

# The Challenges of Developing a Living Arabic Phonetic Dictionary for Speech Recognition System: A Literature Review

Dr. Mohammad Husam Alhumsi\*, Dr. Saleh Belhassen

Saudi Electronic University



\*Corresponding author email: [husam101010@gmail.com](mailto:husam101010@gmail.com)

Received: 04 April 2021 / Revised: 15 April 2021 / Accepted: 22 May 2021 / Published: 03 June 2021

## ABSTRACT

Phonetic dictionaries are regarded as pivotal components of speech recognition systems. The function of speech recognition research is to generate a machine which will accurately identify and distinguish the normal human speech from any other speaker. Literature affirmed that Arabic phonetics is one of the major problems in Arabic speech recognition. Therefore, this paper reviews previous studies tackling the challenges faced by initiating an Arabic phonetic dictionary with respect to Arabic speech recognition. It has been found that the system of speech recognition investigated areas of differences concerning Arabic phonetics. In addition, an Arabic phonetic dictionary should be initiated where the Arabic vowels' phonemes should be considered as a component of the consonants' phonemes. Thus, the incorporation of developed machine translation systems may enhance the quality of the system. The current paper concludes with the existing challenges faced by Arabic phonetic dictionary.

**Keywords:** Arabic Phonetic Dictionary, Speech Recognition System, Phonetics

## 1 Introduction

Speech recognition technology has historically achieved gradual progression since 50 years. It has successful attempts in coining various considerable applications. Several application areas of speech technology need particular type of initial analysis of speech. Such areas involve speaker recognition, speech coding, speech synthesis, and speech recognition (El-Henawy, Khedr, ELkomy & Abdalla, 2014). The phonetic dictionary is regarded as one of the critical parts of the modern large vocabulary speech recognition systems (Young, 1996). In speech recognition systems, this type of dictionary functions as a mediator between Language Model and Acoustic Model. It involves the words existing in the language and the pronunciation with respect to phonemes or the allophones existing in the acoustic model (Ali, Elshafei, Al-Ghamdi, Al-Muhtaseb, & Al-Najjar, 2008). Due to the substantial number of pronunciation exceptions in English, this type of dictionary was basically constructed manually by experts many years ago. In contrast, Arabic language pronunciation is subject to certain rules and patterns if the given text is fully diacritized. Research actually identified a number of these pronunciation rules (Elshafei, 1991; Alghamdi, Almuhtasib, & Elshafei, 2004). Furthermore, it should be noted that the association between the language representation and the acoustic representation is provided by the pronunciation representation. A list of words that have multiple or individual phonetic transcriptions exist in the pronunciation lexicon. Hence, two aspects of textual model represented by phoneme-based model involving short vowels and grapheme-based model that has no short vowels can be utilized in a dictionary (Alsharhan & Ramsay, 2020)

Arabic speech recognition development involves a multi-discipline effort. This process needs to incorporate Arabic phonetic (Alghmadi, 2003) and Arabic speech processing techniques (Elshafei, Almuhtasib, & Alghamdi, 2002). Development of an Arabic speech recognition system has recently been addressed by a number of researchers (Satori, Harti, & Chenfour, 2007; Hiyassat, 2007). Thus, an Arabic phonetic dictionary along with its management tools is highly required by the whole approaches interested in large vocabulary Arabic speech recognition systems.

The purpose of speech recognition research is to generate a machine that will accurately recognize and identify the normal human speech originated from any speaker. However, Arabic speech recognition encounters a number of challenges. One of these problems includes Arabic phonetics. In addition, Arabic has many dialects in a way that words are pronounced differently. According to Elmahdy, Gruhn, Minker, and Abdennadher (2009), the researchers summarized the major challenges in Arabic speech recognition. Such challenges involve Arabic phonetics, diacritization problem, morphological complexity, and grapheme-to-phoneme relation. Hence, additional confusion will be attached to the Arabic Speech Recognition decoder. In addition, a number of challenges for Arabic speech recognition have been noted. These problems include very large lexical variety, no diacritics, and dialectal variants (Lamel, Messaoudi, & Gauvain, 2009).

Moreover, an Arabic speech recognition system development primarily needs the construction of an Arabic Speech Corpus. Speech recognition model training also needs the construction of an Arabic phonetic dictionary along with its management tools. Such dictionary should encompass the entire probable phonetic pronunciations of any given word in the vocabulary system (Elshafei, Al-Muhtaseb, & Al-Ghamdi, 2008). It is important that some studies provide a rule-based technique to produce Arabic phonetic dictionaries including a large vocabulary speech recognition system. The current study aims at reviewing previous studies tackling the challenges experienced by Arabic phonetic dictionaries.

## **2 An Arabic Phonetic Dictionary**

A phonetic dictionary is considered as a basic structure relating to the production of Arabic speech recognition system. Regarding the Arabic phonetics, the significant work began in 2002. Abushariah, Aion, Zainuddin, Elshafei, and Khalifa (2012) conducted a comprehensive study concerning the Arabic phonetics and they revealed almost the most important points about its function. In another study conducted by Ali, Elshafei, Al-Ghamdi, Al-Muhtaseb and Al-Najjar (2009), an Arabic phonetic database was proposed. This proposal gave valuable details of the Arabic sounds articulatory mechanism. It involves more than 46000 files. Hyassat and Zitar (2008) offered a group of Arabic phonological rules. The rules included the phonemes of each letter. They also included the particular phonemes produced from a combination of letters. Another study dealt with Arabic phonetic dictionaries (Abushariah et al., 2012). The authors affirmed that using the rules of linguistic pronunciation in producing dictionaries can develop the result related to phone and word recognition based on Modern Standard Arabic.

Recently, an Arabic speech recognition system development has been tackled by many researchers (Alsharhan & Ramsay, 2020; Satori et al. (2007). For example, Satori et al. (2007) conducted a study relating to the use of Sphinx tools for Arabic speech recognition. The authors explained the use of the tools in order to recognize the isolated Arabic digits. In that study, the data were recorded from 6 speakers. They gained digits recognition accuracy of 86.66%.

Furthermore, Afify, Nguyen, Xiang, Abdou, and Makhoul (2005) made a comparison between grapheme-based recognition system and explicitly modeling short vowels. The researchers concluded that modeling short vowels improves recognition performance. Billa et al. (2002) provided a description of audio indexing system development for Arabic broadcast news. In their study, critical issues involved three main structures of the audio indexing system. These components included automatic speech recognition, named entity identification, and speaker identification. In addition, a study conducted by Hiyassat (2007) interested Arabic pronunciation dictionaries. The researcher cultivated a tool to produce an Arabic pronunciation dictionary. It is interesting to indicate that the generated dictionaries are built on the basis of a small MSA speech corpus containing digits or command as well as control vocabulary.

## **3 Arabic Speech Recognition System Development**

In order to enhance the speech recognition system performance, Alsharhan and Ramsay (2019) determined the appropriate method to well adopt the phonological characteristics of the Arabic language. The researchers noted that the impact of the local context is regarded as one of the major challenges that

confront the processing of Arabic. This local context causes remarkable changes in the phonetic model of a given text, and in turn this induces the misclassification of the recognition engine. To solve such dilemma, the researchers suggested that a group of language-individual grapheme-to-allophone rules should be developed. Thus, particular allophonic variations can be predicted and a phonetic transcription can be therefore provided.

In another study conducted by Masmoudi, Bougares, Ellouze, Estève, and Belguith (2018), the researchers adopted the model of resources and speech tools needed for the Automatic Speech Recognition system development pertaining to the Tunisian dialect. However, their developing system confronted some challenges including the lack of the remarks on tools and resources. Besides, the mispronunciation dictionary required for the development of ASR system noticeably existed. In addition, Zarnoufi, Jaafar, Bachri, and Abik, (2020) noted that using the Moroccan Arabic dialect dramatically causes difficulties in NLP tasks in written communications. Therefore, the researchers adopted a Moroccan Arabic dialect dictionary by which it offered the canonical forms. In their study, they constructed a normalization dictionary referring to as MANorm. Thus, the previous studies outline the challenges confronted by researchers in order to get better solution when adopting a speech recognition system relevant to any Arabic dictionary.

It is important to note that an Arabic speech recognition development has been characterized as a multidiscipline effort; it needs Arabic phonetic integration (Algamdi 2003; Elshafei, 1991), natural language processing (Elshafei, Al-Muhtaseb, & Alghamdi, 2006), and Arabic speech processing techniques (Elshafei et al. 2002). Arabic continuous speech recognition was discussed by Al-Otaibi (2001). The researcher presented a single-speaker speech dataset for MSA. As for labeling Arabic speech, he also suggested a particular technique. Using his technique, he provided a recognition rate for speaker dependent ASR of 93.78%. The Arabic Speech Recognition (ASR) was constructed by the use of the Hidden Markov Model (HMM) tool kit (HTK).

On the other hand, Hyassat and Zitar (2008) provided a description relating to an Arabic speech recognition system on the basis of Sphinx4. Both researchers also suggested an automatic toolkit for constructing phonetic dictionaries for the Holy Qur'an as well as standard Arabic language. In this work, three corpuses were cultivated, namely, the Holy Qura'an corpus (HQC-1) of about 18.5 hour, the Arabic digits corpus (ADC) of less than 1 hour of speech, and the command and control corpus (CAC-1) of about 1.5 hour.

It should be noted that advancements for Arabic speech recognition in the IBM system as a component of the continuous effort for the GALE project have been reported (Soltau et al., 2007). This system contains various stages integrating both vocalized and non-vocalized Arabic speech model. It also includes a training corpus of 1,800 hour of unsupervised Arabic speech. In another study, concerning Arabic spoken digits, Azmi, Tolba, Mahdy, and Fashal (2008) examined the use of Arabic syllables for speaker-independent speech recognition system. The sample for both training and testing involves 44 Egyptian speakers. The researchers found that the recognition rate gained from using syllables outperformed the rate gained from using words, monophones, and triphones by 1.79%, 2.68%, and 1.19%, respectively in a clean environment. However, syllables outperformed the rate gained from using monophones, triphones, and words by 2.09%, 1.5%, and 0.9%, respectively in noisy telephone channel.

A speech-enabled computer-aided pronunciation learning (CAPL) system was described by Abdou et al. (2006). Such system was created for teaching Arabic pronunciations to non-native speakers. In this system, a speech recognizer was used for detection of the errors in user recitation. It is interesting to note that a phoneme duration classification algorithm is activated to discover recitation errors concerning phoneme durations. In the experiment, the performance evaluation that utilized a dataset containing 6.6% wrong speech segments revealed that the system recognized the error correctly in 62.4% of pronunciation errors. It provided "Repeat Request" for 22.4% of the errors. False acceptance of 14.9% of total errors was also made.

In another study, a comparison has been made between dynamic time warping (DTW) recognizer and the polynomial classifier applied to isolated-word speaker-independent Arabic speech (Khasawneh, Assaleh,

Sweidan, & Haddad, 2004). The result showed that the polynomial classifier provided better recognition performance with a faster testing response compared to the DTW recognizer. Additionally, recent developments to English/Iraqi Arabic speech-to-speech translation system were provided in literature (Choi et al., 2008). The system provided wide developments involving machine translation (MT), components dialog manager, user interface (UI), and Arabic Speech Recognition (ASR).

Additionally, a phonetic system development for Arabic speech recognition has been demonstrated (Gales et al, 2007). In order to create and improve a robust Arabic speech recognition system on the basis of a hybrid group of speech characteristics, Shoaib et al. (2003) provided a study based on a novel approach. This hybrid group includes formant frequencies and intensity contours. However, Rambow et al. (2005) revealed that there is a problem tackling the issue of parsing transcribed spoken Arabic. In their study, the researchers investigated three different approaches such as grammar transduction, sentence transduction, and treebank transduction. In general, grammar transduction has a noticeable outperformance rather than the two approaches. It is important to indicate that parsing can be used as a check for the speech recognizer n-best hypothesis in order to have them scored again based on the most syntactically exact one. Nofal, Reheem, El Henawy, and Kader (2004) showed a plan and application of stochastic-based novel acoustic types appropriate for application in addition to control of speech recognition system pertaining to the Arabic language. Park, Diehl, Gales, Tomalin, and Woodland (2009) examined the adaptation and practice of multilayer perceptron (MLP) aspects in Arabic ASRs. It is important that three plans were examined. The first scheme includes the utilization of MLP aspects to integrate short-vowel information into the letter-alphabetic system. The second one involves a quick training method for application with the perceptual linear predictive (PLP). The MLP system was adopted as well. The third plan is the utilization of linear input networks (LIN) adaptation. The last one functions as an alternative to the normal HMM-based linear adaptation was clarified.

It is crucial that a system that matches algorithm on the basis of HMM utilizing field programmable gate array (FPGA) has been carried out (Elmisery Khalil, Salama, & Hamed, 2003). This suggested method was adopted for the sake of isolated Arabic word recognition. It obtained accuracy in comparison with the effective traditional recognition system. In addition, the algorithms and techniques adopted for modeling the acoustic-phonetic structure of Arabic speech recognition via utilizing HMMs have been highlighted in the study conducted by Mokhtar and El-Abddin (1996).

Once again, similar studies have been noticed in literature dealing with affording a system that involves speech features. Sagheer, Tsuruta, Taniguchi, and Maeda (2005) offered a system of novel visual speech aspects. The researchers adopted it so that they can highlight a system of full lip-reading. Another study conducted by Taha, Helmy, and Alez (2007) showed a new agent-based pattern for Arabic speech recognition. The Arabic speech recognition has been introduced as a Multi-agent System; each agent has a certain purpose and only tackles that purpose (Taha et al., 2007).

In order to identify isolated Arabic words, Bahi and Sellami (2001) conducted experiments and their identification system relied on a set of the vector quantization technique at the Markovian modeling and the acoustic level. It should be noted that many scholars examined using neural networks for digits identification and Arabic phonemes (Bahi & Sellami 2003; Shoaib et al. 2003). For instance, to gain constant speech recognition system, Abushariah, Aion, Zainuddin, Elshafei, and Khalifa (2010) highlighted the evaluation, implementation and design of a study for cultivating a high performance of common speaker-individual Arabic constant system of speech recognition. Another study examined classical ASR system for six various kinds of voice disorder patients speaking Arabic digits (Muhammad, AlMalki, Mesallam, Farahat, & Alsulaiman, 2011). They used Gaussian mixture model (GMM)/HMM and MFCC as classifier and features, respectively.

As for generating a model of Arabic pronunciation dictionary, Masmoudi, Estève, Khmekhem, Bougares, and Belguith, (2014) aimed at creating an automatic pronunciation dictionary in the area of rail transport in Tunisia. This has been allocated to the Tunisian Arabic in particular. Using the Tunisian Arabic, two instruments of phonetic vowelized and un-vowelized words have been initiated by the researchers. Their

suggested method relied on rules to automatically establish phonetic dictionaries. The same researchers' work outlined the process to initiate their research corpus 'TARIC', i.e. Tunisian Arabic Railway Interaction Corpus. They discussed in detail the phonological and phonetic exceptions of the Tunisian Arabic. In addition, they demonstrated certain rules utilized for building the phonetic dictionary.

Similarly, a study conducted by Elshafei, Al-Ghamdi, and Al-Muhtaseb (2008) showed a rule-based technique to initiate phonetic dictionaries used for an extended vocabulary Arabic speech recognition system. This system adopted popular pronunciation rules of Modern Standard Arabic, classical Arabic pronunciation rules, and some popular dialectal situations. Their study provided full clarification of these rules in addition to their own formal mathematical presentation. The rules were adopted to establish a dictionary for a 5.4 hours corpus of broadcast news. The phonetic dictionary consists of 23,841 definitions associated with about 14232 words. The initiated dictionary was evaluated on an Arabic speech recognition system. Furthermore, their study reported the findings of the first stage of a research work for constructing an essential performance, individually tackling the natural Arabic speech recognition system of a speaker. The purpose of their work was to develop an Arabic broadcast news transcription system. It is interesting that the Arabic speech recognition system development relied on the Carnegie Mellon university Sphinx tools. Also, the Cambridge HTK tools were also adopted at different testing phases.

It is crucial that the establishment of a dictionary connected with speech needs experts and scholars in phonology and linguistics to initiate the appropriate and relevant dictionary (Labidi, Maraoui, & Zrigui, 2016). However, with regard to Arabic phonemes, the experts in Arabic phonology have got two theories. First, the vowels phonemes do not belong to consonants phonemes. Second, as for vowels, there are no phonemes, whereas consonants only have phonemes and the vowels are included within them. In solving such challenge, Labidi et al. (2016) initiated two phonetic dictionaries in their research. Each one relied on one of the theories mentioned above. These two dictionaries examined by two Arabic automatic speech recognition systems. The findings demonstrated that the dictionaries constructed on the basis of the first theory are much suitable and provide a word error rate less than the other dictionary adopted the second theory.

#### **4 Conclusion**

It is important that the aim of this paper is to identify the challenges confronted by developing an Arabic phonetic dictionary with respect to speech recognition system. Therefore, this study reviews literature related to works on Arabic phonetic dictionary. It has been found that Arabic phonetics is regarded as one of the main problems in Arabic speech recognition. For example, Arabic speech recognition has many research issues that cause challenges such as the presence of compound words constructed by the connection of specific conjunctions, articles, prepositions, absence of short vowels in written text, and prefixes and suffixes to the word stem. Another challenge is the existence of the local context posing noticeable changes in the phonetic model of a given text, leading to the misclassification of the recognition engine in the case of generating a new system. In addition, the mispronunciation dictionary necessary for the development of ASR system does exist in addition to the lack of the remarks on speech tools and model of resources tackling the automatic speech recognition system. It has been recommended that an Arabic phonetic dictionary should not be constructed where the vowels have no phonemes; an Arabic phonetic dictionary should be generated where the Arabic vowels' phonemes should be regarded as a part of the consonants' phonemes. A machine that will accurately recognize and distinguish the normal human speech from any other speaker should be initiated as well. This potentially leads to the belief that the integration of updated, developed machine translation systems probably enhances the quality of the system of translation.

#### **5 Competing Interests**

The authors declared that no conflict of interest exist in this work.



**How to Cite this Article:**

Alhumsi, M. H., & Belhassen, S. (2021). The Challenges of Developing a Living Arabic Phonetic Dictionary for Speech Recognition System: A Literature Review. *Advanced Journal of Social Science*, 8(1), 164–170. <https://doi.org/10.21467/ajss.8.1.164-170>

**References**

- Abdou, S. M., Hamid, S. E., Rashwan, M., Samir, A., Abdel-Hamid, O., Shahin, M., & Nazih, W. (2006). Computer aided pronunciation learning system using speech recognition techniques. In Ninth International Conference on Spoken Language Processing.
- Abushariah, M. A., Aion, R. N., Zainuddin, R., Elshafei, M., & Khalifa, O. O. (2010, May). Natural speaker-independent Arabic speech recognition system based on Hidden Markov Models using Sphinx tools. In International Conference on Computer and Communication Engineering (ICCC'E10) (pp. 1-6). IEEE.
- Abushariah, M. A. A., Aion, R. N., Zainuddin, R., Elshafei, M., & Khalifa, O. O. (2012). Arabic speaker-independent continuous automatic speech recognition based on a phonetically rich and balanced speech corpus. *Int. Arab J. Inf. Technol.*, 9(1), 84-93.
- Afify, M., Nguyen, L., Xiang, B., Abdou, S., & Makhoul, J. (2005). Recent progress in Arabic broadcast news transcription at BBN. In Ninth European conference on speech communication and technology.
- Alghamdi, M. (2003). KACST arabic phonetic database. In the Fifteenth International Congress of Phonetics Science, Barcelona (pp. 3109-3112).
- Alghamdi, M., Almuhtasib, H., & Elshafei, M. (2004). Arabic phonological rules. *King Saud University Journal: Computer Sciences and Information*, 16, 1-25.
- Ali, M., Elshafei, M., Al-Ghamdi, M., Al-Muhtaseb, H., & Al-Najjar, A. (2008, December). Generation of Arabic phonetic dictionaries for speech recognition. In 2008 International Conference on Innovations in Information Technology (pp. 59-63). IEEE.
- Ali, M., Elshafei, M., Al-Ghamdi, M., Al-Muhtaseb, H., & Al-Najjar, A. (2009). Arabic phonetic dictionaries for speech recognition. *Journal of Information Technology Research (JITR)*, 2(4), 67-80.
- Al-Otaibi, F. (2001) speaker-dependant continuous Arabic speech recognition. M.Sc. thesis, King Saud University.
- Alsharhan, E., & Ramsay, A. (2019). Improved Arabic speech recognition system through the automatic generation of fine-grained phonetic transcriptions. *Information Processing & Management*, 56(2), 343-353.
- Alsharhan, E., & Ramsay, A. (2020). Investigating the effects of gender, dialect, and training size on the performance of Arabic speech recognition. *Language Resources and Evaluation*, 54(4), 975-998.
- Azmi, M., Tolba, H., Mahdy, S., & Fashal, M. (2008). Syllable-based automatic arabic speech recognition in noisy-telephone channel. *WSEAS Transactions on Signal Processing*, 4(4), 211-220.
- Bahi, H., & Sellami, M. (2001, June). Combination of vector quantization and hidden Markov models for Arabic speech recognition. In Proceedings ACS/IEEE International Conference on Computer Systems and Applications (pp. 96-100). IEEE.
- Bahi, H., & Sellami, M. (2003, July). A hybrid approach for Arabic speech recognition. In ACS/IEEE International Conference on Computer Systems and Applications, 2003. Book of Abstracts. (p. 107). IEEE.
- Billa, J., Noamany, M., Srivastava, A., Liu, D., Stone, R., Xu, J., & Kubala, F. (2002, May). Audio indexing of Arabic broadcast news. In 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing (Vol. 1, pp. I-5). IEEE.
- Choi, F., Tsakalidis, S., Saleem, S., Kao, C. L., Meermeier, R., Krstovski, K., & Natarajan, P. (2008, December). Recent improvements in BBN's English/Iraqi speech-to-speech translation system. In 2008 IEEE Spoken Language Technology Workshop (pp. 245-248). IEEE.
- El-Henawy, I. M., Khedr, W. I., ELkomy, O. M., & Abdalla, A. Z. M. (2014). Recognition of phonetic Arabic figures via wavelet based Mel Frequency Cepstrum using HMMs. *HBRC Journal*, 10(1), 49-54.
- Elmahdy, M., Gruhn, R., Minker, W., & Abdennadher, S. (2009, October). Modern standard Arabic based multilingual approach for dialectal Arabic speech recognition. In 2009 Eighth International Symposium on Natural Language Processing (pp. 169-174). IEEE.
- Elmisery, F. A., Khalil, A. H., Salama, A. E., & Hamed, H. F. (2003, December). A FPGA-based HMM for a discrete Arabic speech recognition system. In Proceedings of the 12th IEEE International Conference on Fuzzy Systems (Cat. No. 03CH37442) (pp. 322-325). IEEE.
- Elshafei, M. (1991). Toward an Arabic text-to-speech system. *The Arabian Journal for Science and Engineering*, 16(4B), 565-583.
- Elshafei, M., Al-Muhtaseb, H., & Al-Ghamdi, M. (2002). Techniques for high quality Arabic speech synthesis. *Information sciences*, 140(3-4), 255-267.
- Elshafei, M., Al-Muhtaseb, H., & Alghamdi, M. (2006). Statistical methods for automatic diacritization of Arabic text. In The Saudi 18th National Computer Conference. Riyadh (Vol. 18, pp. 301-306).
- Elshafei, M., Al-Muhtaseb, H., & Al-Ghamdi, M. (2008, December). Speaker-independent natural Arabic speech recognition system. In The International Conference on Intelligent Systems ICIS, Bahrain.
- Gales, M. J., Diehl, F., Raut, C. K., Tomalin, M., Woodland, P. C., & Yu, K. (2007, December). Development of a phonetic system for large vocabulary Arabic speech recognition. In 2007 IEEE Workshop on Automatic Speech Recognition & Understanding (ASRU) (pp. 24-29). IEEE.
- Hiyassat, H. (2007). Automatic Pronunciation Dictionary Toolkit for Arabic Speech Recognition Using SPHINX Engine (Doctoral dissertation, Department of Computer Information system, Arab Academy for Banking and Financial Sciences).
- Hyassat, H., & Zitar, R. A. (2008). Arabic speech recognition using SPHINX engine. *International Journal of Speech Technology*, 9(3), 133-150.
- Khasawneh, M., Assaleh, K., Sweidan, W., & Haddad, M. (2004, July). The application of polynomial discriminant function classifiers to isolated Arabic speech recognition. In 2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No. 04CH37541) (Vol. 4, pp. 3077-3081). IEEE.
- Labidi, M., Maraoui, M., & Zrigui, M. (2016, September). New birth of the Arabic phonetic dictionary. In 2016 International Conference on Engineering & MIS (ICEMIS) (pp. 1-9). IEEE.

- Lamel, L., Messaoudi, A., & Gauvain, J. L. (2009). Automatic speech-to-text transcription in Arabic. *ACM Transactions on Asian Language Information Processing (TALIP)*, 8(4), 1-18.
- Masmoudi, A., Bougares, F., Ellouze, M., Estève, Y., & Belguith, L. (2018). Automatic speech recognition system for Tunisian dialect. *Language Resources and Evaluation*, 52(1), 249-267.
- Masmoudi, A., Estève, Y., Khmekhem, M. E., Bougares, F., & Belguith, L. H. (2014). Phonetic tool for the Tunisian Arabic. In: *The 4<sup>th</sup> International Workshop on Spoken Language Technologies for Under-Resourced Languages*.
- Mokhtar, M. A., & El-Abddin, A. Z. (1996, October). A model for the acoustic phonetic structure of Arabic language using a single ergodic hidden Markov model. In *Proceeding of Fourth International Conference on Spoken Language Processing. ICSLP'96 (Vol. 1, pp. 330-333)*. IEEE.
- Muhammad, G., AlMalki, K., Mesallam, T., Farahat, M., & Alsulaiman, M. (2011, March). Automatic Arabic digit speech recognition and formant analysis for voicing disordered people. In *2011 IEEE Symposium on Computers & Informatics (pp. 699-702)*. IEEE.
- Nofal, M., Reheem, E. A., El Henawy, H., & Kader, N. A. (2004, September). The development of acoustic models for command and control Arabic speech recognition system. In *International Conference on Electrical, Electronic and Computer Engineering, 2004. ICEEC'04. (pp. 702-705)*. IEEE.
- Park, J., Diehl, F., Gales, M. J. F., Tomalin, M., & Woodland, P. C. (2009, April). Training and adapting MLP features for Arabic speech recognition. In *2009 IEEE International Conference on Acoustics, Speech and Signal Processing (pp. 4461-4464)*. IEEE.
- Rambow, O., Chiang, D., Diab, M., Habash, N., Hwa, R., Sima'an, K., ... & Shareef, S. (2005). Parsing Arabic Dialects. Final Report. In *2005 JHU Summer Workshop*.
- Sagheer, A., Tsuruta, N., Taniguchi, R. I., & Maeda, S. (2005, December). Hyper column model vs. fast DCT for feature extraction in visual Arabic speech recognition. In *Proceedings of the Fifth IEEE International Symposium on Signal Processing and Information Technology, 2005. (pp. 761-766)*. IEEE.
- Satori, H., Harti, M., & Chenfour, N. (2007). Introduction to Arabic speech recognition using CMUSphinx system. arXiv preprint arXiv:0704.2083.
- Shoab, M., Rasheed, F., Akhtar, J., Awais, M., Masud, S., & Shamaail, S. (2003, December). A novel approach to increase the robustness of speaker independent Arabic speech recognition. In *7th International Multi Topic Conference, 2003. INMIC 2003. (pp. 371-376)*. IEEE.
- Soltau, H., Saon, G., Kingsbury, B., Kuo, J., Mangu, L., Povey, D., & Zweig, G. (2007, April). The IBM 2006 Gale arabic ASR system. In *2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07 (Vol. 4, pp. IV-349)*. IEEE.
- Taha, M., Helmy, T., & Alez, R. A. (2007, November). Multi-agent based Arabic speech recognition. In *2007 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology-Workshops (pp. 433-436)*. IEEE.
- Young, S. (1996). A review of large-vocabulary continuous-speech. *IEEE signal processing magazine*, 13(5), 45.
- Zarnoufi, R., Jaafar, H., Bachri, W., & Abik, M. (2020, December). MANorm: A Normalization Dictionary for Moroccan Arabic Dialect Written in Latin Script. In *Proceedings of the Fifth Arabic Natural Language Processing Workshop (pp. 155-166)*.

**Publish your research article in AIJR journals-**

- ✓ Online Submission and Tracking
- ✓ Peer-Reviewed
- ✓ Rapid decision
- ✓ Immediate Publication after acceptance
- ✓ Articles freely available online
- ✓ Retain full copyright of your article.

Submit your article at [journals.aijr.org](http://journals.aijr.org)

**Publish your books with AIJR publisher-**

- ✓ Publish with ISBN and DOI.
- ✓ Publish Thesis/Dissertation as Monograph.
- ✓ Publish Book Monograph.
- ✓ Publish Edited Volume/ Book.
- ✓ Publish Conference Proceedings
- ✓ Retain full copyright of your books.

Submit your manuscript at [books.aijr.org](http://books.aijr.org)