

# MOBILE DICTIONARY

By

Siti Nabilah Ibrahim

Dissertation Submitted in partial fulfillment of  
the Requirements for the  
Bachelor of Technology (Hons)  
(Information and Communication Technology)

DECEMBER 2005

Universiti Teknologi Petronas  
Bandar Seri Iskandar  
31750 Tronoh  
Perak Darul Ridzuan

t  
PE  
1628.5  
.S35  
S622  
2005

1) English language -- Dictionaries  
2) IT/IS -- Thesis

# CERTIFICATION OF APPROVAL

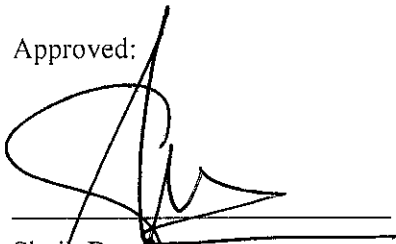
## MOBILE DICTIONARY

by

Siti Nabilah Ibrahim

A project dissertation submitted to the  
Information Communication and Technology Programme  
Universiti Teknologi PETRONAS  
in partial fulfillment of the requirement for the  
Bachelor of Technology (Hons)  
(Information and Communication Technology)

Approved:



Shuib Basri  
Project Supervisor

UNIVERSITI TEKNOLOGI PETRONAS  
TRONOH, PERAK

December 2005

## **CERTIFICATION OF ORIGINALITY**

This is to certify that I am responsible for the work submitted in this project, that the original work is my own except as specified in the references and acknowledgements, and that the original work contained herein have not been undertaken or done by unspecified sources or persons.



---

SITI NABILAH IBRAHIM

## **ABSTRACT**

Having difficulties in understanding the meaning of particular word in a particular language is unavoidable. The solution for this concern is either paper-based or electronic dictionary. With growing technological advancement, electronic dictionary keeps on emerging. Current existing electronic dictionaries are limited with functionality to translate for multiword correctly. This paper aims to discover a searching algorithm that can read single and two words (multiword) in a Malay or English language and translate it correctly to another language, either Malay or English. The objectives of the study is to develop a Mobile Dictionary system, namely MobileDic which able to cater for single and multiword translation in Malay and English language in J2ME-enabled mobile phones. The system incorporates a searching algorithm, by using Hashtable function. The system prototype should be able to recognize whether the input word is a single or multiword and intelligently translate the word matched by Hashtable key and value. The multiword will have its own separate database or corpus listing all the possible second word after the first word of that multiword. The output of this system will be evaluated by users who already have experienced in using existing electronic dictionaries based on MobileDic accuracy as well as its usefulness. As a conclusion, MobileDic is able to improve current situation and can be a useful look-up source especially to students and tourist.

## **ACKNOWLEDGEMENT**

First and foremost, I would like to thank God, for His blessing that helps me to complete this Final Year Project. To be honest, the process to get the project done was tough and difficult, but in the same time it was also fun and interesting. There was at a time I felt like giving up this project, but I received a lot of motivations from people around me to continue it. Thus therefore I would like to take an opportunity here to give appreciations them for all the supports that have been given to me.

A special thanks goes to my supervisor, Mr Shuib Basri for suggesting an interesting topic and choosing me as her FYP supervisee. He was always there to listen and give advice to me. He helps me to complete the writing of this dissertation and the challenging research that lies behind it. Not forgotten to the IT/IS FYP Coordinator and Information System Department of Universiti Teknologi Petronas for the chance and guideline given to complete the project.

In addition, I would like to extend my gratitude to all my friends that helped me during the period of completion especially Mr.Mohd Khairul Zarir who has helped me a lot from the beginning of the project till the end. Million of thanks also go to the rest of my people who directly or indirectly contributed in giving ideas and opinions regarding to this project.

Finally, I would like to thank all lecturers and all my colleagues who are involved either directly or indirectly in completing the FYP. All ideas, advices and compassions to me are greatly appreciated.

## TABLE OF CONTENTS

<b>CERTIFICATION OF APPROVAL</b> .....	<b>ii</b>
<b>CERTIFICATION OF ORIGINALITY</b> .....	<b>iii</b>
<b>ABSTRACT</b> .....	<b>iv</b>
<b>ACKNOWLEDGEMENT</b> .....	<b>v</b>
<b>LIST OF TABLES</b> .....	<b>viii</b>
<b>LIST OF FIGURES</b> .....	<b>ix</b>
<b>LIST OF ABBREVIATIONS</b> .....	<b>x</b>
<b>CHAPTER 1 INTRODUCTION</b> .....	<b>1</b>
1.1 Background of Study .....	1
1.2 Problem Statement.....	3
1.2.1 Problem Identification.....	3
1.2.2 Significant of the Project.....	4
1.3 Objective and Scope of Study .....	5
1.3.1 Objectives.....	5
1.3.2 Scope of Study .....	5
<b>CHAPTER 2 LITERATURE REVIEW</b> .....	<b>6</b>
2.1 Dictionary .....	6
2.1.1 Today's Trend of Dictionary.....	7
2.2 Mobile Technology.....	8
2.2.1 Mobile Phone Market in Malaysia.....	8
2.2.2 Mobile Phones Are Portable .....	9
2.2.3 Mobile Platform Limitations.....	9
2.3 Dictionary Look-Up Approaches .....	10
2.3.1 Related Works.....	11
2.3.2 HashTable as Searching Component .....	14
<b>CHAPTER 3 METHODOLOGY / PROJECT WORK</b> .....	<b>16</b>
3.1 Introduction .....	16
3.2 Research Methodology .....	18
3.2.1 Requirement Analysis and Definition.....	18
3.2.1.1 Questionnaires .....	19
3.2.1.2 Observation .....	20
3.2.2 Research and Feedback Analysis .....	20

3.3 Development Methodology .....	21
3.3.1 System Design.....	21
3.3.1.1 Designing System Workflow .....	21
3.3.1.2 Designing Searching Algorithm .....	23
3.3.1.3 Designing Database Record .....	28
3.3.1.4 Designing User Interface.....	30
3.3.2 System Construction .....	31
3.3.3 Testing and Integration .....	31
3.3.4 System Closing .....	32
3.4 Tools Required .....	32
3.4.1 Hardware.....	32
3.4.2 Software .....	32
<b>CHAPTER 4 RESULT AND DISCUSSION .....</b>	<b>33</b>
4.1 Overview .....	33
4.2 Functional Testing.....	33
4.3 Integration Testing.....	37
4.4 Evaluation.....	39
4.4.1 Pre Analysis Users' Evaluation.....	39
4.4.1.1 Mobile Phone Usage .....	39
4.4.1.2 Emphasis of Existing Dictionaries on Multiword (two- word) translation.....	40
4.4.2 Post Analysis Users' evaluation.....	40
4.5 Screen Capture and Description .....	44
<b>CHAPTER 5 CONCLUSION AND RECOMMENDATION .....</b>	<b>49</b>
5.1 Conclusion.....	49
5.2 Recommendation.....	50
5.2.1 Expansion of multiword.....	50
5.2.2 Suggested most similar meaning if no match is found .....	50
5.2.3 Translator in Mobile Phone.....	50
<b>REFERENCES .....</b>	<b>51</b>
<b>APPENDICES .....</b>	<b>54</b>

## LIST OF TABLES

Table 4.1 Test Results of Each Module.....	33
Table 4.2 Test Results of Integration Testing.....	38



## LIST OF FIGURES

Figure 2.1 Mobile Phone Ownership in Malaysia.....	9
Figure 3.1 Project Methodology.....	17
Figure 3.2 MobilDic system workflow .....	22
Figure 3.3 MobileDic system Searching Methodology.....	24
Figure 3.4 Database containing only the letter A. ....	29
Figure 3.5 The separation of multiword expression in different text file.....	29
Figure 4.1 Chart on Percentage of Respondents owning a mobile.....	39
Figure 4.2 Chart on Percentage of Preferences of having Dictionary in Mobile Phone .....	39
Figure 4.3 Chart on Percentage of emphasis of existing dictionaries on multiword (two-word) translation.....	40
Figure 4.4 Chart for MobileDic ability to correctly translate multiword and single word .....	41
Figure 4.5 Chart showing the percentage whether MobileDic system is useful .....	42
Figure 4.6 Direct translation of word 'kedai makan' through querying each single word .....	43
Figure 4.7 Multiword translation through querying the multiword directly .....	44
Figure 4.8 MobileDic Splash Screen.....	44
Figure 4.9 MobileDic Main Menu Screen.....	45
Figure 4.10 MobileDic Enter Malay Word and Result Screen.....	45
Figure 4.11 MobileDic Enter English Word and Result Screen.....	46
Figure 4.12 MobileDic Word Not Found Screen.....	47
Figure 4.13 MobileDic About Screen.....	47
Figure 4.14 Mobile Dictionary Help Screen.....	48

## **LIST OF ABBREVIATIONS**

### **J2ME**

Java 2 Platform, Micro Edition

### **PDA**

Personal Digital Assistant

### **SMS**

Short Message Service

### **Multiword**

The combination of more than one word to form one meaning of word

### **MIDS**

Micro Information Devices

### **UTP**

Universiti Teknologi Petronas

### **Corpus**

A large collection of words

### **Lexicon**

Dictionary

# CHAPTER 1

## INTRODUCTION

### 1.1 Background of Study

Dictionary is considered as one of the important reference tools in people's daily life. The evolution of dictionary has helped large group of users in the whole world, especially students, travelers and tourists. There are several paper-based Malay-English or English-Malay dictionaries, one of which is Malaysia's best selling bilingual dictionaries, Kamus Dwibahasa Oxford Fajar Inggeris-Melayu and Melayu-Inggeris [15]. As the world becoming more global and technology becomes widen, electronic dictionary keep on emerging. People became less dependent on hard copy dictionaries. There are a number of Malay electronic dictionaries can be accessed online, one of which is Dr.Bhanot's Malay-English Cyber Dictionary [16] and Online Little Malay Dictionary Search, small searchable online trilingual dictionary for Malay-English-Finnish [17]. These dictionaries take advantage of their electronic form to allow searches from either English or Malay.

From Personal Computer (PC) platform, electronic dictionary has moved forward to be a part of portable devices functions, such as mobile phone, handheld devices (pocket dictionary) and PDA. Bilingual machine readable dictionaries are considered important and indispensable information resources for cross-language information retrieval, machine translation and so on. Recently, these cross-language informational activities have begun to focus on specific academic or technology domains. A Pocket Malay Dictionary was introduced to the market by Zuraidah Omar [21], which enables the users to look up words from Malay to English or English to Malay. Apart from pocket dictionary, a downloadable dictionary in mobile phone has become a trend. [22] Malay-English Mobile Dictionary can be easily downloaded to any compatible phones and users can easily retrieve the desired word translation.[18]

eBuzz Active Dictionary has also come out with dictionary in a mobile format, in which the users are required to send via SMS to the service centre or the processing server in order to get the translated word.

However, all of the electronic dictionaries mentioned above either web-based, pocket-based or mobile-based, seem to focus on converting only the single word or specifically doing direct translation.

This study will focus on how the dictionary can be developed in mobile phone and proof the concept that through a proposed searching algorithm, it does not just provide direct translation, but it can perform a correct translation in another language if two distinguished words, which are type of noun, in which if combined, will result in a meaningful noun, are entered. The system, namely MobileDic will perform two way bilingual translations, which is Malay - English or English - Malay, depending on users' selection. Even though there are many limitations of mobile phones compared with personal computers (such as small display screen, low processing power, small keypad and much more), but mobile phones own two very important advantages over the pc which are mobility and small in size. These criteria are vital especially for students and tourists by enabling them to carry it around. Therefore using mobile phone as the platform is the best option.

Why not using PDA, laptop or other handheld devices? They are mobile gadgets too. The reason is not many students or tourists own PDA and laptop is too bulky. The use of handheld devices such as Pocket Dictionary would cost them more. Why need to spend money to buy PDA or handheld devices when you already own a mobile phone? Why need to carry a bulky laptop wherever you go when mobile phone can be easily slipped inside your pocket?

## 1.2 Problem Statement

### 1.2.1 Problem Identification

Availability of electronic dictionary in World Wide Web, PDA, handheld devices (specifically Pocket Dictionary) and mobile phones is deniable. In fact, most dictionaries for Japanese, South African and French languages are already available in the market. South Africa for example, has come out with Mobidic the mobile dictionary [10], which requires querying via SMS in order to get intended word in English language. Japan also has emerged with Electronic Dict Lingo Global 29, a pocket dictionary with their special features, ability to translate up to 29 languages in the world [19]. In Malaysia, there has been several works on electronic Malay – English dictionary. The most popular online dictionary is Dr.Bhanot’s Malay-English Cyber Dictionary [16]. Kamus Melayu-Inggeris (KAMI) was also developed in order to build a lexicon primarily for use in NLP applications, in particular, machine translation [20]. A Pocket Dictionary by Zuraidah Omar [21] called, A Pocket Malay Dictionary, which perform the translation of words from Malay to English and English to Malay is also available in the market.

The problem arose when most of dictionaries intended to translate from Malay words to English and vice versa, ignores the possibility of Malay or English noun which consist of more than one word (multiword). Some of the dictionary, such as Bhanot’s Malay English dictionary [16] just provided a single word translation, which means word to word translation. What if the users need to obtain the meaning of more than one distinguished words? In addition to that problem, some of common Malay noun consist of two words and after the combination of these two words; it will be translated to a single word in English. The best example to illustrate this is one of the common words in Malay [23], which is “ubat gigi”. The correct translation for this word in English would be “toothpaste”, not “tooth medicine”. The same case happens to English word as well. Some of the English words are made up of multiword as well. For example, the word “shopping complex” is supposed to be translated as “pusat membeli belah” in Malay, not “membeli belah rumit”.

In order to overcome this, several translators are developed especially for phrase or sentences translation. Yes, some of them are able to do a very good translation for sentences and long phrase from both languages. But, most of them did not really emphasize the possibility of users entering two words. Most of the translators will actually do a direct mapping if two words are entered. One of the examples is Affendi Malay-English Translator [11], when the word “gigi bongsu” is written as an input and the actual output came out was “the youngest teeth”. This is obviously not the correct translation. It is supposed to produce “wisdom tooth”.

From here, comes the idea of developing dictionary running on mobile phones that will focus on translating maximum of two possible words. Mobile phone is chosen since it will ease the targeted users better than stand alone system as it is being considered as a must have item by 90% of Malaysia’s population [8]

### **1.2.2 Significant of the Project**

The final product will be able solve the problem with existing electronic dictionaries regards to multiword translation, whereby most of them are only doing single word translation and direct mapping whenever two words are entered. Apart from that, the project will also emphasize the presence of various multiwords in English and Malay language being used in daily life. It also able to assist targeted people mostly students, travelers and tourists. Looking up the word using mobile dictionary will ease them without the need to carry large dictionary or spend additional cost on pocket dictionary.

### **1.3 Objective and Scope of Study**

#### **1.3.1 Objectives**

The objectives of this study are as follows :

- To develop a mobile dictionary, namely MobileDic that will be able to perform single and multiword (maximum of two words) translation that fall under categories of noun.
- To study on the approaches in translating multiword (maximum of two words) from Malay language to English language or vice versa
- To come out with an efficient method of searching that can read multiword and translate them correctly.

#### **1.3.2 Scope of Study**

- Focusing on method of searching that can read multiwords which consist of two words only.
- Focusing on searching algorithm to be implemented using Mobile Development Platform , J2ME
- Focusing on common Malay and English words and multiword ( maximum of two words) which fall under noun category
- The final product would be a dictionary application in a mobile phone.

## **CHAPTER 2**

### **LITERATURE REVIEW**

This chapter reviews all the studies, theories and research facts done by credible professionals in their respective field which are published in world wide journals. These facts are being used as the basic of this project. All studies, theories and research facts discussed are based on identified sources and not merely on personal opinion

#### **2.1 Dictionary**

Dictionary is defined as a list of words with their definitions, a list of characters with their glyphs, or a list of words with corresponding words in other languages [3]. A good dictionary is an invaluable tool for anyone who uses the written word and would like to do so more effectively and correctly [1]. From this definition, it can be said that dictionary can be very useful whenever an individual is having difficulties in understanding unfamiliar words in some languages. Apart from that, it can also be used as a spell-checker, pronunciations and derivations for words.

Most traditional dictionaries are prepared in book form. After decades since dictionary was introduced, people are looking forward to a better dictionary and even a more sophisticated one. Good paper-based dictionaries are too fat. That is the reason why students leave them at home, why teachers do not carry sets from classroom to classroom; and why all but the most enthusiastic users deal with only one dictionary at a time, rather than pooling the resources of several different volumes [6]. For example, Webster Dictionary, although it uses light weighted papers, people still considered its size. It is designed to be very large and thick; causing it is difficult to carry around



### **2.1.1 Today's Trend of Dictionary**

Parallel with the development moving towards paperless world, dictionary has also been recognized as one of the reference tools that should go computerized. In order to overcome some of the problems with traditional dictionary, which is its size and weight, **electronic dictionary** has been introduced. [4] Electronic dictionary is a dictionary which is stored on computer and can be accessed by programs, so that definitions can be looked up and displayed on screen. There are three types of electronic dictionary, which are [6] hand-held (portable) electronic dictionaries, dictionaries on CD-ROM, and dictionaries on the Internet. [2] It is more effective to use computerized dictionary information than the traditional dictionaries in book form due to the following reasons :-

#### ***High speed searching for dictionary's information***

By searching the dictionary's information by a computer, a user can obtain necessary information more quickly than turning the pages of a dictionary [2].

Like nowadays, using traditional dictionary would be time consuming. Restraining eyes to look just for one word would be very tiring and normally will take around 1 minute. With electronic dictionary, word can just be looked up within 1 to 2 seconds.

#### ***Modifying and revising the dictionary's information easily***

Because of the very large cost, a dictionary which is written on paper has few opportunities for revisions. But, if the dictionary's information is dealt with on a computer, it can be revised every day [2].

Newly information needs to be updated in the dictionary. The cost of maintaining a paper form dictionary would be much expensive as the whole dictionary needs to be re-built again. But not with electronic dictionary, the new words just need to be added to the database and can be displayed straight away.

## 2.2 Mobile Technology

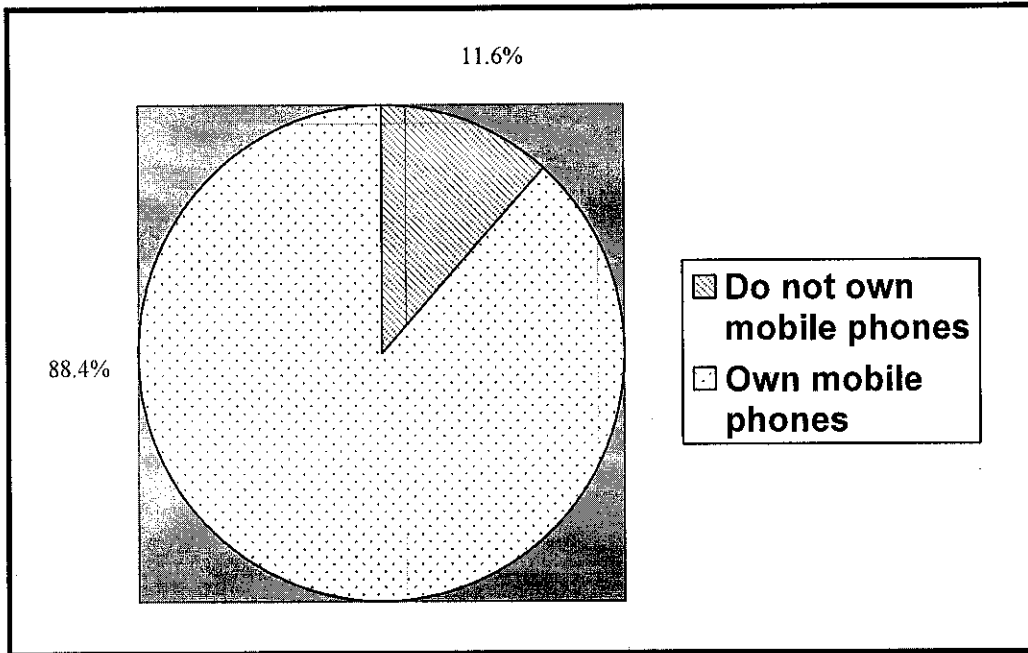
*"...we believe that after the desktop, the handheld device is probably the most important critical piece of equipment, more than a laptop, and could become the second most prevalent piece of IT equipment in and out of the office...the introduction of wireless handheld devices will lower the total cost of ownership of IT equipment, increase productivity, and lead to an increased ROI for enterprise technology." [7]*

The statement proved that mobile technology inventions extend the daily activities boundary far beyond before. When talking about mobile technology, it would relate to things such as PDA, Laptop, and mobile phones. As approximately 90% of Malaysia's population owns their own mobile phone [8], this gadget is something that considered as must-have item. Today mobile phones are not just for makes or receive calls. Many features have been integrated with mobile phone, including dictionary itself.

### 2.2.1 Mobile Phone Market in Malaysia

Mobile phone usage in Malaysia keep on increasing as technology expands. Research study conducted by NST [24] stated that there is a huge number of Malaysian carrying mobile phone i.e 12.3 million users which is almost half the total population. Those up to 19 years of age accounted for 12.3 percent of the total mobile phone subscribers while adults (users from 20 to 49 years of age) accounted for 78.8 percent. Majority of the students and working adults agreed that they "cannot live without their mobile phone". They always bring mobile phone just to make sure that they did not miss any calls. This shows that mobile phone has the highest preference over other mobile gadgets since it is considered as a part of people's daily lives.

Figure 2.1 on the next page illustrates the percentage of mobile phone ownership in Malaysia conducted by Yahoo [25]. The chart highlighted that almost 90% respondents currently own hand-phone and only 11.6% respondent do not own mobile phone.



**Figure 2.1 : Mobile Phone Ownership in Malaysia**

### **2.2.2 Mobile Phones Are Portable**

One of the advantages of having features in mobile phones is because of its convenience small size and portability. [9] With access at your fingertips anywhere you go, you can do many things and not be entirely dependent of your PC to manage your activities. It gives you the convenience of carrying only your hand phone, not your bulky laptop. In short, having a dictionary feature in mobile phone gives convenience on the move.

### **2.2.3 Mobile Platform Limitations**

Designing application on mobile platform differs from an application that works well on the bigger systems. In designing a dictionary in mobile platform, there are few things to be considered with regard to the platform issues. The following issues will look at the mobile platform issues [26].

### ***Screen Limitations***

MID screens are small LCD-based displays typically in the range of 100-200 square pixels. This makes screen real estate a scarce commodity [26]. So, in designing this dictionary application, we cannot expect the interface to look as interesting as in the desktop-based. All that is needed is just a simple interface that can be easily navigated.

### ***Processor Power***

Due to size and power consumption constraints, MIDs typically have low-end CPUs. [26] The slow performance is accentuated by the lack of floating-point hardware because extra time is needed to be spent to do that work as well. Corresponding to this, the users cannot expect to have an application which has faster performance than high-end CPUs. From designing perspective, the only possible way to encounter this issue is through minimizing the requirements and having some methods to reduce the CPU requirements for code.

### ***Storage Limits***

Memory constraints play a big role and the first to encounter in developing mobile application. Modern MIDS can have quite a range of memory capabilities, anything between 100KB and 500KB is common [26]. Another type of storage called, non-volatile RAM used by the RMS (Record Management System). Unlike PC hardware, it has a virtually unlimited space in which to store perhaps saved games or other state information. However, corresponding to this essential limitation, this dictionary application is not expected to have a list of all possible words in this world. Due to this, common words used in daily lives only are selected to be listed.

## **2.3 Dictionary Look-Up Approaches**

Dictionary look-up methods are the concept that is being centralized in this study. Several works have been identified and it involved the development of dictionary ranging from desktop-based to mobile-based. There are several approaches that have been adopted in developing those dictionaries. However, some of them only focus on

single word searching, which resulting the multiword to be ignored and the tendency to produce incorrect translation is quite high. Some of the related works are discussed below ;

### **2.3.1 Related Works**

Most existing electronic dictionaries, mobile dictionaries or translators used different kind of techniques and approaches to gain the correct results.

#### **Malay-English Mobile Dictionary**

There has been work going on Mobile Dictionary. One that has been developed is a Malay-English Mobile Dictionary developed by Unreal Mind Interactive Bhd [22]. This dictionary has the capability of translating words from Malay to English only. However, this dictionary is only concentrating on translating a single word, which might have been listed in the conventional dictionaries only. It does not encounter the possibility of Malay and English words that make up a noun when two distinguished words are combined.

#### **SMS-based Mobile Dictionary**

Some countries such as South Africa has implemented a Mobile Dictionary, which is called MobiDic. By simply sending a word for which the user needs an explanation to a premium - rated short code number, 34744, they will instantly receive a definition in return via short message service (SMS) [10]. eBuzz Active Dictionary, developed by AKN Messaging Technology also use the same SMS feature in order to get the word in either English or Malay [18]. TIM Mobile Dictionary [27] also is an SMS-based dictionary which supports 7 different languages, which are Greek, English, German, French, Italian, Spanish and Albanian. Users are charged per SMS sent to the service centre.

Having this kind of approach is not really efficient as it will burden the mobile phone users through the SMS charges. In addition to that, those SMS based Mobile Dictionary only performs direct translation (word to word). The project that is going to be developed will not charge the users, instead, the dictionary feature is going to be as a part of mobile phones function.

### **KAMI (Kamus Melayu-Inggeris)**

Kamus Melayu-Inggeris (KAMI) is similar to above approaches, but it is not developed in a mobile platform. This electronic dictionary was introduced to satisfy three criteria, to develop and document detailed syntactic features, to use a well-developed semantic ontology and to get as wide cover as possible [20]. Each entry consists of nine major fields, which include Malay Index Word, Malay Root Word, Part-of-Speech, Syntactic features, Semantic features, English Translation, English Definition, Chinese Translation and Meta Tags. KAMI allows users to look up any field of the dictionary. Semantic information can be used to disambiguate words that have same spelling but different meaning. For example, *perang* has two distinctive meanings : ‘brown (as in color) or “war”’.

Although KAMI successfully categorized each possible word under semantic classes as well as syntactic features, it treated multiword expressions as an index word. For example, “*tidak boleh dibahagi*” is stored as an index word and converted to “*indivisible*”. In the proposed approach for this project, each entry in index word field (primary record) would consist of only one word. In order to cater for multiword expressions (maximum of two words), each possible word that might come after the first word, and combining those two words would cause meaningful expressions, will be listed in secondary record.

### **ALT- J/M: the Automatic Language Translator – Japanese to Malay**

ALT-J/M used method of combining two bilingual dictionaries to make a third, using one language as a pivot. It combines a Japanese-English dictionary with a Malay-English dictionary, to produce a Japanese-Malay dictionary suitable for use in a

machine translation system [14]. The overall flow of this system is based on crossing the dictionaries. Once Japanese word is entered, all the possible meaning in English will be listed and then the Malay translations of these English translations. For each Malay word, they then look up all of its English translations, and see how many match the English translations of the Japanese word. The more matches there are, the better the candidate is. If entry in Malay word matches more than one word, then that Malay word will be most suitable match (output) for the Japanese word (input). For example, the word 'in' in Japanese matches with English word of 'mark', 'seal' and 'stamp'. Then, from Malay-English pairs, 'tera', 'mohor' and 'anjing laut' shares the same word. Since 'tera' found four possible matches within English word list, ALT J/M will assign the Malay word to the corresponding Japanese word (input).

The usage of crossing the dictionary is suitable to develop other languages, such as Japan and French by using English language as intermediate language. This project aims to have only two languages involved, which is English and Malay languages, so there is no need to cross against other languages.

### **Affendi Malay-English Translator**

One of the translator available in world wide web is Affendi Malay-English Translator[11]. Affendi Malay-English Translator is an online language translator system that translates English text to Malay text. It focuses on translating word, text, email messages or websites. Even though it is under category of translator and not dictionary, it can also be evaluated whether it focused on the presence of multiword since a translator should produce a correct translated word regardless whether the word is a single word or multiword.

Affendi Malay-English Translator able to translate long text, email or websites effectively. Nevertheless, if two words are entered, the system failed to associate the relationship between those two words and translated those two words by using direct translation. One of the obvious example would be the word "gigi bongsu", which is converted to " the youngest tooth"

The above four approaches did not really concentrate on possibility of two words entered by users. KAMI (Kamus Melayu-Inggeris) did allow multiword expression to be a part of data in the record; however, this may reduce processing time as there are a lot of possibilities of multiword used in daily lives. As a result, the data grows bigger and bigger, and the system needs to find line by line until the result is obtained and will in turn, reducing the processing time.

The suggested approach in this study is through splitting the multiword and assigning each of them as a single index word. Each multiword will have its own file and all the possibility of the second word or next word after the occurrences of the first word will be stored inside that particular text file. This can prevent the record of words to become a large record. By separating the multiword expressions, each possibility for the second word will be listed in another record, and the system may look up to that record only if two words are entered. This can ensure the system to perform faster.

### 2.3.2 HashTable as Searching Component

Hash tables offer a method for quickly storing and accessing data based on a key value. A hashing algorithm is used to determine the location of the associated record. [11].

[30]Essentially, a hash table is an implementation by memorization of a series of partial functions; a hash table can also be regarded as a mutable set of pairs. More informally, a hash table provides an implementation of a *dictionary* interface. Hash table implementations vary as to their Application Programming Interface (API), but most resemble the following methods, called on an instance of a hashtable, which is denoted by HT.

GET(key) function will return the value currently associated with key, if (key,value)  $\in$  HT.

PUT(key,value) causes Hashtable becomes a set

$$(HT - \{(k,v) \mid k = \text{key}\}) \cup \{(\text{key},\text{value})\}$$

REMOVE(key) causes Hashtable to become  $(HT - \{(k,v) \mid k = \text{key}\})$



Hashtable function can be used as a searching component for dictionary lookup. The key of the hashtable will be determined by the input given by users and the value will be determined by the translated word in different language of the key, as illustrated as the following : -

Dictionary (wordm, worde), whereby Dictionary is the Hashtable itself, wordm is the Malay word and worde is the English word. The same thing will happen if the key is the English word and value is the Malay word. The multiword will be recognized prior before the word look up is been done in order to recognize whether the input word is a single or multiword.

Hashtable is suitable in implementing lookup or pairing application such as dictionary since it provides faster matching process. This is possible whereby the key will act as the input word and the value will be returned as a matched word to that corresponding input [30].

## **CHAPTER 3**

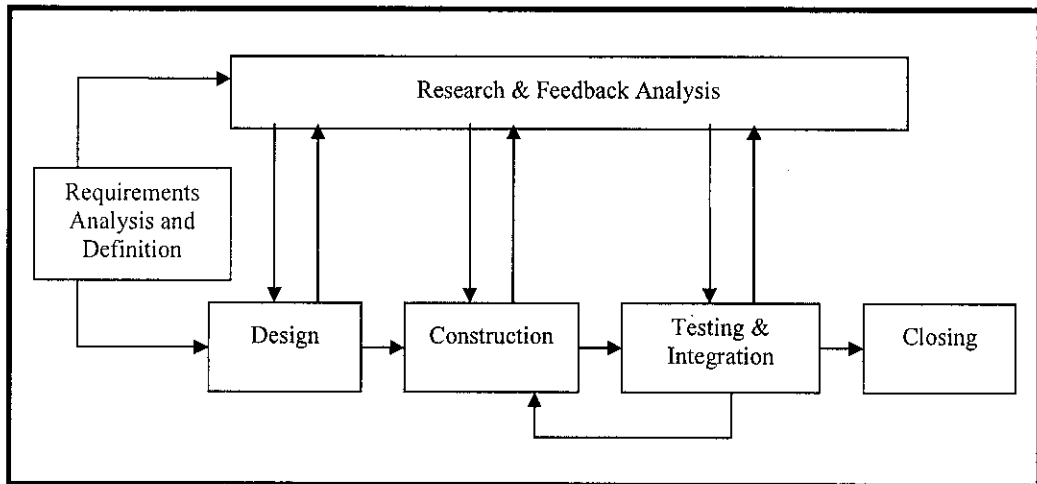
### **METHODOLOGY / PROJECT WORK**

#### **3.1 Introduction**

Methodology is defined at the early stage in order to maintain consistent system development life cycle of the proposed system. Several methodologies have been adopted and mostly used in various software environments, such as waterfall model, formal systems development, evolutionary development and reuse-based development.

Since methodology plays an important role, it has to be carefully chosen. An initiative to develop a wholly own methodology are also encouraged. For this project, several researches have been done in order to decide what kind of methodology is the best for this MobileDic.

The methodology that is used for this project is based on available methodology and is combined to best suit the requirement and time constraint of the project. Basically, it consists of two parts, which are research methodology as well as development methodology. Research methodology comprises of all research-based work, ranging from questionnaire, observation as well as literature study. Research methodology is essential in determining requirements of MobileDic system and obtaining the required functionality based on observation, questionnaires for respondents and literature study. Research will be done concurrently with development of the project. The development methodology consists of all the phases involved with designing MobileDic system, actual development of the MobileDic, the integration and the testing.



**Figure 3.1 : Project Methodology**

Based on Figure 3.1, there are five procedures identified earlier in a form of diagrammatic flow. After the requirement analysis and definition phase, it will feed to the design phase as well as research and feedback analysis. The research and feedback analysis will be an on-going phase along the development of the project and works concurrently with other phases. This is done in order to revise and re-evaluate MobileDic system so that the functionalities and efficiency can always be improved from time to time. By having concurrent activities development as well, the time frame on developing MobileDic system can be reduced and any redundancy or flaws in any significant procedures can be avoided.

The integration and testing phase will be several times going back to the previous phase. This is because MobileDic will be divided into sub systems or modules, and each sub modules will be constructed individually. After the sub module is developed, it will be tested and integrated. This process will be on-going until MobileDic system is finally completed. System closing deals with the deployment of the final product. In this phase, MobileDic system is officially closed and any attempt for enhancement will be made in later version.

## **3.2 Research Methodology**

Research methodology involves with two phases in the project methodology described in Figure 3.1, which are requirement analysis and definition phase as well as research and feedback analysis

### **3.2.1 Requirement Analysis and Definition**

This phase concentrates more on analyzing and obtaining better understanding on the problems encountered. During this phase, studies will be conducted extensively to gather information on the requirement and defining MobileDic system scope. This includes technical and theoretical aspect of the proposed system. In this phase as well, research areas and hardware and software requirements have been identified, objectives of the projects and scopes of project are clearly stated through considering time and other constraints. Although there are large number of words as well as multiword available in Malay and English, an introductory for MobileDic system by including only the common words would be more appropriate.

In addition, functional and non-functional requirements are also identified in the early stage of the project life cycle. This is done in order to further clarify the requirements of the proposed system. Questionnaires are one of the tools that is used to gather those requirements. It will then be supported with observations and literature study that will be conducted throughout the research process. Functional requirements will be emphasized more towards the technical and functionalities of MobileDic system, whereby non-functional requirements will highlight the quality and performance desired by MobileDic system.

Functional requirements of MobileDic system include

- i. Able to translate single word and multiword (maximum of two words)
- ii. Able to translate from Malay to English and English to Malay
- iii. Able to be transferred to J2ME-enabled mobile phones
- iv. Able to be installed and run effectively in J2ME-enabled mobile phones

- v. Simple navigation menu.

Non-functional requirements of MobileDic system include

- i. Simple and user friendly interface for ease of use.
- ii. Fast response time to users during dictionary lookup.
- iii. System will not be suspended and able to give appropriate results if no match is found and continue operation to handle another request.
- iv. Provide easy navigation and appropriate help/assistant menu.

#### *3.2.1.1 Questionnaires*

Questionnaire is used to gather requirements from the users regarding the project. There will be two stages of questionnaires that will be distributed, which are pre questionnaire and post questionnaire. Pre questionnaire is distributed before development of product occurs. The questionnaire is distributed to 30 respondents which comprised of UTP students and have experienced in using electronic or mobile dictionary before. The data that is through this questionnaire will be based on:-

- 1) Mobile phone usage
  - This section asks the users on the usage of mobile phones and their preferences on having dictionary in mobile phones
- 2) Current Electronic/Mobile Dictionary functionalities
  - This section gathers responds from the users with regard to the stress and ability of multiword translation in existing dictionary
  - It also gathers several responds which dealt with either functional / non-functional requirements.

From this, all of the gathered data will be served as specification for design and construction phase. It also will feed to the research and feedback analysis phase. Post questionnaire will be distributed after the product is developed in order to evaluate the effectiveness of MobileDic system and whether it meets the objectives.

### 3.2.1.2 *Observation*

As reinforcement, observations will be conducted throughout the research process. The need of observation is done to reinforce more and validate the requirements. Throughout this task, several activities have been made such as trying out and observing current electronic or mobile dictionaries functionality. For instance, observation on English-Italian mobile dictionary shows that there is no multiword has been included for translations. It only provides translation for single word. Another observation done is on Malay-English electronic dictionaries such as Dr.Bhanot's Malay-English Cyber Dictionary[16] shows that only single word can be translated in the dictionary.”

### 3.2.2 **Research and Feedback Analysis**

On-going research will be conducted concurrently with the development of MobileDic system. The idea of having concurrent activities along with the development phase is to cut down the development time as well as having the ability to revise the content of research. Feedback analysis also is carried out throughout the development phase to gain feedbacks on how to enhance MobileDic system, based on the integrated sub-systems.

The research and feedback analysis include findings from internet, published journals, questionnaires, magazines and books. Most of the published journals are downloaded from digital library websites such as ACM Digital Library and IEEE Electronic library. Several articles which are relevant to MobileDic system are also being identified and are referred.

During this phase, feedbacks from several users are collected and re-evaluated regularly to increase the efficiency of MobileDic system. For example, after MobileDic system is showed to a bunch of friends (students), they suggested that it should have a simpler navigation than the existing one. Corresponding to this, the sub-module is identified and is enhanced to have a simpler navigation and interface and integrated back.

Extensive research along the development phase also plays an important role. Ongoing questions asked in forums, reference to several Java Programming books make sure that MobileDic system can be running efficiently. For instance, upon designing MobileDic system, it turned out that depending solely on RMS (Record Management System) to do the searching method and storing record will not be efficient enough and would reduce the processing time. Some of the users in the forum suggested to use a hashtable instead of relying solely on RMS. A more in-depth research is done on the usage of hashtable after the recommendation.

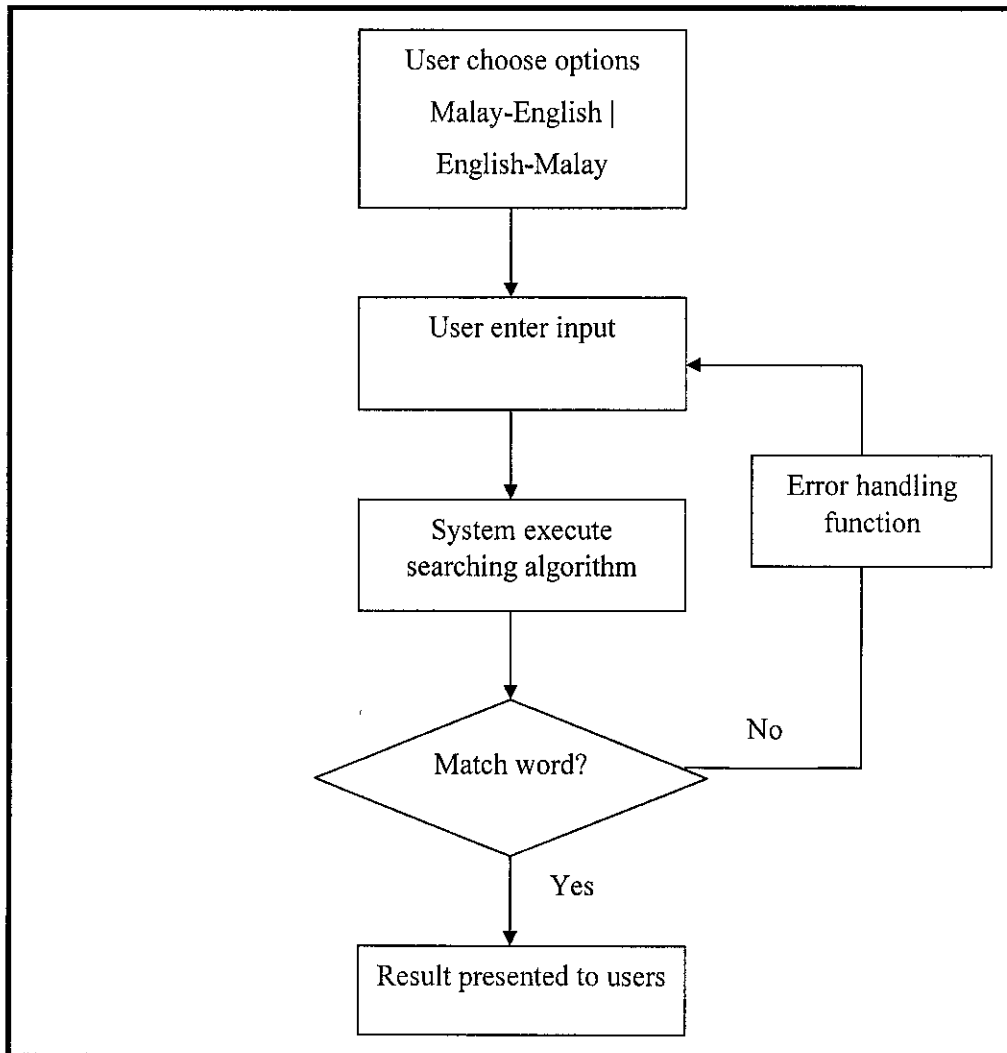
### **3.3 Development Methodology**

#### **3.3.1 System Design**

Proper design needs to be made before constructing the actual system. The design will facilitate the construction phase by providing the grid for the proposed system. For this phase, the focus is more towards designing system architecture and flow in order to identify and describe the fundamental software system abstractions and their relationships. MobileDic system is intended to find the matching word in another language of a given input word. Due to this, a proper searching algorithm and matching is also planned during this phase through considering the constraint of the development platform. A good database design also needed to be done in this phase. Poor database design may result in poor processing time and sometimes may introduce to redundancy and flaws. The following subsections on the next page will further expand each of these tasks in this phase.

##### *3.3.1.1 Designing System Workflow*

In designing MobileDic system workflow, few actors and corresponding activities involved in the system need to be identified. For this system, the actor involved is only the user of MobileDic. So, the resulting MobileDic system workflow will be as shown in Figure 3.2 as below:



**Figure 3.2 : MobileDic system workflow**

In MobileDic system workflow, the first task occurred would be the selection by users, whether they want the dictionary in Malay-English mode or English-Malay mode. After the selection, the user will be provided a text input box in which the users are required to key in the intended word to be translated. Given the input, MobileDic system will perform the searching algorithm and match with the correct output in the database. If there is no match found, MobileDic system will use error handling function to prompt the users indicating that the word has not been found. The users may have the option to search again for another word. If the match of a given input word is found in another language, the result will be presented to user, and the user may have the option to search again for another word.



### 3.3.1.2 *Designing Searching Algorithm*

Figure 3.3 shows the searching algorithm with main modules of MobileDic system. Before designing searching algorithm, an in depth review had been conducted on the tools that are available in order to select the most appropriate tools for the development of MobileDic system which can cater for multiword searching. Since the dictionary needed to be implemented in mobile phones, J2ME language is identified to be chosen for the development of MobileDic. By using J2ME, the usage of Hashtable function is recognized to be used for searching algorithm purpose. Through the usage of Hashtable, a searching algorithm as described in Figure 3.3 is proposed. All of the main modules identified for this project are developed using J2ME language. There will be no interaction with the server or external communication medium for the retrieval of the results. More details on other uses of hardware and software will be stated in Section 3.4 in this chapter. The following sections will describe the sub module which makes up the resulted overall searching algorithm for MobileDic system.

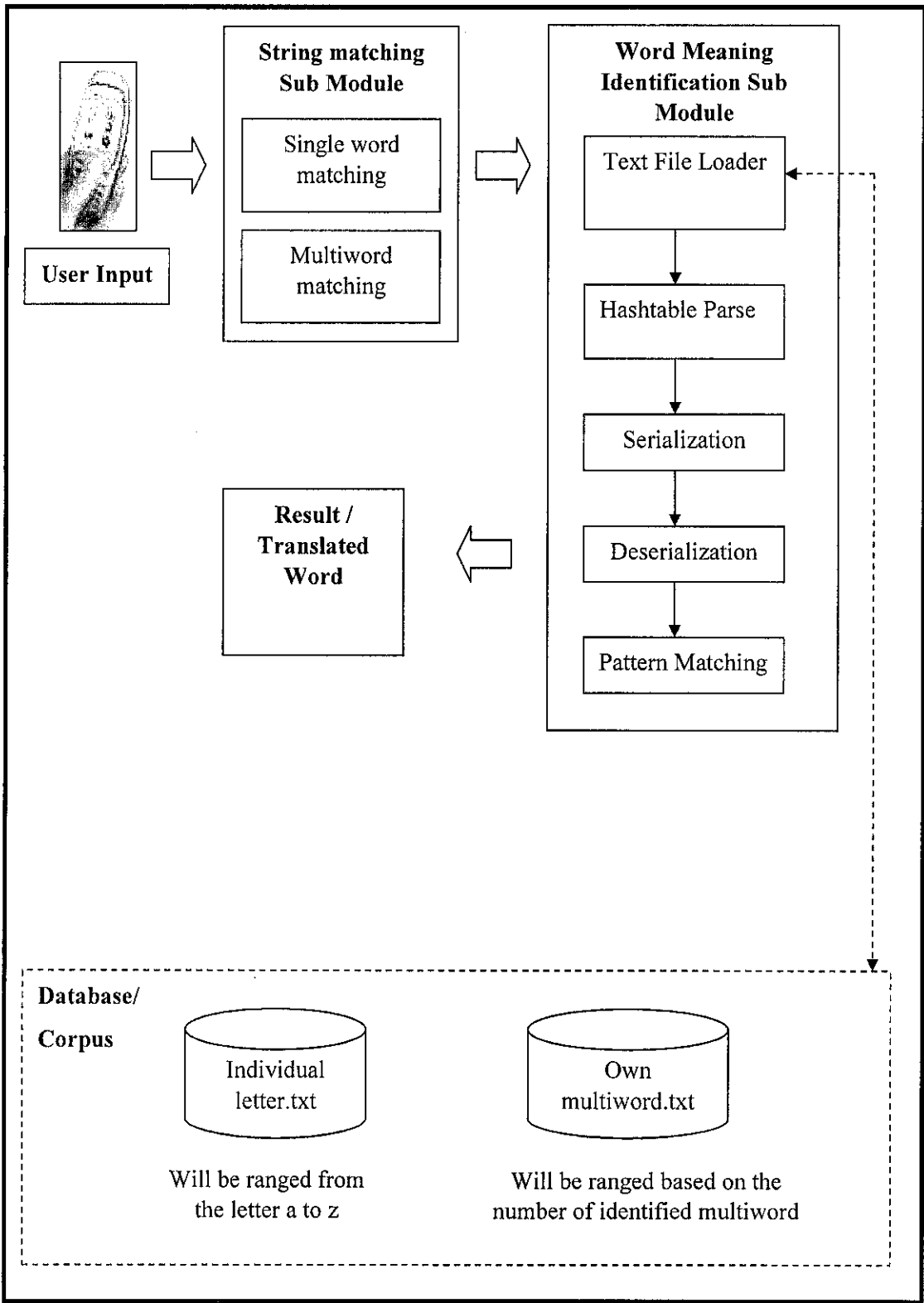


Figure 3.3 : MobileDic system Searching Algorithm

## **String Matching Sub Module**

String matching Sub Module will apply comparison of string from the given user input. If the given input word is a single word, MobileDic system will detect it as a single word and send to the next sub module to proceed, which is Word Meaning Identification Sub Module. If it is a multiword, MobileDic system will detect based on the availability of space. The following subsections will explain the process in further detail.

### ***i) Single word matching***

MobileDic system will detect whether the given input is a single word if there is no space presence in the corresponding text field. After the recognition, MobileDic will extract the first letter of the string to identify which text file / database to be loaded. If the given string starts with capital “a”, MobileDic will send the letter “a” to next sub module to proceed with the rest of the process.

### ***ii) Multiword matching***

MobileDic system will detect whether the given input is a multiword if there is availability of space in the given input string. After the recognition of space, MobileDic will extract the first substring inside the compound string. To do this, the built-in method substring [28] is used. It first detects the occurrences of space and then trims the input string based on space. As a result, the left side of string will be obtained. The left side word will act as a determiner to the next sub module to determine which text file to be loaded. For example, if the multiword is “kedai makan”, it will extract the word “kedai”, and the word “kedai” will serve as a key for the next sub module to determine which text file to be loaded.

## **Word Meaning Identification Sub Module**

Word Meaning Identification Sub Module will go through several processes, before the actual meaning of the input word can be presented to users. The components

under this sub module are including Text File Loader, Hashtable Parse, Serialization, Deserialization as well as Pattern Matching. The following sections discuss in further detail of each component under Word Meaning Identification Sub Module.

#### *i) Text File Loader to Hashtable*

This component will act as loader. In determining which text file to be loaded, it will take the input given from the previous sub module, which is String matching sub module. To load the file the available loading function in J2ME will be used, which is getResourceAsStream function [28]. In case of single word, MobileDic will identify which alphabet is submitted by previous sub module and will search for database or text file containing the alphabet letter. In case of multiword, MobileDic will identify the left substring and this substring will be used to check whether is there any text file with the corresponding word. If match is found, the file that contains the left substring word will be opened and loaded

#### *ii) Hashtable Parse*

After all data have been finally loaded to Hashtable, the Hashtable may play its role to do the parsing. Since Hashtable will play an important role in terms of pattern matching later, all the data that have been loaded will be save into hashtable. Parsing is the process of splitting up a continuous stream of characters (read from a file or keyboard input, for example) into meaningful tokens, and then building a parse tree from those tokens. The name is by analogy with the usage in grammar and linguistics [29]. To do this, again, the function substring is used. It will first search the whole record that has been loaded, and will first search for the first line of the record. For this record line, MobileDic system will perform substring function and split those word|meaning form based on the separator "|". MobileDic then assign word as the key and meaning as the value and this record is stored in Hashtable. The process will be continued until the end of the loaded text file data. At the end, the Hashtable, which stores in **Hashtable(key,value)** format will have the corresponding form

Hashtable(word1, meaning1)

Hashtable(word2, meaning2)

.....

Hashtable(wordn,meaningn)

### ***iii) Serialization***

Serialization component acts as a component to serialize the data inside Hashtable to be converted to byte array. This process is done in order to store the data inside Record store (Record Management System). J2ME MIDP defines a simple record-oriented database system called a Record Management System (RMS) [26]. Even though the data is already stored in Hashtable, adding record to Record Management System is needed to ensure faster processing later since the data is already loaded into the memory. The serialization process will first convert the data to byte and store them in first record of the store in Record Management System.

### ***iv) Deserialization***

Deserialization process will restore back all the data from byte array stored inside Record store of Record Management System (RMS). This is done in order to load the data back to Hashtable. During this process, it will read and get the first data that is stored in the first record of the store. After getting the record based on record id (1), the data will be loaded back to Hashtable for the Hashtable to proceed with pattern matching process.

### ***v) Pattern Matching***

This component will work to provide matching between the given keyword and the resulting value. The input word that is passed to Hashtable will be matched one by one until it reaches the correct and corresponding value. If the key found match with the input word, the value will be extracted by Hashtable and the result will be presented to users. If no match is found, MobileDic will alert the users that no translation is found for that word.

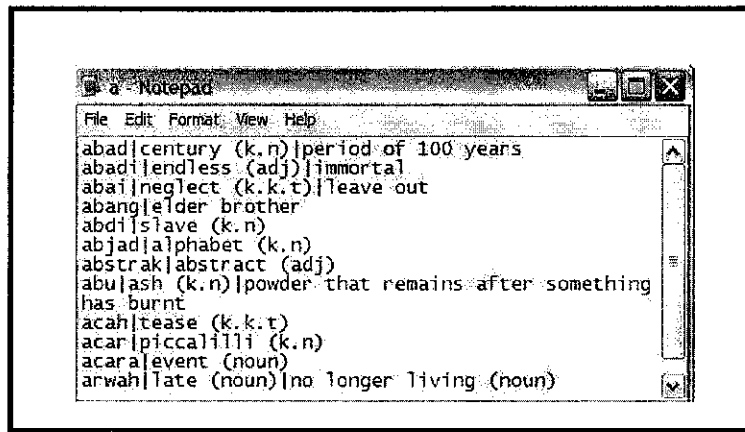
### 3.3.1.3 Designing Database Record

All the possible Malay and English single word as well as multiword (two words) as well as their translations are listed and organized in a document. This is done in order to ensure that only the selected words will be recorded in the application database. After this process is completed, the application database needs to be designed so that it can provide efficient and fast lookup during searching.

The database is designed to be organized based on alphabetical. Each word starting from 'a' to 'z' will be stored separately. This means, each word starting with alphabet 'a' will be stored in text pad file 'a.txt'. The rest of the letters will follow until 'z'. The method of dividing those letters in different text pad is done in order to provide fast response and searching time. Since this dictionary may contain around 5000 words, it will be efficient enough to separate it to a different text file. Each word is organized accordingly as illustrated in Figure 3.4. The separator that is used to separate between Malay and English word is '|'. Thesaurus of the meaning will also be listed if any of the thesaurus is found in referenced dictionary. Thus, the text file will contain a list of data in a form word|meaning|thesaurus (if any) pair and the resulting of the data in text file will have the following format :

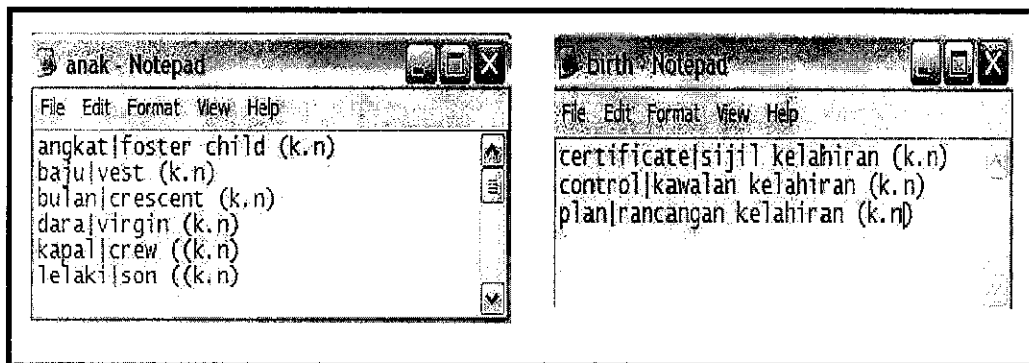
```
word1|meaning1|thesaurus (if any)
word2|meaning2
.....
wordn|meaningn
```

The intention is that, when the system browses through the file, it will find the corresponding separator, extract the word and assigning them to each variable.



**Figure 3.4 : Database containing only the letter A.**

In handling multiword expressions translation, those words will be listed in a separate text file. For example, the word ‘anak’ in Malay may have other next possible word after it, in which, when combined, will result in different meaning in English language. The word ‘anak’ may have ‘anak patung’, ‘anak tiri’, ‘anak angkat’, and when all of these words are converted to English, it needs to have a proper and correct translation such as anak patung > dolls, anak tiri > step daughter, anak angkat > foster child. Corresponding to this, all of these multiword will be stored in another separate text file. Since the word ‘anak’ has many other possibilities of the next word, it will have its single text file. The example is illustrated in Figure 3.5. The word ‘birth’ in English also illustrated



**Figure 3.5 : The separation of multiword expression in different text file.**

All the words to be stored inside the text pad is based on lexicon of Kamus Dwibahasa Oxford Fajar[31]. In order to cater for multiword, another lexicon is used,

which is Easy Learning Bilingual Dictionary English-Malay and Malay-English [32]. This dictionary will further detail on the list of two words or multiword.

#### *3.3.1.4 Designing User Interface*

Designing user interface also is a part of software design phase. In designing user interface, a number of elements needed to be considered since the application is to be developed in mobile phones. Some considerations that need to be highlighted here are, processing power of mobile phones, small keypad and small display screen. These constraints demands a very good interface for dictionary, especially mobile phones only provided small keypad to press, and thus resulting to difficulty of multiple keystrokes.

Based on the requirements collected on the earlier phase, the users preferred the interface for MobileDic system to be simple and have easy navigation. Responding to this, the interface designed for MobileDic system is simple and very straightforward to navigate. MobileDic system is kept at a beginner level, as it is difficult to navigate a complex interface especially when the system is implemented in mobile phones.

Since the dictionary function will only be deal with a single interface, additional interfaces are added in order to make the dictionary looks more interesting. The interface for Splash screen at the initialization phase of MobileDic is added. Help and About menu is also added as one of the features in the dictionary.

On the whole, MobileDic system will have the following interfaces :

- Splash screen
- Main Menu screen
- Malay-English Search Word Screen
- Malay-English Result Screen
- English-Malay Search Word Screen
- English-Malay Result Screen



- About Screen
- Help Screen

### **3.3.2 System Construction**

During this phase, the actual construction of MobileDic system will begin. The implementation of database/text file, construction of interface based on the designed interface and the actual coding of the system will start. The implementation of database will start by adding around 100 words first into the text file for use of testing while implementation of coding. Since all interfaces in J2ME could not just be dragged like any other programming languages, all the components of the interface needed to be coded. For example, in order to create a text field or a list box, they need to be appended to the form through some commands. Each Sub Module will be developed one by one and as one is complete tested, it will be integrated and this process will be ongoing until full system has finally been integrated. For example, the String Matching Sub Module will be coded first because it is one of the primary tasks in MobileDic system. After it is done and successfully tested, the phase continues with another sub module, which is Word Meaning Identification Sub Module, in which this later will be integrated with String Matching Sub Module. Since this task requires a bit tedious job, it tends to consume most of the time along development phase. In case of any problems or difficulties faced during actual coding process, research and asking feedback from other users in the forum on some websites help much.

### **3.3.3 Testing and Integration**

This phase will work closely with the previous phase in the project methodology, which is system construction. Sub Module will be developed one by one during construction phase and every time the sub module is completed, it will undergo the testing phase first. Any errors during this phase will be resolved. After it is successfully completed, it will be integrated and this process will be ongoing until all the sub modules developed have been integrated. After the full integration, MobileDic system is tested again as a whole to ensure it works correctly as expected. In addition, a group of 10 users will be selected to evaluate the output of MobileDic

and test whether it is able to provide functionality as being specified in objectives of the project. If MbileDic is unable to perform as expected or producing error while testing, it will need to be debugged and refined.

### **3.3.4 System Closing**

During this phase, the project is finally completed. This phase will be dealt with preparing and delivering final project report as well as presenting MobileDic system. Any enhancement to be made to MobileDic will be made in later version.

## **3.4 Tools Required**

### **3.4.1 Hardware**

- J2ME enabled mobile phones, with minimum phone or memory card capacity : 200KB

### **3.4.2 Software**

- JDK (Java Development Kit – J2SE SDK) Version 1.4.2
- MIDP (Mobile Information Device Profile) Version 1.0.3
- J2ME Wireless Toolkit Version 2.2
- Forte for Java CE Version 4.0

## CHAPTER 4

### RESULT AND DISCUSSION

#### 4.1 Overview

MobileDic system prototype performs as expected. There will be four parts in this chapter, which is the result of functional testing; secondly the result of evaluation of MobileDic system by several users after the presentation of the prototype. Finally, the screen captures of the MobileDic system will be included as the result. Each of the results presented will be discussed and justified.

#### 4.2 Functional Testing

As being highlighted earlier, MobileDic system will need to go through functional testing in order to detect if there are any bugs upon each function provided by the system. Upon doing this, MobileDic system needs to be tested in development platform, which is Personal computer since the testing results are intended to be showed in command prompt window. The module or components being tested using emulator in development platform are described as the following.

**Table 4.1 : Test Results of Each Module**

<b>Module/Component</b>	<b>Expected Test Result</b>	<b>Actual Test Result</b>
<b><u>Start Button (Splash screen)</u></b>	To ensure that the screen will be directed to Main Menu screen without waiting Splash screen to finish initializing	- Screen successfully redirected from Splash screen to Main Menu screen.
<b><u>Select button (Main</u></b>	To ensure that the screen will	- Screen successfully

<b><u>Menu screen)</u></b>	be directed to each screen depending on users' selection from the list in Main Menu, either Malay-English, English-Malay, Help and About.	directed to <ul style="list-style-type: none"> <li>▪ Enter Malay Word screen for Malay-English selection</li> <li>▪ Enter English Word screen for English-Malay selection</li> <li>▪ Help screen for Help selection</li> <li>▪ About screen for About selection</li> </ul>
<b><u>Go button (Enter Malay Word screen)</u></b>  <b>String Matching Module</b>  - Single word matching - Multiword matching	<ul style="list-style-type: none"> <li>- Detect input word as a single word</li> <li>- To show testing results on command prompt window</li> <li>- Detect input word as multiword based on spaces</li> <li>- To show testing results on command prompt window</li> </ul>	<ul style="list-style-type: none"> <li>- Successfully detect single word</li> <li>- Successfully show the test results on command prompt window.</li> <li>- Successfully detect multiword</li> <li>- Successfully show the test results on command prompt window.</li> </ul>
<b>Word Meaning Identification Module</b>  - Text File Loader	<ul style="list-style-type: none"> <li>- Load text file in Malay-English folder based on each initial letter of the word in case of single word</li> <li>- Load text file in Malay-</li> </ul>	<ul style="list-style-type: none"> <li>- Successfully load text file based on the initial letter of the word</li> <li>- Successfully load text file based on the left substring of the</li> </ul>

	<p>English folder based on the left substring of the multiword.in case of multiword</p> <ul style="list-style-type: none"> <li>- To show testing results on command prompt window.</li> </ul>	<p>multiword</p> <ul style="list-style-type: none"> <li>- Successfully show the test results on command prompt window</li> </ul>
- Hashtable Parse	<ul style="list-style-type: none"> <li>- To parse or split up the word meaning form based on the separator ' '</li> <li>- To store the split words and meanings into Hashtable, mapped by (key,value)</li> <li>- To show testing results on command prompt window</li> </ul>	<ul style="list-style-type: none"> <li>- Successfully parsed the word meaning from loaded text file based on separator ' '</li> <li>- Successfully stored the split words and meanings into Hashtable mapped by (key,value)</li> <li>- Successfully show the test results on command prompt window</li> </ul>
- Serialization	<ul style="list-style-type: none"> <li>- To convert data in Hashtable to byte array</li> <li>- To store data to first record of record store (Record Management system).</li> <li>- To show testing results on command prompt window</li> </ul>	<ul style="list-style-type: none"> <li>- Successfully convert to byte array</li> <li>- Successfully store data to first record in record store</li> <li>- Successfully show the test results on command prompt window</li> </ul>
- Deserialization	<ul style="list-style-type: none"> <li>- To restore back the Hashtable from byte array by getting the first record of record store stored in Record Management</li> </ul>	<ul style="list-style-type: none"> <li>- Successfully restore back the Hashtable from byte array</li> <li>- Successfully getting the first record from record</li> </ul>

	<p>System</p> <ul style="list-style-type: none"> <li>- To show testing results on command prompt window</li> </ul>	<p>store</p> <ul style="list-style-type: none"> <li>- Successfully show the test results on command prompt window</li> </ul>
- Pattern Matching	<ul style="list-style-type: none"> <li>- To find the keyword in Hashtable based on given input word</li> <li>- To match the found keyword with the value stored in Hashtable</li> <li>- To produce and present the correct result for single and multiword input to Malay-English Result screen.</li> <li>- To ensure that message is prompted if the system does not find the match record</li> <li>- To show testing results on command prompt window</li> </ul>	<ul style="list-style-type: none"> <li>- Successfully find keyword based on given input word</li> <li>- Successfully match the keyword with the corresponding value stored in Hashtable</li> <li>- Successfully produce and present the correct result of single and multiword input in Malay-English Result screen.</li> <li>- Successfully prompting message if no word is found.</li> <li>- Successfully show the test results on command prompt window</li> </ul>
<b>Back Button (Enter Malay Word screen)</b>	<ul style="list-style-type: none"> <li>- To redirect the screen to Main Menu Screen.</li> </ul>	<ul style="list-style-type: none"> <li>- Successfully redirecting to screen Main Menu</li> </ul>
<b>OK Button (Malay-English Result screen)</b>	<ul style="list-style-type: none"> <li>- To redirect the screen to Enter Malay Word Screen again.</li> </ul>	<ul style="list-style-type: none"> <li>- Successfully redirecting to screen Enter Malay Word.</li> </ul>
<b>Menu Button (Malay-English Result screen)</b>	<ul style="list-style-type: none"> <li>- To redirect the screen to Main Menu Screen.</li> </ul>	<ul style="list-style-type: none"> <li>- Successfully redirecting to screen Main Menu</li> </ul>
<b>Go Button (English-Malay Result screen)</b>	<ul style="list-style-type: none"> <li>- All the steps are the same as Malay-English, except that the folders referred for</li> </ul>	<ul style="list-style-type: none"> <li>-</li> </ul>

	database would be English-Malay folder and the result screen would be English-Malay Result screen	
<b>Back Button (Enter English Word screen)</b>	- To redirect the screen to Main Menu Screen.	- Successfully redirecting to screen Main Menu
<b>OK Button (English-Malay Result screen)</b>	- To redirect the screen to Enter English Word Screen again.	- Successfully redirecting to screen Enter English Word.
<b>Menu Button (English-Malay Result screen)</b>	- To redirect the screen to Main Menu Screen.	- Successfully redirecting to screen Main Menu
<b>Done button (Help screen)</b>	- To ensure the screen will be redirected to Main Menu back	- Successfully redirecting to screen Main Menu
<b>Done button (About screen)</b>	- To ensure the screen will be redirected to Main Menu back	- Successfully redirecting to screen Main Menu
<b>Exit Button (Main Menu screen)</b>	- To ensure the system is closed	- The system is successfully close

Each of the modules listed above are functioning as expected. Overall, during the testing, MobileDic system was performing well without much redesign has been made.

### 4.3 Integration Testing

Functional testing is concentrating on testing based on functionalities of each module. Integration testing is mainly serves for testing MobileDic system after each sub module has been integrated. Each sub module is integrated one by one at a time and is tested accordingly to make sure each of it is well-functioning. The test result obtained for integration of sub modules testing is as below :

**Table 4.2 : Test Result of Integration Testing**

<b>Module / Component</b>	<b>Expected Test Result</b>	<b>Actual Test Result</b>
<p><b><u>Linkage</u></b> Sub Module Integration</p>	<ul style="list-style-type: none"> <li>- To ensure that each sub module is successfully working after each of it been integrated in order to make up the whole system</li> </ul>	<ul style="list-style-type: none"> <li>- Each sub module is successfully integrated to the integrated system and working correctly as expected.</li> </ul>
<p><b><u>Testing of Whole System</u></b></p> <ul style="list-style-type: none"> <li>- <b>Development Platform (emulator)</b></li> <li>- <b>Mobile Phone Platform</b></li> </ul>	<ul style="list-style-type: none"> <li>- To ensure system as a whole is successfully functioning in development platform (emulator)</li> <li>- To ensure the whole complete system is successfully packaged using J2ME Wireless Toolkit</li> <li>- To ensure the system as a whole is correctly functioning if transferred to Mobile Phone platform (Nokia 6680) via Bluetooth</li> </ul>	<ul style="list-style-type: none"> <li>- System is successfully functioning as a whole inside development platform (emulator)</li> <li>- System is successfully packaged as .JAD file</li> <li>- System is correctly functioning inside mobile phones(Nokia 6680)</li> </ul>

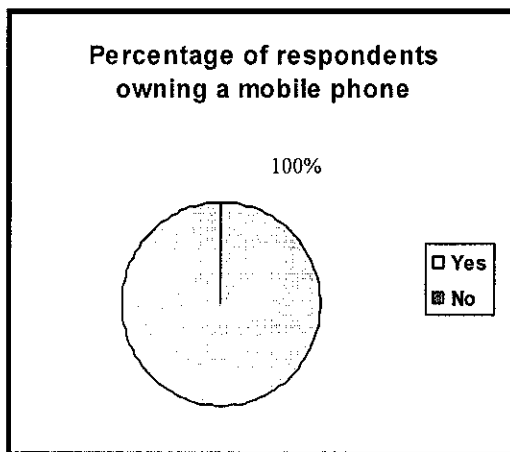


## 4.4 Evaluation

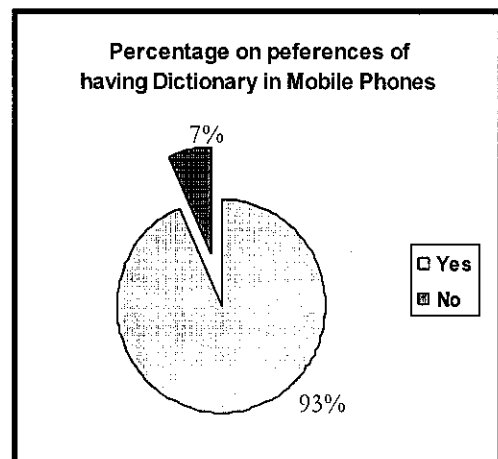
### 4.4.1 Pre Analysis Users' Evaluation

Pre analysis has been done at the early stage of project which is Requirement Analysis and Definition phase in order to understand and gain requirements from several respondents. The analysis is done based on two categories, which are mobile phone usage as well as emphasis of existing dictionaries on multiword translation. This includes the users respond and evaluation of current dictionary functionality. The questionnaire was distributed to 30 UTP students. The analysis is being discussed as the following :-

#### 4.4.1.1 Mobile Phone Usage



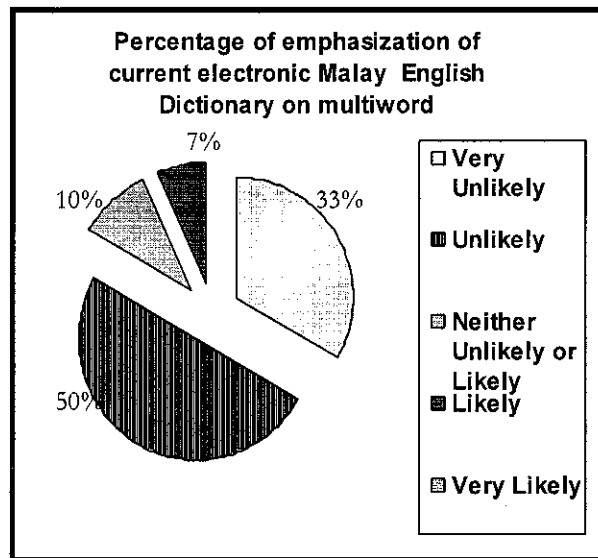
**Figure 4.1 : Chart on Percentage of Respondents owning a mobile**



**Figure 4.2 : Chart on Percentage of Preferences of having Dictionary in Mobile Phone**

As illustrated in Figure 4.1, 100% of the respondents owned a mobile phone. This shows that it is worth implementing the dictionary inside mobile phone since most of them owned it. Having extra dictionary feature would ease them since it is easily carried and is one of their necessities. Figure 4.2 shows the percentage of respondents' preferences of having dictionary inside their mobile phones. Most of them agreed on having this extra feature, which is shown by percentage result of 93%, whereby 7% of the respondents do not agree in having this feature.

#### 4.4.1.2 Emphasis of Existing Dictionaries on Multiword (two-word) translation



**Figure 4.3 : Chart on Percentage of emphasis of existing dictionaries on multiword (two-word) translation**

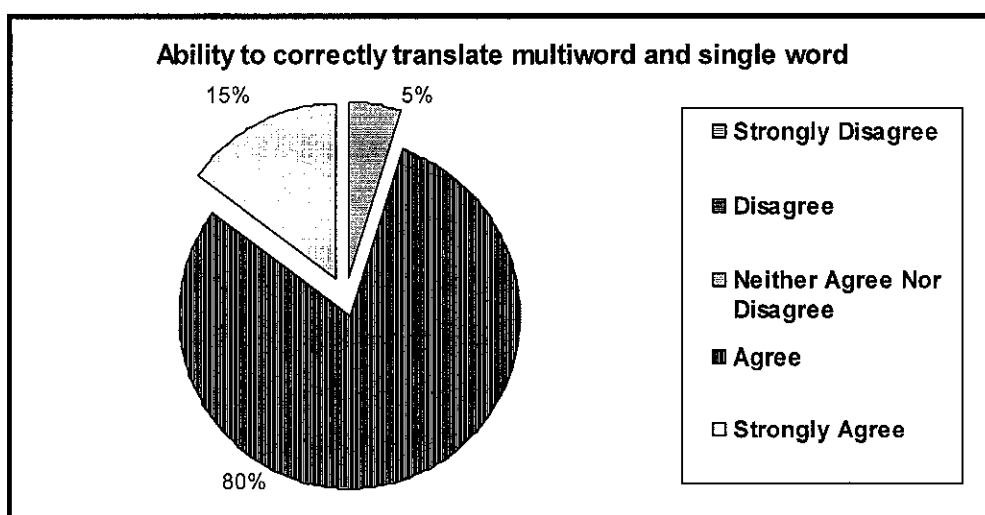
Figure 4.3 illustrates the percentage of emphasis of existing dictionaries on multiword (two-word) translation. As illustrated, 33% of the respondents responded that current dictionaries are very unlikely to focus on multiword translation. Followed by this is 50% of respondents saying that those dictionaries unlikely to focus on multiword translation. 10% of the respondents saying that those dictionaries are neither unlikely or likely to emphasize on multiword translation. Only 7% said that those dictionaries are likely to focus on multiword translation.

#### 4.4.2 Post Analysis Users' evaluation

Post Analysis Users' Evaluation is done after the development of the project is completed. In this evaluation, users are asked to evaluate the functionality of the developed system and whether it meets the objectives. Users' evaluation that is used here is type of intrinsic evaluation, in which it is a subjective evaluation. The criteria in judging MobileDic system is based on whether it is able to translate single word and multiword correctly as well as the usefulness of MobileDic system in terms of handling multiword translation.

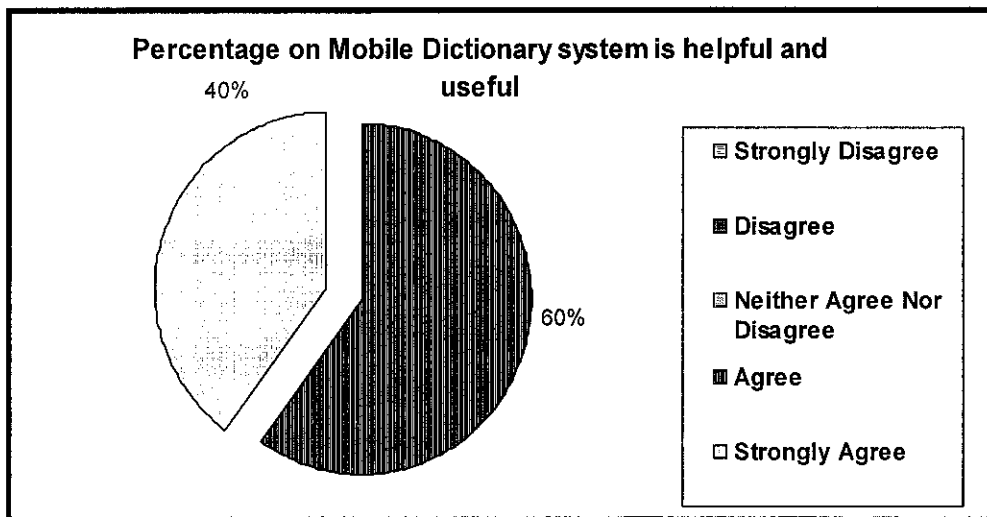
In doing this, a number of 20 users, which comprised of UTP students who have experienced in using Electronic / Mobile Dictionary, are chosen to evaluate on MobileDic system. The users are given a set of questionnaire and the application is displayed to the users so that they are able to do hands-on evaluation to MobileDic system. Each user who performs the evaluation is mainly from students, which consists of local students as well as foreign students. The users need to grade MobileDic system by responding to the questionnaires given.

Based on the evaluation and questionnaire that has been done, referring to Figure 4.4, it is found out that 15 % of the users strongly agree that MobileDic system is able to translate single word as well as multiword correctly. 80% from the users agree with the statement that MobileDic system is functioning well in translating multiword and single word. Meanwhile, only 5% of the users neither disagree nor agree with the statement. The users might be having doubt whether they are the correct translations or not. This can conclude that MobileDic system is able to translate multiword as well as single word accurately.



**Figure 4.4 : Chart for the MobileDic ability to correctly translate multiword and single word**

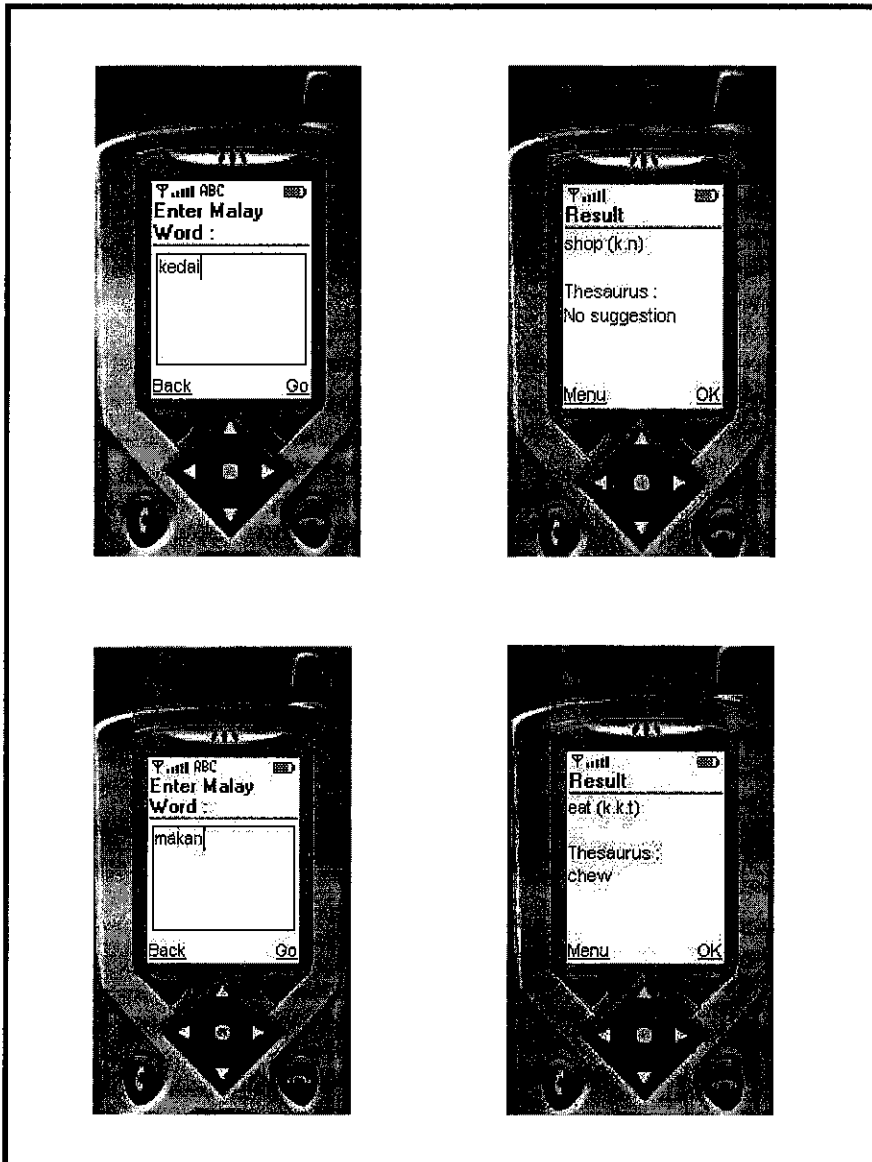
The users also asked to judge whether MobileDic system is a helpful and useful system. As shown in Figure 4.5, 60% of the users strongly agree that MobileDic is a helpful and useful system in terms of multiword translation. Meanwhile 40 % agree with the statement. Simply stated, majority of the users agree that MobileDic system is a helpful and useful system. This shows that the developed searching algorithm in MobileDic system can improve current situation



**Figure 4.5 : Chart showing the percentage whether MobileDic system is useful**

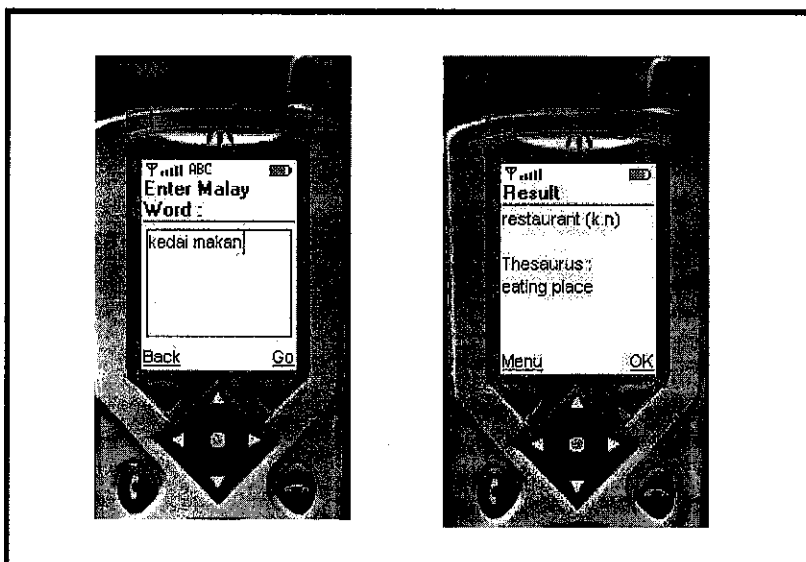
After reviewing the users' evaluations, 80% of the respondents strongly agreed on MobileDic accuracy to translate multiword as well as single word and almost all respondents agreed that MobileDic is a helpful and useful system.

In general, users perceive MobileDic to have an ability to translate multiword as well as single word correctly through proposed searching algorithm. The users also feel that students as well as tourists would benefit besides having a convenience system by having an extra feature in their mobile phones. This can conclude that the proposed searching algorithm in MobileDic is able to improve the current situation. The result can be supported by the following illustration. If the users use existing dictionary which have not supported for multiword translation, they will tend to translate the word one by one and directly as shown in Figure 4.6 by taking example the word to be translated is 'kedai makan'. The resulting answer will be incorrect and cause the users to interpret the result as 'shop eat'.



**Figure 4.6 : Direct translation of word 'kedai makan' through querying each single word**

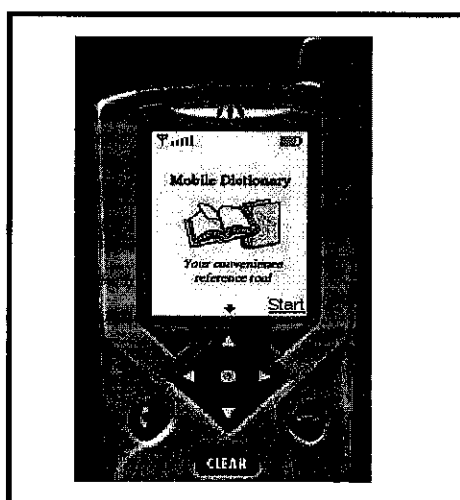
However, through the proposed solution, the users can just enter the multiword directly in the input box and the correct result will appear as shown in Figure 4.7.



**Figure 4.7 : Multiword translation through querying the multiword directly.**

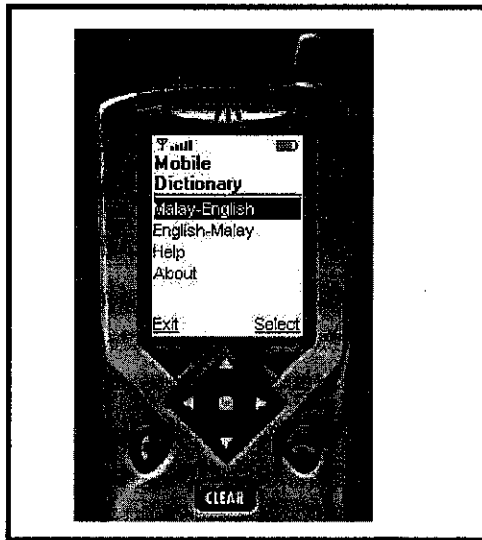
#### 4.5 Screen Capture and Description

The following section will capture screens involved and describe what would be each interface does in MobileDic system.



**Figure 4.8 : Mobile Dictionary Splash Screen**

Figure 4.8 shows MobileDic splash screen. Splash screen here refers to the first screen that is usually being displayed when user launch a mobile application. A good splash screen is necessary to attract user to use the application.



**Figure 4.9 : Mobile Dictionary Main Menu Screen**

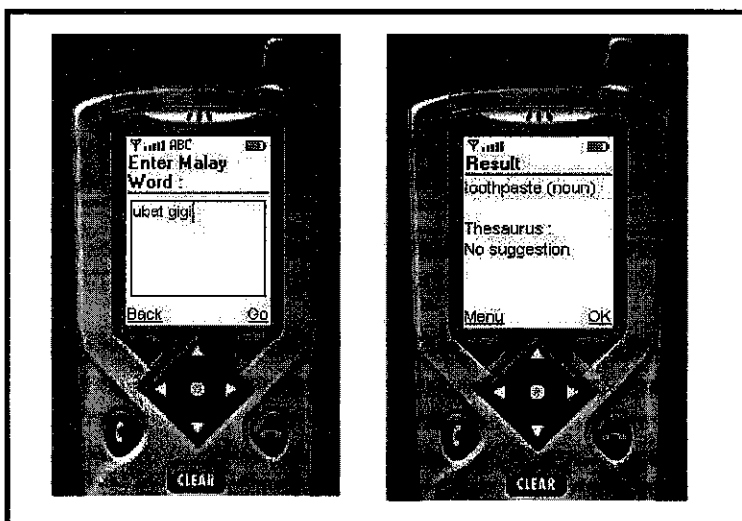
Figure 4.9 shows the main menu screen of the MobileDic. This is the main access to the application. In main menu, there are several links to:

- Malay - English
- English - Malay
- Help
- About

Malay-English option links user to find translation of Malay word to English word.

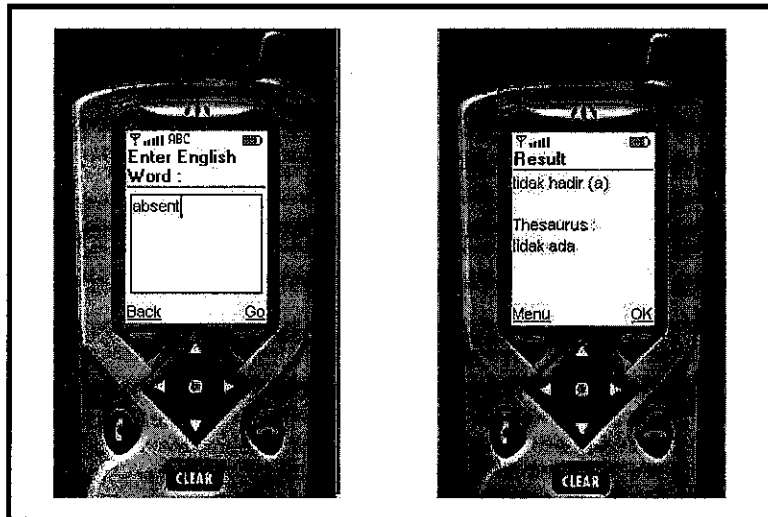
English-Malay option links user to find translation of English word to Malay word.

Meanwhile, Help and About options link user to help and about screen.



**Figure 4.10 : Mobile Dictionary Enter Malay Word and Result Screen**

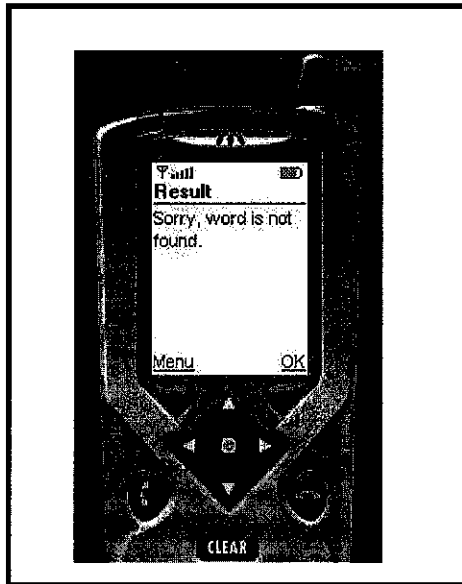
Figure 4.10 shows the screen for Enter Malay Word and its result. Enter Malay Word will appear if the user select the Malay-English option in Main Menu. In this screen, the user will enable to search for particular word to be translated in English. The corresponding output will appear in Result screen. The result will consist of the translated word as well as the thesaurus if any of thesauruses is present for that word. If there is no thesaurus defined, the phrase no suggestion will appear.



**Figure 4.11 : Mobile Dictionary Enter English Word and Result Screen**

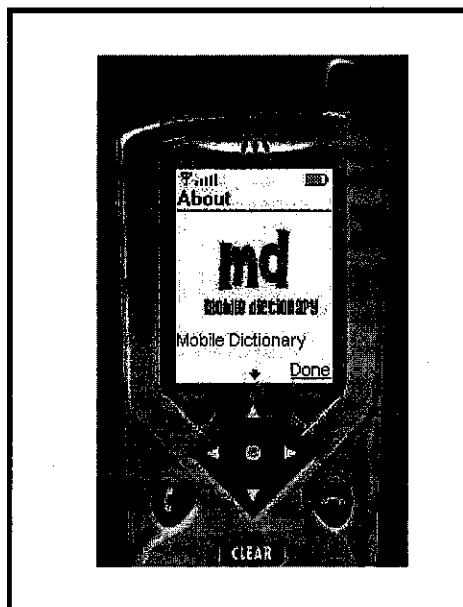
Figure 4.11 shows the screen for Enter English Word and its result. Enter English Word will appear if the user select the English-Malay option in Main Menu. In this screen, the user will enable to search for particular word to be translated in Malay. The corresponding output will appear in Result screen. The result will consist of the translated word as well as the thesaurus if any of thesauruses is present for that word. If there is no thesaurus defined, the phrase no suggestion will appear





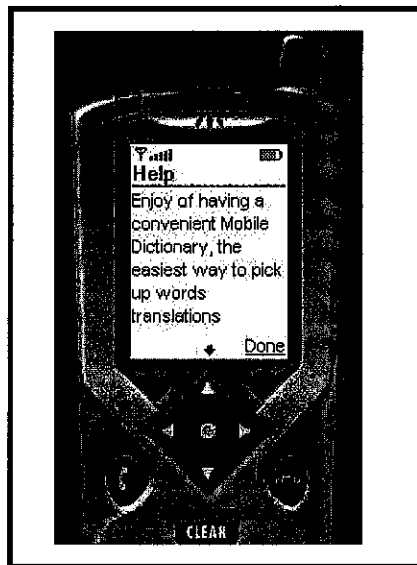
**Figure 4.12 : MobileDic Word Not Found Screen**

In case of the word is not included in the database, the above screen shown in Figure 4.12 will appear. This screen is based on system respond for error handling. The user may have the option to go to main menu or find another word by clicking OK button.



**Figure 4.13 : MobileDic About Screen**

In Figure 4.13, About screen is shown. About screen contains details on what actually the application does as well as the author information.



**Figure 4.14 : MobileDic Help Screen**

Figure 4.14 illustrates Help screen. In case the users find the difficulties in navigating MobileDic, they have the option to go to Help screen to read some abbreviations listed for parts of speech

## **CHAPTER 5**

### **CONCLUSION AND RECOMMENDATION**

#### **5.1 Conclusion**

As a conclusion, MobileDic is significant as most of the Malay-English or English-Malay dictionary or translators ignore the possibility of the multiword. Most of them are just simply providing single index word translation as well as doing direct translation for each input word, which will result in incorrect output.

The proposed searching algorithm in MobileDic serves as the heart of the system. The architecture of searching algorithm in MobileDic has been carefully designed, as it will directly affect the system's output. Overall, the users' evaluation result show that MobileDic is regarded to have produced fairly good result for multiword as well as single word with average users of 80%. Thus, the proposed searching algorithm is thereby fulfills the project objectives.

The idea of producing dictionary in mobile phones is not just simply striving to deliver product by coping with the latest technology. The idea is initiated with the goals of easing the targeted users, mostly students and tourists. Since conventional dictionaries are too bulky and difficult to be carried around, mobile phones is the best option. Mobile phones are considered as basic necessities by most groups of people. Having an extra dictionary function would add an extra advantage to people owning a mobile phone.

## **5.2 Recommendation**

There are a few suggestions and recommendations that can be done to this project and the system, so that it can be expanded in the future to produce more reliable and practical.

### **5.2.1 Expansion of multiword**

Since this project only focus on the possibility of multiword exist in Malay and English word not more than 2 words, there might be a need to cover the possibility of three words exist or multiword that are fall under different categories than noun in Malay and English word . For example, the word “tidak boleh dibahagi” will be translated to the word “indivisible”. As technology keeps on changing, and the requirement of mobile technology keeps on expanding, this feature can be realized in future.

### **5.2.2 Suggested most similar meaning if no match is found**

Since some words may have the possibility to not exist in the database or corpus, suggesting users the most similar words that are related with the meaning of the word will help users. By having this feature, users might have brief idea what does the word mean in particular language they want.

### **5.2.3 Translator in Mobile Phone**

There are various translator exist, either in stand alone platform or in web-based. A development of translator, which will enable users to translate sentences rather than just dictionary, would be very beneficial since there is still no translator being developed in mobile platform. Again, since mobile technology keeps on expanding, for example, size of memory of mobile phone will be increased, development of translator can be realized.

## REFERENCES

1. The history of Dictionary. Retrieved August 4, 2005 from the World Wide Web.
2. Ozawa, H. Isoda, M. & Aiso, H. *DIS: A User Interface System design for the dictionary's database.*. IEEE journal
3. Wikipedia Dictionary. Retrieved August 5, 2005 from the World Wide Web: <http://en.wikipedia.org/wiki/Dictionary>
4. Linguistic Department. Retrieved August 4, 2005 from the World Wide Web: [www.essex.ac.uk/linguistics/clmt/MTbook/HTML/node98.html](http://www.essex.ac.uk/linguistics/clmt/MTbook/HTML/node98.html)
5. Meirs, V, 1999, *Are Wired Schools Failing Our Kids?* , Pc World (U.S.A. Edition), *Pc World Sep. – Dec. Journal*
6. Nesi, H, 1998, *Dictionaries on Computer : How Different Markets Have Created Different Products*, Conference 1998 papers
7. Sachs, G, Unwired Express-Products. Retrieved August 7, 2005 from the World Wide Web: [http://www.unwiredexpress.com/products/products\\_whyhmobile.htm](http://www.unwiredexpress.com/products/products_whyhmobile.htm)
8. Azlyn Abd Rahman. Ericsson survey shows Malaysians are ready for more new mobile services. Retrieved August 5, 2005 from the World Wide Web: <http://www.ericsson.com.my/press/050203.htm>
9. Al-Baker, A, 2004 *Mobile E-mail: Moving Away From Desktop*, Mobile World (Issue 17)
10. Science in Africa. *Mobidic the Mobile Dictionary – just SMS*. Retrieved August 5, 2004 from the World Wide Web:
11. Affendi. *Affendi Malay-English Translator*. Retrieved August 5, 2005 from the World Wide Web: <http://www15.brinkster.com/affendi/m2etranslator/default.asp>
12. Niemann. Thomas, 1999. *Visual Basic Collection and Hash Table*, EPaperPress Journal
13. Ogura, K., Bond, F. & Ooyama, Y, 2002 *ALT J/M. A Prototype Japanese-to-Malay Translation System*, CiteSeer Journal

14. Bond, F., Sulong, Ruhaida., Yamazaki, T. & Ogura, K, 2002 *Design and Construction of a machine-tractable Japanese-Malay Dictionary*, CiteSeer Journal
15. Oxford University Press. *Malay Dictionaries*. Retrieved August 14, 2005 from the World Wide Web:  
<http://www.oup.com.au/content/printer.asp?ContentID=1239>
16. Dr.Bhanot. *Bhanot's Malay-English Cyber Dictionary*. Retrieved August 16, 2005 from the World Wide Web: <http://dictionary.bhanot.net/>
17. Little Malay Dictionary Search. Retrieved August 16, 2005 from the World Wide Web: <http://www21.brinkster.com/database2001>
18. AKN Messaging Technologies Berhad. *eBuzz Active Dictionary*. Retrieved August 14, 2005 from the World Wide Web: <http://www.gsm.com.my>
19. Electronic Dict Lingo Global 29. Retrieved August 15, 2005 from the World Wide Web:  
[www.worldlanguage.com/Products/Japanese/HandheldDictionary/Page1.htm](http://www.worldlanguage.com/Products/Japanese/HandheldDictionary/Page1.htm)
20. Kin Quah, C. Bond, F. Yamazaki, T., 2003 *Design and Construction of a Machine-Tractable Malay-English Lexicon*, CiteSeer Journal
21. Zuraidah Omar, A Pocket Malay Dictionary. Retrieved September 11, 2005 from the World Wide Web  
<http://www.schoenhofs.com/app/detail?0794600573>
22. Unreal Mind Shabox, Malay-English Dictionary. Retrieved September 11, 2005 from the World Wide Web:  
[http://www.shabox.com/sb\\_hp\\_download\\_rich\\_main.asp?param1=j2me\\_files/MA/2000\\_MA\\_103x80.gif&type=19&cat=399&code=5002000](http://www.shabox.com/sb_hp_download_rich_main.asp?param1=j2me_files/MA/2000_MA_103x80.gif&type=19&cat=399&code=5002000)
23. Goh Peng Joo. A basic course on Bahasa Malaysia (also known as Bahasa Melayu or the Malay language. Retrieved September 1, 2005 from the World Wide Web  
[http://pgoh.free.fr/malay\\_course.php](http://pgoh.free.fr/malay_course.php)
24. Yuen, 2004, *An Ergonomic Study of Mobile Phone Usage in Malaysia*.
25. Yahoo!, Inc., 2005. *Mobile Phone Survey-Malaysia*, Yahoo Online Mobile Survey
26. J.Wells, Martin, 2004, J2ME, *Game Programming*, Thomson Course Technology.

27. Tim Italia S.P.A, *TIM Mobile Dictionary*. Retrieved September 11,2005 from the World Wide Web:  
[http://www.tim.com.gr/en/02\\_01\\_01.cfm?PageID=483](http://www.tim.com.gr/en/02_01_01.cfm?PageID=483)
28. Deitel. H.M, Deitel. P.J (2005), *Sixth Edition Java How to Program*, Prentice Hall/ Deitel Publications
29. Wikipedia.Org, Retrieved August 14, 2005 from the World Wide Web:  
<http://en.wikipedia.org/wiki/Parse>
30. Friedman, S. Krishnan, A. Leidenfrost, N. Brodie, B.C. Cytron, R. Niehaus, D, 2003, *Hash Tables for Embedded and Real-Time Systems*, Washington University, USA
31. M.Hawkins, J, 1991, *Kamus Dwibahasa Oxford Fajar Inggris-Melayu, Melayu-Inggris*. Penerbit Fajar Bakti
32. Headstart, C, 2002, *Easy Learning Bilingual Dictionary English-Malay, Malay-English*, HarperCollins Publishers

## APPENDICES

### APPENDIX 1 : PROJECT TIMELINE

Activities	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
<b>Planning</b>																
ResearchTitle for Initial Proposal	■															
Identify problem statement		■														
FYP Briefing			■													
Identify project objectives and project scope				■												
<b>Analysis</b>																
Research (Journals and articles)			■	■	■											
Research on other related work				■	■	■										
Research on available tools						■										
Analyse research materials					■	■										
Literature Review																
<b>Design</b>																
System Architecture				■	■	■										
Database Design																
<b>Development</b>																
Implement design into coding				■	■	■	■	■	■	■	■	■	■	■	■	■
Learning Programming Language																
<b>Testing</b>																
Functional Sub Modules testing				■	■	■	■	■	■	■	■	■	■	■	■	■
Integration testing												■	■	■	■	■
Revision of System																■
Preparation on Final Report/ Dissertation																

**Project Timeline**



## APPENDIX 2 : PRE-QUESTIONNAIRE

### Instruction

The following are the questionnaires related to the mobile phone usage as well as incorporation of dictionary in mobile phones. The purpose of Section A questionnaire is to identify the need to develop a dictionary as one of the mobile applications. Section B purpose is to elicit response on existing dictionary usage as well as its functionalities. Please provide your best answer for each of the questions provided.

### Particulars

Age :

Programme :

### Section A : Mobile Phone Usage

- 1) Do you own a mobile phone?  
a) Yes                      b) No
  
- 2) Does your mobile phone have a Malay-English or English-Malay dictionary application?  
a) Yes                      b) No
  
- 3) Do you prefer the dictionary application to be implemented in mobile phone rather than separately developed in other handheld devices? ( Note : Handheld devices may include PDA, Pocket PC)  
a) Yes                      b) No

### Section B : Dictionary Usage

- 4) Have you ever used Malay-English or English-Malay electronic or mobile dictionary?

Very Unlikely	Unlikely	Neither Unlikely or Likely	Likely	Very Likely
1	2	3	4	5

- 5) If yes, have you tried translating Malay or English multiword (e.g “ubat gigi”, “anak patung” for Malay word or “wisdom tooth”, “shopping complex” for English word) that will produce correct answer?

Very Unlikely	Unlikely	Neither Unlikely or Likely	Likely	Very Likely
1	2	3	4	5

- 6) It is important for a dictionary to include the translation for multiword as described above

Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree
1	2	3	4	5

- 7) Do you think that current Malay-English or English-Malay electronic dictionary is very helpful?

Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree
1	2	3	4	5

- 8) In having a mobile dictionary, do you want the dictionary to have  
a) Fast processing time?

Very Unlikely	Unlikely	Neither Unlikely or Likely	Likely	Very Likely
1	2	3	4	5

- b) Simple interface and easy navigation?

Very Unlikely	Unlikely	Neither Unlikely or Likely	Likely	Very Likely
1	2	3	4	5

~Thank you for your participation. Your co-operation is highly appreciated~

## APPENDIX 3 : POST-QUESTIONNAIRE

### Instruction

The following are the questionnaires related to the accuracy of MobileDic system as well as its usefulness. You are required to fill in this questionnaire right after the hands-on assessment is done first on MobileDic.

### Particulars

Age :

Programme :

- 1) After trying out MobileDic system, do you feel that MobileDic system is accurate at translating multiword as well as single word?

Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree
1	2	3	4	5

- 2) Do you feel that MobileDic system is useful if it is implemented in your mobile phones?

Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree
1	2	3	4	5

~Thank you for your participation. Your co-operation is highly appreciated~