

**UNIVERSITI TEKNOLOGI MARA**

**CONSTRUCTION MODEL OF NEW  
MALAY LANGUAGE LEXICONS  
USING MORPHOLOGICAL  
AFFIXED RULES**

**HARSHIDA BINTI HASMY**

Thesis submitted in fulfillment  
of the requirements for the degree of  
**Master of Science**

**Faculty of Computer and Mathematical Sciences**

July 2016

## **CONFIRMATION BY PANEL OF EXAMINERS**

I certify that a panel of examiners has met on 2<sup>nd</sup> December 2015 to conduct the final examination of Harshida Binti Hasmy on her Master thesis entitled “Construction Model of New Malay Language Lexicons Using Morphological Affixed Rules” in accordance with Universiti Teknologi MARA Act 1976 (Akta 173). The Panel of Examiners recommends that the student be awarded the relevant degree. The panel of Examiners was as follows:

Seripah Awang Kechil, PhD  
Associate Professor  
Faculty of Computer & Mathematical Sciences  
Universiti Teknologi MARA  
(Chairman)

Nurazzah Abd Rahman, PhD  
Associate Professor  
Faculty of Computer & Mathematical Sciences  
Universiti Teknologi MARA  
(Internal Examiner)

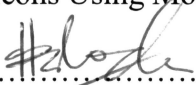
Rabiah Abdul Kadir, PhD  
Senior Lecturer  
Institute of Informatics Visual  
Universiti Kebangsaan Malaysia  
(External Examiner)

**DR MOHAMMAD NAWAWI**  
**DATO' HAJI SEROJI**  
Dean  
Institute of Graduates Studies  
Universiti Teknologi MARA  
Date: 14 July 2016

## AUTHOR'S DECLARATION

I declare that the work in this thesis was carried out in accordance with the regulations of Universiti Teknologi MARA. It is original and is the results of my own work, unless otherwise indicated or acknowledged as referenced work. This thesis has not been submitted to any other academic institution or non-academic institution for any degree or qualification.

I, hereby, acknowledge that I have been supplied with the Academic Rules and Regulations for Post Graduate, Universiti Teknologi MARA, regulating the conduct of my study and research.

Name of Student	:	Harshida Binti Hasmy
Student I.D. No.	:	2011109231
Programme	:	Master of Science in Information Technology and Qualitative Sciences – CS780
Faculty	:	Faculty of Computer and Mathematical Sciences
Thesis Title	:	Construction Model of New Malay Language Lexicons Using Morphological Affixed Rules
Signature of Student	:	 .....
Date	:	July 2016

## **ABSTRACT**

A lexicon is the source of specific knowledge about individual words in the language which also known as the heart of language processing system. This research will focus on construction of a computational lexicon model for Malay Language that involved computational study of the form and behaviour of words. This research also includes study on morphological arrangement of Malay affixation process which includes prefixes, suffixes, circumfixes and infixes with the intention of constructing a collection of new Malay lexicons that will be automatically constructed from a single root word. This research conducted experiments on 2101 root words found in the Malay translated Quranic documents. The words then experimented with Malay affixation rules using the affixed words analyser. Numerous new words are constructed from a single root word with the word classes using the affixed words analyser by adding 52 affixes rules which consists of 20 prefixes, 3 suffixes, 25 circumfixes, and 4 infixes to the root word. Nevertheless, proper names need to be extracted from the list and this is done by recognising and removing particular name entities. Finally, each new word is then compared with current Malay dictionary to ensure whether the word generated is currently being used or it is a new generated new word. Results from this analysis open opportunity to construct new Malay word variant to enrich the Malay lexicon and may help to support more efficient method for any related Malay language computer linguistic analysis particularly any research on Malay Quranic translation documents.

## ACKNOWLEDGEMENT

*Bismillahirrahmanirrahim*

In the name of Allah, The most Gracious, the most Merciful and Him alone is worthy of all praise for giving me strength and showering me an unforgettable experienced thought out writing this thesis.

Firstly, I would like to express my sincere gratefulness and gratitude to my supervisor Professor Dr. Zainab Abu Bakar for her invaluable guidance, encouragement and advice during the course of this thesis. She is also the one who encourage me the freedom of individual essential to fulfilment of my work and gave possible assistance in the pursuit of this study.

To my friends, thanks for all the moral supports that they had given me and other individuals who have directly or indirectly help me in completed my thesis.

Last but never the least, I would like to thank my beloved husband, Mohd Zainal Abidin, both of my children, Muhammad Eyzzull Addeen and Muhammad Aqeell Zuhaer together with my parents who always encouraged and helped me to uplift my morale to stand where I am today. I am so thankful for all their prayers, love, and support during my journey to complete my study. My appreciation also goes to my brothers and sisters and other family members for giving me motivations with positive energy. To my late grandmother, who passed away during the progression of completing this thesis, thank you very much for all the love, guidance, and precious life lessons you have taught me. You will be missed, Al- Fatihah.