UNIVERSITI TEKNOLOGI MARA

SPEAKER INDIVIDUALITY DETERMINATION BY USING PROSODIC FEATURES ANALYSIS

SHUKUR BIN KHALID

Thesis submitted in partial fulfillment of the requirements for Master of Science (Information Technology)

Faculty of Information Technology And Quantitative Science

May 2007

ACKNOWLEDGMENT

I would like to express my gratefulness to Allah S.W.T for giving me strength and wisdom in my thesis work.

First and foremost, I wish to express my sincere appreciation to my project supervisor, Dr. Mazani Abd. Manaf and co-supervisor Mr. Farok Hj. Azmat for their encouragement, guidance and critics. Their ideas enlighten my curiosity. Without their continued support and interest, this thesis would not have been the same as presented here.

Special thanks to all my lecturers, my friends and former course mates who have been very supportive of my decision to return to the academics.

Last but not least, I am greatly indebted to my family for their understanding, patience and support during the entire period of my study. My deepest gratitude goes to my lovely wife and son. Thank you *Nor* and *Danial*, for coming into my life, for all the love and support you have given me and for seeing this thesis through with me. This thesis would not have taken the shape it has now, without your help and support.

Finally, I would like to extend my gratitude and thanks to my dear friends especially Zaidi, Ismail, Rozanudin and Latifah whose valuable assistance given all along the process of writing this thesis is very much remembered.

TABLE OF CONTENTS

ACKNOWLEDGEMENT ii
TABLE OF CONTENTS iii
LIST OF TABLES vi
LIST OF FIGURES vii
LIST OF ABBREVIATIONS ix
ABSTRACT x

CHAPTER 1: INTRODUCTION

1.1 Introduction 1
1.2 Research Background 2
1.3 Problem Statement 3
1.4 Problem Definition 4
1.5 Research Objective 4
1.6 Research Scope 5
1.7 Research Significance 5
1.8 Summary 6

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction 7
2.2 Human Speech Perspectives 7
2.2.1 Articulatory Approach 8
2.2.2 Acoustic Approach 11
2.2.2.1 Spectographic Analysis of Speech 12
2.2.2.2 Source Filter Model 14
2.2.2.3 Segmentals and Suprasegmentals 17
2.2.3 Perception Approach 19
2.3 Speaker Individuality 22
2.3.1 Voice Source 23
2.3.2 Vocal Tract 24

2.3.3 Segmental Differences 24
2.3.4 Suprasegmental Differences 27
2.4 Speech Analysis Tools 27
2.5 Summary 29

CHAPTER 3: RESEARCH METHODOLOGY

3.1 Introduction 30
3.2 Research Approach 32
3.2.1 Study 32
3.2.2 Testing 32
3.2.3 Analysis 33
3.3 Information Gathering 33
3.3.1 Prosodic Parameters 33
3.4 Hardware and Software 35
3.4.1 Hardware 35
3.4.2 Software 35
3.5 Hardware Setup and Software Installation 35
3.6 Configuration 36
3.6.1 Voice File Configuration 36
3.6.2 Praat Configuration 37
3.7 Testing 38
3.7.1 Data Collection 38
3.7.2 Statistics Computation 39
3.7.3 Prosodic Feature Extraction 41
3.7.4 Praat Script Organization 43
3.8 Data Analysis 44
3.9 Summary 47

CHAPTER 4: FINDING AND DISCUSSION

4.1 Introduction 49
4.2 Speaker Individuality Determination Analysis 49
4.2.1 Pitch analysis 51

ABSTRACT

Speech signal conveys information on several levels. It is an outcome of influence of several physiological, psychological and environmental factors. It contains a message generically expressed as a sequence of words, information specific to the speaker that produced the speech, and information about the environment in which the speech was produced and transmitted. Speaker specific information includes the identity of the speaker, the gender of the speaker, the language or dialect of the speaker and possibly the physical and emotional condition of the speaker. That richness of information gives a distinctive "identity" to the speech of different individuals. Suprasegmental refers to the time span of the acoustic analysis. Its parameters are spread over several segments as the name suggests. Suprasegmental parameters are also called prosodic parameters. Prosodic is responsible for controlling the intonation, stress, and rhythmic organization of the speech. Typical acoustic suprasegmental measurements are the energy and *F0* contours. In this research, both prosodic features namely pitch and energy are investigated and analyzed. The purpose is to study how far the normal and disguised speech can be discriminated and to find the best technique for characterizing the individual speakers. Some statistics and prosodic feature extraction computation are conducted through Praat tool and a comparative analysis between normal and disguised utterances collected from ten different subjects is done. Some variations in pitch and energy features are recorded and from those findings, it is concluded that speaker individuality can be determined whether the subjects speak in normal or disguised manner.