ISSN 2007-9737

Arabic Discourse Analysis: A Naïve Algorithm for Defective Pronunciation Correction

Naim Terbeh, Mohsen Maraoui, Mounir Zrigui

Faculty of sciences, Computer Sciences Department, Monastir, Tunisia

{naim.terbeh, maraoui.mohsen}@gmail.com, mounir.zrigui@fsm.rnu.tn

Abstract. This manuscript summarizes a novel study to detect and correct voice anomalies contained in Arabic discourses. These objectives are attained by following some fundamental steps. The first consists in classifying the Arabic produced healthy or pathological vocal signals. Second, the identification of problematic phonemes takes place. The proposition of algorithm allowing the correction of defective pronunciations presents the aim of the latter task. We are satisfied with the obtained results. Indeed, the elaborated algorithm has attained a correction performance of 90% based on 52 Arabic voice sequences covering male and female, healthy and pathological speeches and clustering several areas. Consequently, researchers of topics related to speech processing can benefit from our proposition in the conception and development of their systems.

Keywords. Arabic discourses, healthy speech, pathological speech, problematic phonemes, defective pronunciations, speech correction.

1 Introduction

The human-machine communication presents a large progression and raises new challenges in the objective of comforting the dialogue conditions between humans and machines. Consequently, the literature imposes frequently new applications related to acoustic signal processing like speech recognition, vocal synthesis, speech analysis, voice anomaly detection and assistance to learners of a second language.

In spite of the diversity of techniques comforting the voice communication between humans and machines, the progressive number of people with voice disabilities presents one of the principal obstacles reducing the performance of developed applications. In order not to exclude this population from the man-machine dialogue, this study focusing on the rectification of mispronunciations contained in voice commands has been established. The basic idea is to compare between the referenced phonetic model referring to Arabic healthy speeches and the model proper to the concerned speaker.

The organization of this article is as follows. Section 2 is dedicated to excite the main problematics related to people with voice anomalies during their communication with the machine. In section 3, we cite the famous studies addressing the processing of pathological speeches. Our contribution is detailed in section 4. Section 5 is consecrated to describe the obtained experiment results. The conclusions and perspectives are listed in section 6.

2 Problematics

2.1 Arabic Vocabulary

Arabic is the language spoken by original Arabs. It counts more than 445 million speakers, ranked the 4th in the number of speakers and the 8th in the number of pages that circulate on the Internet [1].

On the one hand, and with its morphological and syntactic properties, Arabic is considered difficult to non-native learners in the area of automatic language processing [2]. On the other hand, degradations contained in produced speeches increase this difficulty. Consequently, the introduction of other means to facilitate the comprehension of defective Arabic discourses

Table 1. Different disabilities with their percentages

Pathologies	Percentage
Mental disabilities	2.3%
Learning problems	3%
Hearing handicaps	0.6%
Visual disabilities	0.1%
Physical disabilities	0.5%
Behavioral disabilities	2%
Language disabilities	3.5%

presents a necessity. Then the automatic correction of pathological speeches will take place.

2.2 Domains of Application

Human-machine communication environments provide the accessibility to information through different modes (keyboard, gestures and voice commands) according to the preferred and/or adequate manner for each user. Voice communication between humans and machines can intervene in different practical areas. We can mention:

- Voice services like the speaking clock, the weather and the reservation tickets.
- Avionics: Several airline companies introduce in the dashboard level the possibility of launching a voice command without interrupting the current task.
- Data entry like voice dictation.
- Training: Children and adults are attracted by endowed speech games (video games, educational games).
- Assistance to people with learning disabilities.

Despite diversity of practical applications involving voice human-machine communication, we mention that several factors prevent an effective dialogue. We distinguish:

- Internal factors: ambiguity [3], implicitness, language model, acoustic model.
- External factors: noise, climate, type of speech (continuous speech or single words).
- Factors due to speakers: voice disabilities, native or non-native speakers.

An erroneous pronunciation can falsify information submitted to the machine which increases the difficulty of a good communication and even falsify the desired results. Therefore, the solution is to introduce a pathological speech corrector in the machine whose functionality is to adjust received data before their treatments.

2.3 People with Language Disabilities

The statistics of the World Health Organization show that 12% of the international population (700 million people) suffers from pathologies where the voice disability is on top of the list with 30% of this population [4]. Thus, the user can communicate with the machine only after adapting this latter to consider the speaker to produce a defective speech. Consequently, the introduction of a speech correction means can provide confidence and safety to the obtained results of communication. In Table 1, we mention the percentage of each disability.

As a recapitulation, the voice communication environments are characterized by:

- The diversity of applications that involve human-machine communication applications
- The diversity of factors generating barriers of human-machine discourses
- The large number of individuals with vocal pathologies
- This population is not immune to human-machine dialogues.

Hence, the introduction of the means whose target is to detect and correct the mispronunciations contained in Arabic discourses (voice commands) can exceed these limits.

3 State of the Art

Not to eliminate people suffering from voice anomalies from vocal dialogues with machines, researchers are fully interested in establishing new methods addressing the numerical accessibility. The development of new platforms focusing on the rectification of defective pronunciations manifested in acoustic signals. From the literature, we can mention the following work:

Voice Anomalies Detection

- A novel methodology invented by Terbeh et al. in [5] whose goal was the detection of voice anomalies appeared in acoustic signals. The basic idea was to compare between different phonetic models to classify the input speech into healthy or pathological.
- In order to offer an important role in early diagnosis, progression tracking, and even the effective treatment of pathological voices, Harar et al. proposed in [11] an original approach whose aim was to detect voice pathologies contained in speech signals. The authors utilized a lot of acoustic features like the raw waveforms, the spectrograms, the Mel-frequency cepstral coefficients and the conventional acoustic features. The obtained results were around 70%.
- In [12], Muhammad et al put forward a novel method based on the interlaced derivation pattern [15], which implicated an n^{th} directional derivative order on a spectral-temporal description of an acoustic signal. The adopted approach showed that the directional information was profitable in the detection of voice pathologies contained in Arabic vocabulary.

Speech Correction

- An advanced approach was elaborated by Ajibola et al. in [6] whose objective is to correct faulty pronunciations due to bad articulation or insufficient air pressure. The proposed method was based on three fundamental tasks. Firstly, the main idea consisted in suppressing the defective frames. Secondly, based on the probabilistic method, suppressed pieces were reconstructed. Finally, the authors launched the algorithm of recognizing reconstructed speeches. According to recognition results, the authors distinguished two cases:
  1. In the case of recognizing a resulting speech, this rephrasing will be saved.
  2. Otherwise, reconstructing the resulting speech can be erroneous.
- On the basis of speech recognition system, a novel approach was suggested in [10] by Bassil et al. An error correction method for a post-editing automatic speech recognition error correction method as well as an algorithm that was in fact based on the online spelling was proposed in this work.

The literature presents a full progression with novel studies addressing the processing of defective acoustic signals, profiting from new technologies of automatic sound processing. However, the pathological Arabic dialogues have not been treated yet. Our contribution consists, for each tagged pathological acoustic signal, in introducing new algorithms whose objective is to correct defective pronunciations.

4 Proposed Methodology

Our method focuses on three main components:

- Classification of acoustic signals into healthy or pathological
- Localization of problematic sounds for pathological speeches
- Correction of defective pronunciations

Figure 1 illustrates our method planning:

4.1 Voice Defect Detection

4.1.1 Probabilistic-Phonetic Model Generation

For this finality, an acoustic model and a large base of healthy Arabic speech are required. After that, this base is preprocessed; frames representing noise and non-Arabic words are suppressed. The next step consists in utilizing the Sphinx-align tool to calculate the vector summarizing the appearance probability of each Arabic bi-

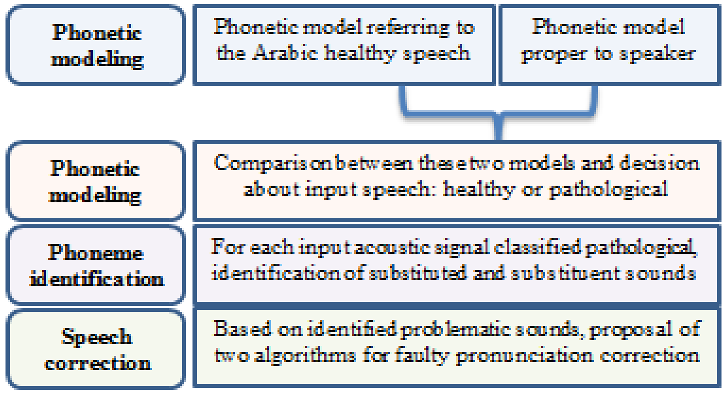

Fig. 1. Voice anomalies correction procedure

phoneme. The resulting vector forms the phonetic model referring to healthy Arabic speeches.

4.1.2 Scalar Product in $\Re^n$

The scalar product in $\Re^n$, noted by <x|y>, is the function which associates to the vectors $x=(x_1, x_2, x_n)\varepsilon\Re^n$ and $y=(y_1, y_2, ..., y_n)\varepsilon\Re^n$ the quantity:

$$\langle x|y\rangle=\sum_{i=1}^{n} x_i y_i. \quad (1)$$

This quantity can also be expressed as follows:

$$\langle x|y\rangle=||x||.||y||.\cos(\theta). \quad (2)$$

We note by:

– θ the angle which separates the two vectors x and y,
– ||x|| the norm of the vector x, $||x||=\sqrt{x_1^2+x_2^2+\cdots+x_i^2+\cdots+x_n^2}$
– ||y|| the norm of the vector y, $||y||=\sqrt{y_1^2+y_2^2+\cdots+y_i^2+\cdots+y_n^2}$

Consequently, we can deduce that:

$$\cos(\theta)=\sum_{i=1}^{n} x_i y_i/||x||.||y||. \quad (3)$$

If we can generate a vectorial representation in $\Re^n$ for each produced speech signal, angle θ (equations (1) and (2)) can be used as a measure of similarity between the speech signals represented by these vectors. Indeed, the similarity between different acoustic signals is proportional to the angular distance (angle θ) which separates between the representative vectors.

4.1.3 Phonetic Distance

In this work, the notion of phonetic distance is invented for the first time. It describes the angular distance that separates two different phonetic models. In order to calculate the phonetic distance, n healthy speech bases are required. These bases are treated as follows:

– For each corpus C_i (i=1-n), we calculate the equivalent probabilistic-phonetic model M_i.
– $S=\{\alpha_{ij};\ 1\leq i,j\leq n \text{ and } i\neq j\}$ contains all angles separating two different models M_i and M_j.
– Max defines the maximum of the ensemble {S}.
– δ defines the standard deviation of {S}.
– Avg defines the average of {S}.
– The phonetic distance is defined by this expression: $\beta=Max+|Avg-\delta|$.

4.1.4 Speech Classification

Each input speech undergoes two different decisions; it will be classified into healthy or pathological. The generated decision depends on the phonetic distance θ which separates between the probabilistic-phonetic model referring the healthy Arabic spoken vocabulary and the model characterizing the input speech recorded by a new speaker (native, non-native, healthy, with voice anomalies). Two cases are envisaged:

- If θ≤β, then the input acoustic signal is heathy and the speaker does not suffer from any voice anomaly.
- Else (θ>β), the input speech is pathological. Accordingly, we launch an appropriate procedure whose aim is to identify problematic phonemes.

4.2 Correction Procedure

Correction procedure is based on two fundamental tasks. The first consists in identifying the problematic sounds (sounds begetting the voice anomalies and replacement ones). Based on identified phonemes in the first task, the second one will be dedicated for the matching between problematic sounds to generate the correct pronunciation corresponding to each faulty one.

4.2.1 Problematic Sound Identification

The current procedure will be applied only if the input acoustic signal contains voice anomalies and classified as defective. The proposed algorithm of identifying problematic sounds generates two main phoneme categories (substituted phonemes and substituent ones).

4.2.1.1 Substituted Sounds

This category of sounds prevents the correct pronunciation and presents the source of voice anomalies for the concerned talkers. The suggested algorithms whose aim is to identify these sounds are based on the hypothesis considering that one substituted phoneme does never figurate in the phonetic transcription summarizing the input speech. Accordingly, all coefficients of probabilistic-phonetic model corresponding to bi-phonemes containing a faulty produced sound will be equal to zero.

On the basis of the previously mentioned hypothesis, a simple comparison between the probabilistic-phonetic model referring to the Arabic healthy speech and the probabilistic-phonetic model proper to speaker can lead to the extraction of sounds posing degraded speeches. Figure 2 illustrates the obtained results for one practical case where the speaker suffers from the pronunciation of the Arabic sounds "خ/x" and "ق/q".

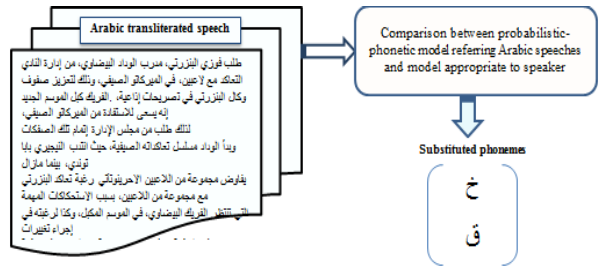

Fig. 2. An example of identification of substituted phonemes

4.2.1.2 Substituent Sounds

This subsection treats the objective of identifying substituent sounds. These latter are phonemes that are pronounced instead of substituted ones. This finality is based on the hypothesis considering that the sum of occurrence probabilities of bi-phonemes containing substituted sounds will be distributed to bi-phonemes covering replacement sounds. Therefore, all coefficients of the probabilistic-phonetic model covering bi-phonemes containing a substituent sound will present a remarkable increase compared to the referenced probabilistic-phonetic model. Figure 3 shows the obtained results for the previous example (in Figure 2):

4.3 Defective Speech Correction

In this subsection, we give details of the proposed algorithms addressing the correction of pathological speeches. We distinguish two different correction procedures according to the type of pathological speech to be corrected: homogeneous or heterogeneous classes of defective speeches.

4.3.1 Correction of Homogeneous Classes

4.3.1.1 Procedure

Homogeneous classes cluster the pronunciations containing just one faulty pronounced phoneme in each one. The correction of one defective pronunciation contained in a homogeneous class necessitates just one possible substitution of the replacement phoneme, which appears in this pronunciation by one of the substituted phonemes. The correction of mispronunciations contained in this class follows Algorithm 1:

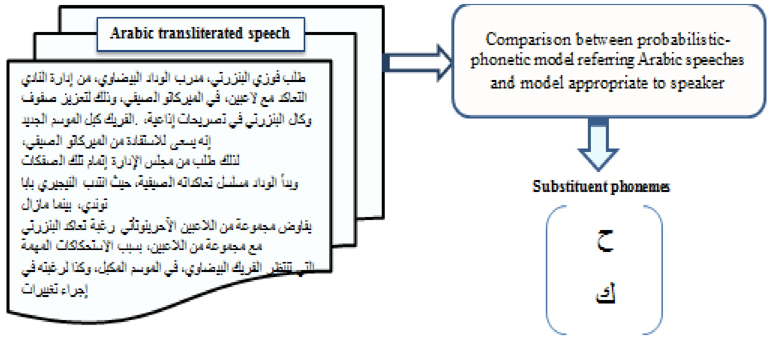

Fig. 3. An example of identification of substituent phonemes

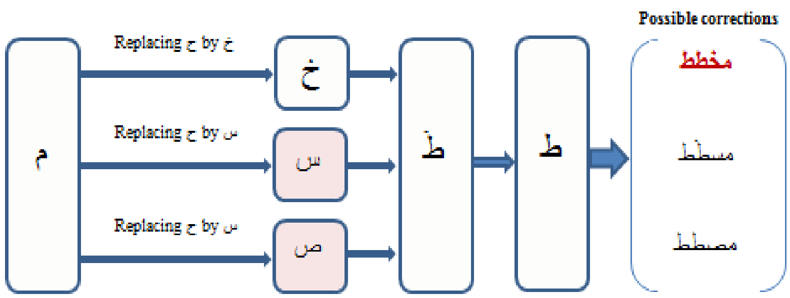

Fig. 4. Correction of the Arabic mispronunciation "**محطّط**"

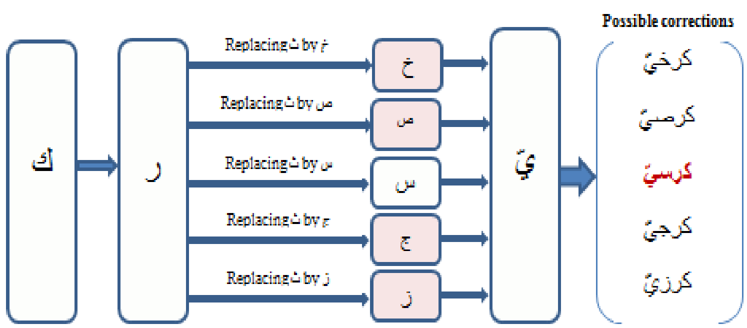

Fig. 5. Correction of the Arabic mispronunciation "**كرثيّ**"

Algorithm 1. Correction of homogeneous class of defective pronunciations

For each pronunciation in C_i (C_i is the class of defective pronunciation that the phoneme P_i is substituent)
1.replace P_i by P_j (P_j is a phoneme from the set of substituted phonemes, $1 \leq j \leq m$, m is the number of substituted phonemes)
2.save the matching which maximizes the number of pronunciations accepted by an Arabic lexicon

4.3.1.2 Practical Examples

First Example

Let us take the example of the correction of the pronunciation "محطّط [muħatˤatˤ]", where:

- M={خ، ص ، س},
- R={ح، ث}.

On the basis of the proposed solution in Algorithm 1, and applying possible matchings between problematic sounds, the pronunciation "مخطّط [muxatˤatˤ]" (planning) is calculated as a correct one which corresponds to the input elocution. The colored pronunciation in Figure 4 presents the accepted one (the other pronunciations are not accepted because they do not belong to the Arabic lexicon).

Second Example

Let us take the correction of the badly pronunciation "كرثيّ [kurθi`yɔn]", considering the following conditions:

– M={ز ،ج ،س ،ص ،خ},
– R={ذ ،ح ،ث}.

Supporting Algorithm 1, we have obtained the tree as shown in Figure 5. The pronunciation "كرسيّ [kursi`yɔn]" (chair) (the colored pronunciation) is calculated as the correct one, which corresponds to the badly input elocution.

4.3.2 Correction of Heterogeneous Classes

Dissimilar classes assemble the mispronunciations covering more than one badly pronounced sound in each one. The correction of a pronunciation contained in a heterogeneous class requires $(n+1)^m-1$ substitutions between the substitute phonemes that appear in this pronunciation and the substituted phonemes. In order to facilitate the management of these matchings, we introduce for this treatment the decision trees. We note by:

– n the number of substituted phonemes
– m the number of replacement phonemes that figurate in the defective pronunciation

In this class, the correction of badly pronunciations is based on Algorithms 2 and 3.

4.3.2.1 Tracing Algorithm

In order to be able to manage the very high number of substitutions between substituted phonemes and substituent ones, we propose a probabilistic-tree structure; the nodes are labeled by the phonemes and the arcs connecting these nodes are weighted by the probabilities of occurrence of the bi-phoneme [Phoneme Father-Phoneme Son]. We suggest the tracing procedure to ensure drawing the probabilistic-tree representing all possible pronunciations coming from the faulty input one:

Algorithm 2. Tracing algorithm

Posing m the badly pronunciation to be corrected
Begin
1.If the first phoneme P_1 is substituent then
-The root will be labeled NULL
-P_1 and its correspondent substituted phonemes $\{P'_1, P'_2, \dots, P'_n\}$ form the sons of this root
-Arcs from root to P_1 and P'_i $(1 \le i \le n)$ will be equal-probably weighted
2.Else
-P_1 is the root of this tree
End if
For $2 \le i \le |m|$ ($|m|$ is the length of the pronunciation m)
3.If P_i is substituent then
-P_i and its correspondent substituted phonemes, $\{P'_{i1}, P'_{i2}, \dots, P'_{ik}\}$ form the sons of P_{i-1}
-Arcs $P_{i-1}P_i$ and $P_{i-1}P'_{ij}$ $(1 \le j \le k)$ will be weighted respectively by $P(P_{i-1}P_i)$ and $P(P_{i-1}P'_{ij})$
End if
End for
End

4.3.2.1 Routing Algorithm

After drawing all possible pronunciations coming from the badly input one, the routing procedure will be used in browsing the probabilistic-tree and in calculating the list of Arabic words that correspond to the falsely pronounced ones. Details of this algorithm are as follows:

Algorithm 3. Routing algorithm

Begin
Depth-first path of probabilistic-tree representing the pronunciation to be corrected (m) by following the arc with the highest weight until arriving at the leaf level.
1.If the obtained path forms one pronunciation m' existing in the lexicon then
-m' is the correct pronunciation of m
2.Else
-go back until you get the branch with largest height,
-delete this branch,
-Repeat the routing from the actual position.
End if
End

4.3.2.1 Practical Examples

First Example

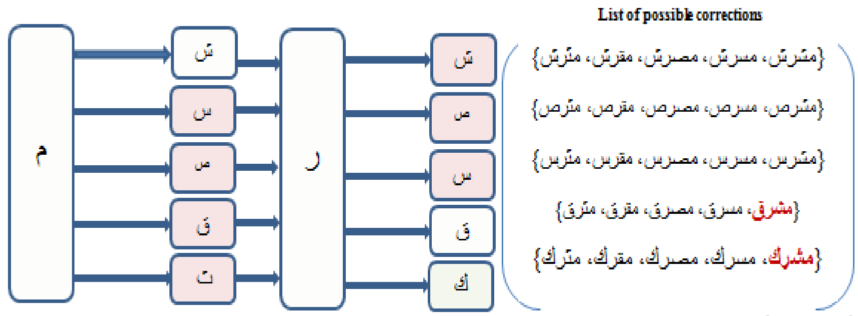

Fig. 6. Correction of the Arabic mispronunciation “مَثْرَك”

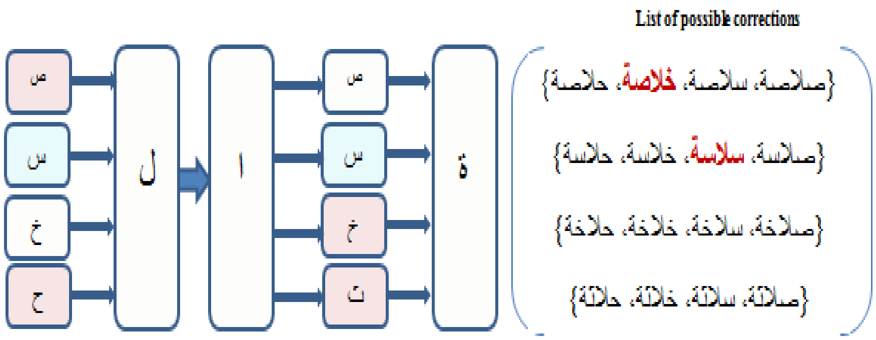

Fig. 7. Correction of the Arabic mispronunciation “خُلاَثَةٌ”

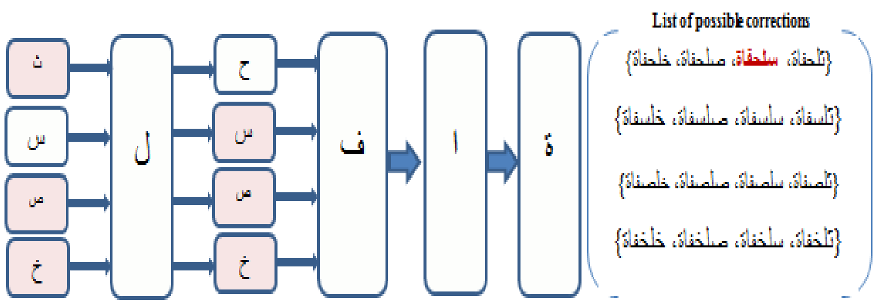

Fig. 8. Correction of the Arabic mispronunciation “ثُلَخْفَاةٌ”

Let us take the correction of the badly pronunciation "مَثْرَكٌ [maθrakɔn]", considering the following conditions:

- Substituted phonemes={ش، س، ص، ق}.
- Substituent phonemes={ث، ك }.
- Possibilities of substitution between substituted phonemes and substituent ones are:

S1: “ث” can be a replacement phoneme of the phoneme “س”.

S2: “ث” can be a replacement phoneme of the phoneme “ص”.

S3: “ث” can be a replacement phoneme of the phoneme “ق”.

S4: “ث” can be a replacement phoneme of the phoneme “س”.

S5: “ك” can be a replacement phoneme of the phoneme “س”.

S6: “ك” can be a replacement phoneme of the phoneme “ص”.

S7: “ك” can be a replacement phoneme of the phoneme “ق”.

S8: “ك” can be a replacement phoneme of the phoneme “ص”.

In the current case, the badly pronunciation “مَثْرَكٌ” contains the substituent sounds “ث” and “ك”. Applying the possible substitutions, the tracing algorithm will generate the probabilistic-tree as depicted in Figure 6:

We apply thereafter the routing algorithm to find two correct pronunciations, accepted by the Arabic lexicon, “مَشْرَقٌ [maʃraqɔn]” (the orient) and “مُشْرِكٌ [muʃrikɔn]” (polytheist). The adequate correction from the obtained ones is the pronunciation following the same rhyme compared to the input badly pronounced one. The other pronunciations are not accepted by the reference lexicon, so they are not selected.

Second Example

Let us take the correction of the badly pronunciation "حُلَاثَةٌ [ħulæθatɔn]", considering the following conditions:

- M={خ، س، ص }.
- R={ح، ث}.
- Possibilities of substitution between substituted phonemes and substituent ones are:

S1: “ث” can be a replacement phoneme of the phoneme “س”.

S2: “ث” can be a replacement phoneme of the phoneme “ص”.

S3: “ث” can be a replacement phoneme of the phoneme “خ”.

S4: “ح” can be a replacement phoneme of the phoneme “س”.

S5: “ح” can be a replacement phoneme of the phoneme “ص”.

S6: “ح” can be a replacement phoneme of the phoneme “خ”.

In this case, phonemes “ح” and “ث” appear in the bad pronunciation “حُلَاثَةٌ”. Based on possible substitutions, the tracing algorithm will generate the probabilistic-tree detailed in Figure 7.

Applying the routing algorithm, we find two possible correct pronunciations “خلاصة [xulæsˤatɔn]” (conclusion) and “سلاسة [sælæsætɔn]” (smoothness). They are accepted by the Arabic lexicon. To select the adequate correction from the obtained ones, we can utilize semantic analysis. It may be possible also to choose the correction with the pronunciation rhyme similar to the input mispronunciation. The other pronunciations are not chosen because they are not accepted by the reference lexicon.

Third Example

Let us take the correction of the badly pronunciation "ثُلَحْفَاةٌ [θuleħfætɔn]", considering the following conditions are:

- M={س ، ص ،خ}.
- R={ث ،ح}.
- Possibilities of substitution between substituted phonemes and substituent ones are:

S1: “ث” can be a replacement phoneme of the phoneme “س”.

S2: “ث” can be a replacement phoneme of the phoneme “ص”.

S3: “ث” can be a replacement phoneme of the phoneme “خ”.

S4: “ح” can be a replacement phoneme of the phoneme “س”.

S5: “ح” can be a replacement phoneme of the phoneme “ص”.

S6: “ح” can be a replacement phoneme of the phoneme “خ”.

The falsely pronunciation “ثُلَحْفَاةٌ” contains the two phonemes “ح” and “ث”. Applying the possible substitutions, the tracing algorithm will generate the probabilistic-tree illustrated by Figure 8.

Thus, we apply the routing algorithm to obtain the correct elocution “سلحفاة [suleħfætɔn]” (tortoise) (the colored pronunciation) corresponding to the input falsely pronounced one.

ISSN 2007-9737

5 Tests and Results

5.1 Data of Tests and Obtained Results

In order to test the proposed method, a large base of Arabic healthy and pathological speeches is required. For this finality, we collaborate with all members of our laboratory[1] to collect 14 hours of pathological speeches and 6 hours of healthy speeches. The collected base covers different areas like technology, education, economy, sports, sciences, migration and politics. The results of pathological speech correction are detailed in Table 2:

The obtained results, shown in Table 2, present the efficiency of our pathological speech corrector. Figure 9 depicts the speech classification and the defective speech correction rates for each gender and for each type of speeches (healthy and pathological).

5.2 Discussion and Limits

Table 2 presents some misclassifications; i.e. a healthy speech classified as a pathological one and vice versa. These bad classifications can be due to the following reasons:

- The concerned speakers do not suffer from voice pathologies, but they present some difficulties in mastering the pronunciation of some phoneme combinations like the succession of two emphatic phonemes, the succession of thin and amplified sounds. In such cases, and the speaker presents some phoneme outlet troubles.
- The speaker suffers from distorted voice pathologies. In such a case, the desired sound is badly pronounced, but the substituent phoneme is nevertheless like the hoped sound. For example, the concerned speaker can produce the defective pronunciation "تاولة" [tæwilə] (with a voiceless /t/) instead of the desired "طاولة" [tˤa:wilə] (with a voiced /t/) (table). In this case, despite the fact that a mispronunciation manifests, the produced speech is understandable, especially in human-human communication.
- For several speakers, we can also justify misclassifications by the rhythm of elocution. Indeed, speakers who talk quickly may merge two or more phonemes, which can influence the proper resulting probabilistic-phonetic model. Consequently, the comparison between the phonation model referring to Arabic vocabulary and the elocution model specific to the speaker will generate erroneous results and the produced speech will be misclassified.

Sometimes the algorithm of defective speech correction generates several possible correct pronunciations for a single erroneous pronunciation. In such a case, semantic analysis is required to select the adequate correction. We also notice that the problematic phonemes for the same speaker are often clustered in the same category (hissing, emphatic, Occlusive, fricative, nasal, etc.)

6 Conclusions

To conclude, we can confirm that the probabilistic-phonetic modeling is a deciding factor in Arabic speech classification into healthy or pathological. This manuscript describes a novel naive algorithm whose aim is to correct bad pronunciations manifested in Arabic discourses.

We can conclude that the comparison between different phonetic models presents a deciding factor to detect vocal pathologies contained in Arabic speeches and to generate an adequate correction for each input mispronunciation. In this paper, we have proposed a novel algorithm to correct mispronunciations contained in an Arabic acoustic signal.

For this purpose, a corpus of 11 hours of healthy Arabic speeches has been prepared for

[1] Http://www.latice.rnu.tn/

Table 2. Summary of test base and obtained elocution-speed for each speaker

Speaker	Area	Gender	Age (year)	Duration (min)	Type of speech	Nbr. of defective pronunciations	Problematic sounds	Correction rate (%)
01		M	32	16	Healthy		Classified as healthy	
02		M	19	33	Pathological	211	س، ص (sˤ, s)	88.62
03		M	22	21	Pathological	147	خ، غ (ɣ, x)	87.75
04	Technology	M	41	25	Pathological	169	س، ص (sˤ, s)	89.34
05		F	26	22	Healthy		Classified as healthy	
06		F	17	19	Pathological	143	ر (r)	87.41
07		F	21	24	Pathological	167	ش، س، ص (sˤ, s, ʃ)	89.82
08		M	25	15	Healthy	62	ذ، ض، ظ (ðˤ, dˤ, ð)	91.93
09		M	23	32	Healthy		Classified as healthy	
10		M	18	16	Pathological	129	ر (r)	93.02
11	Education	M	32	23	Pathological	63	ش (ʃ)	85.71
12		M	39	27	Pathological	137	ق (q)	90.51
13		F	27	24	Healthy		Classified as healthy	
14		F	18	31	Pathological	239	ر (r)	92.46
15		F	20	18	Pathological	117	س، ص (sˤ, s)	87.17
16		M	30	20	Healthy		Classified as healthy	
17		M	21	27	Pathological	59	ث (θ)	88.13
18	Economy	M	27	23	Pathological		Classified as healthy	
19		M	19	30	Pathological	43	ر (r)	93.02
20		F	36	19	Healthy		Classified as healthy	
21		F	22	32	Pathological	227	ر (r)	90.30
22		M	24	14	Healthy	89	س، ص (sˤ, s)	91.01
23		M	20	31	Healthy		Classified as healthy	
24		M	19	33	Pathological	106	ش (ʃ)	93.39
25		M	31	27	Pathological	186	س، ص (sˤ, s)	90.86
26	Sports	M	42	29	Pathological	204	خ، غ (ɣ, x)	92.64
27		M	17	23	Pathological	121	ق (q)	93.38
28		F	21	21	Healthy		Classified as healthy	
29		F	24	26	Pathological	173	س، ص (sˤ, s)	87.86
30		F	38	19	Pathological	141	خ، غ (ɣ, x)	91.49
31		M	20	19	Healthy		Classified as healthy	
32	Sciences	M	16	25	Pathological		Classified as healthy	
33		M	29	19	Pathological	152	ر (r)	92.76
34		M	27	31	Pathological	229	س، ص، ق (q, sˤ, s)	92.13

35		F	31	28	Healthy		Classified as healthy	
36		F	18	23	Pathological	159	خ، غ (ɣ, x)	92.45
37		F	44	25	Pathological	119	ق (q)	89.91
38		M	32	22	Healthy		Classified as healthy	
39		M	25	23	Pathological	46	ث (θ)	89.13
40		M	22	19	Pathological	156	س، ص، ق (q, sˤ, s)	91.02
41	Migration	M	34	28	Pathological	125	ذ، ض، ظ (ðˤ, dˤ, ð)	93.60
42		F	19	27	Healthy		Classified as healthy	
43		F	23	17	Healthy		Classified as healthy	
44		F	18	29	Pathological	204	ش، س، ص (sˤ, s, ʃ)	91.66
45		M	43	22	Healthy		Classified as healthy	
46		M	18	28	Pathological	117	ذ، ض، ظ (ðˤ, dˤ, ð)	89.74
47		M	34	17	Pathological	113	خ، غ (ɣ, x)	92.03
48	Politics	M	40	20	Pathological	89	ق (q)	83.14
49		M	38	18	Pathological	109	ر (r)	91.74
50		M	31	19	Pathological		Classified as healthy	
51		F	25	16	Healthy	127	ر (r)	88.18
52		F	32	15	Pathological	69	ذ، ض، ظ (ðˤ, dˤ, ð)	81.15

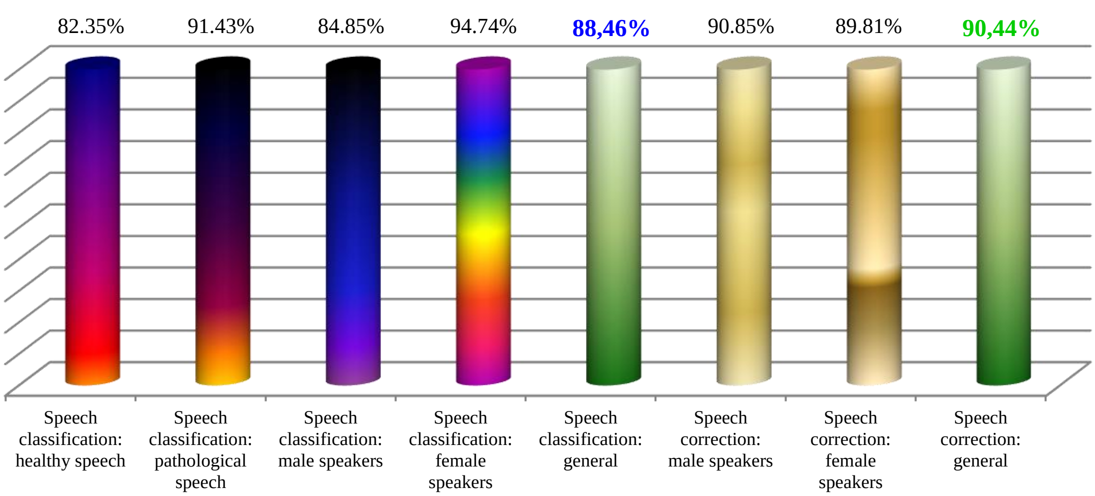

Fig. 9. Speech classification and defective speech correction rates based on our methodology

calculation of the phonetic model referring to the spoken Arabic language.

On the basis of 52 speech sequences recorded by male and female native speakers covering healthy and pathological speeches, the suggested methodology has attained high speech classification accuracy. Indeed, the algorithm whose target is to classify input speeches into healthy or pathological has a classification rate of 88%. For each speech sequence classified as pathological, we launch a naïve algorithm based on testing all possible substitutions that can be applied between substituted and substituent sounds.

ISSN 2007-9737

Table 3. Arabic to IPA mapping (consonants)

Arabic consonants	IPA symbol	English approximation
ا ,ى	a:	f**a**ther
ب	b	**b**ee
د	d	**d**ash
ض	dˤ	like the sound **d** in **d**uck, bu**d**, no**d**
ج	dʒ	**j**am
ذ	ð	**th**ese
ظ	ðˤ	no English equivalent (like voiced **th**)
ف	f	**f**ather
ه	h	w**h**o
ح	ħ	No English equivalent, (like Mexican **j**ota)
ي	j	**y**es
ك	k	s**k**in
ل	l	too**l**
م	m	**m**e
ن	n	**n**o
ق	q	no English equivalent **(**emphatic /**k**/)
ر	r	**r**hythm
س	s	**s**ee
ص	sˤ	**s**on
ش	ʃ	**sh**e
ت	t	**t**able
ط	tˤ	like the sound **t** in bo**tt**le
ث	θ	**th**ink
و	w	**w**e
خ	x	no English equivalent (Spanish **j**ota)
غ	ɣ	no English equivalent **(**Spanish fue**g**o**)**
ز	z	**z**oo
ظ	zˤ	no English equivalent (emphatic /**z**/)
ء	ʔ	no English equivalent (like **ai**m)
ع	ʕ	no English equivalent

We distinguish for each substitution two decisions: If the resulting pronunciation is accepted by an Arabic lexicon, then we will save this matching.

Otherwise, the substitution will not be suitable, so we will try with another analogous one. By its obtained correction rate reaching 90%, the proposed algorithm addressing the correction of defective pronunciations has presented high efficiency.

The contribution detailed in this article presents one of the profitable correctors of defective speeches.

ISSN 2007-9737

Table 4. Arabic to IPA mapping (vowels)

Arabic vowels	IPA symbol	English approximation
◌َ	a	m**a**n
◌ِ	i	h**i**m
◌ُ	u	p**u**t
◌َا، ◌َى	a:	c**a**r
◌ِي	i:	sh**ee**p
◌ُو	u:	r**u**le

Table 5. Transliteration of pronunciations used in this paper

$$\begin{pmatrix} \text{مخطّط} \\ \text{مسطّط} \\ \text{مصطّط} \end{pmatrix} \approx \begin{pmatrix} \text{muxatˤatˤ} \\ \text{musatˤatˤ} \\ \text{musˤatˤatˤ} \end{pmatrix}$$

Pronunciations used in the example depicted in Figure 4

$$\begin{pmatrix} \text{كرخيّ} \\ \text{كرصيّ} \\ \text{كرسيّ} \\ \text{كرجيّ} \\ \text{كرزيّ} \end{pmatrix} \approx \begin{pmatrix} \text{kurxi`yɔn} \\ \text{kursˤi`yɔn} \\ \text{kursi`yɔn} \\ \text{kurdʒi`yɔn} \\ \text{kurzi`yɔn} \end{pmatrix}$$

Pronunciations used in the example depicted in Figure 5

$$\begin{pmatrix} \text{مثرش} & \text{مقرش} & \text{مصرش} & \text{مسرش} & \text{مشرش} \\ \text{مثرص} & \text{مقرص} & \text{مصرص} & \text{مسرص} & \text{مشرص} \\ \text{مثرس} & \text{مقرس} & \text{مصرس} & \text{مسرس} & \text{مشرس} \\ \text{مثرق} & \text{مقرق} & \text{مصرق} & \text{مسرق} & \text{مشرق} \\ \text{مثرك} & \text{مقرك} & \text{مصرك} & \text{مسرك} & \text{مشرك} \end{pmatrix} \approx \begin{pmatrix} \text{maθraʃ} & \text{maqraʃ} & \text{masˤraʃ} & \text{masraʃ} & \text{maʃraʃ} \\ \text{maθrasˤ} & \text{maqrasˤ} & \text{masˤrasˤ} & \text{masrasˤ} & \text{maʃrasˤ} \\ \text{maθraq} & \text{maqraq} & \text{masˤraq} & \text{masras} & \text{maʃras} \\ \text{maθraq} & \text{maqraq} & \text{masˤraq} & \text{masraq} & \text{maʃraq} \\ \text{maθrak} & \text{maqrak} & \text{masˤrak} & \text{masrak} & \text{maʃrak} \end{pmatrix}$$

Pronunciations used in the example depicted in Figure 6

$$\begin{pmatrix} \text{حلاصة} & \text{خلاصة} & \text{سلاصة} & \text{صلاصة} \\ \text{حلاسة} & \text{خلاسة} & \text{سلاسة} & \text{صلاسة} \\ \text{حلاخة} & \text{خلاخة} & \text{سلاخة} & \text{صلاخة} \\ \text{حلاثة} & \text{خلاثة} & \text{سلاثة} & \text{صلاثة} \end{pmatrix} \approx \begin{pmatrix} \text{ħulæsˤatɔn} & \text{xulæsˤatɔn} & \text{sælæsˤætɔn} & \text{sˤulæsˤatɔn} \\ \text{ħulæsatɔn} & \text{xulæsatɔn} & \text{sælæsætɔn} & \text{sˤulæsatɔn} \\ \text{ħulæxatɔn} & \text{xulæxatɔn} & \text{sælæxætɔn} & \text{sˤulæxatɔn} \\ \text{ħulæθatɔn} & \text{xulæθatɔn} & \text{sælæθætɔn} & \text{sˤulæθatɔn} \end{pmatrix}$$

Pronunciations used in the example depicted in Figure 7

$$\begin{pmatrix} \text{خلحفاة} & \text{صلحفاة} & \text{سلحفاة} & \text{ثلحفاة} \\ \text{خلسفاة} & \text{صلسفاة} & \text{سلسفاة} & \text{ثلسفاة} \\ \text{خلصفاة} & \text{صلصفاة} & \text{سلصفاة} & \text{ثلصفاة} \\ \text{خلخفاة} & \text{صلخفاة} & \text{سلخفاة} & \text{ثلخفاة} \end{pmatrix} \approx \begin{pmatrix} \text{xuleħfætɔn} & \text{sˤuleħfætɔn} & \text{suleħfætɔn} & \text{θuleħfætɔn} \\ \text{xulesfætɔn} & \text{sˤulesfætɔn} & \text{sulesfætɔn} & \text{θulesfætɔn} \\ \text{xulesˤfætɔn} & \text{sˤulesˤfætɔn} & \text{sulesˤfætɔn} & \text{θulesˤfætɔn} \\ \text{xulexfætɔn} & \text{sˤulexfætɔn} & \text{sulexfætɔn} & \text{θulexfætɔn} \end{pmatrix}$$

Pronunciations used in the example depicted in Figure 8

doi: 10.13053/CyS-23-1-2820

Indeed, to the best of our knowledge, it is the first attempt addressing the rectification of vocal anomalies contained in the Arabic speeches in the purpose of facilitating the accessibility to the numerical information in the natural language processing domain. As concluding remarks, the obtained results are satisfactory and our proposition can be applied in others related studies.

As any research work, our contribution presents some limits. Consequently, we can put forward a phase of semantic analysis that can assist the defective-speech corrector to select the appropriate correction that corresponds to the desired meaning.

As perspectives, we can benefit from this work to elaborate new systems addressing the assistance to scholar children suffering from voice disabilities [13, 14, 17]. It may be possible also to use other features like the voyellation and the pronunciation rhymes for guiding the correction of degraded acoustic signals.

Acknowledgements

We would like to express our deepest regards to all members of the committee in the Computación y Sistemas journal to accept the evaluation of our work. I extend also my thanks to my supervisors Mr. Mounir Zrigui and Mr. Mohsen Maraoui.

References

1. **American Speech Language Hearing Association. (2014).** http://www.asha.org/public/speech/disorders/ChildSandL.htm.
2. **Aljlayl, O. & Frieder, M. (2002).** On Arabic search: Improving the retrieval effectiveness via a light stemming approach. *11th International Conference on Information and Knowledge Management.*
3. **Merhbene, L., Zouaghi, A., & Zrigui, M. (2013).** An experimental study for some supervised lexical disambiguation methods of Arabic language. *4th International Conference on Information and Communication Technology and Accessibility (ICTA).*
4. **Celikoglu, M., Kamak, H., & Oktay, H. (2010).** *Prevalence and characteristics of supernumerary teeth in a non-syndrome Turkish population: associated pathologies and proposed treatment.* Med Oral Patol Oral Cir Bucal.
5. **Terbeh, N., Maraoui, M., & Zrigui, M. (2015).** Probabilistic Approach for Detection of Vocal Pathologies in the Arabic Speech. *CICLing*.
6. **Ajibola, A. S., Rashid, N. K. B. A. M., Sediono, W., & Hashim, N. N. W. N. (2016*).*** *A novel approach to stuttered speech correction. Jurnal Ilmu Komputer dan Informasi.*
7. **Labidi, M., Terbeh, N., Maraoui, M., & Zrigui, M. (2015).** Semantic indexing of continuous speech. *Information & Communication Technology and Accessibility (ICTA).*
8. **Terbeh, N. & Zrigui, M. (2012).** *Arabic Language Learning Assisted by Computer, based on Automatic Speech Recognition.*
9. **Terbeh, N. & Zrigui, M. (2016).** A novel approach to identify factor posing pronunciation disorders. *International Conference on Computational Collective Intelligence,* pp. 153−162.
10. **Bassil, Y. & Alwani, M. (2012).** Post-Editing Error Correction Algorithm for Speech Recognition using Bing Spelling Suggestion. *International Journal of advanced Computer Science and Applications,* Vol. 3, No. 2.
11. **Harar, P., Galaz, Z., Alonso-Hernandez, J. B., Mekyska, J., Burget, R., & Smekal, Z. (2018).** Towards robust voice pathology detection. *Neural Computing and Applications,* pp. 1−11.
12. **Muhammad, G., Alsulaiman, M., Ali, Z., Mesallam, T. A., Farahat, M., Malki, K. H., & Bencherif, M. A. (2017).** Voice pathology detection using interlaced derivative pattern on glottal source excitation. *Biomedical Signal Processing and Control,* Vol. 31, pp. 156−164.
13. **Terbeh, N., Trigui, A., Maraoui, M., & Zrigui, M. (2017).** Correction of pathological speeches and assistance to learners with vocal disabilities. *Multimedia Tools and Applications.*
14. **Terbeh, N. & Zrigui, M. (2017).** Identification of Pronunciation Defects in Spoken Arabic Language. *International Conference of the Pacific Association for Computational Linguistics*, pp. 355−365.
15. **Duse, E. (2015).** *On Uniformly Random Discrete Interlacing Systems.* Doctoral thesis, KTH Royal Institute of Technology.
16. **Trigui, A., Terbeh, N., Maraoui, M., & Zrigui, M. (2016).** Statistical Approach for Spontaneous Arabic Speech Understanding Based on Stochastic Speech Recognition Module. *Research in Computing Science,* Vol. 117.
17. **Terbeh, N., Trigui, A., Maraoui, M., & Zrigui, M. (2016).** Arabic speech analysis to identify factors posing pronunciation disorders and to assist

learners with vocal disabilities. *IEEE International Conference on Engineering & MIS.*

Article received on 30/10/2017; accepted on 20/04/2018. Corresponding author is Naim Terbeh.