. An HMM/GMM system with 5 states and 8 components Gaussian is sufficient in our application.

Generally, the proposed system has been designed to illustrate the selection of intervention targets as well as specific therapy activities and materials. If the target phoneme is selected, child can use this system with a great number of repetitions to teach the phonetic placement of different sounds. To simplify the use of this system we built a graphical interface with a simple activities and pictures, its performances can be improved by using another type of coefficients including shimmer and jitter features in addition to the MFCC. The shimmer is expressed as the variability of the peak-to-peak amplitude in decibels and the jitter is the cycle-to-cycle variation of fundamental frequency.

References

[1] Abed A. and Guerti M., "Apllication Des HMM A La Substitution Phonémique Dans l'Arabe Parlé," *Journées d'Etudes Algéro-Françaises de Doctorants en Signal-Image and Applications*, pp. 18-23, Algeria, 2012.

[2] Abed A. and Guerti M., "Errors Classification of Phonemic Substitution in Arabic Speech," *in Proceedings of the International Congress on Telecommunication and Application'14*, Bejaia, Algeria, pp. 23-24, 2014.

[3] Alotaibi Y. and Selouani S., "Investigating the Adaptation of Arabic Speech Recognition Systems to Foreign Accented Speakers," *in Proceedings of the 10th International Conference on Information Science, Signal Processing and their Applications*, Kuala Lumpur, pp. 646-649, 2010.

[4] Amrouche A., Debyeche M., Ahmed A., Rouvaen J., and Yagoub M., "An Efficient Speech Recognition Systemin Adverse Conditions Using the Nonparametric Regression," *Engineering Applications of Artificial Intelligence*, vol. 23, no. 1, pp. 85-94, 2010.

[5] Azmi M. and Tolba H., "Syllable-Based Automatic Arabic Speech Recognition in Different Condition of Noise," *in Proceedings of the 9th International Conference on Signal Processing*, Beijing, pp. 601-604, 2008.

[6] Danubianu M., Pentiuc S., and Socaciu T., "Towards the Optimized Personalized Therapy of Speech Disorders by Data Mining Techniques," *in Proceedings of the 4th International Multi-Conference on Computing in the Global Information Technology*, Cannes, pp. 20-25, 2009.

[7] Davis S. and Mermelstein P., "Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences," *IEEE Transaction on Acoustics, Speech and Signal Processing*, vol. 28, no. 4, pp. 357-366, 1980.

[8] Djemili R., Bourouba H., and Korba A., "A Combination Approach of Gaussian Mixture Models and Support Vector Machines for Speaker Identification," *the International Arab Journal of Information Technology*, vol. 6, no. 5, pp. 490-489, 2009.

[9] Elmahdy M., Gruhn R., Minker W., and Abdennadher S., "Survey On Common Arabic Language Forms From Speech Recognition Point of View," *in Proceedings of International Conference on Acoustics*, pp. 63-66, 2009.

[10] Fisher W., Doddington G., and Goudie-Marshall K., "The DARPA Speech Recognition Research Database: Specifications and Status," *in Proceedings of the DARPA Speech Recognition Workshop*, pp. 93-100, 1986.

[11] Kumari R., Nidhyananthan S., and Anand G., "Fused Mel Feature Sets Based Text-Independent Speaker Identification Using Gaussian Mixture Model," *Procedia Engineering*, vol. 30, pp. 319-326, 2012.

[12] Muhammad G., AlSulaiman M., Mahmood A., and Ali Z., "Automatic Voice Disorder Classification using Vowel Formants," *in Proceedings of IEEE International Conference on Multimedia and Expo (ICME)*, Barcelona, pp.1-6, 2011.

[13] Muhammad G., Mesallam T., Malki K., Farahat M., Mahmood A., and Alsulaiman M., "Multidirectional Regression (Mdr)-Based Features For Automatic Voice Disorder Detection," *Journal of Voice*, vol. 26, no. 6, pp. 817.e19-817.e27, 2012.

[14] Praat: doing Phonetics by Computer., available at: http://www.fon.hum.uva.nl/praat, last visited 2014.

[15] Salhi L., Mourad T., and Cherif A., "Voice Disorders Identification Using Multilayer Neural Network," *the International Arab Journal of Information Technology*, vol. 7, no. 2, pp. 177-185, 2010.

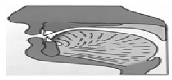
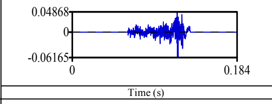
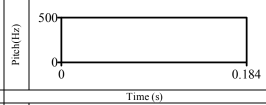
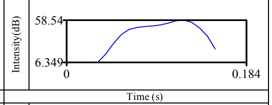
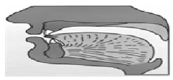
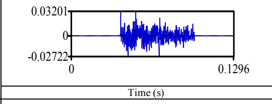
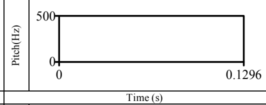
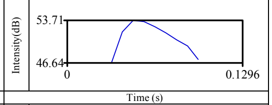
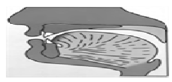
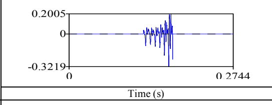
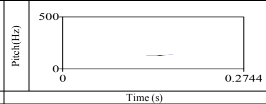
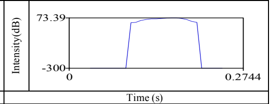
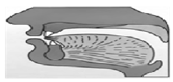
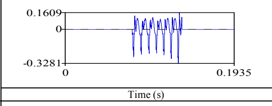
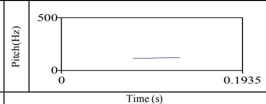
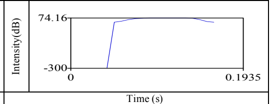
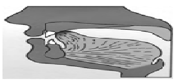
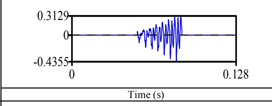
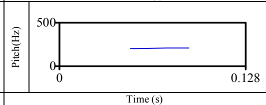
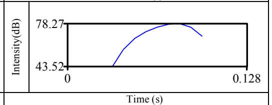
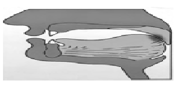
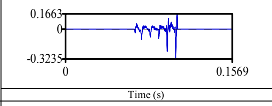
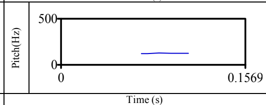
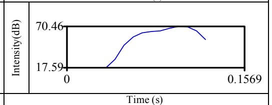
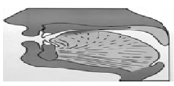
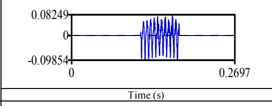
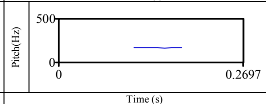
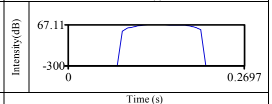
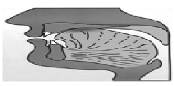
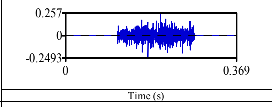
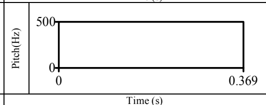
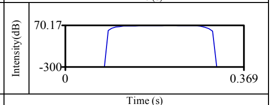
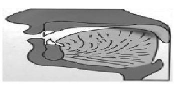
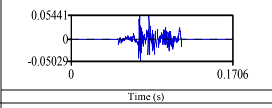
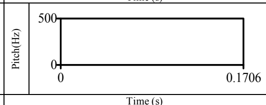
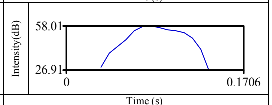
[16] Zeng J., Duan J., and Wu C., "A New Distance Measure for Hidden Markov Models," *Expert Systems with Applications*, vol. 37, no. 2, pp. 1550-1555, 2010.

Abed Ahcène PhD student, ENP Algiers, is a permanent researcher in Scientific and Technical Research Center for Development of Arabic Language (CRSTDLA), University of Bouzareah. He received the engineering and MS degree in 2005 and 2008, from university of M'sila and Military Polytechnics School Algiers, respectively. His research areas include Automatic Arabic Language Processing and Arabic Voice Disorder Assessment Systems.

Mhania Guerti is currently a Professor in the Department of Electronics and the Director of the Signal and Communications Laboratory, National Polytechnic School of Algiers (Algeria). She received her MSc in 1984, from the ILP Algiers in collaboration with the CNET-Lannion (France). She got her PhD from ICP-INPG (Grenoble France), in 1993. She is specialized in Speech and Language Processing. He has supervised a lot of students in Master and PhD degree. Her current research interests include the areas of speech processing, acoustics and audiovisual systems.

Appendix 1

Table 2. Place and manner of articulation of pathological and its substituted sounds.

Sound	Place of Articulation	Manner of Articulation		
		Duration	Pitch	Intensity
[s]				
[θ]				
[z]				
[ð]				
[r]				
[ʁ]				
[ʒ]				
[ʃ]				
[k]				
[t]