Cairo University

A STUDY ON ARABIC PHONEMES TOWARDS AN AUTOMATIC TEACHING SYSTEM FOR THE RECITATION OF HOLY QUR’AN

By

Fatma Shawky AbdEl-Hamid Mohamed Khaled

A Thesis Submitted to the
Faculty of Engineering at Cairo University
In Partial Fulfillment of the
Requirements for the Degree of
MASTER OF SCIENCE
In
Biomedical Engineering and Systems

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2017

A STUDY ON ARABIC PHONEMES TOWARDS AN AUTOMATIC TEACHING SYSTEM FOR THE RECITATION OF HOLY QUR'AN

By

Fatma Shawky AbdEl-Hamid Mohamed Khaled

A Thesis Submitted to the
Faculty of Engineering at Cairo University
In Partial Fulfillment of the
Requirements for the Degree of
MASTER OF SCIENCE
In
Biomedical Engineering and Systems

Under the Supervision of

Prof. Dr. Ahmed M. ELBialy	**Prof. Dr. Ahmed H. Kandil**
………………………………	………………………………
Professor	**Professor**
Biomedical Engineering and Systems	**Biomedical Engineering and Systems**
Faculty of Engineering, Cairo University	**Faculty of Engineering, Cairo University**

Dr. Sahar A. Fawzi

……………………………………

Associate Professor
Biomedical Engineering and Systems
Faculty of Engineering, Cairo University

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2017

A STUDY ON ARABIC PHONEMES TOWARDS AN AUTOMATIC TEACHING SYSTEM FOR THE RECITATION OF HOLY QUR’AN

By

Fatma Shawky AbdEl-Hamid Mohamed Khaled

A Thesis Submitted to the
Faculty of Engineering at Cairo University
In Partial Fulfillment of the
Requirements for the Degree of
MASTER OF SCIENCE
In
Biomedical Engineering and Systems

Approved by the
Examining Committee

____
Prof. Dr. Ahmed Mohamed ELBialy

____
Associate Prof. Dr. Sahar Ali Fawzi

____
Prof. Dr. Ayman Mohamed ElDeib

____
Prof. Dr. Samia Abd El Razek Mashaly
Electronics Research Institute

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2017

Engineer's Name: Fatma Shawky AbdEl-Hamid
Date of Birth: 12/11/1984
Nationality: Egyptian
E-mail: Fatma_khaled@hotmail.com
Phone: 01001910642
Address: 7 Al Falouga St., Nasr city
Registration Date: 01/10/2007
Awarding Date: 2017
Degree: Master of Science
Department: Biomedical Engineering and Systems

Insert photo here

Supervisors:

Prof. Dr. Ahmed Mohamed ELBialy
Prof. Dr. Ahmed Hisham Kandil
Associate Prof. Dr. Sahar Ali Fawzi

Examiners:

Prof. Dr. Ahmed Mohamed ELBialy
Associate Prof. Dr. Sahar Ali Fawzi
Prof. Dr. Ayman Mohamed El Deib
Prof. Dr. Samia Abd El Razek Mashaly, Electronics Research Institute

Title of Thesis:

A Study on Arabic Phoneme Towards an Automatic Teaching System for the Recitation of Holy Qur'an.

Key Words:
Speech recognition; Hidden Markov Model (HMM); Mel Frequency Cepstral Coefficients (MFCC); Classical Arabic (CA)

Summary:

This thesis is part of ongoing integrated studies concerning Classical Arabic recognition for both teaching and learning purposes. The major point of strength is using Al Norania Rule for the first time as training and testing dataset to differentiate between Arabic phonemes based on their exits and characteristics. This presents a substantial contribution summing up recognition models for recognizing different features and tiny details of each letter. This work is a good seed for different speech synthesis or speech recognition projects later on.

Acknowledgments

First and foremost, I thank Allah Almighty for the completion of this Master's thesis.

Thanks for everyone who helped me in completing this work, I would like to express my sincere and profound gratitude to my advisor Dr. Sahar Fawzi for the continuous support of my research, for her patience, motivation, enthusiasm, and immense knowledge.

Her guidance helped me in all the time of research and writing of this thesis.

I would like also to thank Prof. Ahmed Hisham and Prof Ahmed El-Bialy for their encouragement and insightful comments.

Last but not the least, I would like to thank my family for their love and support through my whole life. A special thanks goes to my mother; thank you for your understanding and spiritual support in many moments of crisis in my life.

Dedication

This work is dedicated to my family.

Table of Contents

Acknowledgments i
Dedication ii
List of Tables v
List of Figures vii
List of Abbreviations ix
Abstract x
1 Chapter 1 Introduction 1
1.1 Thesis Objective 1
1.2 Thesis Organization 1
2 Chapter 2 Background and Literature Review 3
2.1 Introduction 3
2.2 Arabic Language 3
2.2.1 Characteristics and Importance 3
2.2.2 Classical Arabic 3
2.3 Al Norania rule 3
2.3.1 Advantages of teaching Al Norania rule 4
2.4 Speech Production 5
2.4.1 Speech Production Organs 5
2.5 Phonemes and syllables 6
2.5.1 Arabic Phonetic Alphabet 6
2.5.2 Arabic Articulation Points (Exits) 7
2.5.3 Characteristics of the Alphabet (Sifaat) 9
2.5.4 Types of Speech Signal 12
2.6 Literature Review 12
3 Chapter 3 Methodology 14
3.1 Introduction 14
3.2 Data Collection 14
3.3 Learning and Testing Subsets 14
3.4 Preprocessing 17
3.5 Feature Extraction 17

3.5.1 Mel Frequency Cepstral coefficients (MFCC) 17
3.5.2 MFCC Derivation 17
3.5.3 Performance of MFCC 24
3.6 Verification 24
3.6.1 Training HMM 26
3.6.2 Recognition 28
3.6.3 Analysis 28
4 Chapter 4 Results and Discussion 30
4.1 Feature Extraction 30
4.2 Acoustic Training 33
4.3 Verification Results 33
4.3.1 Syllables Recognition 33
4.3.2 Phonemes Recognition 39
4.3.3 Recognition of phonemes with sukun 46
4.4 Phoneme duration Results 63
4.5 Discussion 70
5 Chapter 5 Conclusions and Future Work 73
5.1 Conclusions 73
5.2 Future Work 73
References 74

List of Tables

Table 2-1 Classification of the syllables in Arabic *6*
Table 2-2 Arabic Vowels *6*
Table 2-3 Articulation Points of Arabic letters *7*
Table 2-4 Characteristics of Arabic letters with opposites *9*
Table 2-5 Characteristics of Arabic letters without opposites *10*
Table 3-1 Letters sets having same Exits *15*
Table 3-2 Letters sets having same Characteristics *15*
Table 4-1 Parameters of Feature Extraction *30*
Table 4-2 Percentage of Accuracy of Speech Verification experiments results; comparing effect of different number of Cepstral Coeffients on Phonemes with sukun having same Characteristics with TARGETKIND MFCC_0_D_A_Z *32*
Table 4-3 Percentage of Accuracy of Speech Verification experiments results; comparing effect of different number of Cepstral Coeffients on Phonemes with sukun having same Characteristics with TARGETKIND MFCC_E_D_A_Z *32*
Table 4-4 Percentage of Accuracy of Speech Verification experiments results comparing effect of different number of Cepstral Coeffients on Syllables having same Exits with TARGETKIND MFCC_0_D_A_Z *34*
Table 4-5 Percentage of Accuracy of Speech Verification experiments results comparing effect of different number of Cepstral Coeffients on Syllables having same Exits with TARGETKIND MFCC_E_D_A_Z *35*
Table 4-6 Percentage of Accuracy of Speech Verification experiments results comparing effect of different number of Cepstral Coeffients on Syllables having same Characteristics with TARGETKIND MFCC_0_D_A_Z *36*
Table 4-7 Percentage of Accuracy of Speech Verification experiments results comparing effect of different number of Cepstral Coeffients on Syllables having same Characteristics with TARGETKIND MFCC_E_D_A_Z *37*
Table 4-8 Percentage of Accuracy of Speech Verification experiments results, comparing effect of different number of Cepstral Coeffients on Phonemes having same Exits with TARGETKIND MFCC_0_D_A_Z *39*
Table 4-9 Percentage of Accuracy of Speech Verification experiments results; comparing effect of different number of Cepstral Coeffients on Phonemes having same Exits with TARGETKIND MFCC_E_D_A_Z *41*
Table 4-10 Percentage of Accuracy of Speech Verification experiments results; comparing effect of different number of Cepstral Coeffients on Phonemes having same Characteristics with TARGETKIND MFCC_0_D_A_Z *43*
Table 4-11 Percentage of Accuracy of Speech Verification experiments results; comparing effect of different number of Cepstral Coeffients on Phonemes having same Characteristics with TARGETKIND MFCC_E_D_A_Z *44*
Table 4-12 Percentage of Accuracy of Speech Verification experiments results comparing effect of different number of Cepstral Coeffients on Phonemes with sukun having same Exits with TARGETKIND MFCC_0_D_A_Z *46*

Table 4-13 Percentage of Accuracy of Speech Verification experiments results; comparing effect of different number of Cepstral Coeffients on Phonemes with sukun having same Exits with TARGETKIND MFCC_E_D_A_Z *48*
Table 4-14 Percentage of Accuracy of Speech Verification experiments results comparing effect of different number of Cepstral Coeffients on Phonemes with sukun having same Characteristics with TARGETKIND MFCC_0_D_A_Z *49*
Table 4-15 Percentage of Accuracy of Speech Verification experiments results comparing effect of different number of Cepstral Coeffients on Phonemes with sukun having same Characteristics with TARGETKIND MFCC_E_D_A_Z *51*
Table 4-16 Average total accuracy for all data sets having same Exits examining different schemes for different targetkinds *53*
Table 4-17 Average total accuracy for all data sets having same characteristics examining different schemes for different targetkinds *53*
Table 4-18 Best Verification Models for different Exit Sets *54*
Table 4-19 Best Verification Models for different Characteristics Sets *55*
Table 4-20 Arabic phonemes versus MFCCs to be investigated for best recognition values *61*
Table 4-21 Phoneme duration results for Phonemes attached to vowels in lesson 4 of Al Norania Rule . *63*
Table 4-22 Phoneme duration results for phonemes (Consonant part of syllables) from second lesson of Al Norania Rule *64*
Table 4-23 Phoneme duration results for Phonemes with sukun in lesson 10,11 of Al Norania Rule for Farouk AlRa'y *66*
Table 4-24 Phoneme duration results for Phonemes with sukun in lessons 10 and 11 of Al Norania Rule for Osama Kare *67*
Table 4-25 Phoneme duration results for Phonemes with sukun From Holy Quran Recitation by Al Hossary *68*
Table 4-26 Minimum, Maximum and average phonemes duration *68*
Table 4-27 Accuracy results for phonemes with sukun recorded by a male trainee. *69*

List of Figures

Figure 2-1 The human vocal organs 5

Figure 2-2 Emission points for Arabic Letters 7

Figure 3-1 Extracting MFCC feature vector and its derivatives 18

Figure 3-2 A spectral slice from the vowel [aa]; (a) before pre-emphasis and (b) after pre-emphasis 19

Figure 3-3 Windowing Process, A figure by Bryan Pellom 20

Figure 3-4 Window in time domain, Rectangular (a) and Hamming (b) 20

Figure 3-5 Mel Filter bank. Triangular filter collects energy from a given frequency range......... 22

Figure 3-6 Summation process of filter bank on the input spectrum......... 22

Figure 3-7 Log Energy Computation 23

Figure 4-1 Accuracy Results for Syllables having same Exits with TARGETKIND MFCC_0_D_A_Z........ 34

Figure 4-2 Accuracy Results for Syllables having same Exits with TARGETKIND MFCC_E_D_A_Z........ 35

Figure 4-3 Accuracy Results for Syllables having same Characteristics with TARGETKIND MFCC_0_D_A_Z 36

Figure 4-4 Accuracy Results for Syllables having same Characteristics with TARGETKIND MFCC_E_D_A_Z 37

Figure 4-5 Accuracy Results for Phonemes having same Exits with TARGETKIND MFCC_0_D_A_Z 40

***Figure 4-6 Accuracy Results for Phonemes having same Exits with TARGETKIND MFCC_E_D_A_Z** 41*

Figure 4-7 Accuracy Results for Phonemes having same Characteristics with TARGETKIND MFCC_0_D_A_Z 43

Figure 4-8 Accuracy Results for Phonemes having same Characteristics with TARGETKIND MFCC_E_D_A_Z 44

Figure 4-9 Accuracy Results for Phonemes with sukun having same Exits with TARGETKIND MFCC_0_D_A_Z 46

Figure 4-10 Accuracy Results for Phonemes with sukun having same Exits with TARGETKIND MFCC_E_D_A_Z 48

Figure 4-11 Accuracy Results for Phonemes with sukun having same Characteristics with TARGETKIND MFCC_0_D_A_Z 49

Figure 4-12 Accuracy Results for Phonemes with sukun having same Characteristics with TARGETKIND MFCC_E_D_A_Z 51

Figure 4-13 Extracted MFCC features for Baa (ب) phoneme with sukun in cluster set 4 with MFCC_E_D_A_Z and only 6 coefficient........ 57

Figure 4-14 Extracted MFCC features for Ttaa (ط) phoneme with sukun in cluster set 4 with MFCC_E_D_A_Z and only 6 coefficient........ 57

Figure 4-15 Extracted MFCC features for Daal (د) phoneme with sukun in cluster set 4 with MFCC_E_D_A_Z and only 6 coefficient........ 58

Figure 4-16 Extracted MFCC features for Ddaad (ض) phoneme with sukun in cluster set 9 with MFCC_E_D_A_Z and only 6 coefficient........ 58

Figure 4-17 Extracted MFCC features for Zzaa (ظ) phoneme with sukun in cluster set 9 with MFCC_E_D_A_Z and only 6 coefficients 59

Figure 4-18 Extracted MFCC features for faa (ف) phoneme with sukun in cluster set 10, with MFCC_0_D_A_Z and only 6 Coeffiecients........ 59

Figure 4-19 Extracted MFCC features for Seen (س) phoneme with sukun in cluster set 10 with MFCC_0_D_A_Z and only 6 Coeffiecients 60

List of Abbreviations

ASR	Automatic speech recognition
C	Consonant
CA	Classical Arabic
DFT	Discrete Fourier Transform
DCT	Discrete Cosine Transform
HMM	Hidden Markov Models
HTK	Hidden Markov Model Toolkit
IDFT	Inverse Discrete Fourier Transform
MSA	Modern Standard Arabic
MLF	Master Label Files
MFCC	Mel Frequency Cepstral Coefficients
NUMCEPS	Number of Cepstral Coefficients
V	Vowel

Abstract

This thesis investigates the sensitivity of MFCC features to recognize and differentiate fine differences between uttered Classical Arabic phonemes.

Correct pronunciation of Classical Arabic phonemes depends on their exits and characteristics. Al Norania rules were compromised to be used as a tool for teaching nonnative Arabic speakers the correct pronunciations of Classical Arabic phonemes.

A model was built based on MFCC features to investigate its sensitivity to different changes with respect to reference pronunciation

A speech corpus of syllables and phonemes was collected from audio signals of guaranteed readers of Al Norania rule. The majority of collected data was used as training data set; the remaining part was used as testing data set.

These syllables and phonemes were clustered according to their exits and characteristics. Then multiple features were extracted from each cluster using Mel Frequency Cepstral Coefficients (MFCC) -as the chosen feature extraction technique- examining different coefficients combinations and converting analogue form of speech signals into a parametric representation, from which parameters of Hidden Markov Models (HMM) were estimated for both training and testing processes.

Statistical results showed recognition accuracy of different models applied on the same cluster set. By comparing these results, it was found that the highest recognition accuracy for each cluster set was obtained using a specific MFCC recognition model, extracting features of very tiny details of this cluster.

Duration of pronunciation of tested phonemes was measured. Results showed different values according to belonging to one of three categories: explosive, in-between and softness phonemes (حروف شديدة. حروف توسط. حروف رخوة), keeping a constant ratio between median values of the three categories. To achieve correct pronunciation, this ratio should be kept for different persons and different methods of recitations for Holy Qur’an.

These findings can be summed up as a rigid frame for implementing an automatic system for both teaching and testing recitation of Holy Qur’an, and many educational projects preserving correct pronunciation of eloquent Arabic.

Chapter 1
Introduction

The Arabic Language serves as a powerful symbol of Arab national identity; ranks sixth in the world's league table of languages with an estimated 186 million native speakers. This language is considered one of the major languages of the world reflecting not only its number of speakers but also the important role that it played in history and is still playing especially in the development of Arab-Muslim Society [1-3].

There are three main variants of Arabic Language; one is Classical Arabic (CA) which was originated in the Arabian Peninsula. When literary talent and eloquence of Arabs were at their peak, Holy Qur'an was handed down to Prophet Muhammad raising an intimate relationship between Arabic language and Islam. Being complete true revelation, Holy Qur'an which contains the message of Islam is itself a miracle [1].

One of the traditions attributed to the prophet Muhammad is "the best among you is the one who learns and teaches the Qur'an“. Al Norania rule (Al Qaeda Al Norania) as a systematic method developed by Sheikh Nour Muhammad Hakany to teach the right pronunciation of Arabic phonemes and syllables (speech units either separated or concatenated). It is composed of 16 lessons, each of them introduces at least one pronunciation rule. Lessons of Al Qaeda Al Norania are very suitable for children and non-native Arabic readers as they introduces the separated speech units then proceed with the composite words [2-3].

Mel-Frequency Cepstral Coefficients (MFCC) extracted from phonemes and syllables of Al Qaeda Al Norania are chosen to be the speech features to be used with Hidden Markov Model Toolbox (HTK) for the aim of Speech Verification in the learning process of Al Qaeda Al Norania.

1.1 Thesis Objective

Correct pronunciation is an essential component in learning any language. The objective of this thesis is to generate a verification model for Arabic phonemes and syllables extracted from Al Qaeda Al Norania Lessons; to detect how close the trainee's pronunciation is to the reference reader.

The Verification task is implemented through three main steps:

1. Preprocessing stage
2. Feature Extraction
3. Pattern Recognition

This study is part of an integrated set of studies on achieving proficiency in Arabic and the Holy Quran.

1.2 Thesis Organization

The thesis consists of five chapters organized as follows: