

UNIVERSITI TEKNOLOGI MARA

**GRAMMAR GUIDED AUTOMATIC SPEECH
RECOGNITION SYSTEM FOR MALAY
LANGUAGE**

RADDE IDZWAN BIN ABDUL RAZAK

MSc

APRIL 2009

ACKNOWLEDGEMENTS

Bismillahirrahmanirrahim

In the name of Allah, The most Gracious and the most Merciful

First and foremost, I would like to pay my gratitude to Allah S.W.T for giving me the strength to complete this research. Thank you for giving me the guidance and courage for me to be able to finish my research paper.

A special thank to *Associate Professor Nurazzah Abdul Rahman* who act as my supervisor, for her patience, dedications, guidance and also who has been contributing so many ideas for my research paper.

My utmost appreciation also goes to *Dr. Siti Salwa binti Salleh, Associate Professor Rosmah Abd. Latif, Associate Professor Normaly Kamal Ismail* and other lecturers for their help, support and encouragement. Without their support and help, I would not be able to complete this work

I would also like to thank my beloved family and my supportive friends for giving me their support and advises towards the process of developing my final research.

Thank you.

ABSTRACT

Speech recognition is difficult because of the analog nature of the speech itself. In speech recognition, a system is trained to receive human speech, recognize each word in the speech and transform it into text.

There are many existing Automatic Speech Recognition (ASR) system that able to recognize English language speech with high accuracy. Unfortunately, an ASR system normally focuses on only one language. Today, there are no or lack of Malay language recognizable ASR system. Thus, an ASR system that is capable of recognizing Malay words would be beneficial especially in Malaysia.

In order for the ASR system to transform speech signal into a specific word, the system requires a database which store list of words with each word pronunciation. This database is called a dictionary.

A dictionary is use by an ASR system as references to find words match to the received speech signal. In order to be able to recognize spoken Malay words, a dictionary must contain a list of Malay words and each word must be associated with its pronunciation. Thus, this research is focusing on building Malay dictionary for an ASR system.

The dictionary built in this research is meant to be used for ASR system based on Sphinx-4 framework. Each word is tested for recognition accuracy on the ASR system prototype. Trial and error method was used to produce word's phonemes that are most accurate.

Well defined words in the dictionary do not guarantee high recognition accuracy. Word Error Rate (WER) typically increases when the size of dictionary is increase. Grammar implementation would reduce the number of words to be combined to construct a sentence. Thus, it increases the chances of accurate recognition with large dictionary for continuous speech.

The finding of this research shows that Sphinx-4 is a good framework to be use with Malay language ASR. The outcome of this research proved that grammar able to increase recognition accuracy of Malay language ASR.

TABLE OF CONTENTS

	Page
Thesis Submitted In Fulfillment	i
Declaration	ii
Approval	iii
Acknowledgements	iv
Abstract	v
Table of Contents	vi
List of Tables	x
List of Figures	xiii
List of Abbreviations	xiv
1.0 Introduction	1
1.1 Background of the Project	1
1.2 Problem Statement	2
1.3 Objectives of the Project	3
1.4 Scope of the Project	4
1.5 Significant of the Project	6
1.6 Summary of Dissertation	7
2.0 Literature Review	8
2.1 Introduction	8
2.2 Dictionary	9
2.3 Generic Speech Recognizer Architecture	10
2.4 Selected ASR Framework: Sphinx-4	11
2.5 Sphinx-4 Architecture	12
2.5.1 The front-end module	12
2.5.2 Linguist	14
2.5.3 Decoder	15
2.6 Other ASR System Framework / Tools	15
2.6.1 CSLU Toolkit	16

CHAPTER 1

INTRODUCTION

1.1 Background of the Project

Automatic Speech Recognition (ASR) is a technology that allows a computer to identify the words that a person speaks into a microphone or telephone (Satori *et al.*, 2009). Basically, this technology converts spoken words to machine-readable input. This technology allows people to write a paper, compose an e-mail message, and open programs without ever touching the keyboard.

There have been numerous researches on ASR since many years ago. Despite many researchers efforts to improve the quality of ASR application, current ASR application does not perform satisfactorily (Munteanu *et al.*, 2006). Among the factors that affect the ASR quality are poor acoustic conditions, diverse speakers and large vocabularies.

ASR has a wide area of applications, some of it are command recognition, dictation, interactive voice response, simple data entry, speech-to-text processing. Examples of speech-to-text processing are word processors or emails. For command recognition, the example is pilot command in aircraft cockpit. This technology is getting more attraction from the public and large company such as Microsoft as it brings benefit to users. Microsoft has integrates ASR functionality to their word processing product, Microsoft Word (Moskowitz, 2003). ASR is capable of becoming key or added functionality to word processing, which would be able to distinguish them from the rest of the competitor. It is a useful tool in Microsoft Word as user will be able to just speak to the speaker and Microsoft Word is capable of transforming the spoken words to text, which is much more convenience compares to have to manually type those words to a keyboard.