DEVELOPMENT OF AN ELECTRONIC ARABIC DICTIONARY

SITI NOORMAYA BINTI BILMAS

UNVERSITI TEKNOLOGI MALAYSIA

DEVELOPMENT OF AN ELECTRONIC ARABIC DICTIONARY

SITI NOORMAYA BINTI BILMAS

A project report submitted in partial fulfilment of the requirements for the award of the degree of Master of Electric - Electronic and Telecommunication

Faculty of Electrical Engineering
Universiti Teknologi Malaysia

NOVEMBER 2009

Dedicated to my beloved father, mother and family.

ACKNOWLEDGEMENT

Alhamdulillah, after sometimes I have made it to finish up this project report. In preparing this project report, there are so many persons gave support and help.

First of all, I would like to convey my sincerest appreciation to my supervisor, Dr. Muhammad Nasir bin Ibrahim who has given guidance throughout the process of doing this project.

I also would like express my deepest appreciation to my family, especially for my mother for all their moral and financial support within these years.

To all my friends and colleagues thank you for your kindliness and understanding. Hope our friendship lass forever.

Last but not least, I am deeply appreciation to those individuals who involve indirectly throughout the process of doing this project.

Thank you so much to all.

"Gratitude is the memory of the heart."

Jean Baptiste Massieu

ABSTRACT

Currently, the arrangements of common Arabic dictionaries are based on the first letter of the words according to the Arabic alphabet sequence. For this 'Electronic Arabic Dictionary' it is developed based on the arrangement of Hans Wehr Dictionary edited by J. M. Cowan. Since published in 1964, this dictionary is the only one that has been arranged according to the Arabic root word. So far, there is no electronic version for this dictionary format. Therefore, the objective of this project is to develop the software format for Hans Wehr dictionary and as the first step is to design the database structure based on the root word arrangement concept. The content of this electronic dictionary is fully based on the Hans Wehr dictionary content. Besides providing the meaning for each word, this dictionary also gives information about the root word derivation forms, derivation words and their usage examples.

ABSTRAK

Kebiasaannya kamus bahasa Arab disusun mengikut huruf pertama perkataan berdasarkan susunan huruf Arabnya. Namun ‘Kamus Elektronik Bahasa Arab’ ini dibangunkan berdasarkan tatasusunan dalam kamus Hans Wehr yang disusunatur oleh J.M. Cowan. Sejak penerbitannya pada tahun1964, kamus ini adalah satu-satunya kamus yang mengenengahkan susunan perkataan berdasarkan kata dasar. Buat masa ini kamus ini masih tidak ada dalam bentuk elektronik. Maka, tujuan utama pembangunan ‘Kamus Elektronik Bahasa Arab’ ini adalah untuk membangunkan kamus Hans Wehr ini dalam bentuk elektronik dengan langkah pertama adalah untuk merekabentuk struktur pengumpulan data yang sesuai dengan konsep susunan perkataan berdasarkan kata dasar. Maklumat yang terkandung dalam kamus elektronik ini adalah sepenuhnya berdasarkan kamus Hans Wehr yang mana selain memberi makna perkataan, maklumat seperti kata terbitan berformat, kata-kata terbitan daripada kata dasar dan contoh-contoh penggunaannya turut disertakan.

TABLE OF CONTENT

CHAPTER	TITTLE	PAGE
	DECLARATION	ii
	DEDICATION	iii
	ACKNOWLEDGEMENT	iv
	ABSTRACT	v
	ABSTRAK	vi
	TABLE OF CONTENTS	vii
	LIST OF TABLES	ix
	LIST OF FIGURES	x
	LIST OF APPENDICES	xiii
1	**INTRODUCTION**	1
	1.1 Project Introduction	1
	1.2 Objectives	2
	1.3 Problem Statement	2
	1.4 Scope Research	2
	1.5 Research Background	3
2	**LITERATURE REVIEW**	4
	2.1 Arabic Language	4
	2.1.1 Hans Wehr Dictionary	9
	2.2 Unicode	11
	2.2.1 Unicode Coding	15
	2.3 Database	16
	2.3.1 Microsoft Access 2007	18

3	**METHODOLOGY**		23
	3.1	Research Stage Methodology	23
	3.2	Designing Stage Methodology	24
4	**DATABASE DESIGN**		27
	4.1	Database Structure	27
	4.2	Table	29
		4.2.1 Table Field	29
		4.2.2 Table Properties	33
	4.3	Data Entry	37
		4.3.1 Database Entry	38
	4.4	Relationship	41
	4.5	Primary Key	42
	4.6	Query	44
5	**RESULT & DISCUSSION**		49
	5.1	Result	49
	5.2	Discussion	55
6	**CONCLUSION & FUTURE WORKS**		62
	6.1	Conclusion	62
	6.2	Future Works	63
REFERENCES			64
APPENDICES			65

LIST OF TABLE

TABLE NO.	TITLE	PAGE
2.1	Arabic Alphabet	6
2.2	List of ‘Form’	7
2.3	Derivation Words	8
2.4	The Example of the Root Word, Forms & Derivations	8
2.5	Unicode for Arabic Standard Character	13
2.6	Unicode for Arabic Supplement Character	14
2.7	Example of Unicode Coding with Spacing Coding	15
2.8	Example of Unicode Coding without Spacing	15
2.9	Example of Unicode Coding for a Word	16
4.1	List of Table Field	29
4.2	Data No. Format	30
4.3	Character Sequence	31

LIST OF FIGURE

FIGURE NO.	TITLE	PAGE
2.1	Example of Hans Wehr Dictionary Entry	10
3.1	Research State Flowchart	24
3.2	Design State Flowchart	25
4.1	Flowchart of the Database Process	28
4.2	Example of DATA NO	32
4.3	Properties for Table 1	33
4.4	Properties for Table 2	33
4.5	Properties for Table 3	34
4.6	Properties for Table 4	34
4.7	Example of Table 1	35
4.8	Example of Table 2	36
4.9	Example of Table 3	36
4.10	Example of Table 4	36
4.11	Keyboard Layout for PC Tablet	37
4.12	Form for Table 1	38
4.13	Form for Table 2	39
4.14	Form for Table 3	39

4.15	Form for Table 4	40
4.16	Relationship One – to – Many	41
4.17	Database Tool	42
4.18	Primary Key Toolbar	43
4.19	Query Toolbar	44
4.20	Query Wizard Option	44
4.21	Example of Simple Select Query	46
4.22	Example of Created Query with Criteria	47
4.23	Dialog Box	47
4.24	Example of Query include Criteria using Multiple Table	48
5.1	Main Menu	50
5.2	Sample of User Manual	51
5.3	Sample of Root Word List	52
5.4	Dialog Box Filled with Root Word	52
5.5	Display for ROOT WORD MEANING Application	53
5.6	Display for FORM Application	53
5.7	Display for LIST OF EXAMPLE Application	54
5.8	Display for DERIVATION Application	54
5.9	Query for Root Word List	55
5.10	Query for Root Word Meaning	56
5.11	Query for Form	56
5.12	Query for Example	57
5.13	Query for Derivation	57
5.14	Report Toolbar	58
5.15	Report Wizard Menu	58

5.16	Control Toolbar	59
5.17	Main Menu Display in Design View	60
5.18	Properties of Command Button	60
5.19	Choose Builder Menu	61
5.20	Macro Function	61

LIST OF APPENDICES

APPENDIX	TITLE	PAGE
A	SCHEDULE SEM 1/2009	65
B	SCHEDULE SEM 2/2009	66

CHAPTER 1

INTRODUCTION

1.1 Project Introduction

Dictionary is a reference containing an alphabetical list of words with information about them. Arabic dictionary is a reference that lists all the Arabic words with their meaning. Currently, there are many Arabic dictionaries are being developed. It has been develop to be on line usage and off line usage. Various types of platform have been used in process to produce this kind of dictionary. But all these dictionaries have a similarity at the aspect of the data arrangement. All the words are arranged according to its first character of the word.

Due to this reason, this project will propose a new arrangement for this Arabic dictionary. Actually, the arrangement that will propose is an adaptation from a dictionary that written by Hans Wehr originated from Germany around 1964. This dictionary is arranged the Arabic words according to its root word. So far, this is the only dictionary that used this kind of system arrangement. And this kind of system is still not produced in software format. As far as this project concern is to produce a database structure that based on this system. Root word base system.

1.2 Objectives

The main objective of this project is to study and create a new database structure for the Arabic dictionary that is different in aspect of arrangement and improvement of the information of each word.

1.3 Problem Statement

Nowadays, Arabic dictionaries are having been developed using various platforms and various programming languages. All of it has the same data structure. All these Arabic dictionaries have the arrangement of database based on their first character of the words. Usually, the arrangement of the word in the database will start with ا and end with ي. With this kind of arrangement the user will never know the exactly the root word of these word. Or in other word user will never know from where those word are derived. Thus, in this project will propose a database that shows the information of the meaning of the root word to the derivation words of each root word in a form of software format.

1.4 Scope Research

This research will concern about the database structure for Arabic words entries. This research will cover up how the words will be arranged according its root word. Beside that, this research also concern about the detail of the root word information.

1.5 Research Background

Arabic words are derived from the root word. This root word is usually consisting two to four consonant. When it join with the other vowel is becomes can perform many words that give many different meaning. "Dictionary of Modern Written Arabic: Arabic – English" that written by Hans Wehr from Germany around year 1964, have been list out all the Arabic words according to its root word. This dictionary which has been edited by J. Milton Cowan is the only one dictionary that used this arrangement. Further more, this dictionary has been only produced in a form of book and never been produced in software format yet.

In order to produce the software format according the dictionary, a research towards the suitable format of database should be thinking of. The structure of the database should be suitable with the arrangement of the Arabic language in order to organize it according to its root word. Due to the large amount of information and the load of data, the software that will be used is also taking into account.

How to put the data into database format and how to relate the information in the database is also considered. How to make computer recognise the Arabic character?

These aspects which are the language and database structure are the basic things that must be searched in deep for this whole research of "Development of an Electronic Arabic Dictionary".

REFERENCE

[1]. Hans Wehr. *"Dictionary of Modern Written Arabic: Arabic-English"* edited by J.M.Cowan, 3th Edition, New York, 2004

[2]. G.W. Hansen, J.V. Hansen, *"Database Management and Design"*, 2nd Edition, New Jersey, 1996.

[3]. *"Arabic, Range: 0600-06FF"*, retrieved from www.w3c.org.ma/Talks/opstic3/data/arabicfomb.pdf on January 2009.

[4]. *"Arabic, Range: FB50-FDFF"*, retrieved from www.w3c.org.ma/Talks/opstic3/data/arabicfoma.pdf on January 2009.

[5]. *"Chapter 4: Language Analysis and Literature Review Findings: Arabic"*, retrieved from http://www.ncela.gwu.edu/restabout/research/nonroman/ch4_ARABIC.pdf on January 2009.

[6]. *"Database Basic Concept"*, retrieved from http://www.geekgirls.com/database_dictionary.htm on February 2009.

[7]. *"An Introduction to Unicode"*, retrieved from http://scripts.sil.org/cms/scripts/page.php?site_id=WSI_Guidelines_Sec_6_2 on February 2009.

[8]. William Black, Sabri Elkateb, *"Introducing the Arabic WordNet Project"* retrieved from www.globalwordnet.org/AWN/meetings/GWApaper.pdf on February 2009.

[9]. Mark Van Mol, Hans Paulussen, *"AraLat: A relational database for the Development of Bilingual Arabic Dictionaries"*, retrieved from www.kueleuven-kortrijk.be/~hpauluss/DOC/asialex.pdf on February 2009

[10]. Khirulnizam, Syuria dll, *"Solving the Arabic UTF-8 Character Transaction Issues in an Online Malay-Arabic Dictionary"*, retrieved from http://myais.fsktm.um.edu.my/1898 on February 2009.

[11]. *"What is Unicode?"* retrieved from http://www.unicode.org/standard/WhatIsUnicode.html on February 2009.

[12]. *"Arabic Grammar,"* retrieved from http://scripts.sil.org/cms/scripts/page.php?side_id=nrsi&item_id=WSI_Guidelines_s_Sec_6_2 on Mac 2009

[14]. *"Microsoft Access 2007 Top 10 Benefits"*, retrieved from http://office.microsoft.com/en-us/access/HA101650211033.aspx on April 2009.

[15]. *"Convert Microsoft Access Database Files in Legacy Encodings to Unicode"*, retrieved from http://vietunicode.sourceforge.net/howto/convertaccessdb.html on May 2009.

[16] *"Step-by-Step Guide to Build Sample"*, retrieved from http://support.microsoft.com/kb/193540 on Jun 2009.

[17]. Wahidudin Richard Shelquist, *"Brief List of Arabic Root"*, retrieved from http://wahiduddin.net/words/arabic_glossary.htm on Jun 2009.

[18]. "What is root?", retrieved from http://www.studyquran.co.uk/Qwiar.htm on May 2009.

[19]. *"Database Management System"*, retrieved from http://www.hermit.cc/teach/ho/dbms/dbms.htm on May 2009.

[20]. *"Arabic Language"*, retrieved from http://en.wikipedia.org/wiki/Arabic_language on April 2009.