Detecting Sentences Types in the Standard Arabic Language

Ramzi Halimouche and Hocine Teffahi
Laboratory of Spoken Communication and Signal Processing, Electronics and Computer Science Faculty, University of Sciences and Technology Houari Boumediene, Algeria

Abstract: *The standard Arabic language, like many other languages, contains a prosodic feature, which is hidden in the speech signal. The studies related to this field are still in the preliminary stages. This fact results in restraining the performance of the communication tools. The prosodic study allows people having all the communication tools needed in their native language. Therefore, we propose, in this paper, a prosodic study between the various types of sentences in the standard Arabic language. The sentences are recognized according to three modalities as the following: declarative, interrogative and exclamatory sentences. The results of this study will be used to synthesize the different types of pronunciation that can be exploited in several domains namely the man-machine communication. To this end, we developed a specific dataset, consisting of the three types of sentences. Then, we tested two sets of features: prosodic features (Fundamental Frequency, Energy and Duration) and spectrum features (Mel-Frequency Cepstral Coefficients and Linear Predictive Coding) as well their combination. We adopted the Multi-Class Support Vector Machine (MC-SVM) as classifier. The experimental results are very encouraging.*

Keywords: *Standard arabic language, sentence type detection, fundamental frequency, energy, duration, mel-frequency cepstral coefficients, linear predictive coding.*

Received January 19, 2017; accepted August 23, 2017

1. Introduction

The speech signal contains much more information than the words themselves. This information, known as "extralinguistic," is hidden in the speech signal. Researchers have used prosodic features in several domains such as detecting the emotional state of the speaker [3, 16, 28], sentence boundary detection [12, 15, 18, 23], dialog act detection [4, 19, 25, 27] etc., the classification and detection type of sentences could be seen as a special case of the dialogue acts' classification. This work aims at studying the prosodic difference between the sentence types before, so that to recognize it automatically in: questions, exclamations and declarative sentences. Such an information then may be used to enrich the man-machine dialogue in standard Arabic language. Several previous work of the various systems used in the question detection with the prosodic model and the lexical model have been made in this regard in different languages. In order to justify our choice of feature vectors and used classifier. In order to place our work over the already made studies, we consider some work in different languages being ought to be considered.

In English language, the authors of the laboratory "Spoken Language Processing, Department of Computer Science-Columbia University, New York, USA" have solved the sentence type detection problem in the context "Intelligent Tutoring Spoken Dialog System" [17, 29]: Software for educational use is developed to assist students and substitute teachers in their study process. However, these systems do not take into account the type of sentence. The authors have studied how to develop a system that automatically detects a type of sentence [17]. The authors used lexical and prosody parameters extracted from a dataset of student-teacher dialogues. They classified questions with accuracy equal to 79%, using the Fundamental Frequency as parameter. Recently another study [6] was conducted in the same field.

For the same language, another work [7] has presented an automatic approach to the questions detection among English utterances from multiparty conversational speech. The author conducted several experiments from three distinct classes: lexico-syntactic, turn-related, and Fundamental Frequency related to particular interest in the use of parse tree information in classification. In conclusion, the author stated that lexico-syntactic features seem to enable the classifier to identify correctly the cues that indicate a question.

For the French language, the author [24] attempts to detect the type of sentence with only the lexical approach, using a Memory Based Learning (simple learning method). The author used a dataset composed of human-human's dialogues; it's recorded in a banking call center. The rate of correct detection of dialog acts equal to 84% using the implementation

IB1-IG timbl software with Manhattan method for distance measuring.

In work [31], the authors replicated the old experiences namely [17, 20, 29] that used the lexical and prosodic model to detect the type of sentence. In order to quantify the distinct interrogative sentence' types in different domains for European Portuguese, the authors note that the use of the lexical cues only, results that they are strongly correlated with the detection of a specific type of interrogatives (namely wh-questions). Therefore, the authors concluded that once the acoustic and prosodic features are added, the results get improved significantly.

In Chinese language, the author Yuan and Jurafsky [31] used the prosodic and the lexical approach to detect the type of sentence. However, for the lexical model, the author chose to use the last word in each sentence, because in the Chinese language suffix ('ma') could be added to the end of any affirmative sentence to change it in interrogative sentence, and he used the decision tree for an automatically classification [31, 32]. The author found that the intonation curve is not sufficient to detect the type of sentences in Chinese. The final classification system developed, obtained a correct detection of interrogative sentences equal to 85.1%.

In our previous work in Berber language [11], we studied the prosodic difference between sentences (yes/no question). We developed an automatic detection that uses prosodic Feature (Fundamental Frequency, Energy). These features were used as an input for two different classifiers (Support Vector Machine and neural networks), in order to classify each sentence into either interrogative or affirmative sentence. We classified questions with accuracy equal to 93%. A further feature-specific analysis reveals that energy and fundamental frequency (F0) features are mainly responsible of discriminating between question and affirmative sentences.

Languages other than Arabic have received a lot of attention in this regard, but in recent years some researchers started working in this search path specifically the study [14], in this analysis the author focused on choosing the best feature vectors and the best classifiers to differentiate between interrogative and affirmative sentences. The author failed to consider the prosodic difference between these two types of sentences, as well as he did not take into consideration the type of exclamatory sentences.

To our knowledge, our work is the first attempt made in the domain of automatic detection of exclamatory sentences among the affirmative and interrogative sentences in standard Arabic language it shows the prosodic difference between these three sentences types as several researchers have done in other languages.

2. Feature Set

In order to estimate automatically the nature of sentences, it is possible to analyze the speech signal directly, with no need to the lexical result of an automatic speech recognition engine, by using the prosodic and spectrum features. The target is to make a universal algorithm able to be applied on the other languages, especially poor endowed language.

2.1. Fundamental Frequency (F0)

Pitch is the fundamental frequency of speech signal. Fundamental frequency is an estimation of the periodicity of sound. In the power spectrum, it is the lowest common denominator of the harmonic peaks; Fundamental Frequency is considered the most important set of features in determining types of sentences [14]. There are a number of techniques presented in [30] for pitch extraction.

2.2. Energy (E)

Contrary to the fundamental frequency, energy is considered as the easiest prosodic parameter to calculate. However, it is among the most important parameter in speech signal [26]. The estimation of the energy of a sampled signal $x(t)$ is defined by [26]:

$$E = \sum x(\mathrm{t})^2 \tag{1}$$

2.3. Duration (D)

The duration in the speech processing area is considered as a prosodic feature; it is considered to be a significant prosodic curve for the detection of the type of sentences. Several studies were made in this regard [31].

2.4. Mel-Frequency Cepstral Coefficients (MFCC)

The use of Mel Frequency Cepstral Coefficients could be considered as one of the standard method for feature extraction [21]. MFCC are based on human hearing perceptions which cannot perceive frequencies over 1 KHz. In other words, this parameter is based on a known variation of the human ear's critical bandwidth with frequency [10, 22].In speech processing, 10-12 coefficients are often considered to be sufficient for coding speech [10]. The most notable downside of using MFCC and its sensitivity to noise due to its dependence on the spectral form.

2.5. Linear Predictive Coding (LPC)

Linear Predictive Coding is one of the most powerful speech analysis techniques. It has gained popularity as a formant estimation technique since it was introduced [22]. The 2.5. Linear Predictive Coding (LPC) calculates a power spectrum of the signal. It is used for formant analysis. Its operating principle is related to the speech model in which speech is modeled as the output of a linear, time varying system excited by either quasi-periodic pulses or random noise [2].

3. Prosodic Study

Prosody is an essential element in the domain of sentence type detection. Therefore, our first study is to observe the prosodic differences between the three types of sentences. Finding out this observation will confirm if the sentence in the Arabic standard language conveys extra-linguistic information, as well as the exact location of the helpful information to decode the type of sentence.

The sentences used in the prosodic study are causing a compilation of Compact Disc Read-Only Memory (CDROM) for learning the standard Arabic language for foreigners [1]. It comprises speech read by professional speakers in noiseless conditions. The conversations were recorded in Pulse Code Modulation (PCM) format, 16 kHz, 16 bit, mono, 5 speakers and about 1 hour of speech. We extracted from this speech 20 sentences of each type of pronunciation (declarative, interrogative and exclamatory sentences).The peculiarity of these selection is that all sentences have same words in same location; we removed the lexical indicators of sentences types (نعم، لا، لما، ما، هل، لماذا) ➔ (Why, Do you, where, What, No, Yes) etc., This choice allows us to see the differences between these three sentences types without influencing of the words:

- Interrogative sentence: " **هل** الجوّ جميل ؟ " ➔ In English: "The weather is nice ?"
- Affirmative sentence: "**نعم** الجوّ جميل" ➔ In English: "**yes,** weather is nice."
- Exclamatory sentence: "الجوّ جميل!" ➔ In English: "The weather is nice!"

The sentences are presented in Table 1

Table 1. Dataset of standard Arabic language.

	Standard Arabic Language Sentences	Translation to English
01	الجوّ جميل(./؟/!)	**The weather is nice (. /?/!)**
02	المنزل بعيد(./؟/!)	**The house is far (. /?/!)**
03	هذا المشهد جميل (./؟/!)	**This scene is beautiful (. /?/!)**
04	السماء زرقاء (./؟/!)	**The sky is blue (. /?/!)**
05	الأستاذ هنا (./؟/!)	**The teacher is here (. /?/!)**
06	السوق مكتظ (./؟/!)	**The market is crowded (. /?/!)**
07	أصلحت السيارة (./؟/!)	**The car was repaired (. /?/!)**
08	توقف تساقط الأمطار(./؟/!)	**It stopped raining (. /?/!)**
09	الطبق لذيذ (./؟/!)	**The dish is delicious (. /?/!)**
10	المعلم بدأ الدرس (./؟/!)	**The teacher began the lesson (. /?/!)**
11	الرحلة جيدة (./؟/!)	**The travel is good (. /?/!)**
12	الفلم طويل (./؟/!)	**The movie is long (. /?/!)**
13	الدرس ممل (./؟/!)	**The lesson is boring (. /?/!)**
14	أكل كل الصحن (./؟/!)	**He finished all the plate (. /?/!)**
15	أولئك جيراننا (./؟/!)	**They are our neighbors (. /?/!)**
16	سجل اللعب هدف (./؟/!)	**The player has scored a goal (. /?/!)**
17	هذه الأرض خصبة ؟(./؟/!)	**This land is fertile (. /?/!)**
18	الشجرة مثمرة (./؟/!)	**The tree is fruitful (. /?/!)**
19	نضجت الثمار (./؟/!)	**The fruits are ripe (. /?/!)**
20	تحصلت على الشهادة (./؟/!)	**I have obtained the certificate (. /?/!)**

3.1. Fundamental Frequency Study

By observing the F0 contour of the different sentences in our dataset, we noticed an important point, for most sentences, the intonation contour of the last word or the second half of it seems to be falling for affirmative sentences (A, B), rising for interrogative sentences (C, D), rising and suddenly falling for exclamatory sentences (E, F). The Figure 1 (A, B, C, D, E, and F) shows F0 contour associated with the corresponding speech signal. We used the praat software to achieve this figure. We note: A: Affirmative sentences. Q: Interrogative sentences. E: Exclamatory sentences.

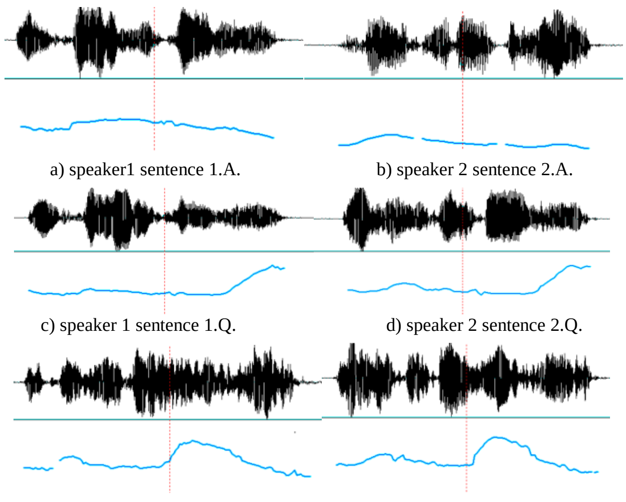

Figure 1. F0 contour with the speech signal corresponding.

3.2. Duration Study

In this experiment, we calculate the average duration for each sentence type. The results in figure 2 confirms that there is a minor difference in duration of utterance between affirmative and interrogative sentences, but there is a significant difference between these two types of sentences and exclamatory sentences. Figure 2 presents the average duration in seconds for each type of sentences

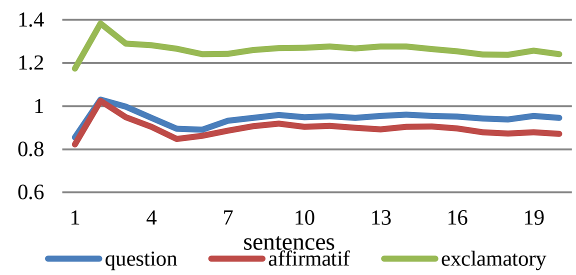

Figure 2. Average duration in seconds for each type of sentences.

3.3. Energy Study

As well; for energy parameter, we calculate the average energy for each sentences type. The results in the Figure 3 confirm that there is significant difference in the energy between different sentences' types, with a significant difference between exclamatory sentences and other types of sentences

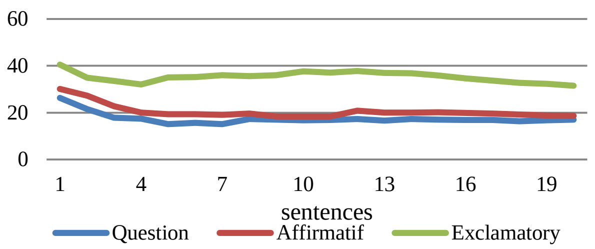

Figure 3. Average duration in (DB) for each type of sentences.

Results confirm that there is a difference between the three sentences types and reveal that the prosody of the sentences in standard Arabic language conveys extra-linguistic information similar to non-tonal languages (English or French as examples), which allows the sentences types identification (interrogative, affirmative and exclamatory).

The next step is to implement an automatic recognition system of sentences type in order to confirm our observations in this section.

4. Automatic Recognition System

In this work, our model for automatic speech recognition system is composed by four major phases: speech input, feature extraction, classification and sentences type output. Figure 4 shows the different phases of our automatic speech recognition system.

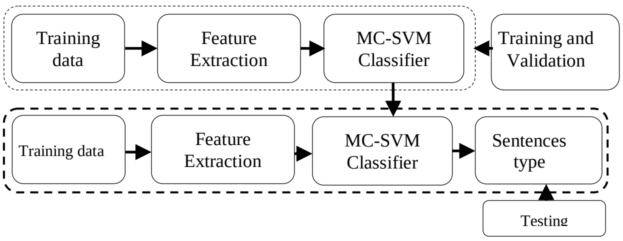

Figure 4. Automatic recognition system.

4.1. Dataset

The standard Arab language is among the languages which often suffers from gaps in linguistic work, among these difficulties lack of dataset. To cross this difficulty, we took again the sentences extracted from CDROM for the training of the standard Arab language of section 3, we recorded a dataset with the same features design (PCM format, 16 KHz, 16 bit, mono). As we have already mentioned, the particularity of these dataset is that all the sentences selected from test have the same words in the same location. These sentences were embedded in meaningful dialogue, so that their pronunciations are as natural as possible, in this way, we eliminate all the phenomena of co-articulation that could interfere with our prosodic analysis. Five Arabic native speakers (2 males and 3 females) who repeats each dialogue five times, we have chosen five speakers to reproduce the original dialogue, this gives us a dataset consisting of 1500 sentences (500 affirmatives sentences, 500 interrogatives sentences and 500 exclamatory sentences).

4.2. Features Generation

The extraction of the best parametric representation of acoustic signals is an important task to produce a better recognition performance. The efficiency of this phase is important for the classification phase. The features extracted include: Energy, Fundamental Frequency, Duration, Linear Prediction Coding and Mel-Frequency Cepstral Coefficients. In this work, the energy of a signal is calculated on a short-time basis, by windowing the signal at 30ms, squaring the values and taking the average and we exploited the duration of the whole utterance, the average duration of syllables and the flow of each sentence. We used the software Praat to extract Energy and Fundamental Frequency and the software MATLAB to extract Duration, MFCC and LPC.

4.3. Classification

4.3.1. Review of Multi Class Support Vector Machine Classifier (MC-SVM)

Initially, SVMs were developed to perform binary classification (two classes) [8, 13]. However, applying the binary classification is very limited for example; in our case, we have three classes (declarative, interrogative and exclamatory sentences). In literature, there are many approaches to generate MC-SVMs [13] from binary SVMs have been. In our case, we adopted one against one approach. The result of study [8] concluded that one against one approach gives better results. This method consisted for G classes, there will be binary classifiers. The output from each classifier in the form of a class label is obtained. The class label occurring the most is assigned to that point in the data vector. In case of a tie, a tie-breaking strategy may be carrying out; this strategy is to select randomly one of the class labels that are tied [8, 13]. This classifier uses learning and testing dataset. In our case, learning dataset presents (60%) of dataset, and testing dataset presents (40%) of dataset. There is no intersection between the testing speakers and training dataset.

4.3.2. Validation System

MC-SVMs have many qualities that distinguish them from many other machine learning algorithms, including the non-existence of local minima, the speed of calculation, and the use of only two tuning parameters [5, 9]. These two parameters are defined by the cross validation method, in order to control the compromise between the complexity of the machine and the number of non-separable data.MC-SVM classifier used in this work is implemented in the MATLAB software

5. Results and Discussion

In this section, the performances of the various features vectors used here for the task of affirmative, interrogative and exclamatory identification. Namely F0, Energy, duration, LPC and MFCC are compared to each other using criteria based on classification accuracy. Thus, we tested several segments of the sentence in order to deduce the useful information location.

5.1. Prosodic Features

Table 2 shows the results of the average rate of correct identification using prosodic features we note that:

- A.R.C.I.A: Average Rate of Correct Identification of Affirmations Sentences.
- A.R.C.I.I: Average Rate of Correct Identification of Interrogative Sentences.
- A.R.C.I.E: Average Rate of Correct Identification of Exclamatory Sentences.

However, the results of the automatic study, presented in Table 2, confirm the hypothesis developed in the part 3. The useful information is located in the end of the sentence “The last syllable of the sentences” with an overall rate of 76.10 %.Furthermore, the prosodic features distinguish significantly the type of sentences and the curve of the fundamental frequency for interrogative sentences having a growing edge at the end of the sentence (85.5%). Falling for affirmative sentences (88%), and growing and suddenly falling for exclamatory sentences (93.5%). We notice that exclamatory sentences are better identified with respect to affirmative and interrogative sentences; this is mainly due to the particular pronunciation in the last syllable to other types of sentences. It seems therefore confirmed that the macro-prosody influences the overall prosody sentence in the Standard Arabic language. In addition, we found that the fundamental frequency is more significant with an average rate of correct identification equal to 89 %, compared to the energy (69.16%) and the duration (70.16%).This study allowed us to conclude that in the case of the Standard Arabic language, the slope of the evolution of the last syllable curve is a decisive factor in the sentences detection

Table 2. Average rate of correct identification using prosodic features.

The whole sentences				
	A.R.C.I.A (%)	**A.R.C.I.I (%)**	**A.R.C.I.E (%)**	**A.R.C.I.G (%)**
F0	83.5	81	88.5	84.33
E	71.5	63.5	66.5	67.16
D	53.5	77	75	68.5
average classification	69.5	73.83	76.66	73.33
The first half of the sentences				
	A.R.C.I.A (%)	**A.R.C.I.I (%)**	**A.R.C.I.E (%)**	**A.R.C.I.G (%)**
F0	51.5	48	49.5	49.66
E	54	53.5	48	51.83
D	50.5	41.5	44.5	45.5
average classification	52	47.66	47.33	49
The second half of the sentences				
	A.R.C.I.A (%)	**A.R.C.I.I (%)**	**A.R.C.I.E (%)**	**A.R.C.I.G (%)**
F0	83.5	80	90.5	84.66
E	73	59.5	69	67.16
D	52.5	77.5	78.5	69.5
average classification	69.66	72.33	79.33	73.77
The last syllable of the sentences				
	A.R.C.I.A (%)	**A.R.C.I.I (%)**	**A.R.C.I.E (%)**	**A.R.C.I.G (%)**
F0	88	85.5	93.5	89
E	74.5	64	69	69.16
D	53	78	79.5	70.16
average classification	71.83	75.83	80.66	76.10

5.2. Spectrum Features

Table 3 shows the results of the spectrum features tested on our database.

Table 3. Average rate of correct identification using spectrum features.

The whole sentences				
	A.R.C.I.A (%)	**A.R.C.I.I (%)**	**A.R.C.I.E (%)**	**A.R.C.I.G (%)**
LPC	71	68.5	83	74.16
MFCC	93	89.5	99.5	94
average classification	82	79	91.25	84.08
The first half of the sentences				
	A.R.C.I.A (%)	**A.R.C.I.I (%)**	**A.R.C.I.E (%)**	**A.R.C.I.G (%)**
LPC	50.5	48	49	49.16
MFCC	39.5	52	41.5	44.33
average classification	45	50	45.25	46.75
The second half of the sentences				
	A.R.C.I.A (%)	**A.R.C.I.I (%)**	**A.R.C.I.E (%)**	**A.R.C.I.G (%)**
LPC	72.5	69.5	85	75.66
MFCC	94.5	94	99.5	96
average classification	83.5	81.75	92.25	85.83
The last syllable of the sentences				
	A.R.C.I.A (%)	**A.R.C.I.I (%)**	**A.R.C.I.E (%)**	**A.R.C.I.G (%)**
LPC	73.5	73	88	78.16
MFCC	94.5	95	99.5	96.33
average classification	84	84	93.75	87.25

Likewise, we noticed also that the useful information is located in the end of the sentence and the exclamatory sentences in this case are better identified (93.75%) compared to affirmative (84%) and interrogative sentences (84%). we also found that MFCC is the best element to detect the type of sentence in our case. We can see from Table 3 that, the spectrum features have given very satisfactory results in this work.

5.3. Fusion of Feature Vectors

To exploit the strengths of each parameter, we opted for the hybrid method of prosodic and spectrum features. This method is based on the idea that certain characteristics of the speech signal are further highlighted by certain parameter sets. The interest is to exploit the advantages of each parameter.

Table 4, shows the rate of correct identification by hybrid method using MC-SVM classifier. We tested the first three parameters that gave the best recognition rates namely F0, LPC and MFCC

We note that:

- Model 1: (F0 + MFCC)
- Model 2: (F0 + LPC)
- Model 3: (MFCC + LPC)
- Model 4: (F0 + MFCC+ LPC

Table 4. Correct identification of the hybrid method using MC-SVM classifier.

	A.R.C.I.A	A.R.C.I.I	A.R.C.I.E	A.R.C.I.G
Model 1	95.50	95.50	99	96.66
Model 2	91.50	89	93	91.16
Model 3	93	88	94.50	91.83
Model 4	94.5	96.5	97	96

To evaluate results of the fusion method of feature vectors as well as to find the best combination of parameters, we will plot comparative graphs. Figures 5 show the comparison of the average rate of correct identification by the hybrid method.

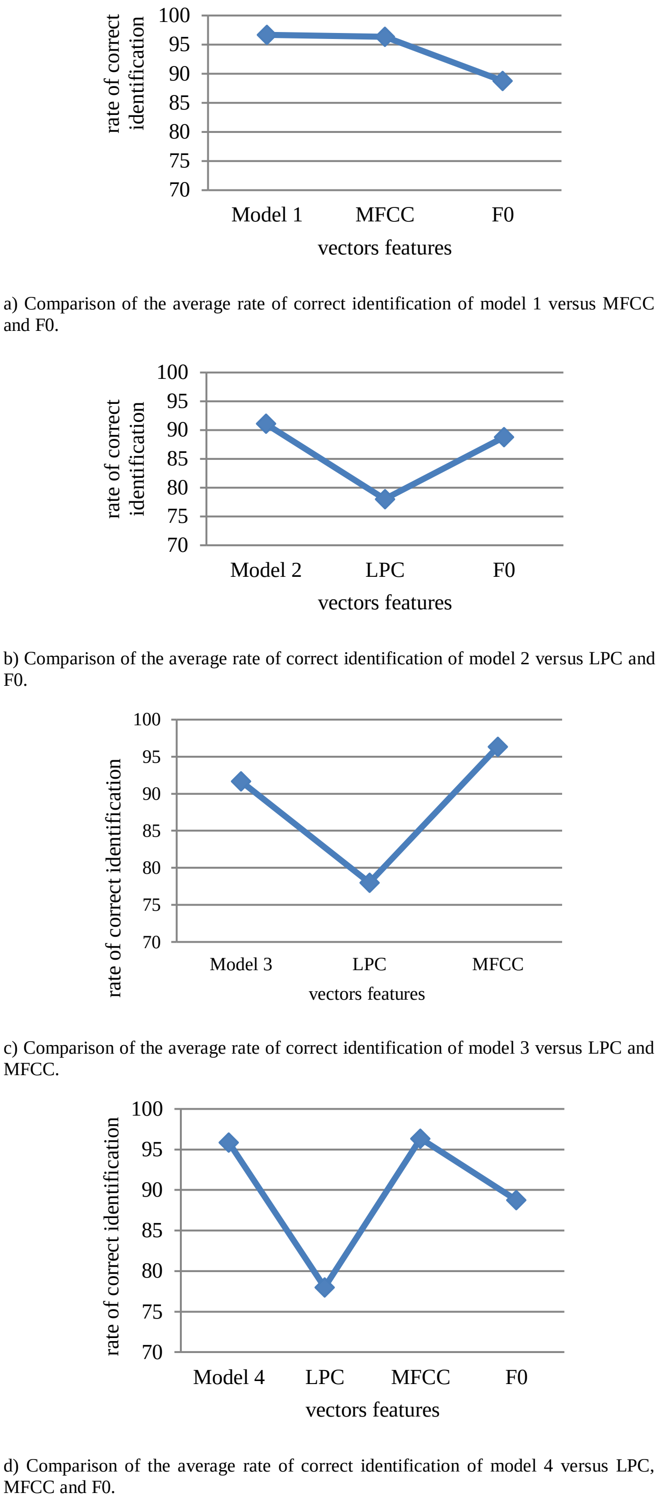

Figure 5. Comparison of the average rate of correct identification.

The results confirm that the fusion of feature vectors improves significantly the results such as model 1 (F0 U MFCC) and model 2 (F0 U LPC). But this is not always the case; the fusion parameters with a large gap of correct identification such as model 3 (MFCC U LPC) and model 4 (F0 U MFCC U LPC) gives rates of correct identification overall lower compared to the parameters included in the models (F0 and MFCC). The results obtained in this work confirm that the prosodic and spectrum features give satisfactory results in this language.

In the Standard Arabic language, to our knowledge, no statistical study on a specific dataset which includes the three types of sentences (interrogative, exclamatory and affirmative); except [14] who have worked on the detection of question in Standard Arabic language. For this we have not been able to compare the rate of correct identification.

6. Conclusions and Future Work

In this paper, we presented our study of sentences type detection in Standard Arabic language. In a first step, our analyses helped to understand the prosodic difference between the types of pronunciation. After, we elaborated a specific dataset containing 1500 affirmative, interrogative and exclamatory sentences. We used this dataset for extract the feature vectors based on several parameters (F0, F, D, MFCC and LPC). Besides, we used the concatenation of feature vectors in order to improve our results. In a last step, we have chosen the MC-SVM to classify our sentences.

The novelty introduced in this work is the introduction of an automatic method for detecting exclamatory sentences in Standard Arabic language, using the prosodic features and spectrum features in Standard Arabic language, using ~~a~~ special dataset, (the three sentences are identical except pronunciation).

The results confirm that the prosody of the sentence Standard Arabic language conveys extra linguistic information similarly to non- tonal languages (including English or French as examples), which allows the identification of types of sentences.

In continuation to the present work, the next objective consists of exploring the lexical parameters to improve outcomes for that language and applying the most recent and most appropriate methods to improve results. With this in mind, we started recording other datasets such as regional languages in order to study the prosodic differences for the poorly endowed languages.

References

[1] American Univ in Cairo, *Modern Standard Arabic Vocab Clinic*, American University in Cairo Press, 2005.

[2] Awasthy N., Saini J., and Chauhan D., "Spectral Analysis of Speech: A New Technique," *International Journal of Electrical, Computer, Energetic, Electronic and Communication Engineering*, vol. 2, no. 7, pp. 1517-1526, 2008.

[3] Bänziger T. and Scherer K., "The Role of Intonation in Emotional Expressions," *Speech Communication*, vol. 46, no. 3-4, pp. 252-267, 2005.

[4] Bazillon T., Maza B., Rouvier M., Bechet F., and Nasr F., "Speaker Role Recognition Using Question Detection and Characterization," *in Proceedings of INTERSPEECH*, Florence, pp. 1333-1336, 2011.

[5] Bishop C., *Pattern Regression and Machine Learning*, Springer, 2006.

[6] Blanchard N., Donnelly P., Olney A., Samei B., Ward B., Sun X., Kelly S., Nystrand M., and D'Mello S., "Identifying Teacher Questions Using Automatic Speech Recognition in Classrooms," *in Proceedings of the SIGDIAL*, Los Angeles, pp. 191-201, 2016.

[7] Boakye K., Favre B., and Hakkani-tur D., "Any Questions? Automatic Question Detection in Meetings," *in Proceedings of IEEE Workshop in Automatic Speech Recognition and Understanding*, Merano, pp. 485-489, 2009.

[8] Chamasemani F. and Singh Y., "Multi-Class Support Vector Machine (SVM) Classifiers-An Application in Hypothyroid Detection and Classification," *in Proceedings of Bio-Inspired Computing: Theories and Applications 6th International Conference Proceedings*, Penang, pp. 351-356, 2011.

[9] Cristianini N. and Shawe-Taylor J., *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*, Cambridge University Press, 2000.

[10] Hagen A., Connors D., and Pellm B., "The Analysis and Design of Architecture Systems for Speech Secognition on Modern Handheld-Computing Devices," *in Proceedings of 1st IEEE/ACM/IFIP International Conference*, Newport Beach, pp. 65-70, 2003.

[11] Halimouche R., Teffahi H., and Falek L., "Detection of Questions in Berber Language Using Prosodic Features," *in Proceedings of International Conference on Multimedia Computing and Systems*, Marrakech, pp. 197-200, 2014.

[12] Hasan M., Doddipatla R., and Hain T., "Multi-Pass Sentence-End Detection of Lecture Speech," *in Proceedings of INTERSPEECH*, Singapore, pp. 2902-2906, 2014.

[13] Hsu C. and Lin C., "A Comparison of Methods for Multiclass Support Vector Machines," *IEEE Transactions on Neural Networks*, vol. 13, no. 2, pp. 415-425, 2002.

[14] Khan O., Al-Khatib W., and Cheded L., "A Preliminary Study of Prosody-Based Detection of Questions in Standard Arabic Language Speech Monologues," *The Arabian Journal for Science and Engineering*, vol. 35, no. 2C, pp. 167-181, 2010.

[15] Kolář J. and Liu Y., "Automatic Sentence Boundary Detection in Conversational Speech: A Cross-Lingual Evaluation on English and Czech," *in Proceedings of IEEE International*

Conference on Acoustics, Speech and Signal Processing, Dallas, pp. 5258-5261, 2010.

[16] Lee C. and Narayanan S., "Toward Detecting Emotions in Spoken Dialogs," *IEEE Transactions on Speech and Audio Processing*, vol. 13, no. 2, pp. 293-303, 2005.

[17] Liscombe J., Venditti J., and Hirschberg J., "Detecting Question-Bearing Turns in Spoken Tutorial Dialogues," *in Proceedings of the International Conference of Spoken Language Processing*, 2006.

[18] Liu Y., Shriberg E., Stolcke A., Hillard D., Ostendorf M., and Harper M., "Enriching Speech Recognition With Automatic Detection of Sentence Boundaries and Disfluencies," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 14, no. 5, pp. 1526-1540, 2006.

[19] Margolis A. and Ostendorf M., "Question Detection In Spoken Conversations Using Textual Conversations," *in Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Short Papers*, Portland, pp. 118-124, 2011.

[20] Moniz H., Batista F., Trancoso I., and Mata A., "Analysis of Interrogatives in Different Domains," *in Proceedings of the 3rd International Training School Conference on Toward Autonomous, Adaptive, and Context-Aware Multimodal Interfaces: Theoretical and Practical Issues*, Caserta, pp. 134-146, 2010.

[21] Motlíček P., *Feature Extraction in Speech Coding and Recognition*, Oregon Graduate Institute of Science and Technology, 2002.

[22] Muda L., Begam M., and Elamvazuthi I., "Voice Recognition Algorithms Using Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time Warping (DTW) Techniques," *Journal of Computing*, vol. 2, no. 3, pp. 138-143, 2010.

[23] Raj S., Rehman Z., Rauf S., Siddique R., and Anwar W., "An Artificial Neural Network Approach for Sentence Boundary Disambiguation in Urdu Language Text," *The International Arab Journal of Information Technology*, vol. 12, no. 4, pp. 395-400, 2015.

[24] Rosset S., Tribout D., and Lamel L., "Multi-Level Information and Automatic Dialog Acts Detection in Human-Human Spoken Dialogs," *Speech Communication*, vol. 50, no. 1, pp. 1-13, 2008.

[25] Shriberg E., Stolcke A., Jurafsky D., Coccaro N., Meteer M., Bates R., Taylor P., Ries K., Martin R., and Dykema C., "Can Prosody Aid the Automatic Classification of Dialog Dcts in Conversational Speech?," *Language and Speech*, vol. 41, no. 3-4, pp. 443-492, 1998.

[26] Sinha P., *Speech Processing in Embedded Systems*, Springer, 2010.

[27] Stolcke A., Coccaro N., Bates R., Taylor P., Ess-Dykema C., Ries K., Shriberg E., Jurafsky D., Martin R., and Meteer M., "Dialogue Act Modeling for Automatic Tagging and Recognition of Conversational Speech," *Journal Computational Linguistics*, vol. 26, no. 3, pp. 339-373, 2000.

[28] Varpe D., Sawant N., Shah S., Shivarkar S., and Pawar P., "Improve HCII by Using Gender and Emotion Recognition Using Speech Signal: A Survey," *International Journal of Advanced Research*, vol. 4, no. 3, pp. 1344-1347, 2016.

[29] Venditti J., Hirschberg J., and Liscombe J., "Intonational Cues To Student Questions in Tutoring Dialogs," *in Proceedings of INTERSPEECH*, Pittsburgh, pp. 549-552, 2006.

[30] Ververidis D., Kotropoulos C., and Pitas I., "Automatic Emotional Speech Classification," *in Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, Montreal, pp. 1-593, 2004.

[31] Yuan J. and Jurafsky D., "Detection of Questions in Chinese Conversational Speech," *in Proceedings of IEEE Workshop on Automatic Speech Recognition and Understanding*, San Juan, pp. 47-52, 2005.

[32] Yuan J., Shih C., and Kochanski G., "Comparison of Declarative and Interrogative Intonation in Chinese," *in Proceedings of Speech Prosody*, Aix-en Provence, 2002.

Ramzi Halimouche is PhD student in telecommunication and information processing at the university of Science and Technology Houari Boumediene in Algiers Algeria he has a Master's degree in Intelligent and communicates systems. Currently he is working on the influence of prosody on dialogue acts. In particular, the automatic classification of interrogative and affirmative sentences, the analysis parameters and adequate classifiers.

Hocine Teffahi received his Magister and doctoral in electrical engennering from the University of Science and Technology Houari Boumedienne, Algiers, Algeria. Currently, he is teacher-researcher as a full professor. His research interest on prosody of the speech signal. He is responsible of doctoral formation. He co-authored many paper in international peer-reviewed journal and conferences.