Algerian Arabic Speech Database (ALGASD): Corpus design and automatic speech recognition application

Article *in* ARABIAN JOURNAL FOR SCIENCE AND ENGINEERING · December 2010

3 authors:

Ghania Droua-Hamdani
Centre de Recherche Scientifique et Technique pour le Développement de la Lang…

Sid Ahmed Selouani
Université de Moncton

Malika Boudraa
University of Science and Technology Houari Boumediene

ALGERIAN ARABIC SPEECH DATABASE (ALGASD): CORPUS DESIGN AND AUTOMATIC SPEECH RECOGNITION APPLICATION

Ghania Droua-Hamdani*[1]
Speech Processing Laboratory, CRSTDLA, Algiers, Algeria

Sid Ahmed Selouani[2]
LARIHS Laboratory, University of Moncton, Canada

and Malika Boudraa[3]
Speech Communication Laboratory, USTHB University, Algiers, Algeria

الخلاصـة:

يتناول هذا المقال تصميم قاعدة بيانات صوتية باللغة العربية الفصحى المنطوقة في الجزائر (ALGASD). وقد نُطقت نصوص هذه القاعدة من قبل 300 متكلم جزائري اختيروا من 11 منطقة من الجزائر. إن أحد أهداف هذه المدونة الصوتية هو تمثيل اللهجات الإقليمية للغة العربية الفصحى الحديثة المتداولة في الجزائر. كما تحتوي على معلومات مفيدة عن المتكلمين مثل العمر والجنس ومستوى التعليم. ويتعرض هذا المقال أيضا لتقارير عن نتائج تطبيقية في الاستكشاف الآلي للكلام المستمر مع استعراض أهم الخطوات المتبعة في تصميم أنموذج أحادي الصوتيات لمعالجة التغير اللغوي. معدل نتائج الاستكشاف الآلي للكلام لهذا النظام المرجعي مرض جدا ويمكن أن يُستخدم كأساس لأنظمة الاستكشاف الآلي للكلام المخصصة للغة العربية.

***Classification*:** electrical engineering, signal processing, automatic speech recognition, speech database

---

*Corresponding Author:
E-mail: gh.droua@post.com

Paper Received July 11, 2010; Paper Revised December 6, 2010; Paper Accepted December 15, 2010

ABSTRACT

This paper presents the Algerian Arabic Speech Database (ALGASD), a Modern Standard Arabic (MSA) speech corpus composed of utterances pronounced by 300 Algerian native speakers selected from eleven regions of Algeria. One of the objectives of this corpus design is to be representative of the regional accents of MSA spoken in Algeria. Useful information related to the speakers, such as gender, age, and education level, is provided. This paper also reports the results of the Automatic Speech Recognition (ASR) application of the corpus and outlines an original global monophone recognition model designed to handle linguistic variability. The global phone recognition rate for this ASR reference system is satisfactory and may constitute a useful baseline ASR system dedicated to MSA.

***Key words*:** speech corpus, Algerian speakers, modern standard Arabic, automatic speech recognition, hidden Markov models

ALGERIAN ARABIC SPEECH DATABASE (ALGASD): CORPUS DESIGN AND AUTOMATIC SPEECH RECOGNITION APPLICATION

1. INTRODUCTION

Oral corpora are required in various applications of speech and language processing. For instance, modern Automatic Speech Recognition (ASR) systems use statistical models that are trained on corpora of relevant speech. Besides this, the development of ASR systems requires speech resources from a large number of speakers to achieve acceptable performance. Arabic speech corpora are relatively less numerous compared to the linguistic resources devoted to other major spoken languages in the world. In recent years, the Language Data Consortium (LDC) published the first public MSA speech corpus that was designed for speech recognition experiments. This corpus, called West Point, contains speech data that was collected and processed by members of the Department of Foreign languages at the United States Military Academy at West Point. The original purpose of this corpus was to train acoustic models for automatic speech recognition that could be used as an aid in teaching Arabic to West Point cadets. Another LDC resource provides spontaneous telephone dialogs collected from Egyptian, Syrian, Palestinian, Lebanese, and Jordanian speakers [1]. BBN/AUB is a corpus proposed by LDC and concerns the Levantine Arabic which is spoken by ordinary people in Lebanon, Jordan, Syria, and Palestine. It is significantly different from MSA Arabic in that it is a spoken rather than a written form. ELRA is another institution that provides relevant corpora for both MSA and colloquial Arabic. ELRA is involved in the NEMLAR project, which produced a broadcast news speech corpus consisting of about 40 hours of Standard Arabic news broadcasts [2]. The GlobalPhone project also provides an oral corpus from political and economic newspapers [3]. An important package of speech resources concentrating on both MSA and modern colloquial Arabic is OrienTel, which is driven by an international industrial and academic consortium. Orientel is recorded in many Arab countries: Saudi Arabia, UAE, Morocco, Tunisia, Egypt, and Palestine. All OrienTel databases are recorded from fixed and mobile networks via ISDN lines and multiple channels [4]. Another relevant Arabic speech resource collected on a telephone network is the Saudi Accented Arabic Voice Bank (SAAVB) [5]. This corpus is an example of a normal life speech corpus. The speech data was acquired from 1,033 Saudi native speakers and recorded over the Saudi mobile and fixed-line telephone network, covering all regions of the kingdom.

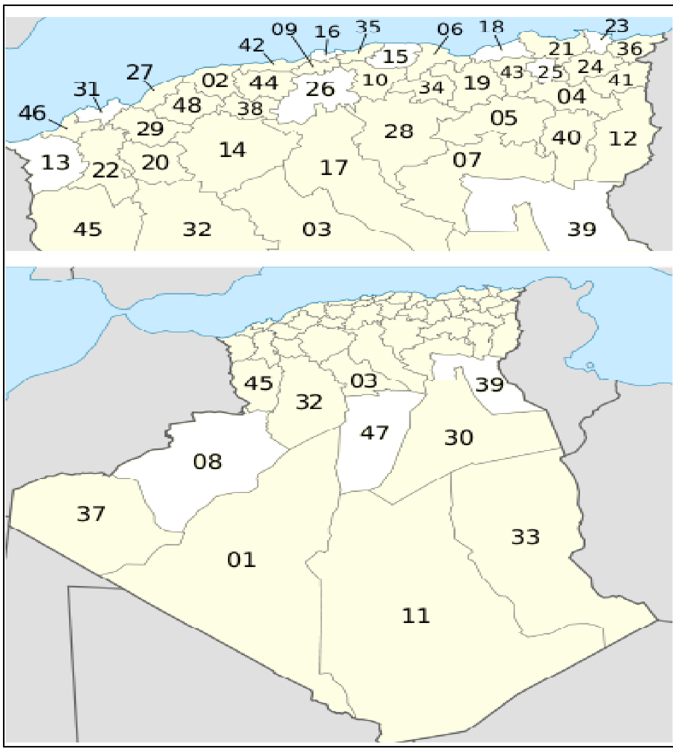
In this paper, we present a new MSA linguistic resource, the *ALGerian Speech Database* (ALGASD), a speech collection which mirrors the main features of pronunciation related to regional and social variations of Algerian speakers. We describe the design and structure of the corpus and we evaluate its usefulness for the continuous speech recognition of the Arabic language. We begin with an overview of regional linguistic variations of Arabic in Algeria, and then present experiments that test an original global monophone recognition model. Finally, we conclude our work and give some perspectives on research applications of the corpus.

2. ARABIC LANGUAGE BACKGROUND

Arabic is an official language in more than 22 countries. The estimated number of Arabic speakers is about 300 million. However, a greater number of speakers have a passive knowledge of Arabic since it is the language of instruction in Islam. Recent approaches in language and speech processing categorize the Arabic language as Modern Standard Arabic (MSA) and Modern Colloquial Arabic (MCA). Modern Standard Arabic is the form of Arabic that is used in education, media, and formal talks. Colloquial Arabic is what is spoken in everyday conversation and varies considerably not only across countries, but also within the same country. It has many differences when compared with Indo-European languages. Some of the differences include unique phonemes and phonetic features, and a complex morphological word structure. The phonetic system of MSA has basically 34 phonemes: 6 vowels (3 short vowels with 3 opposite long ones) and 28 consonants. Among these consonants, there are two distinctive classes, which are named pharyngeal and emphatic phonemes. In addition to the [madd] (long vowels) and emphasis [tafχim], there is an important Arabic characteristic, which is the *Gemination* [taʃdid] [6].

3. MSA VARIATIONS IN ALGERIA

Algeria is located in North Africa, bordering the Mediterranean Sea. It also shares borders with Morocco, Mauritania, Mali, Niger, Libya, and Tunisia. It extends over an area of 2,380,000 km^2 divided into 48 provinces (Wilayas) as illustrated in Figure 1. About 34.8 million people live in Algeria.



Code	Wilaya	Code	Wilaya
01	Adrar	**25**	**Constantine**
02	Chlef	**26**	**Médéa**
03	Laghouat	27	Mostaganem
04	Oum-El-Bouaghi	28	M'Sila
05	Batna	29	Mascara
06	Béjaia	30	Wilaya d'Ouargla
07	Biskra	**31**	**Oran**
08	**Béchar**	32	Wilaya d'El Bayadh
09	Blida	33	Wilaya d'Illizi
10	Bouira	34	Bordj-Bou-Arreridj
11	Tamanrasset	35	Boumerdès
12	Tébessa	36	Wilaya d'El-Tarf
13	**Tlemcen**	37	Tindouf
14	Tiaret	38	Tissemsilt
15	**Tizi-Ouzou**	**39**	**El-Oued**
16	**Alger**	40	Khenchela
17	Djelfa	41	Souk-Ahras
18	**Jijel**	42	Tipaza
19	Sétif	43	Mila
20	Saida	44	Aïn-Defla
21	Skikda	45	Naâma
22	Sidi-Bel-Abbès	46	Aïn-Témouchent
23	**Annaba**	**47**	**Ghardaia**
24	Guelma	48	Relizane

Figure 1: Wilayas (provinces) of Algeria [7]

MSA is the official language of Algeria. It is the language used in schools and administrative institutions (government, media, justice, *etc.*). However, MSA differs substantially from the Algerian spoken dialects. Indeed, approximately 72 % of inhabitants daily use the *Darija,* which is the Algerian Arabic dialect, and 28% of them have a second mother tongue called *Tamazight,* which is a Berber language. These two native languages constitute the basis of oral communication between Algerians. When we refer to dialects, we mean regional variations of spoken Arabic. Tamazight is considered independent from the Arabic language (MSA) and its dialects.

Algerian dialects are variants of MSA stemming from ethnic, geographical, and colonial influences of Spanish, French, Turkish, and Italian. For instance, the MSA spoken in Algiers is largely influenced by Berber and Turkish; the Constantine dialect is affected by Italian; and the Oran dialect by Spanish. As a result, there are significant local variations (in pronunciation, grammar, *etc.*) of spoken Arabic in Algeria, and many of its varieties can be encountered across the country. Examples of well-known differences are observed in the Wilaya of Jijel where the phoneme [q] is replaced by [k]; [ð] is pronounced as [d]; [dˁ] and [ðˁ] are pronounced as [tˁ]; and [θ] and [t] are pronounced as [ts]. In the Wilaya of Tlemcen, [q] is produced as glottal stop [?]. In the spoken Arabic of Oran, [q] is replaced by [g] [8,9].

4. ALGASD DESIGN AND STRUCTURE

Regional linguistic variations of MSA in Algeria have been the focus of only a very small number of studies. These studies are based on partial sociolinguistic and dialectological surveys, and they identify several regional differences. In order to elicit the main features of MSA pronunciations that are influenced by regional and social variations, we considered the major dialect regions of Algeria in the design of the ALGASD corpus. A speaker's dialect region is a geographical distribution within the Algerian mainland. Speakers involved in the project lived during their childhood years in the same area.

4.1. Speakers and Regions

ALGASD aims to be a resource that reflects the main pronunciation variations due to regional and social differences in Algeria. Thus, the regional coverage corresponds to the major dialect groups. The distribution of speakers according to their number and gender is proportional to the population of regions (with 5% tolerance) and with a minimum of speakers per region (ONS, 1998 Census [10]).

The recordings are collected from 300 speakers recruited in eleven distinct regions. Information about gender, age, and education level of speakers is also provided. There are 3 age groups: younger speakers between 18 and 30, middle-aged speakers between 30 and 45, and speakers aged more than 45 years of age. Three categories of education level are also considered: middle group (primary to secondary school), graduate group (university), and post-graduate group. Speakers are native and lived in their selected localities. They come from different socioeconomic backgrounds (doctors, teachers, students, unemployed persons, *etc.*). Table 1 shows the distribution of speakers in ALGASD by gender and region.

Table 1. Speakers' Distribution in Different Regions of ALGASD

Regions	Algerian Regions	Number of Female Speakers	Number of Male Speakers	Total Speakers/Regions
R1	Algiers	40 (50%)	40 (50%)	80 (27%)
R2	Tizi Ouzou	17 (50%)	17 (50%)	34 (11%)
R3	Médéa	13 (52%)	12 (48%)	25 (8%)
R4	Constantine	13 (52%)	12 (48%)	25 (8%)
R5	Jijel	09 (50%)	09 (50%)	18 (6%)
R6	Annaba	09 (52%)	08 (48%)	17 (6%)
R7	Oran	19 (50%)	19 (50%)	38 (13%)
R8	Tlemcen	13 (50%)	13 (50%)	26 (9%)
R9	Bechar	04 (52%)	03 (48%)	07 (2%)
R10	El Oued	08 (50%)	08 (50%)	16 (5%)
R11	Ghardaïa	07 (50%)	07 (50%)	14 (5%)
TOTAL	**11**	**152 (51%)**	**148 (49%)**	**300 (100%)**

4.2. Speech Material

The text material in the ALGASD corpus consists of 200 Arabic Phonetically Balanced (APB) sentences [11]. The elaboration of these phonetically balanced sentences is mainly based on statistics using the $\chi 2$ test. This test is performed thanks to Moussa's statistical analysis that considers the occurrence frequencies of vowels and consonants composing the Arabic word roots [12]. Therefore, the APB sentences reflect the distribution of phonemes used in the whole MSA language. These sentences were divided into 10 lists of 20 sentences. Each sentence contains 208 diphones [CV]. The words used to build the sentences are fairly common and are not part of a specialized lexicon.

APB sentences are originally used to perform speech coding tests and thus, they are not adapted to characterize the variations in pronunciations. In order to cope with this problem, we performed minor modifications that consist of moving some sentences from their original lists to new ones. As a result, two representative "dialect" sentences were selected from the corpus. They were used to elicit specific regional dialect features such as the pronunciation differences of [q], [$d^ʕ$], and [t]. These two sentences were read by all 300 speakers. From the remaining sentences (198), we select 62 as shared sentences. These sentences were read by 86 speakers. The 136 remaining sentences are randomly assigned to be read once by each speaker. Each speaker reads either 2 or 3 or 6 different sentences:

- the two dialect sentences;
- the two dialect sentences and one individual sentence;
- or the two dialect sentences and four of the shared sentences.

The total number of ALGASD recordings reached 1080 controlled read sentences: 600 recordings of the two dialect sentences, 344 of the shared sentences, and 136 of the individual sentences. Table 2 shows the number of sentences read by speakers by region.

Table 2. Numbers of Read Utterances and Speakers by Region

Regions	Number of Read Sentences			Total Speakers
	6	3	2	
R1	23	36	21	80
R2	10	15	9	34
R3	7	11	7	25
R4	7	11	7	25
R5	5	8	5	18
R6	5	8	4	17
R7	11	17	10	38
R8	7	12	7	26
R9	2	5	0	7
R10	5	7	4	16
R11	4	6	4	14
Total Speakers/ Sentences	**86**	**136**	**78**	**300**

The interviewers were students in a Master's degree program who were either native speakers or lived in the different selected regions where data was collected. Before the recording session, each participant was explained the purpose of the ALGASD speech corpus. Recordings were made in quiet environments that were familiar to the speakers. The same conditions of sound recording and protocols were followed for all regions. Instructions and guidelines were given to the interviewers regarding the recording room dimensions, the equipment that permitted a high quality of recording, and the protocols of reading and deleting sentences that contained hesitations.

4.3. Corpus Structure

The ALGASD structure is similar to that used in the TIMIT corpus, but the number of sentences assigned in TIMIT for each speaker was set at 10 [13]. This is not the case in the ALGASD corpus because the number of sentences (200 sentences) is less than that of the speakers (300 speakers). In order to make corpus use easier, a document containing specifications is attached to each speaker directory. These specifications provide information about speakers, namely their region, initials, gender, age, and education level. The texts of read sentences and phonetic transcriptions are also provided.

The wave files are sampled at 16 kHz and stored at .wav format. A text file containing the text of read sentences (.txt), a time-aligned orthographic word transcription file (.wrd), and two time-aligned phonetic transcriptions (.phn) and (.lab) files, are also stored in the same directory. The segmentation and labeling are carried out by experienced people using spectrograms, listening, and visual examination of waveforms, and other parameters such as energy, formants, and fundamental frequency. The transcription of the sentences is performed through the use of *Speech Assessment Methods Phonetic Alphabet* (SAMPA). Table 3 gives a transcription of Arabic phonemes using the SAMPA alphabet [14]. An example of ALGASD specifications is given in Table 4.

Table 3. SAMPA Transcription of Modern Standard Arabic Phonemes

SAMPA	Keyword	English glossary	Orthography	SAMPA	Keyword	English glossary	Orthography
b	ba:b	door	باب	x	xit`a:b	letter	خطاب
t	tis?`	nine	تسع	G	Garb	west	غرب
d	da:r	home	دار	X\	X\ilm	dream	حلم
t`	t`a:bi?`	stamp	طابع	?` (?\)	?`alam	flag	علم
d`	d`arab	he hit	ضرب	h	hawa:?	air	هواء
k	kabi:r	large	كبير	m	ma:l	money	مال
?	?akl	food	أكل	n	nu:r	light	نور
q	qalb	heart	قلب	r	rima:l	sand	رمال
f	fi:l	elephant	فيل	l	la:	no	لا
T	Tala:T	three	ثلاث	w	wa:hid	one	واحد
D	Dakar	male	ذكر	j	jawm	day	يوم
D`	D`ala:m	darkness	ظلام	i	shadow		ظل
s	sa?`i:d	happy	سعيد	a	solution		حل
z	zami:l	colleague	زميل	u	age		عمر
s`	s`aGi:r	small	صغير	i:	feast		عيد
S	Sams	sun	شمس	a:	money		مال
Z	Zami:l	beautiful	جميل	u:	beans		فول

Table 4. ALGASD Specifications

ALGASD- R1-mHAX0-cc1.wav	
ALGASD	Database name
R1	Region 1 which refers to Algiers
mHAX0	Speaker gender (male (m) /female (f)); Speaker initials (HAX). X is replaced by the third initial if the full name is composed of more than two names. Digit from 0-9. It is used to distinguish between speakers that have the same initials.
cc1	Code of the sentence to read
.wav	File format (wave)
.wrd	Time-aligned orthographic word transcription
.phn	time-aligned phonetic transcription

5. AUTOMATIC SPEECH RECOGNITION APPLICATION

Compared to other languages, there has been relatively little speech recognition research devoted to the Arabic language. Previous studies have been carried out on Arabic alpha-digits recognition; isolated Arabic vowels and isolated Arabic word recognition; and more recently on the development of continuous speech recognition systems [15–19].

The ALGASD corpus is intended primarily for the development of original recognition systems for modern standard Arabic spoken in Algeria. In this section, we outline one such system and then discuss experiments that evaluate how well the system performs. Data of 165 speakers corresponding to 570 recordings extracted from R1, R2, R5, R9, R10, and R11 subsets are used throughout the experiments. Both the automatic speech recognition system (described below) and the parameterization method were designed by using the Hidden Markov Model Toolkit (HTK) which runs on a Windows platform [20]. The HTK toolkit is a general-purpose tool designed for the creation of Hidden Markov Models (HMM) [21]. It lends itself perfectly to the creation and evaluation of automatic speech recognition engines.

5.1. Parameterization Model

The parameterization used to provide front-end features consists of extracting Mel-Frequency Cepstral coefficients (MFCCs). These coefficients were calculated from the log filterbank amplitudes $\{m_j\}$ using Discrete Cosine Transform as in the following formula:

$$c_i = \sqrt{\frac{2}{N}} \sum_{j=1}^{N} m_j \cos\left(\frac{\pi i}{N}(j - 0.5)\right) \tag{1}$$

where N is the number of the filterbank channels. The filters used were triangular and were equally spaced along the Mel-scale. The model uses MFCCs, logarithmic energy, and delta coefficients that are the derivatives of the MFCCs calculated through the use of regression analysis according to the following equation:

$$d_t = \frac{\sum_{\theta=1}^{\delta} \theta\left(c_{t+\theta} - c_{t-\theta}\right)}{2\sum_{\theta=1}^{\delta} \theta^2} \tag{2}$$

where d_t was the delta coefficient at time t computed by using the static coefficients $c_{t-\theta}$ to $c_{t+-\theta}$ with a possible delay of δ. The same formula was applied to the delta coefficients to obtain acceleration (delta-delta) coefficients. The performance of the speech recognition system is improved by adding derivative coefficients to static parameters [20].

In our ASR, we used the following configuration:

- 39 dimensional standard feature vector composed of 12 MFCCs and the normalized energy with their delta and acceleration coefficients;
- first-order digital filter with a transfer function $H(z) = 1 - kz^{-1}$ was used for the pre-emphasis processing with $k=0.97$;
- 16 kHz sampling frequency;
- 25-millisecond Hamming window duration with a step size of 10 milliseconds;
- 22 as the length of cepstral liftering;
- 26 filterbank channels;
- is set to 3 in computation of delta coefficients (Equation 2).

5.2. Dictionary

The development of an effective ASR system requires a dictionary which includes all possible phonetic pronunciations of any word in the corpus. Texts of ALGASD were transcribed using the SAMPA alphabet, but some of the phonemes were renamed for machine convenience. Thus, we replaced them with symbols that were proposed in the West Point modern standard Arabic database by LDC [1] with minor modifications. Table 5 gives an overview of the Arabic phonemes list used throughout our experiments.

Table 5. Dictionary Transcription

Transcription	Grapheme	Transcription	Grapheme	Transcription	Grapheme	Transcription	Grapheme
Q_	ء	r	ر	G_	غ	y	ي
b	ب	z	ز	f	ف	ih	الكسرة (ِ)
t	ت	s	س	q	ق	ah	الفتحة (َ)
th	ث	sh	ش	k	ك	uh	الضمة (ُ)
g	ج	S_	ص	l	ل	iy	ـيِ
H_	ح	D_	ض	m	م	ae	ـا
x	خ	T_	ط	n	ن	uw	ـو
d	د	DH_	ظ	h	ه	التنوين	
dh	ذ	C	ع	w	و	in/un/an	ٍ / ٌ / ً

5.3. ASR Experiments and Results

From the speech corpus, we built two subsets to train and test the Arabic recognizer. These two subsets are composed of speaker recordings extracted from six selected regions: R1, R2, R5, R9, R10, and R11. These regions correspond to southern and northern regions of Algeria. In this study, we focus only on these regions because they are characterized by more homogeneous phonetic features and fewer phonologic differences to achieve a high recognition rate. The specificities of the regions located in Eastern (close to Tunisia) and Western (close to Morocco) Algeria are characterized by important phonological differences, compared to MSA. The inclusion of these regions may lead to a decrease in the recognition performance and, thus, the HMM-based ASR system cannot be used as a reference system for further studies.

The system was trained using a total of 413 sentences, which represents two-thirds of the total sentences used for the development of the ASR system. The number of speakers used for the training was 115 participants. HMM Toolkit was used to train the acoustic models of all thirty-four Modern Standard Arabic (MSA) phonemes, to which we added a model of silence (sil). A short pause (sp) model is created from the silence model and tied to it. All the models were context independent, 5-state HMM (first and fifth states were non-emitting) left to right without skip state, all with one Gaussian mixture (diagonal covariance) per state. The optimization of the monophone models was performed through nine Baum–Welch re-estimations [22,23]. An alignment of speech data was done after the seventh re-estimation using the Viterbi algorithm [22,23]. A bigram language model was built from labels used in the training process. Using the same front-end as for the training phase, the test corpus includes 50 speakers representing all selected regions (R1, R2, R5, R9, R10, and R11). The test set is composed of 157 sentences that had not been used in the training phase. The speaker-independent ASR performance is 92.11%, which is a very satisfactory result. The accuracy rate is 91.65 %. The total number of labels is 431 and the number of substitutions is 28, while the number of deletions is 6 and the number of insertions is 2.

6. CONCLUSION AND PERSPECTIVES

The availability of relevant spoken corpora is an important step towards the development of reliable speech-enabled applications. This paper has introduced ALGASD, a speech corpus dedicated to MSA speech processing applications, especially speech recognition of MSA spoken in Algeria. ALGASD contains recordings of 300 Algerian native speakers selected from eleven regions of the country. Information about speakers (gender, age, and education level) is provided within the corpus. This information and ALGASD data are a potential resource for acoustic-phonetic and socio-phonetic studies that examine MSA accent variations. The global monophone speech recognition model described in the present study is successful and might constitute a useful baseline model for further studies using complex ASR systems dedicated to MSA. The ALGASD corpus has numerous strengths: large number of speakers, high quality recordings, useful information about speakers, and integration of regional phonetic variations. All these characteristics provide an interesting set of perspectives for speech and language processing applications, such as automatic speech recognition, acoustic phonetic analysis, perceptual experiments to study classification of the different regional varieties, and a comparison of Algerian pronunciation of MSA with other Arabic accents such as Maghreb Arabic or the eastern ones. Nowadays, the ALGASD corpus is used in different research fields that are related to acoustical and statistical studies of qualitative and quantitative vocalic variations according to the education level of the speakers [24], and in a prosody research initiative targeting Arabic rhythm classification [25].

REFERENCES

[1] Linguistic Data Consortium (LDC): http://www.ldc.upenn.edu.

[2] K. Choukri, M. Nikkhou, and N. Paulsson, "Network of Data Centres (NetDC) BNSC - An Arabic Broadcast News Speech Corpus", in *Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC04), Lisbon, Portugal,* May 2004.

[3] T. Schultz and A. Waibel, "GlobalPhone, Das Projekt GlobalPhone: Multilinguale Spracherkennung Computers, Linguistics, and Phonetics Between Language and Speech", in *ISBN Proceedings of the 4th Conference on NLP - Konvens-98, Bonn, Germany*, October, 1998, ed. Bernhard Schroder et al. Berlin: Springer, 1998.

[4] Language Resources Association: http://www.icp.grenet.fr/ELRA.

[5] M. Alghamdi, F. Alhargan, M. Alkanhal, A. Alkhairy, M. Eldesouki, and A. Alenazi, "Saudi Accented Arabic Voice Bank", *Journal of King Saud University: Computer Sciences and Information*, Riyadh, 1429H, 2008, pp. 43–58.

[6] J. F. Bonnot, "Étude Expérimentale de Certains Aspects de la Gémination et de l'Emphase en Arabe", *Travaux de l'Institut Phonétique de Strasbourg*, **11**(1979), Strasbourg, France, pp. 109–118.

[7] Algerian map downloaded from: http://maps.nationalgeographic.com/maps.

[8] P. Marçais, *Textes Arabes de Djidjelli.* Presse Universitaire de France, 1954.

[9] D. Caubet, "Questionnaire de Diactologie du Maghreb (D'après les Travaux de W. Marçais, M. Cohen, G. S. Colin, J. Cantineau, D. Cohen, P. Marçais, S. Levy, *etc.*) ", *Estudios de Dialectologia Norteafricana y Andalusi*, 2000–2001, pp. 73–92.

[10] Algerian National Office of Statistics: http://www.ONS.dz

[11] M. Boudraa, B. Boudraa, and B. Guerin, "Twenty Lists of Ten Arabic Sentences for Assessment", *ACUSTICA, ACTA-ACUSTICA*, **86(5)**(1998).

[12] A. H. Moussa, "Statistical Study of Arabic Language Roots in Moεjam Al-Sehah", Kuwait University, 1973.

[13] Texas Instruments and Massachusetts Institute of Technology corpus (TIMIT): Acoustic-Phonetic Continuous Speech Corpus, DMI, 1990, p. 165.

[14] The Speech Assessment Methods Phonetic Alphabet (SAMPA): http://www.phon.ucl.ac.uk/home/sampa/arabic.htm

[15] Y. A. Alotaibi, S.-A. Selouani, and D. O'Shaughnessy, "Experiments on Automatic Recognition of Non-Native Arabic Speech", *EURASIP Journal on Audio, Speech, and Music Processing*, Article ID 679831, 9 pages, doi:10.1155, 2008.

[16] M. Elshafei, H. Al-Muhtaseb, and M. Al-Ghamdi, "Speaker-Independent Natural Arabic Speech Recognition System", *The International Conference on Intelligent Systems ICIS 2008, Bahrain*, December 2008.

[17] M. Elmahdy, R. Gruhn, W. Minker, and S. Abdennadher, "Modern Standard Arabic Based Multilingual Approach for Dialectal Arabic Speech Recognition", *International Conference on Speech and Language Processing (Interspeech), Makuhari, Japan,* September 2010.

[18] D. Vergyri, K. Kirchhoff, K. Duh, and A. Stolcke, "Morphology-Based Language Modeling for Arabic Speech Recognition", in *INTERSPEECH- 2004*, 2004, pp. 2245–2248.

[19] Y. A. Alotaibi, M. Alghamdi, and F. Alotaiby, "Speech Recognition System of Arabic Alphabet Based on a Telephony Arabic Corpus", *Lecture Notes in Computer Science*, **6134**(2010), pp. 122–129, DOI: 10.1007/978-3-642-13681-8_15.

[20] S. Young, G. Evermann, D. Kershaw, G. Moore, J. Odell, D. Ollason, V. Valtchev, and P. Woodland, "The HTK Book (for HTK Version. 3.4)", Cambridge University Engineering Department, 2006, http:///htk.eng.cam.ac.uk/prot-doc/ktkbook.pdf, p. 165.

[21] F. Jelinek, *Statistical Methods for Speech Recognition*. Massachusetts Institute of Technology (MIT), second printing, 1999.

[22] L. Rabiner and B. H. Juang, *Fundamentals of Speech Recognition*. Englewood Cliffs, New Jersey 07632: Prentice Hall PTR, 1993.

[23] X. Huang, A. Acero, and H-W. Hon, *Spoken Language Processing, A Guide to Theory, Algorithm and System Development*, Upper Saddle River, New Jersey 07458: Microsoft Research, Prentice Hall PTR, 2003.

[24] G. Droua-Hamadani, S. A. Selouani, M. Boudraa, and B. Boudraa, "ALGASD PROJECT: Statistical Study of Vocalic Variations According to Education Levels of Algiers Speakers", *Intonational Variation in Arabic Conference IVA09, September 28–29, York, England*, 2009.

[25] G. Droua-Hamdani, S. A. Selouani, M. Boudraa, and W. Cichocki, "Algerian Arabic Rhythm Classification", *Tutorial and Research Workshop on Experimental Linguistics, Exling 2010, International Speech Communication Association (ISCA), August 25–27, Athens, Greece*, 2010.