IMPROVING BAHASA MELAYU TEXT-TO-SPEECH SYSTEM USING UNIT-SELECTION METHOD: UTeM HEALTH CENTRE CASE STUDY

NURUL ANISA SRI WINARSIH

MASTER OF COMPUTER SCIENCE
(SOFTWARE ENGINEERING AND INTELLIGENCE)

2017

Faculty of Information and Communication Technology

IMPROVING BAHASA MELAYU TEXT-TO-SPEECH SYSTEM USING UNIT-SELECTION METHOD: UTeM HEALTH CENTRE CASE STUDY

Nurul Anisa Sri Winarsih

Master of Computer Science in Software Engineering and Intelligence

2017

IMPROVING BAHASA MELAYU TEXT-TO-SPEECH SYSTEM USING UNIT-SELECTION METHOD: UTeM HEALTH CENTRE CASE

NURUL ANISA SRI WINARSIH

A thesis submitted
in fulfillment of the requirements for the degree of Master of Computer Science
(Software Engineering and Intelligence)

Faculty of Information and Communication Technology

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2017

DECLARATION

I declare that this thesis entitled “Improving Bahasa Melayu Text-to-Speech System using Unit-Selection Method: UTeM Health Centre Case Study” is the result of my own research except as cited in the references. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Signature : ……..

Name : ……..

Date : ……..……

APPROVAL

I hereby declare that I have read this thesis and in my opinion this thesis is sufficient in term of scope and quality for the award of Master of Computer Science (Software Engineering and Intelligence).

Signature	:	……..
Supervisor Name	:	……..
Date	:	……..……

DEDICATION

This thesis are dedicated to the Almighty God

Prophet Muhammad SAW

To my beloved parents, Mr. Slamet and Mrs. Husnul Khotimah

To my lovely sisters, Martini Puji Astuti and Ratih Nur Hanifah Sholekhah.

To my beloved partner, Galuh Wilujeng Saraswati, S. Kom, M. CS

To my precious country, Indonesia

To my scholarship, Beasiswa Unggulan from Kementerian Pendidikan dan Kebudayaan

ABSTRACT

The main objective of UTeM health center is to provide the best medical facilities for all the community on campus. Nowadays UTeM Health Center already implements a doctor-patient consultation system. In that consultation system has a calling system using Text-to-Speech (TTS) to call patient name to go to doctor's room but lack of natural sounding speech. That is way the research wants to make improvement for calling system UTeM Health Center with a TTS using Unit-Selection Synthesis (USS) method with Malaysian native speaker for the corpus database. The corpus consists of 200 famous Malaysian names and 1500 sentences. Those corpus segmented into word, syllable, and diphone based. TTS system consists of Natural Language Processing (NLP) and Digital Signal Processing (DSP). NLP is a module to produce a phonetic transcription of the read text. In this study, NLP process consists of text normalization and phonetizer. Then followed by DSP process using Unit-Selection Synthesis (USS) to convert phonetic transcriptions from NLP module into speech synthesis. This TTS has been tested by Comparison Category Rating (CCR) from listening test. The first test is to compare the naturalness of the current calling system and the proposed calling system. The proposed system can solved the problem of naturalness because its result is higher than current system. The second testing is to test the improvement of TTS method for Bahasa Melayu. The proposed system give improvement because the total average result score is 2.55 over 3.

ABSTRAK

Objektif utama pusat kesihatan UTeM ialah menyediakan kemudahan perubatan terbaik kepada semua komuniti di kampus. Kini, pusat kesihatan UTeM telah meletakkan sistem perundingan doktor pesakit. Dalam sistem perundingan tersebut terdapat sistem panggilan menggunakan teks kepada pertuturan (TTS) untuk memanggil nama pesakit untuk pergi ke bilik doktor tetapi kekurangan pertuturan berbunyi semulajadi. Itulah kehendak kaedah penyelidikan untuk menambah baik sistem panggilan pusat kesihatan UTeM dengan TTS melalui kaedah Sintesis Pemilihan Unit (USS) dengan penutur asal Malaysia untuk pangkalan data korpus. Korpus meangandungi 200 nama Malaysia dan 1500 ayat yang terkenal. Korpus ini akan disegmenkan berdasarkan perkataan, sebutan, dan difonetik. Sistem TTS mengandungi pemproses bahasa semulajadi (NLP) dan pemproses isyarat digital (DSP). NLP merupakan modul untuk menghasilkan transkripsi fonetik daripada teks yang dibaca. TTS telah diuji dengan perbandingan penarafan kategori (CCR) daripada ujian pendengaran. Ujian pertama adalah untuk membandingkan kesemulajadian sistem panggilan kini dan sistem panggilan cadangan. Sistem panggilan cadangan mampu menangani masalah kesemulajadian kerana keputusan sistem cadangan lebih tinggi berbanding sistem kini. Ujian kedua adalah untuk menguji kaedah TTS yang ditambah baik dalam bahasa Melayu. Sistem cadangan memberikan penambahbaikan kerana markah purata keseluruhan ialah 2.55 daripada 3.

ACKNOWLEDGEMENTS

Alhamdullilahi Robbil'Alamnin, first of all the researcher would like to express my deepest praise and gratitude to Allah SWT and Prophet Muhammad SAW who have given blessing and merciful to complete this thesis. Second, I would like to express my deepest appreciation to my beloved supervisor Professor Dr. Mohd Khanapi Abd Ghani for his guidance, suggestion, advice and encouragement throughout this study. I would also like to express my appreciation to Dean Faculty Information and Communication Technology, Prof. Dr. Burairah Bin Hussin and also Prof. Nanna Suryana Herman for their support and help towards my postgraduate studies.

I am especially grateful to Ministry of Education and Culture Indonesia, Ministry of Research Technology and Higher Education Indonesia and also Dian Nuswantoro Universisty (UDINUS) for providing scholarship Beasiswa Unggulan, without the scholarship, I could not have stated this master study in UTeM Malaysia.

Last but not least, thanks to my all friends Syaifur, Naufal, Tosan, Maulana, Wendy, Masduki, Aditya, especially Nuur Farhani, Firdaus, Safira, Erba, Zia'ul Haq, Nanang Arsyad, Kadek, all friends in Lab. Biocore and others for their support.

Finally, my deepest gratitude goes to my parents, Bapak Dr. Muljono, Ibu Ninuk's family and also Galuh Wilujeng Saraswati for their love, prayers and encouragement. To those who indirectly contributed to this research, my proud thanks for their kindness.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

TTS	-	Text-to-Speech
UTeM	-	Universiti Teknikal Malaysia Melaka
NLP	-	Natural Language Processing
DSP	-	Digital Signal Processing
NL	-	Natural Language
HMM	-	Hidden Markov Model
HTS	-	HMM-based speech synthesis
USS	-	Unit Selection Synthesis
CCR	-	Comparison Category Rating
NLG	-	Natural Language Generation
PAT	-	Parametric Artificial Talk
SAMPA	-	Speech Assessment Methods Phonetic Alphabet

CHAPTER 1

INTRODUCTION

1.1 Preamble

Language is the ability to communicate someone's mind through of a set of signs or texts, gestures, and sounds (Kamble and Kagalkar, 2014). All this time, speech through sound is the most important tool of communication and interaction between human (Mattheyses and Verhelst, 2015). Currently, communication not only between human being but also between human and computer. More than 50 years, speech-processing technology has been a mainstream area of research. The field of study from speech-processing are text to speech synthesis (TTS), speech recognition, and machine translation. TTS is the process of converting written text into speech, speech recognition is opposite with TTS, it is the process of converting speech into text, and machine translation is the process of converting writing or speech in one language into another language (Taylor, 2009). This study discuss about speech-processing technology that focus on TTS.

According to (Dutoit, 1997), TTS is a system that can make the machine automatically can speak text input. According to (Heggtveit, 2003) explain that TTS is a method of generating part of the communication mimic natural speech. Meanwhile, according to (Govind and Prasanna, 2013) TTS is the process of converting the message in text form into a message in the form of an oral equivalent. The main aim of the TTS is to generate natural-sounding speech synthesis (Muljono, 2016).

TTS unwitting very useful for life, especially for disabled people such as the blind and dyslexia. The blind only communicates through oral method, screen reader is TTS application can help the user navigate around computer systems. TTS also can read aloud the e-book called Digital Talking Book. For dyslexia TTS could help in learning to read the words, letters, sentences, and other symbols.

TTS system can be applied to many sectors as well. Such as in education (used to a teaching tool to talk to children and tourists), in computer gaming (used to voice characters automatically.), in business entertainment (used to reading weather reports, news stories, navigation, narration, email or short message service (SMS) playback and travel directions), in global business (used to call centers for the banking industries, warning/ alarm systems, the provider company, financial, airlines), and others.

1.2 Overview of Research and Its Context

Universiti Teknikal Malaysia Melaka (UTeM) was established on 1 December 2000. UTeM is a public university that using “Practice and Application Oriented” teaching and learning method for technical education in Malaysia. The university offers for diploma, bachelor, master, and Doctor of Philosophy (PhD) academic program level spread in seven faculties, one center, and one institute. The faculties are Electronics and Computer Engineering, Electrical Engineering, Mechanical Engineering, Manufacturing Engineering, Information and Communications Technology, Technology Management and Technopreneurship, Engineering Technology while the center is the Languages and Human Development, and lastly the institute is The Technology Management and Entrepreneurship (UTeM, 2016).

UTeM has their own health centre with 3 branches in the Technology Campus, Main Campus, and City Campus. There are doctor-patient consultation service, pharmacy

service, ward service, and dentist service to serve all the community on campus such as students, lecturers, and staffs. All that services directly under the management of office of UTeM's Student Affairs (UTeM, 2016).

Initially, UTeM health centre using patient calling system and doctor-patient consultation process manually. Now, the system change conventional approach become computer integrated with all three branches. The system using hybrid architecture that can operate during online and offline situation. This system can reduce time taken and improves response time for the doctor to input patient health record dramatically (Mukhtar, 2016). Although the system overall is good and runs well, the patient calling system that using TTS technology still using US pronunciation and robotic sound. It makes the sound aloud not naturally. Naturalness is one of the challenges in the TTS system beside speech intelligibility, cost effectiveness, and expressivity.

The process of TTS system divided by two processes, first step Natural Language Processing (NLP), capable of producing a phonetic transcription of the text read and Digital Signal Processing (DSP), transforms the symbolic information it receives into speech (Dutoit, 1997). There are five speech methods to generate TTS system, such as articulatory synthesis, formant synthesis, concatenation synthesis, statistically synthesis, and unit-selection synthesis.

1.3 Problem Statements

UTeM health centre already implements a doctor-patient consultation system. One of the main system is patient calling system that using Text to Speech technology to call the patient's name and the doctor's room or pharmacy service. The main problem of the current system is that the TTS is not accurate according to Malaysian Language. The

system still using US pronunciation and robotic sound that makes the sound aloud not naturally. Hence, made the patient did not recognize that her/ his name called by system.

Naturalness is one of the challenges in the TTS system beside speech intelligibility and cost effectiveness and expressivity. The purpose of this research is to propose an improvement method in order to provide a good calling system. The proposed method will be effective to make the system speech aloud naturally.

1.4 Research Question

Based on the background and problem statements, the research questions in this study as follows:

1. What is the suitable method to handle naturalness in TTS?
2. How does the method improve TTS method for Malaysian language?

1.5 Research Objective

In this research, there are two objectives to be achieved based on the research problems. The objective are:

1. To propose Unit-selection based to handle naturalness in TTS.
2. To propose Malaysian native speaker to improve TTS method for Malaysian language.

1.6 Significant of Study

The research is expected to produce an improvement of patient calling system in Malaysian Language. The system will present to the UTeM Health Centre for the deployment or further development for industrial usage either in Malaysia as well as South East Asian Country, such as Brunei, Indonesia, Singapore, Thailand, etc. The data and

findings from Malaysian speech recorded that will be the source of TTS sounds. By having this improvement, the patient calling system is able to solve the problem of naturalness.

1.7 Scope of Study

The scope of this research is focusing more on patient calling system using TTS technology. Under domain of computer science, this research covers the field of Artificial Intelligent. This research focus on solving the problem of naturalness. The scope of this research is shown on the underline words in Figure 1.1.

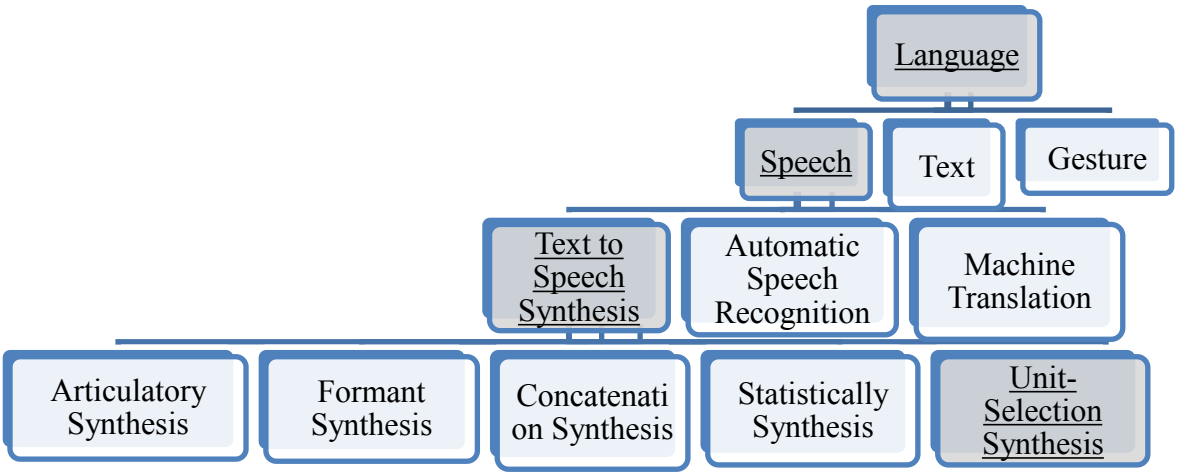

Figure 1.1: Scope of Text to Speech Research

The Unit-selection synthesis (USS) will be built for Malaysian people. The language is using Bahasa Melayu instead of Bahasa Kedah, Kelantan, etc which are more specific in the term of its region. The process is using Malaysian native speaker to provide the naturalness.

1.8 Organization of The Report

This research provides five chapters of the project report. The report organization structure is as follow:

Chapter 1 is the introduction part of this study. This chapter provides information about the origins of the research. This chapter also provides a brief outlining of background study, and problem statement that is followed by research question, research objective, significant of study and scope of study.

Chapter 2 is the literature review part of this study. This chapter presents related literature and studies on the research problem. Moreover, this chapter also describes Bahasa Melayu, Text to Speech Synthesis and its method.

Chapter 3 is the methodology. This chapter discusses the methodology used to achieve the research objectives, including the research method, research design, and proposed methodology.

Chapter 4 is conceptual design of Text to Speech. This chapter elaborates about general the work of this study. The result of the product also been presented in this chapter.

Chapter 5 is result. This chapter elaborates the case study in the previous chapter. Lastly describes the results that have been validated.

Chapter 6 is conclusion. This chapter summarizes the research attainment, concludes the research observations, limitations, and constraints and future works.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter begins with the overview of UTeM Health Center continue with Bahasa Melayu as Malaysian language. This chapter also provides a detail overview, analysis, and evaluation of Text-to-Speech in the natural language processing process, several methods, and the assessment test. The discussion of this chapter is to identify the suitable method that could be used for the Text-to-Speech.

2.2 Overview of UTeM Health Center

UTeM Student Clinic was established on 9 June 2004 and re-branded to UTeM Health Center on 16 October 2014. Its main objective is to provide the best medical facilities for all the community on campus such as students, lecturers, and staffs in 3 branches (Technology Campus, Main Campus, and City Campus). There are four Medical Officers, one Dental Officer, and professionally trained paramedics to make sure the medical services delivered safely and efficiently (UTeMHealthCenter, 2016).

UTeM Health Center have 21 staffs; 5 doctors, 1 dentist, 3 pharmacists, 4 medical assistants, 4 nurses, 1 secretary, and 1 administrative assistants. UTeM have 11,763 students and 899 staffs of local and international as at 31 December 2016. There are averages of 100 to 200 person that come there to every day. The UTeM Health Center's organization is shown in Figure 2.1

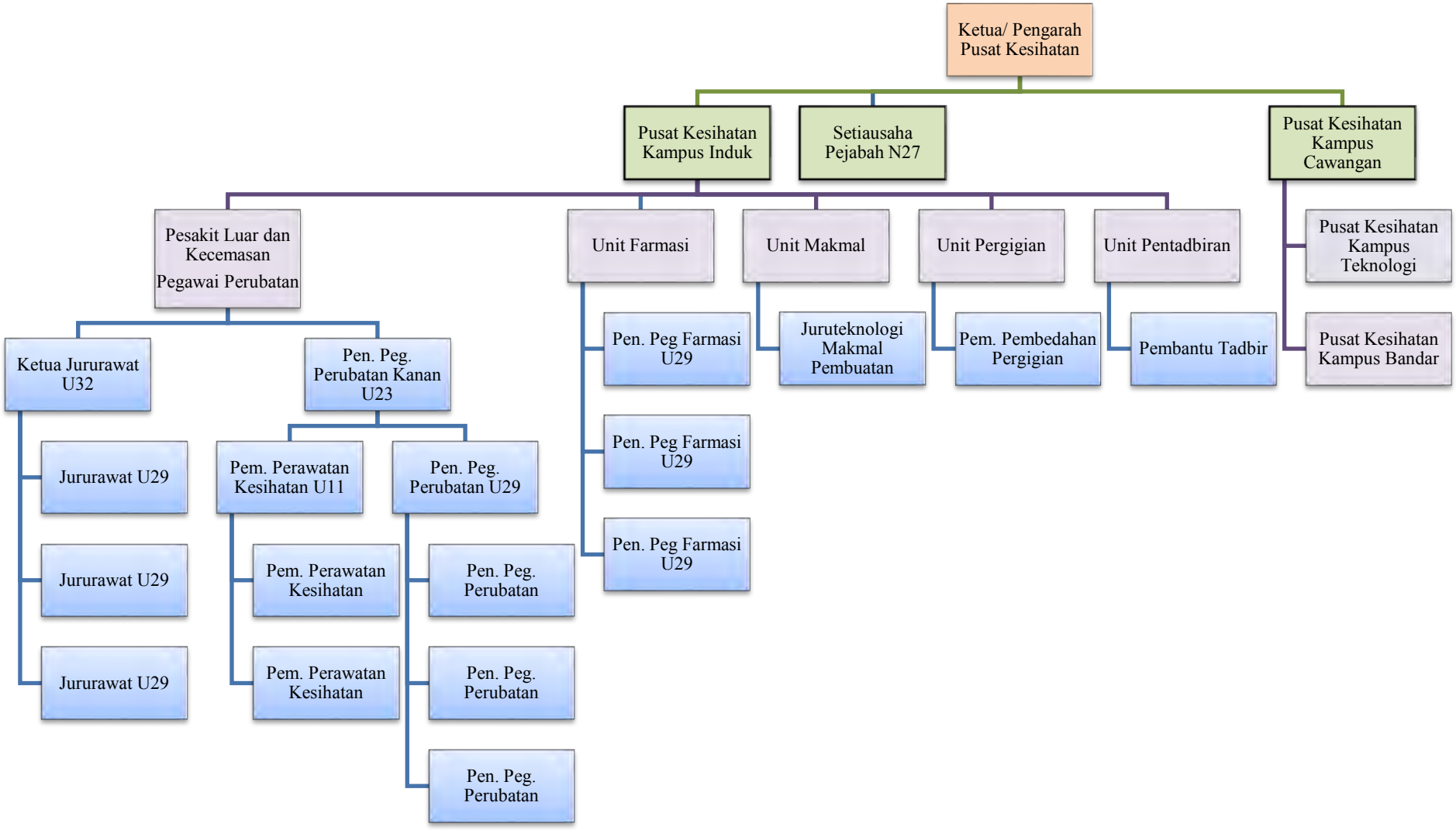

Figure 2.1: UTeM Health Center's Organization