Using Automatic Speech Recognition to Evaluate Arabic to English Transliteration

GHADEER KHALIL

A thesis submitted in partial fulfillment of the requirements of Nottingham Trent University for the degree of Doctor of Philosophy

July 2013

Abstract

Increased travel and international communication has led to an increased need for transliteration of Arabic proper names for people, places, technical terms and organisations.

There are a variety of available Arabic to English transliteration systems such as Unicode, the Buckwalter Arabic transliteration, and ArabTeX. The transliteration tables have been developed and used by researchers for many years, but there are only limited attempts to evaluate and compare different transliteration systems.

This thesis investigates whether or not speech recognition technology could be used to evaluate different Arabic-English transliteration systems. In order to do so there were 5 main objectives: firstly, to investigate the possibility of using English speech recognition engines to recognize Arabic words; secondly, to establish the possibility of automatic transliteration of diacritised Arabic words for the purpose of creating a vocabulary for the speech recognition engine; thirdly, to explore the possibility of automatically generating transliterations of non diacritised Arabic words; fourthly to construct a general method to compare and evaluate different transliteration; and finally, to test the system and use it to experiment with new transliterations ideas.

A novel testing method was found to evaluate transliteration rules and an automatic application system has been developed. This method was used to compare five existing transliteration tables: UN, Qalam, Buckwalter, ArabTeX and Alghamdi tables. From the results of these comparisons, new rules were developed in order to improve transliteration performance; these rules achieved of score 37.9% transliteration performance which is higher than the 19.1% score achieved using Alghamdi's table which was the best performing of the existing transliteration tables tested. Most of the improvement was obtained by changing letter(s) for letter(s) transliterations, further improvements were made by more sophisticated rules based on combinations of letters and diacritics.

Speech recognition performance is not a direct test of transliteration acceptability, but does correlate well with human judgement, and offers consistency and repeatability. The issues surrounding the user of English ASR for this application are discussed, as are proposals to further improve transliteration systems.

Acknowledgements

This Ph.D. thesis is the outcome of collective efforts from all those contributed to it directly or indirectly. Therefore, I would like to acknowledge them all for their love and support.

Praise be to Allah (God), the most gracious and the most merciful, without his blessing and guidance my accomplishments would never have been possible.

I would like to express acknowledgement to Dr. Graham Tranfield, for accepting me as a Ph.D. student and introducing me to this ever challenging field of transliteration and speech recognition and for his patience, constant support from the beginning of my PhD studies, especially the understanding shown during my first year and his guidance throughout my research program at Nottingham Trent University. I also would like to express my appreciation to Dr. Taha Osman and Dr Tony Allen, for their guidance and support.

Subsequently, I would like to dedicate this to my husband Bashar and my son Abdulrahman.

The unconditional love and encouragement provided by my family served as a secure anchor during the hard and easy times; thank you Dad (Ismail), Mum (Fatima) and my loving grandmother (Haya), may her soul rest in peace!

I also would like to thank my sisters (Khairya, Haya, Shahad, and Zain), brother (Mohammed) and auntie (Huda) who are always there for me. I cannot thank you enough for all you have done.

Finally, I wish to express my gratitude to the University of Bahrain for its support and sponsorship during my post-graduate studies which greatly contributed to the successful completion of this study.

Thank you all, for everything!

Table of Contents

CHAPTER 7 IMPROVEMENTS TO ALGHAMDI'S TRANSLITERATION TABLE

List of Figures

List of Tables

CHAPTER1

Introduction

This research proposes a novel systematic approach to evaluating Arabic to English transliteration systems with the aid of speech recognition technology.

Mubarak et al., (2005) define transliteration as "transcribing a word or text written in one writing system into another writing system".

Transliteration retains the original sound of the word, so when a person attempts to pronounce the transliterated word, they make the same sound as the native speaker pronouncing the word written in the original language.

A lot of words like proper names for people, places, technical terms and organisations are rarely translated because they don't have a meaning. Instead they are transliterated. For example a name like "Shaheen", would be spelled in a French influenced country as "*Chahine*". Also the name *Antonio* could have an English equivalent e.g. Anthony, but that is not really his name, so it should be transliterated as *Antonio* to preserve the pronunciation of the phrase (Knight and Graehl, 1998).

It is common for language pairs that use the same script like Spanish -English to use the original spelling , For example *Antonio* gets transliterated as *Antonio* and Paris is used in English and in French. Nevertheless, "for language pairs that use different alphabets and sound systems, such as Japanese/English and Arabic/English the situation is more complicated" as stated by Knight and Graehl (1998).

The history of transliteration goes back to ancient times. Ancient maps and documents show names of cities written in Latin script which is clearly not the native

way of writing of the inhabitants; also early transliteration of Hebrew occurred with the contact between the Romans and the Jews.

As for the history of rules for transliteration, in 1885 the American Library Association (ALA) created a system for representing Cyrillic characters (Slavic information literacy, 2012). No diacritics were used and reverse transliteration was not considered. Diacritics are marks, or glyphs, sometimes called accesnt. They can appear above or below a letter, or sometimes in other positions such as within the letter or between two letters. The effect of diacritical marks is to change the sound of the letter to which they are attached. In very general terms, linguistic oriented publications tend toward systems with diacritics, while literature and cultural publications tend toward systems without diacritics; hence there was a need to develop a system that would incorporate both with and without diacritics. In 1905 the Library of Congress created their system for representing Cyrillic, which is almost identical to what is used today (Slavic information literacy, 2012). The British Academy created their own system in 1917 after appointing a committee to consider and draw up a practical scheme for transliteration into English of words and names belonging to Russian and other Slavonic languages and the languages of the Nearer East (UOAL, 2010; The British Academy, 1917).

The need for Arabic transliteration technology is increasing and this is derived from the major role it is playing in many applications, for example cross language information retrieval, airline, tickets, medical records, and a range of security applications such as terrorist watch lists, named entity recognition for instance the passport, as the principle information (name and place of birth) cannot be translated.

The biggest complication for Arabic transliteration seems to be that some sounds in Arabic may not exist in the target language. An obvious example is the problem for Arabic to English transliteration; Only eight out of 28 Arabic letters have an obvious equivalent in the Roman alphabet: B, F, K, L, M, N, R, and Z. Moreover Arabic has two

distinct consonants that are close to the sound of S. The same applies to D, H and T. Also, there are two glottal sounds that do not obviously correspond to any Roman letter (Al-bab, 2009).

The Deutsche Morgenländische Gesellschaft scheme was most likely the earliest attempt at standardization of Arabic- English data; in 1936 the system was approved by the International Convention of Orientalist Scholars (Whitaker, 2002). It is also used in the Hans Wehr Arabic dictionary. In 1971 another Arabic-English transliteration standard was adopted at a conference of Arab experts in Beirut (see Al-bab, 2009).

Two international symposiums were held in 2003 and 2006 by the Saudi academia and authority. The purpose was to customize the transliteration of Arabic names into the English Alphabet (Alghamdi, 2009). Both symposiums ended by developing a standardized Romanization table and algorithms.

1.1 Overview and contribution of this work

Nowadays, the increased demand for travel has led to an increased need for transliteration, for example the passport and ID; as the principle information (name and place of birth) cannot be translated. Similar issues arise for birth certificates, driving license and airline tickets. Increased travel means more requirements for transliteration of place names and addresses; increased world trade prompts similar demands (need to translate shipping locations, etc. and also the names of companies and people involved in transactions).

If documents such as those mentioned are used in a country that speaks a different language and the lettering system is different, problems can arise. For example when an Arabic traveller visits a foreign country, entering the data from the Arabic person's passport into the foreign country's system is impossible unless an appropriate transliteration system is available.

Even though there are spelling principles, there isn't one "correct" spelling for an English-Arabic transliteration as declared by AbdulJaleel and Larkey, (2003). For example Whitaker (2008) classifies about 32 different English spellings for the name of the Libyan leader Muammar Al-Gaddafi. The name of the Indian capital was at one time transliterated as "Bombay" in English; at present the official transliteration "Mumbai" is used. Another example is the Chinese capital Beijing which was formerly transliterated as Peking.

If someone needs to do a search for flights to Mumbai in a database, then any information that has been entered about Bombay will be missed as it is different data but means the same thing. There is therefore a need for a consistent method of transliteration.

The Arabic language is represented in 28 letters, which differs from the 26 letters of the Roman alphabet in which English is written. A word in Arabic may appear in different forms in English as there is no standard way of transliterating the letters from Arabic to English. For example, the name "غدير" can be transliterated: Ghadeer, Ghadir, Ghader… This can create a misunderstanding for officials and employees. The main reason for this is was the absence of standards as stated by Alghamdi (2009).

For Arabic/English transliteration there are a variety of schemes, such as The Buckwalter Arabic transliteration (Habash et al., 2007), SATTS, ISO 233 , Qalam (Becker, 1987), and ArabTeX (Lagally, 2004). These have been developed for the Arabic language. Transliteration systems have been developed for many other languages, such as Serbian and Russian (UNESCO, 2006a, 2006b).

However, there is still the question of how effective any particular transliteration system is. Alghamdi (2009) stated that "the used Arabic-English transliteration systems are inconsistent, inappropriate, or unsystematic. These difficulties caused concerns for the security and legal authorities." (p1). The systems are available but the question of how to evaluate them still remains. While there has been a developing interest in transliteration, the technology of speech recognition has also

been developing. Speech recognition systems deal with ways of representing and recognising sounds.

In recent years significant advances have been made in the field of speech recognition. It is now well established that accurate systems have been developed (Doe, 1998).

According to Wadhwani et al. (2011) "speech recognition is the ability to listen to (input in audio format) spoken words and classifying various sounds present in it, and recognizing them as words of some known language". For transliteration, when an Arabic name is converted into English script, the resulting transliteration is a word which should produce a close match of the original sound when spoken in the foreign language. This basically means choosing the best letters in the foreign language to represent the sound of the word in its original language (Sherif and Kondrak, 2007). When a word is spoken in a microphone, the voice analog signal gets converted into digital chunks of data that the computer must analyze. It is from this data that the computer must extract enough information to guess the spoken word (Haque et al., 2010).

A word consists of sounds or linguistic units known as phonemes. The speech recognition engine can match the segments to phonemes in the appropriate language. Many factors can affect how phonemes are converted into words like the speaker accent and age and the surrounding phonemes.

According to Deb et al., (2010) "English uses about 40 phonemes to convey the 500,000 or so words it contains, making them a relatively good data item for speech engines to work with". Transliteration tables are basically representations of every letter in the chosen language and the phonetic representation for this letter (how this letter should be pronounced).

That is why phonemes are studied in this research to allow the use of English engines to recognize Arabic words, for the purpose of comparing and testing transliteration tables.

There are several commercially available speech recognition systems such as Dragon Naturally Speaking (Nuance, 2006) and IBM ViaVoice (IBM, 2006). Peissner, (2002) states that the majority have been developed for the English language although there are several speech engines that have been developed for other languages.

Research into English speech recognition is becoming more intensive, and work on other languages, such as Farsi (Saleem, 2008; Srinivasamurthy and Narayanan, 2003), Vietnamese (Viet-Bac, 2007), and Arabic (Alghamdi, 2003) is steadily catching up.

The main aim of the research is to demonstrate a novel systematic way for evaluating currently published transliteration systems and to identify ways for improving these systems with the aid of speech recognition technology.

The quality of the transliterations could be tested using English speech recognition engine by matching the transliterated English words with their original Arabic words. If the recognised word matches the spoken word, this means that the transliterated word matches the original word and therefore the transliterated word is a good representation of the original word.

Automating the testing process ensures the repeatability and consistency of measuring the accuracy of the transliterations which contributes to the field and complements the existing evaluation methods of relying on subjective judgments.

Research Question and Hypothesis

Research Question

What is the relationship between transliteration and speech recognition technology?

Hypothesis

Comprehensive transliterated vocabulary and speech recognition technology could be used to implement an application to construct a novel general method to test different ways of performing transliteration; this could be used to evaluate currently published tables.

The initial objectives of this work can be summarized as follows:

1. Determine whether it is possible, for English speech recognition engines to recognize Arabic words with aid of Arabic transliteration.
2. Establish whether it is possible to automatically transliterate diacritised words for the purpose of creating a vocabulary for the speech recognition engine that could be used to evaluate transliteration tables.
3. Explore the possibility of automatically generating transliterations of non diacritised Arabic words (words without short vowel marks that provide a phonetic guide) and using speech recognition technology to evaluate transliteration tables.
4. Determine whether it is possible to construct a novel method to test and compare transliteration rules.
5. Establish whether it is possible to experiment with new novel transliteration ideas to find improvements in the transliteration rules.

1.2 Outline of the thesis

This thesis consists of 8 chapters which are as follows:

Chapter 1 provides a general introduction to the subject area within which this project is set in addition to discussing the motivation and contribution of this work.

Chapter 2 presents a literature survey of previous and recent work in the field of transliteration including Arabic transliteration and speech recognition technology. The problems associated with transliteration are also introduced in this chapter.

Chapter 3 provides a detailed discussion regarding the feasibility of using English speech engines to recognize Arabic transliterated words. This was achieved by manually finding a set of words which could be used to represent the 28 characters of the Arabic language. The English speech engine tries to match the English sounds that it believes that the Arabic speaker made against the transliterated vocabulary. The expected output is a string of English letters that would be pronounced like the input, these English letters represent the transliterations of the original words.

The possibility of using an American or British engine and the difference between them will also be covered. Issues surrounding the design and implementation of a letter identifier application that recognizes Arabic manually transliterated words will be provided. Finally experimental results are given.

Chapter 4 explores the idea that appropriate English vocabularies could be produced by automatically transliterating and generating words from diacritised Arabic for the purpose of creating a vocabulary for the speech recognition engine that could be used to evaluate transliteration tables. Moreover, the methodology and the effect of automating the process on recognition rates will be discussed. And finally the results for the automatic process based on diacritised Arabic will be mentioned.

Chapter 5 considers the possibility of generating transliterations of non diacritised Arabic words; the proof that this is impractical is clearly stated.

Chapter 6 presents a detailed description of the structure, development and implementation of a novel proposed system to test and compare transliteration tables. This involves identifying a comprehensive Arabic vocabulary as a research infrastructure which would also be available for Arabic researchers to stimulate further research in this field and its application.

Chapter 7 reports on using the proposed system to compare and improve currently published transliteration tables using new novel transliteration ideas like changing the letter for letter transliterations, then more sophisticated rules where different transliterations for letters depending on whether they are adjacent to specific diacritics also the use of more complex rules based on the novel concept of the diacritem has been explored.

Chapter 8 concludes the work by summarizing the major achievements and weaknesses of this study as well as discussing the contribution of the work and the potential avenues for further work.

CHAPTER 2 Literature review

In this chapter, detailed discussions of transliteration, the differences between Arabic and English, and speech recognition are presented.

2.1 Introduction to Transliteration

The increased need for Arabic transliteration technology is derived from the major role it is playing in a variety of applications, for example machine translation, cross language information retrieval, a range of security applications such as anti-money laundering and terrorist watch lists, named entity recognition (for instance the passport, as the principle information (name and place of birth) cannot be translated; airline tickets, medical, financial, and educational records). Moreover increased travel means more requirements for transliteration of place names and addresses, similarly for increased world trade (need to translate shipping locations, etc. and also the names of companies and people involved in the transactions).

Many words like proper names for people, places, technical terms and organisations are hardly ever translated because they don't have a meaning. Instead they are transliterated.For example; Arabic proper names such as Fatima are generally transliterated into the English script. There is enormous unpredictability in the Arabic representation of foreign words, in particular named entities. Even though there are spelling principles, there isn't one approved spelling for Arabic (AbdulJaleel and Larkey, 2003).

Alghamdi (2005) searched for the name "Mohammed" using the Google search engine and summarised the results in Table 2.1. Similar results were obtained when searching for other Arabic names.

No.	Transliteration	Frequency
1	Muhammad	2.280.000
2	Mohammed	2.000.000
3	Mohamed	1.600.000
4	Mohammad	1.150.000
5	Muhammed	388.000
6	Mohamad	264.000
7	Muhamed	69.100
8	Muhamad	44.600

Table 2.1: Statistics of the occurrence of the Arabic name "Mohammed" with different diacritics on the Internet using the Google search engine.

Although Mohammed is one of the most common Arabic names, it is clear from the above table that people transliterate it differently because there is no commonly accepted scheme for transliteration that everyone agrees on.

Moreover, the absence of standards leads to difficulty in transliterating names for officials, employees and name carriers. An Arabic letter may appear in different forms. For example, "عبدالرحمن" can be transliterated: Abdulrahman, Abdalrahman, Abdelrahman… This creates confusion for officials, employees and name carriers (Alghamdi et al., 2006).

Transliteration can generally be defined as changing (letters, words, etc.) into the equivalent characters of another language or alphabet, a good example is transliterating *the Greek X as* Ch (dictionary.com, 2010).

Other definitions for transliteration exist like "the process of obtaining the phonetic translation of names across languages. A source language word can have more than one valid transliteration in the target language" (Shishtla et al., 2009, p.40).

It can also be described as "The process of converting a word from one orthography into another" according to AbdulJaleel and Larkey, (2003, p.1). Additionally Mubarak, Al Sharqawy, and Al Masry (2005) identify transliteration as "transcribing a word or text written in one writing system into another writing system". Best candidates for transliteration include people names, locations and organizations in addition to words borrowed into the language.

The process of transliterating Arabic to a Roman script representation is called transliteration; it is also called *Romanization*, due to the fact that the target language uses the Roman alphabet. The opposite operation of transliterating non Arabic script into Arabic is called *Arabization* according to Halpern (2007).

Transcribing spoken language phonetically is very straightforward as stated by Atkielski (2005), when a word is spoken, the phonetic symbols that correspond to the sounds of the spoken word are written. It is easier to recognize the sound if you understand the language but it isn't necessary, as long as you can recognize sounds and transcribe them.

According to Dobrovolsky and Katamba (2008) phonetic transcription is a system for transcribing sounds that occur in a language. It attempts to represent each sound of speech with a single symbol. These symbols are enclosed in brackets [ ] to indicate that the transcription is phonetic and does not represent the spelling system of a particular language. For example, the sound spelled th in English this is transcribed as [ð] (pronounced eth, as in weather).

Two types of transcription exist, *broad transcription* that gives only a basic idea of the sounds of a language; in some cases this may be equivalent to a phonemic transcription. A close transcription, representing specific details of the sounds, is called a *narrow transcription (*Du Bois et al., 1993).

A number of systems have been developed for writing the sounds of the world's languages such as Alexander Melville Bell's Visible Speech (Duchan, 2006) and IPA (IPA, 2005). Many of the early workers made their own systems because there was no agreed standard or knowledge of the complete speech sound inventory (Hieronymus, 1993).

The most widely known system of phonetic transcription, the International Phonetic Alphabet (IPA) has hundreds of symbols, but only about fifty corresponding to the number of sounds used in English (Atkielski, 2005). IPA was developed in 1888 and revised several times to its present form. It represents putting a symbol to each sound in all of the known languages in the world; it is an exact one to one correspondence between written symbols and spoken sounds (Hieronymus, 1993). Refer to appendix A for the IPA full chart.

The main drawback of using Phonetic transcription is that it requires the user to be familiar with whatever system of transcription is used.

AbdulJaleel and Larkey (2003) argue that a lot of mystery surrounds the terms *transliteration* and *transcription*, with the first frequently used misleadingly in the sense of the second even in academic papers.

According to them, "*Transliteration* is a representation of the script of a source language by using the characters of another script. Ideally, it unambiguously represents the graphemes (spelling), rather than the phonemes (sounds), of the source language, whereas, *Transcription is a* representation of the source script of a language in the target script in a manner that reflects the pronunciation of the original, often ignoring graphemic correspondence" (p.1).

Additionally, Zhang and Li, (2012), indicated that "a transliteration may be almost the same as a transcription if the relations between letters and sounds are similar in both languages. Also some mixed transliteration/transcription systems exist. In a broader sense, the word *transliteration* may be used to include both transliteration in the

narrow sense and transcription" (p.1). This research will look into transliteration as it attempts to transcribe from one orthography to another, so that the word, when read, sounds the same.

As described in Al-Onaizan and Knight, (2002), two types of transliteration exist, *forward transliteration* and *backward transliteration.* Forward Transliteration is the transliteration of a foreign name into English or another language. Typically, there are several acceptable transliteration candidates.

Backward Transliteration is the reverse transliteration process used to obtain the original form of an English name that has already been transliterated into the foreign language. In this case, only one transliteration is retained, for example Graham is the original transliteration for غراهام, and جراهام, other transliterations like Garaham, and Jraham are acceptable but not correct.

This research will concentrate on forward transliteration only, since this is what is needed for testing and comparing transliteration systems.

According to Whitaker (2008), "Transcribing Arabic into the Roman alphabet is fraught with difficulty. And in an age of electronic text, search engines and databases, the problem is only going to get worse".

When transliterating between two languages with many phonemic incompatibilities, such as English and Arabic, this is particularly true.

All of the previously mentioned definitions of transliteration and transcription are vague. The problem is that the various definitions are not the same and this could cause confusion when these terms are used interchangebly.

The word transliteration will be used in this research as it describes an attempt to transcribe from one orthography to another, so that the word, when read, sounds the same.

2.1.1 Transliteration Schemes

Habash et al., (2007) referred to transliteration as a way of using English letters and other symbols to represent Arabic letters in a one-to-one way by using letters and phonetic symbols. Their scheme highlighted the transliteration and transcription of Arabic letters to English.

Although this scheme defined transliteration more specifically, it cannot be used by officials for writing peoples' names for example, because of the usage of phonetic symbols like θ and ð. The Habash et al., scheme is based on the Buckwalter transliteration scheme. The main advantages of the Buckwalter transliteration is that it is written in ASCII characters. However, the Buckwalter transliteration is not easy to read. Hence, the Habash et al., scheme avoided this problem by extending the Buckwalter transliteration scheme to include non-ASCII characters of which the pronunciation is easier to remember.

Table 2.2 below illustrates Habash et al.,'s transliteration and transcription scheme.

Arabic	(Habash et al., 2007) Transliteration	Pronunciation As in		Arabic	(Habash et al., 2007) Transliteration	Pronunciation As in
ب	B	Ball		ط	T	Emphatic t
ت	T	Tree		ظ	Ď	Emphatic D
ث	Θ	Three		ع	ς.	Sounds like a sharp a
ج	J	Jordan		غ	γ	Parisian French r
ح	H	Sounds like a sharp h		ف	f	Film
خ	X	Scottish Loch		ق	q	Sounds like a deep k
د	D	Door		ك	k	Kite
ذ	Đ	The		ل	l	Cool
ر	R	Road		م	m	Man
ز	Z	Zoo		ن	n	New
س	S	Sue		ه	h	Hot
ش	Š	Shoe		و	w	Would
ص	S	Emphatic s		ي	y	Yoke
ض	D	Emphatic d				

Table 2.2 Habash et al's (2007) transliteration and transcription scheme.

Other schemes exist like the ISO 233 code, which is an established system for Arabic transliteration Romanization that was completed in 1984.

Table 2.3 below presents the ISO 233 scheme (Pedersen, 2008).

Arabic alphabet	ا	ب	ت	ث	ج	ح	خ	د	ذ	ر	ز	س	ش	ص	ض	ط	ظ	ع	غ	ف	ق	ك	ل	م	ن	ه	و	ي
ISO 233	ʾ	B	t	ṯ	ǧ	ḥ	ḫ	d	ḏ	r	z	S	š	ṣ	ḍ	ṭ	ẓ	ʿ	ġ	f	q	k	l	m	n	h	w	y

Table 2.3 ISO 233 transliteration scheme.

The ISO 233 scheme contains some symbols (punctuation marks and Latin letters with a caron) like ʿ, ʾ and ǧ; this is why it cannot be used for the transliteration of names by officials.

Buckwalter (2002) developed the Buckwalter Arabic transliteration system which follows the standard encoding preferences prepared for representing Arabic characters for computers. The key advantages of this transliteration system are that it is written in ASCII characters and is a strict one to one transliteration. Nevertheless, the Buckwalter transliteration is not necessarily easy to read. Habash et al., (2007) mentioned that The Buckwalter transliteration has been used in a lot of natural language processing publications and at the Linguistic Data Consortium (LDC) resources.

Arabic alphabet	ا	ب	ت	ث	ج	ح	خ	د	ذ	ر	ز	س	ش	ص	ض	ط	ظ	ع	غ	ف	ق	ك	ل	م	ن	ه	و	ي
Buckwalter	A	b	t	V	j	H	x	d	*	r	Z	s	$	S	D	T	Z	E	g	f	q	k	l	m	n	h	w	y

Table 2.4 Buckwalter's (2002) transliteration scheme.

Like ISO233 and Habash et al., (2007) transliteration schemes, symbols are found in the Buckwalter scheme like $ and * which makes it not suitable for use by people.

SATTS is another transliteration standard which is one to one mapping to Latin Morse equivalents (Arabic Transliteration wiki, 2012).

In Morse code the signals are sequences of short and long pulses (dots and dashes) and the significates are the 26 letters of the English alphabet, the digits 0-9, and certain punctuation marks (Krauss, 2002).

2.1.2 Problems with the available schemes and proposed solution

All of the previous standards contain some nonstandard English letters (symbols) such as $,*, Ɵ and '; these standards use these symbols to represent the equivalent letters with one character only. Hence, Alghamdi et al., (2006) indicated that due to the lack of standard English letters transliteration, transliterating Arabic names is difficult for officials, employees, etc... Thus Saudi academia and authority held two international symposiums in 2003 and 2006 to try and solve this problem of standardising the transliteration of Arabic proper names into the English alphabet and the transliteration of foreign proper names into Arabic. Experts were invited to participate in the two events.

The first symposium was titled "Standardizing Arabic Names Transliteration: Security Dimensions". Topics including writing Arabic names in the Arabic alphabet, existing Romanization systems of Arabic names and the problems of the existing Romanization methods were covered. The symposium ended by developing a standardized Romanization table. This table didn't solve the problem completely as some problems remain, such as parsing and compound names.

Three years later, another symposium was held under the name: "Transliteration between Languages: Romanization of Arabic Proper Names". The outcomes of the symposium were a transliteration table and algorithms. This table solved the problems in the first table, like No difference between capital and small letters; for example not using the capital A for The letter ع and the small a for the letter أ, parsing should not be included in the transliteration; Muhammad not Muhammadan and compound names to be treated as one; Abdulrahman not Abdul Rahman (Alghamdi, 2009).

Based on these outcomes the door was opened for software engineers and phoneticians to implement systems that can aid in areas like health records, security, immigration, travel agencies and educational institutions.

Alghamdi et al., (2006) introduced a transliteration table that uses only plain Roman Alphabets that can be processed and printed easily, so that ordinary people can read the transliterations. His transliteration scheme differs from the previously mentioned schemes in that his scheme uses only plain letters (no symbols are used), and he uses more than one character to represent some letters whereas the previous schemes used only one character to represent each letter.

King Abdulaziz City for Science and Technology (KACST) supported a project in 2006 to develop a software system that can transliterate any name in Arabic into English based on the standards from the outcome of the two symposiums. The software system plus a collection of more than 70,000 Arabic proper names were developed. The system has been available and used since then (Alghamdi et al., 2006). However the look up system is only capable of transliterating Arabic names that are part of the 70,000 names in the system, it is not capable of transliterating anything else.

A lot of transliteration schemes are available and continue to emerge (Habash et al., 2007), and (Wikipedia, 2010e). However, Halpern (2007) believes that, in spite of the importance of Arabic transliteration, it has not been the subject of sufficient studies.

2.1.3 Difference between Arabic and English

"Arabic language faces some challenges like dialects, contrast between written and spoken language, gender differences in speech and vowelling" (Tomokiyo et al., 2003, p.1).

According to Frankfurt International School, (2012) the main differences between Arabic and English can be summarized as, 28 letters represent Arabic language, 10 of

these do not exist in the English language. The letters P, V, X, Ch, G do not exist in Arabic (except in certain dialects) and are replaced by the Arabic sounds B, F, and KS. Arabic is written from right to left, the opposite of the English writing system, which is written from left to right. Arabic is written cursively, and the letter appearance changes when occurring in the beginning, middle, or end of a word or when written alone. Additionally, there are major differences between male and female in pronouns, verbs, words, and sentence structure. Conjugation in Arabic is not the same as English. All verbs stem from a root verb and conjugate depending on number and gender. The root verbs conjugate to make different meanings as well, if you know the root word you can almost always guess what the conjugated verb means. The grammar structures are very complex, but systematic and contain few exceptions. It is an orthographically regular language, unlike English, which is irregular.   There aren't really any silent letters except in a few rare cases (Frankfurt International School, 2012).

There are many different Arabic dialects, which vary according to the speaker's city, district or country. There are three main classes of Arabic dialects: the Eastern dialects of Egypt and Sudan, the Middle East and the Western dialects of North Africa. All dialects are commonly understood among all Arabs, with the exception of the Western dialects of Tunisia, Algeria, and Morocco (Tomokiyo et al., 2003). Major phonological differences are apparent between these groups of dialects (Gulf, Levantine, Egyptian/Sudanese, and Maghrebi); the main differences are in the pronunciation of specific phonemes, such as the *qaf*, the *jim*, and the *tha* and *dha*.
In the eighth century AD, Sibawayh recognized and approved the scheme and place of pronunciation of each Arabic sound, in his famous book "Al-Kitaab" *(*Abdilmun'im*, 1993)*. Alghamdi (2006) continued Sibawayh's work by fully analysing, clarifying and describing Arabic sounds.

King Abdulaziz City for Science and Technology (Alghamdi, 2003) has published a detailed and comprehensive database called KACST Arabic Phonetics Database

(KAPD). KAPD contains more than 46000 files, and gives almost all the details of the articulatory mechanism of Arabic sounds. This database is very rich and is considered an important resource for all researchers in the field.

Arabic differs from English as it faces some challenges like Arabic dialects are essentially spoken varieties; also gender differences in speech and vowelling and the contrast between written and spoken language. These differences should be taken into account prior to transliteration.

2.1.4 The state of the art of computer transliteration

Noeman (2009) states that *"Most prior work in Arabic-related transliteration has been for the purpose of machine translation and for Arabic/English transliteration"* (p200).

Many transliteration systems appear to be included along with online translation (Ajeeb, 2010). Transliteration for many language pairs like English and Arabic (Al-Onaizan et al., 2002) and English and Korean (Lee and Choi, 1998) has been the focus of many research projects.

A simple Arabic/English transliteration system has been implemented by Al-Onaizan and Knight, (2002); they have also evaluated the reasonableness of their transliterations according to human judges. They report the overall accuracy of their transliteration algorithm using a phonetic- based model is 37.16% whereas the spelling-based model achieved 56.88%.

Arbabi et al., (1994) developed an algorithm at IBM for automatic forward transliteration of Arabic personal names into their Roman equivalent by vowelizing the given Arabic name by inserting the appropriate short vowels. Then the vowelized Arabic name is converted into its phonetic Roman representation using parser and table look up. The phonetic representation is then used in a table look up to produce

the spelling. This method applies only to Arabic names that follow strict morphololigical rules which limit the applicability of this approach since many organization and person names do not conform to these rules.

Ben Sassi, Braham, and Balghith, (2001) implemented a system where letters to sound rules are specified in a neural network based diaphone system manually.

The traditional methods like synthesis by rule and synthesis by concatenation of pre-recorded sounds used for this haven't given good results. Hence, neural networks were used because they have the potential to give better results thanks to their property of interpolation and their capacity of generalisation.

Tomokiyo et al., (2003) described a synthesis system for Modern Standard Arabic (MSA) that uses diaphones and definite subsyllable units. They automatically produce vowels. Their general-domain Arabic synthesizer runs 7 times faster than real time with a 9MB footprint and has an accuracy of 84.7 for sentences. The only limitation is that this system was designed for handheld devices only.

In modern Arabic text, there are no diacritics, which make it very difficult for the computer to process it, because the pronunciation of Arabic words cannot be fully determined by their spelling characters only. It could happen that two different words have identical spelling whereas their pronunciations and meanings are totally different. To remove this ambiguity, diacritics should be applied to determine the correct pronunciation. If a non diacritised word was to be used as a part of the vocabulary for a speech recognition engine, all the possible diacritics would be applied to all letters in different places and thus more possibilities would be generated to ensure that all words with identical spelling but different meaning are covered. The more words the vocabulary contains, the more load is placed on the system to attempt to recognise the spoken word.

Providing the computer with algorithms to copy the human ability in identifying the proper diacritics of the text is crucial. This tool can form the basis for text to speech

applications, Automatic Translation (Trost, 1991) and Arabic data mining applications (Hussein, 1998).

As an example of how frequent an Arabic word is diacritised in modern writing, Alghamdi et al., (2010) searched for the word "ورق" with different diacritics using the Google search engine and summarised the results in Table 2.5. Similar results were obtained when searching for other Arabic words. The frequency of the occurrence of the Arabic undiacritised version of the words is always higher than the occurrence of the diacritised words. This proves that in modern Arabic diacritics aren't always included.

Arabic Word	English Meaning	Frequency	Percentage
ورق		1,380,000	99.91
وَرَق	Paper	962	00.07
وَرِق	Silver	258	00.02
وَرّق	coming out Leaves	1	00.00
Total		1,381,221	100

Table 2.5: Statistics of the occurrence of the Arabic word "ورق"with different diacritics on the Internet using the Google search engine (Alghamdi et al., 2010).

Diacritisation can be defined as the process of adding the correct diacritics to an unmarked text. According to Tomokiyo et al., (2003) "Diacritics representing the correct Classical Arabic vowels appear in religious texts and children's literature, and are identified as the *vowelling* or the *vocalization*"(p.2).

On the other hand Elshafei et al., (2002) insists on using the term diacritization instead of vowelling for the reason that, the missing symbols do not represent vowels only but also, *shaddah* (consonant doubling), lack of vowels *sukoon (*written as a small circle as in ڣْ) and *Tanween* (the doubled case ending diacritics are vowels used at the end of the words to mark case distinction, which can be considered as a double short vowels) (Zitouni et al., 2006) which make it more comprehensive, and that is why the term diacritisation will be used in this thesis.

This dilemma of Arabic diacritization in general has been addressed by El-Imam, (2003), Zitouni et al., (2006), Habash and Rambow, (2007), and Elshafei et al., (2006); all trying to handle this problem using statistical approaches but they tend to handle the case ending diacritic mark in the same way they used to handle the internal (any letter but the last) diacritics. This is a problem according to Shaalan et al. (2009), because they believe that the detection of case-ending diacritics is a syntactic based problem whereas detecting the internal diacritics is a morphological-based problem so the two should be dealt with differently.

Obviously, determining the correct diacritics is a major consideration for Arabic recognition systems. Kirchhoff et al., (2003) describes an approach to automatic romanization for natural speech recognition that achieves 80% accuracy in generating the correct diacritisation as predicted by comparing it with manual diacritisation. This is an enormous improvement over the 50% accuracy measured for commercially-available diacritisers, which are targeted toward Modern Standard Arabic.

There are a few systems for diacritisation that are available in the market like that of Sakhr (2011), and RDI (Al Badrashiny, 2009). However, they are not open source and usually are integrated with other systems. Researchers who are interested in this area have tried their own technique and different methods have been applied for the diacritisation of Arabic text, for example AbdulJaleel and Larkey, (2003) managed to develop an n-gram based statistical system for romanising Arabic. Their system

achieved an error rate of 10%-20%. Automatic Diacritisers using Hidden Markov Models were developed by Elshafei et al., (2006), Ananthakrishnan et al., (2005), Kirchhoff et al., (2004) and Nelken and Shieber (2005), in addition to rule-based automatic diacritisers (El-Imam, 2004), example-based, hierarchical (Emam and Fischer, 2005), morphological and contextual-based were also developed by Kirchhoff et al., (2004). El-Sadany and Hashish, (1989) treated diacritisation as a machine translation problem.

The main disadvantage of these systems is the difficulty to keep the rules consistent, up-to-date and extend them to other Arabic dialects. All of the previous methods managed to score high accuracy rates (70-90%) but the most successful seems to be using the HMM approach to solve the problem of automatic generation of the diacritical marks of the Arabic text, the use of a preprocessing stage and trigrams for selected number of words and articles may improve the performance to about 2.5% error rate. Further improvement may require some knowledge-based tools involving morphology-syntax analysis (Elshafei b et al., 2006).

A number of the transliteration systems and diacritisers have been developed and used by researchers and users (Habash et al., 2007). Although human judges have been used to evaluate transliteration systems, the question of finding other ways to evaluate these systems remains as depending on human judges to evaluate transliteration systems is not reliable nor consistent.
The readily available transliteration systems still cause some concerns for security and legal authorities because they are inconsistent, and inappropriate as stated by Alghamdi (2009).

2.1.5 Transliteration Evaluation

Lawson (2008) evaluated 6 famous transliteration schemes, ISO 233-2, Qalam, SATTS, Arabic chat alphabet, Buckwalter, and ALA-LC/UNGENG based on phonetic and spelling accuracy and usability.

Lawson's definition of accuracy examines how close the pronunciation is to the original Arabic letter, this could be very tricky because there are some letters in Arabic that have no direct English equivalent like the letter ض.

The usability part of the evaluation investigated each tables' adherence to ASCII standards i.e. non use of symbols (non letters). Usability measures how accurate the representation of the transliterated word is. According to Bevan et al., (1995) usability is defined as "the effectiveness, efficiency and satisfaction with which specified users can achieve specified goals in a particular environment" (p.2). Hence in this part Lawson tried to measure how usable people found the representation of the transliterated word, and whether they found it effective, efficient and whether they were satisfied with it.

On the other hand accuracy assesses how close the pronunciation of the word is to the original Arabic word and does it match with the transliteration. "Accuracy is a faithful measurement or representation of the truth; correctness; precision" (Collins English Dictionary, 2003). Thus Lawson used accuracy to assess the pronunciation of the words and how close they were to the original Arabic word.

Lawson scored each system from 0 to 120, with 60 points allotted to phonetic accuracy and 60 points allotted to usability. Of the latter 60, 30 points come from ASCII compatibility and 30 points come from maintaining non-use of diacritics native to Arabic (Lawson, 2008).

The results of his evaluation method were that Qalam transliteration scheme emerged as the winner scoring 110 out of the possible 120 points, and ALA-LC claimed

a very respectable second scoring 107 points. ISO 233-2 and SATTS both scored 94 points which makes them the lowest.

This is a good method to evaluate transliteration tables as there are clear criteria to measure and test different letters and tables, using usability and accuracy measures.

Karimi et al., (2007) stated that the accuracy of the results of transliteration experiments are evaluated using a metric called word accuracy which quantifies the proportion of transliterations that are correct based on the available test corpus:

$$A = \frac{\text{number of correct transliterations}}{\text{number of test words}}$$

Since Arabic transliteration is playing a gradually more important role in a variety of practical applications, it is necessary to pursue efforts to research this field more.

At the same time as the interest in transliteration has been developing, the technology of speech recognition has been developing as well, which is concerned with ways of representing and recognising sounds.

2.2 Speech recognition

Over the last two decades, speech technology has witnessed a steady improvement. Today, speech technologies are commercially available for an unlimited and interesting range of tasks (Vimala and Radha, 2012).

Speech recognition can be defined as "the process of extracting the message information in a voice signal in order to write the spoken text or control the machine actions in response to spoken commands" (Doe, 1998, p.7).
It can also be described as the method by which a computer identifies spoken words as stated by Cook, (2002).

Speech recognition systems have numerous applications; for example, they provide a helpful tool to learn a new language known as computer aided learning system, also they provide aid to handicapped people (Medical Disabilities) like users with carpel tunnel syndrome and people with visual impairment, in addition to their use in Talking books and toys and Telecommunication services plus multimedia.

Users nowadays can easily input text and data into a computer or smart phone orally.

This modern technology allows users to speak commands so as to carry out tasks. Speech recognition software can be used together with a PC or Mac and with the aid of a microphone headset (Doe, 1998).

Determining whether the spoken words are interpreted as dictated text or commands is one of the most basic distinctions in speech recognition. Commands are easier than dictated text to implement for the reason that the number of recognizable words is limited. In contrast, dictation tends to recognize any spoken words.

In the past years speech recognition has made significant advancement. Systems continue to emerge with remarkable accuracy.

According to Vimala and Radha, (2012) most systems try to overcome restrictions such as 1) small vocabulary 2) isolated words (discrete speech), and 3) speaker dependence. The most difficult constraint for systems to overcome has been found to be speaker independence.

There are two types of automatic speech recognition, continuous or discrete. In the continuous type words are spoken in a natural manner whereas the discrete type requires the user to speak with a pause between each word. For example the words "recognize speech" can easily be confused as "wreck a nice beach".

Speech recognition systems can also be categorised as speaker dependent and speaker independent systems.

Speaker dependent systems require the speaker to train the system before reasonable performance can be anticipated. On the other hand speaker independent systems do not require any former training by the users as stated by Kemble, (2001).

2.2.1 The evolution of speech recognition

The evolution of speech recognition goes back to the 1857 when Frenchman Léon Scott invented the phonoautograph which is the earliest known device for recording sound (Cho, 2005). In 1952, a system that could recognize digits 0 to 9 from a single speaker was developed in Bell laboratories. This system had an accuracy of 97%-99%. (Jurafsky and Martin 2009). In the 1970s, most speaker dependent systems required the user to train the system for long hours, besides the system limitation of handling small vocabularies. Currently, speech recognition systems are able to deal with continuous dictation and to handle large vocabularies.

Research is being conducted and systems continue to emerge in order to equip speech recognition systems to recognize natural language even in difficult conditions (e.g. recognize speech with eliminating noise in noisy environment).

The diagram in figure 2.1 shows the speech recognition system parts.
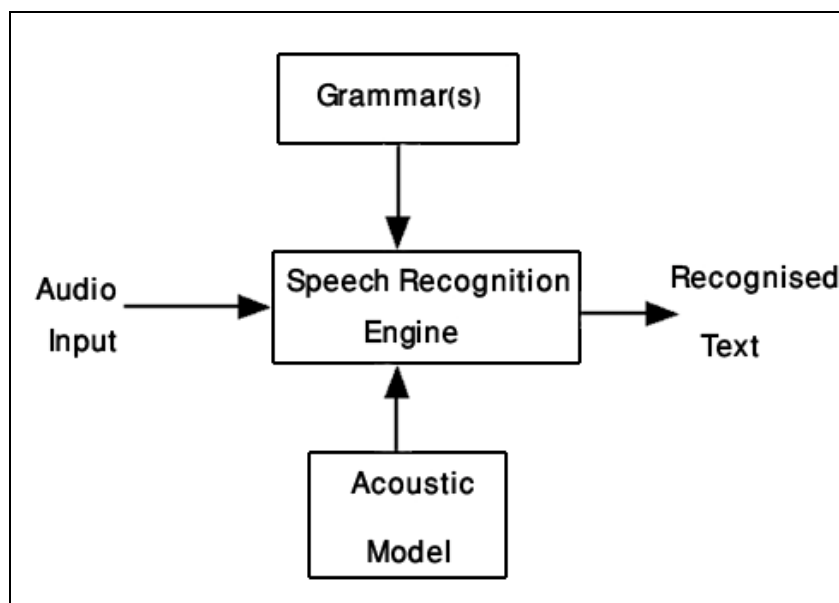
Figure 2.1 The speech recognition system parts (Kemble, 2011).

Speech recognition systems consist of five main components (the Speech Recognition Engine, Audio Input, Grammar(s), Acoustic Model and Recognized text).

In order for the speech recognition engine to take raw audio input and translate it to recognized text, "it utilizes all sorts of data, statistics, and software algorithms. Its first job is to process the incoming audio signal and convert it into a format best suited for further analysis" as declared by Kemble, (2001, p.5).

Then the engine takes into consideration its knowledge of the environment in which it is operating e.g. telephony environment etc. (provided in an acoustic model form) in conjunction with the words it knows about the (vocabulary) and tries to find the best match, when the speech data is in the proper format. When a match is found it gets returned as a string.

Speech recognition engines require two types of files to recognize speech. An acoustic model, which is created by taking audio recordings of speech and their transcriptions (taken from a speech corpus), and 'compiling' them into statistical representations of the sounds that make up each word and describe the sound of language. They also require a language model or grammar file which describes how words are distributed in spoken language. A language model is a file containing the probabilities of sequences of words. The speech engine's language model and acoustic model enable it to process spoken variations of the pronunciations specified in its lexicon, as well as new words (Mansikkaniemi, 2010).

A pronunciation lexicon is a collection of words or phrases together with their pronunciations, which consist of letters and characters from a supported phonetic alphabet.

Speech recognition engines have an internal lexicon that identifies which word in a language can be recognized. The lexicon indicates how the engine expects a word to be pronounced using characters from a single phonetic alphabet (MSDN, 2012).

According to MSDN (2012) phones make up a phonetic alphabet that contains combinations of letters, numbers and characters. Phones describe the spoken sounds of one or more human languages, and characterize the valid set of tokens that can be used to define the pronunciations of words using phonetic spellings. Similar to those used in dictionaries, phonetic spellings in lexicons describe how words should be pronounced for speech recognition. Hence, the speech recognition engine listens for pronunciations of words that correspond to phonetic spellings that are specified in its internal lexicon.

"A speech recognition engine can also create pronunciations on-the-fly for words it encounters that are not included its lexicon" (MSDN, 2012). To improve the accuracy of speech recognition engine, the default lexicon can be supplemented by creating an application specific lexicon. Though it is often not necessary because a speech engine can find and create pronunciations for both common and uncommon words in a language.

Although some promising solutions are available for speech recognition, most of them are tuned to English. The acoustic and language models for these systems are for the English language. If new words or words that are not included in the default lexicon like Arabic names (names and business names, or words that are specific to specialized areas of business, education, or medicine) were to be guessed this will be challenging for the speech engine because some of the letter in Arabic have no equivalent in English for example the letter (خ). For these cases a custom pronunciation must be specified, that may improve the recognition accuracy for the specialized vocabulary in the application also specifying a new pronunciation that replaces the predefined pronunciations; for example adding pronunciations to cover dialects and slang, also specifying multiple phonetic pronunciations (spellings) for a word.

For this research a new transliteration table containing each letter of the Arabic alphabet and its English equivalent will be specified and this can solve the problem of

the difference between Arabic and English. In the cases of Arabic letters that have no equivalent in English, an equivalent will be chosen for each of these letters, the chosen representations will be similar to the pointed sounds, but are not the same.

It is easy to distinguish among a small set of words, but error rates naturally increase as the vocabulary size grows (Tebelskis, 1995). Thus these letters will be included in a pre-defined, distinguishable, and small sized vocabulary, since even a small vocabulary can be hard to recognize if it contains confusable words.

Vibrations are created when a person speaks, then an analog to digital converter (ADC) converts analog signals into digital form. Digitization of sound takes place by measuring it at regular intervals. The sound then is filtered into different frequency bands and normalized; so that it attains a constant volume level. The sound is then checked whether it matches the stored sound templates. After that analog signals gets divided into segments that range from a few hundredths to thousands of a second. This helps in identifying plosive consonant sounds like "t" and it can be matched to phonemes that are already stored in the system (Vollmann et al, 2000).

The statistical modeling systems, which use mathematical systems and probability, are used to determine or predict the outcome after a particular phoneme. It becomes easier to guess where a specific word begins and ends.

Hidden Markov Models are the most commonly used speech recognition algorithms and are commonly used in speech applications (Roe and Wilpon, 1993). This is due to their ability to characterize the speech signal in a mathematically tractable way.
Rabiner and Juang,(1986) refer to HMM as "a doubly stochastic process with an underlying stochastic process that is hidden, but can only be observed through another set of stochastic processes that produce the sequence of observed symbols" (p.2) which means that using HMMs requires using random modeling to decode a sequence of symbols.

"Hidden Markov model is the simplest model that can be used to model sequential data, specifically data samples that are not independent from each other" (Edward et al., 2007) (p.3). Other options are rule based, neural networks and template matchers.

There are some difficulties that can be expected in spontaneous speech for example, out of vocabulary words, false starts, disfluency and lip smacks.
In order for the speech recognition system to work effectively it is supposed to identify the meaningful keywords embedded in fluent speech and ignore all the other speech events according to Lin et al., (2002).

Other issues that can influence the performance of a speech recognition system include the format for talking (isolated or connected inputs and continuous speech), the speaking environment and transmission conditions, type and amount of semantic and syntactic information and finally whether the system is speaker independent or speaker trained.

Vocabulary size can vary from 2 words to more than 40,000 words. Small vocabularies can force restrictions like out of vocabulary error on the naturalness of communication, however large vocabularies have more errors in speech recognition accuracy, as stated by Bazzi, (2000).
Grammar rules that classify how words can be spoken in context, often limit the vocabulary (Peacocke, 1990).

Large vocabulary can create some problems and limitations. As a system's vocabulary increases, the number of confusable words (i.e., the words that the system might mistake for others because they have the same pronunciation) increases (Doe, 1998).

According to Franz, (2002) a speaker-independent system was deployed by Google Labs as a demo of a telephone interface for its popular search engine. Nevertheless, their system is limited. The user can only say a word, but not a full question.

Together Carnegie Mellon University and Sun Microsystems developed a speaker-independent speech recognizer: Sphinx (Walker, 2004). Continuous speech using a large vocabulary can be recognised. The complication of the grammatical structure in the sentences can affect recognition results; SPHINX achieves speaker independent word recognition accuracies of about 71-96% on a vocabulary of over 21,000 words, depending on the complexity of the grammatical structure in the sentences (Agaram et al., 2001).

There has been a lot of research that highlights that some systems use information from the context to identify words, but that is not what this research is intending to cover.

There are systems that rely on learning the user's voice like the desktop dictation systems. These systems are speaker-dependent (i.e. IBM Via Voice (IBM, 2006), Philips Dictation Systems (Philips, 2005)). Since they operate using very large vocabularies, dictation systems perform much better when the speaker has spent the time to train the system to his/her voice. Generally speaker-dependent systems are reasonably accurate for the trained speaker, but much less accurate for other speakers. They also assume the speaker will speak in a steady voice and tempo. This is not what this research is focusing on.

Systems that comprehend isolated word recognition have been in existence for many years. Continuous speech research thrives because only through continuous speech can desired speed and naturalness of man machines communications be achieved as stated by Lee, (1990), Lewin et al., (1993), Glass and Hazen, (1998), Seneff, (2002) and Baumgarten, (2000).

Words spoken in a natural pace can be recognized by continuous speech systems rather than isolated words. On the other hand non-continuous speech systems require a calculated pause between each word. Grasso, (2005) states that even though continuous systems are more attractive, continuous speech is harder to

process, because of the complexity in identifying word boundaries. Such continuous systems are to be used for people who are intending to recognize dictation, but this is not the case for this research as it is aimed at recognizing isolated words.

Research work and developed systems on English speech recognition is becoming more intensive than before and other lagging languages like, Farsi (Saleem, 2008) and (Srinivasamurthy and Narayanan, 2003), Vietnamese (Viet-Bac, 2007), Indonesian (Sakti, 2007), Spanish (Niculescu, 2008), Estonian (Alumae, 2004) and Arabic (Alghamdi, 2003) are steadily catching up.

2.2.2 Arabic speech recognition

Kirchhoff, (2003), and Kirchhoff et al., (2004) state that although Arabic is currently one of the most extensively spoken language in the world, there has been fairly little speech recognition research on Arabic compared to the other Languages.
Although lagging behind other languages, research work on Arabic speech recognition is becoming more thorough than before, and a number of papers on the topic have been published (Ismail and bin Ahmad, 2004).

As stated by AbuZeina and Elshafei (2012) the development of an Arabic speech recognition is a multidisciplinary effort, which requires integration of Arabic phonetic (Algamdi, 2003), Arabic speech processing techniques (Elshafei et al., 2007) and natural language processing (Elshafei et al. 2006). Development of an Arabic speech recognition system has recently been addressed by a number of researchers.

A speech dataset for (MSA) Modern Standard Arabic that can be used as a main resource for researchers in the speech recognition field has been provided by Al-Otaibi, (2001). As well as offering a new technique for labeling Arabic speech, it achieved a recognition rate for speaker dependent ASR of 93.78%. Alotaibi (2003) also reported achieving high performance Arabic digits recognition using recurrent networks.

Sagheer et al. (2005) presented a novel visual speech features representation system. They used it to comprise a complete lip-reading system.

Research issues for Arabic speech recognition and the problems of indexing of Arabic news broadcast were addressed by Billa et al., (2002).

The problem of parsing transcribed spoken Arabic was addressed by Rambow et al. (2006). They examined three different approaches: sentence transduction, treebank transduction, and grammar transduction. Overall, grammar transduction outperformed the other two approaches. Parsing can be used to check the speech recognizer n-best hypothesis to rescore them according to most syntactically accurate choice (AbuZeina and Elshafei, 2012).

When developing an Arabic Speech recognition application, Modern Standard Arabic (MSA) must be taken into consideration, which is a formal linguistic standard used throughout the Arabic speaking world and is employed in the media.

2.2.3 Summary

Even though speech recognition technology is one of the most complex areas, this technology for English language has basically reached the point of technical maturity. It is commonplace to be able to create systems to recognise specific words in a list.

There are several commercially available voice recognition systems such as Dragon Naturally Speaking (Nuance, 2006) and IBM Via Voice (IBM, 2006). These systems are fast and aid in creating or editing documents and emailing them without typing and have a recognition rate of 99%, which makes them nearly perfect (Petrie, 2003).

Voice dialing applications like Call Home are the most developed and widely used. Peissner (2002) states that the majority have been developed for the English language although there are several speech engines that have been developed for other languages.

Research showed the availability of well-developed systems that are capable of recognising English words that may be part of the language or proper names and these proper names may have their origins in English or any other language

2.3 Discussion

Referring to the previous definitions of transliteration, this can be summarized as basically finding the right orthography in another language to represent an original pattern so that it would be pronounced correctly. This is done by developing a transliteration scheme; which is basically using letters and phonetic symbols to represent letters of a specific language. These representations should sound the same as the original letter.

The simplest way to test a transliteration scheme is to let an English person read transliterated words and ask an Arabic native speaker evaluate the test. This could be time consuming and it may be very inconsistent and difficult to repeat the results. Hence using an Arabic speech recognition engine to test Arabic transliterated words could be an alternative.

Speech recognition systems try to listen to an input in audio format, classify various sounds present in it and finally recognise them as words of some language according to Gupta, (2005). In the transliteration case, when an Arabic name is converted into English script, the resulting transliteration is hopefully a close match of the original sound produced in the foreign language. This basically means choosing the best sounds to represent the word (Pouliquen et al., 2005).

Speech recognition technology which is now well developed could be used to evaluate transliterations because both technologies are concerned with the sounds of words.

According to Sugumaran (2013) when we speak the sound comes out in phonemes, each phoneme resonates at a fundamental frequency and harmonics of it and

therefore have high energy at those frequencies. The first three harmonics are known as formant frequencies and have significantly high energy levels. Each phoneme has a unique fundamental frequency and hence unique formant frequencies and it is this feature that enables the identification of each phoneme at the recognition stage.

Reference templates of phonemes or words with which input speech is compared and the closest word or phoneme is given out are stored in the speech recognition systems.

When the user reads the transliterated words; the spoken phonemes or sounds should match the phonemes stored in the speech recognition system. Both transliterated words and reference template of phonemes are based on the same transliteration scheme, and hence speech recognition systems are ideal to test transliteration schemes. When the speech recognition engine recognises a word correctly this could indicate that this word is transliterated correctly and hence the transliteration scheme is good.

The recent Arabic speech technology is not as well developed as the more mature English speech technology also Arabic speech engines weren't available to the researcher. Therefore this research will concentrate on using an English speech engine as an alternative to recognize Arabic transliterated words for the purpose of evaluating transliteration schemes.

The English speech engine tries to match the English sounds that it believes that the English speaker made against the sounds it hears from native speakers. The recognition accuracy depends on the accuracy of rules of what sound should represent (Grammar rules). If the rules are good the sound would be equivalent and vice versa.

The only problem with Arabic is that some sounds in Arabic don't exist in English or any other language like the letter (ظ) which is unique to Arabic.

In this case a good transliteration means that the English text generates sounds that are generated as close as possible to the original Arabic sound and they should be part of a distinguishable vocabulary.

2.4 Chapter Summary

This section presents an overview of transliteration and also highlights the speech recognition problem.

Speech recognition technology which is now well developed could be used to evaluate transliterations because both technologies are concerned with the sounds of words (phonemes).

This research intends to investigate the possibility for using a well-developed speech recognition technology like English speech recognition technology to help test transliteration rules. Also the option of using the English speech recognition engine to recognize Arabic words will be studied.

CHAPTER 3

Recognising Arabic words using an English speech recognition Engine

This chapter discusses using English speech recognition technology to aid transliteration of Arabic vocabularies; this is identified in Chapter 1 as the first step that is required in this research.

A second motivation for this work was to see if this provided a practical alternative to the development of Arabic speech recognition engines.

3.1 Introduction

The quality of transliteration could be simply tested by transliterating a small text and evaluating the result.

If Arabic to English transliteration was to be tested, a word in Arabic for example (أرنب) would be transliterated into English as (Arnab), and an English native speaker would be asked to read the word, both the transliteration and testing process should be done by a linguist specialized in both languages. The main goal of the transliteration is to provide nonnative speakers with the correct pronunciation of the word, so that when an English native speaker reads the word (Arnab) it should be close to how the Arabic native speaker pronounces it, the matching of the two sounds determines the quality of the transliteration, in this case judged by the human expert.
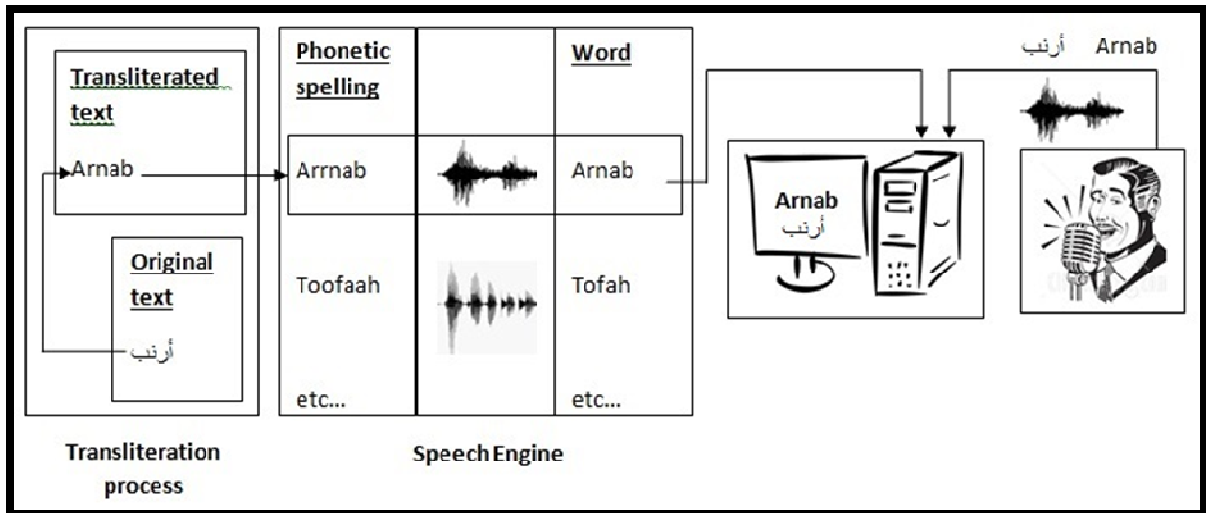
Figure 3.1 Transliteration evaluation process

Evaluating Arabic to English transliteration manually is time consuming and may be inconsistent and difficult to repeat the results; it also requires the presence of the human expert. Hence using an Arabic speech recognition engine to test Arabic transliterated words could be an alternative. An Arabic speech engine wasn't available to the researcher and therefore an English Speech engine will be used, taking into account the problems that might accompany using an English engine to recognize Arabic speech, such as recognizing accents and sounds that are unique to the Arabic language and don't exist in English or any other language, like the letter (ظ). For these cases identifying a custom pronunciation might help improve recognition accuracy. A good transliteration means that the English text generates sounds that are as close as possible to the original Arabic sound and they should be part of a distinguishable vocabulary.

For this research the human testers will be replaced with an English speech recognition engine to save time and ensure the consistency of the results. Using an English Engine to recognize Arabic transliterated words for the purpose of evaluating transliteration tables has potential since it has the ability to give information about the accuracy or quality of transliteration. For example, as previously mentioned if the word (Arnab) is used in the lexicon of the English speech engine, the engine ought to

match the sound that is internally created for this word with the spoken word (أرنب)
pronounced in Arabic. If the sounds match, this could indicate that the transliteration
is good, and vice versa poor transliteration could lead to misrecognition of sounds and
words.

In other words speech engines compare spoken sounds with sounds created internally
from the text of words stored in their lexicon. Refer to figure 3.1.

When a word is spoken the Speech engine selects the closest matching word from a
list of words that match the spoken word to some degree, nevertheless this doesn't
exactly give a direct evaluation of the match between the spoken and written word.
In a sense the accuracy of the similarity of the spoken and written word can be
verified by selecting the correct word from the list.

How good the recognition rate is does however depend to some extent on the
number and nature of other words in the lexicon – if there are a lot of words, and if
there is a high degree of similarity amongst the words, recognition rates will be lower.

The transliterated word's recognition rate contrasted with the spoken words in the
original language could provide an easy and effective way to measure the accuracy of
transliteration schemes. While there are limitations to this approach, as already
mentioned, it does provide a more efficient indication of transliteration accuracy than
that of using human judges.

In order to explore this proposal an initial simple application was created and tested
which was to recognise the 28 names of the letters of the Arabic Alphabet in a similar
way to that used in the Civil Aviation Organisation code to identify letters of the
English alphabet (Alpha, Bravo…..Zulu).

Actually, no such similar code exists for Arabic letters (except for a names code that
was used by the Iraqi Army which was not available to the author), so it was first
necessary to create a code by choosing words that would be familiar to Arabic

speakers, but that would be sufficiently different from one another to be easily distinguished by the application in the recognition stage. Work was then undertaken to identify the best English spelling to represent the phonetic structure of these Arabic words.

The methodology for this work is shown in Fig 3.2.
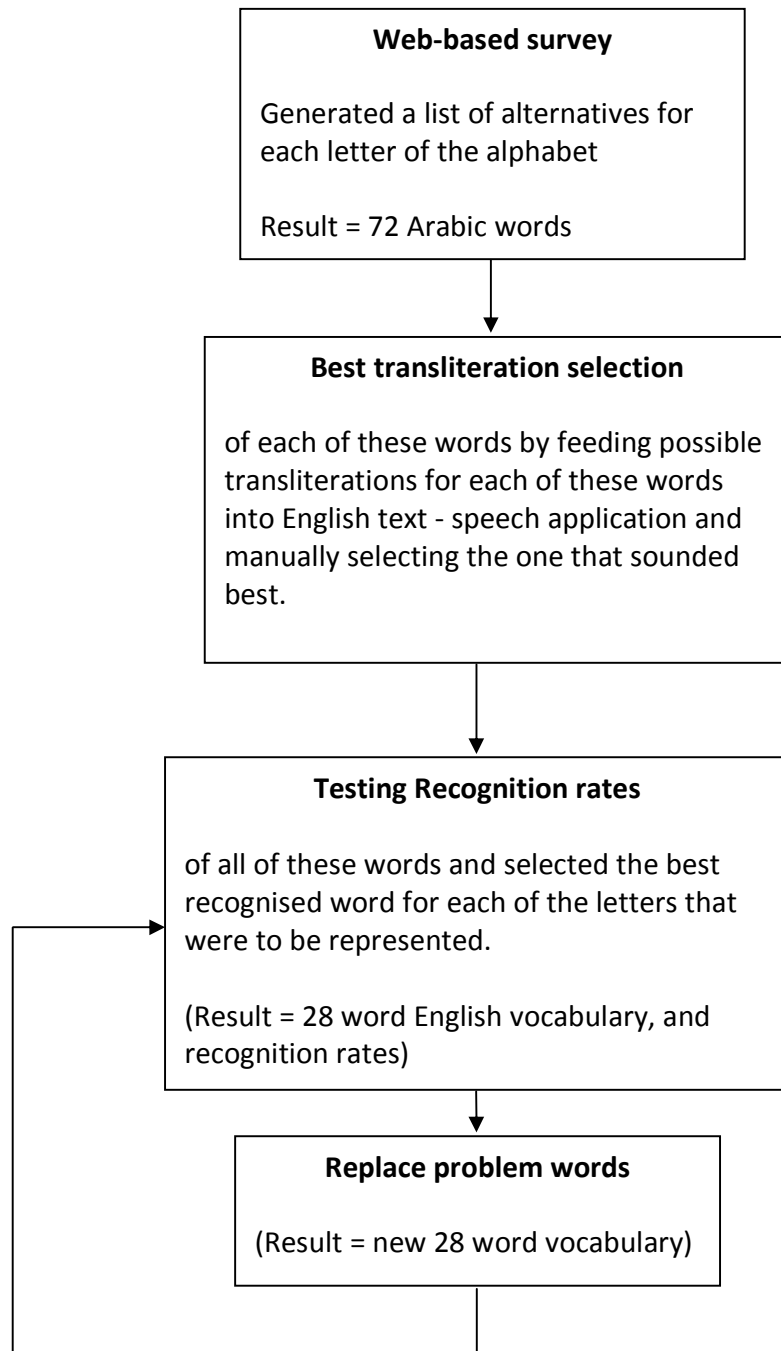
Figure 3.2: Recognising Arabic words using an English Speech engine

3.2 Initial Word Selection

The initial selection of words was made by publishing a web-based survey. Friends, family and first year computing students at the University of Bahrain were invited to fill in the questionnaire and 100 people took part.

A version of FreeOnlineSurveys.com was used to construct a Web-based survey for the evaluation. The survey simply provided a space by each letter of the alphabet (arranged vertically on the page in Arabic alphabetical order), with instructions appropriate for the specific experiment (to provide any Arabic word that starts with each Arabic alphabet).

After receiving the e-mailed invitation, participants clicked a link in the message that brought up the page containing the survey that the participant was to take. The participants read an introduction explaining the purpose of the survey, then completed and submitted the form. Refer to appendix B for a copy of the survey and a summary of the results.

Word	FCW%	Word	FCW %	Word	FCW %	Word	FCW %
أ Arnab* Asad	59 34	د Dob Deek*	52 39	ض Dhifdaaa Dha baaab Dhameer*	90 2 2	ك Kalb Korrssay*	75 21
ب Batta Boostan* Baab	42 28 21	ذ Thora The a bab *	63 2	ط Taawela Taa era Teen*	62 27 9	ل Laimoon Lail Lee bas*	63 23 5
ت Toofah* Toot Tem sah	74 9 3	ر Roomaaan Reeesh*	47 21	ظ The laam* Tharf	47 42	م Maawz Madrasa*	45 19
ث Thaalab Thoor Thoom*	43 21 8	ز Zahraa Zarafa Zak kaah*	74 16 1	ع Ayn Asal Aali*	84 10 4	ن Nasr* Naml	51 40
ج Jamal Jazar Jowz*	74 21 3	س Samaaka Samak Sakan*	56 34 5	غ Gazal* Ghoraab	65 30	هـ Hood hood* Herra	85 7
ح Hemar Ham mama* Hessan	44 14 6	ش Shams* Shabaka	48 22	ف Feeel Fanoos*	87 2	و Wa rdda Wadi* Wet waat	77 9 2
خ Khaa roof Khawkh Kho soof*	78 7 2	ص Sagor Soorah* Sadeeq	72 18 3	ق Galam Galb Gassi*	44 41 3	ي Yas meen* Yad Yam mama	83 10 2

Table 3.1 Initial Code Word Selection
FCW= frequency of the chosen words.
*=The words with the best recognition rate i.e. chosen words. (Refer to section 3.5)

Table 3.1 shows the 3 most frequently suggested words for each letter of the alphabet as chosen by the participants. Also included in the table are the approximate number of people who chose each of the words (unsuitable words and blank spaces were omitted from the results).

3.3 Manual Transliteration

The next task was to find the most effective spelling for each of the words that were being considered. Being Arabic words, there is no 'correct' English spelling to be used in the vocabulary for the speech recognition application. For example the Arabic word جوز can be spelt in many different ways but the spelling that sounds closer to the Arabic pronunciation is jows as judged by the author and the Arabic expert.

A variety of potential English Spellings were considered and these were then typed into a text to speech program called Free Natural Reader. This software comes with natural sounding voices and is easy to use.

Other text to speech software exist like ReadPlease, ispeech, eSpeak, and Dspeak, most of these software have the same features and work in the same way, but Free Natural Reader was found easier to use and it was recommended by some of the websites like dyslexia (2011). Most file formats can be read directly, including Microsoft Word, PPT, Outlook, PDF and images and the interface is very easy to use, it comes with a toolbar that that can be inserted in the browser, it also provides both male and female voices in a range of accents and languages that work with the software like Spanish, French, English (with US or British voices), German and Arabic, etc (Natural Reader, 2012). The Arabic version is able to work with Arabic script and comes with the voices of Youssef and Salma.

Two or more different spellings for each word were tried, where different alternatives for each letter where possible were tested. Words were then chosen on the basis of how close each spelling sounded, when read by the text to speech program, compared to Modern Standard Arabic pronunciation as judged by the author and an expert in Arabic literature from the University of Bahrain.

The best transliterations of the chosen words as chosen by the author and an Assistant Professor of Arabic Language and linguistics from the University of Bahrain and compared with the original Arabic words sounds are shown in Table 3.1.

The next phase was then to select the most appropriate word to represent each letter of the alphabet.

3.4 Selection of Words

It is crucial at this stage to check if English speech engines could be used to recognise Arabic words, hence a simple speech recognition application has been developed in Microsoft Visual Basic (VB) and uses the Microsoft Speech SDK V5.1 to create an interface to the Microsoft English (U.S.) V6.1 Recognizer speech recognition engine.

There are many different regional accents in both Britain and American. According to Qiu (2011) the most important differences between 'standard' American and 'standard' British speech are as follows:

1. Stressed vowels are often lengthened more in American English than in British.

2. Vowels are often nasalized in American English; that is to say, air comes out through the nose and mouth at the same time. Vowels are not nasalized in most British pronunciations, so this makes the two accents sound very different.

3. Most vowels are pronounced a little differently in British and American English.

   · The vowel **/ɒ/** (as in pot) is pronounced in American words without lip-rounding, and sounds like the vowel **/ɑ:/** (as in palm).
   · Many British people pronounce **/ɑ:/** (a back vowel) in some words where Americans pronounce **/æ/** (a front vowel). Examples: can´t, castle, fast, glass, class, staff, after, pass, example.

4. The most obvious difference that the average English speaker or learner might notice between the two dialects is that American English is what is known as

P a g e | **47**

The next phase was then to select the most appropriate word to represent each letter of the alphabet.

3.4 Selection of Words

It is crucial at this stage to check if English speech engines could be used to recognise Arabic words, hence a simple speech recognition application has been developed in Microsoft Visual Basic (VB) and uses the Microsoft Speech SDK V5.1 to create an interface to the Microsoft English (U.S.) V6.1 Recognizer speech recognition engine.

There are many different regional accents in both Britain and American. According to Qiu (2011) the most important differences between 'standard' American and 'standard' British speech are as follows:

1. Stressed vowels are often lengthened more in American English than in British.

2. Vowels are often nasalized in American English; that is to say, air comes out through the nose and mouth at the same time. Vowels are not nasalized in most British pronunciations, so this makes the two accents sound very different.

3. Most vowels are pronounced a little differently in British and American English.

   · The vowel **/ɒ/** (as in pot) is pronounced in American words without lip-rounding, and sounds like the vowel **/ɑ:/** (as in palm).
   · Many British people pronounce **/ɑ:/** (a back vowel) in some words where Americans pronounce **/æ/** (a front vowel). Examples: can´t, castle, fast, glass, class, staff, after, pass, example.

4. The most obvious difference that the average English speaker or learner might notice between the two dialects is that American English is what is known as

rhotic, a voicing of words that pronounces hard R's, while British English is generally non-rhotic and rolls over R's silently (Citizendium, 2012).

In standard British English, r is only pronounced before a vowel. In American English, r is pronounced in all positions in a word, and it changes the quality of a vowel that comes before it. So words like car, turn, offer sound very different in British and American speech. *Non-rhoticity*, meaning the *r* at the ends of words isn't pronounced (*mother* sounds like "muhthuh") in British English (Dialect Blog, 2012).

An Arabic word like (أرنب) is transliterated from Arabic to English as (Arnab); an American would pronounce it (a**r**nab) which is nearly how Arabic native speakers pronounce it because in standard American English, r is pronounced in all positions, whereas a British speaker would pronounce it as (Aanab) because in British English, r is only pronounced before a vowel.

American standard accent pronunciation of some letters is similar to Arabic when compared to British English. This is the main reason for choosing an American English speech engine.

Most importantly, any engine can be chosen because they can be altered by changing the transliterations (to create the lexicon) to match the chosen accent or language through choosing the closest pronunciation to Arabic, and if the chosen language does not have some of the letters of Arabic the closest letters will be chosen to match the pronunciation. For example, the letter (خ) which is pronounced like Scottish Lo(ch) can be transliterated as (kh) which is not an exact match but close enough.
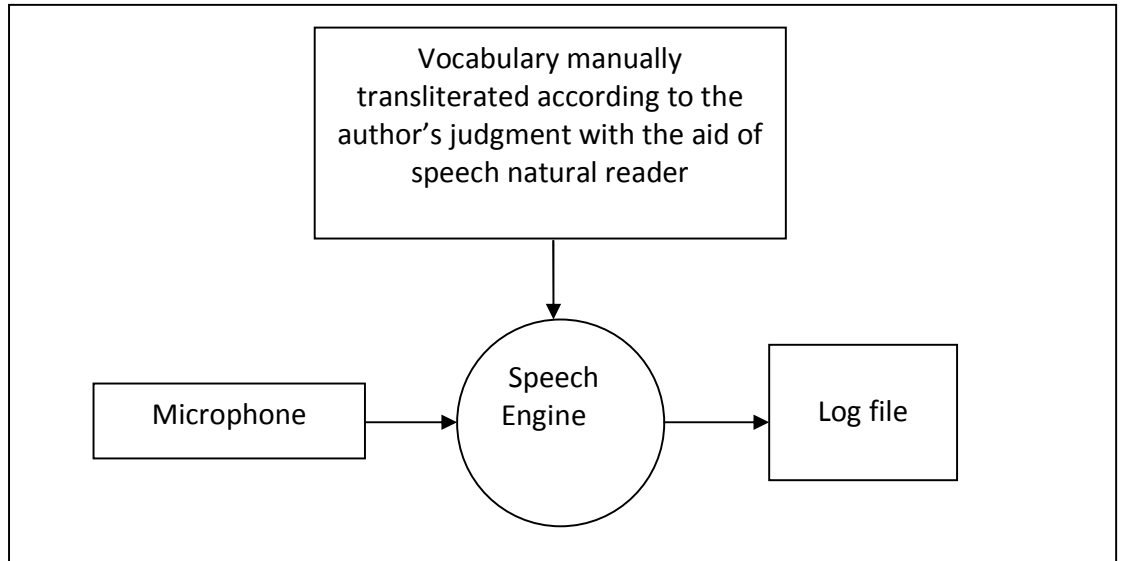
Figure 3.3 Diagram of the experimental methodology

The 72 manually transliterated words in table 3.1 are fed in to the VB application (as part of the code) to form the vocabulary (or lexicon) of the speech recognition engine and then when a user speaks in a microphone as he/she tries to read the list of Arabic diacritised vocabulary and the English transliterated versions were also presented. All readers read and spoke English. The recognized words get recorded into a log file (written file). The user was informed to wait for a nod from the examiner as he/she reads each Arabic diacritised word from the list clearly using a microphone, the English transliterated version is also available to the reader (the examiner nods after the recognised word gets displayed on screen), one after the other, and the recognised words are displayed on screen in Arabic (diacritised) and English, and the recognised words (in both diacritised Arabic and English) for each user are saved into a log file, the recognised words saved in the log file are then compared with the words the users read from the list, and recognition rates are calculated. The log file is a spread sheet that allows subsequent analysis. It consists of 3 columns, the original 28 words, the recognised words and yes/no column to check whether the two columns match. Refer to figure 3.3.

(Refer to appendix C for the application code).

The recognition rates are used to determine the most appropriate words to represent the 28 words.

For the purpose of testing the transliterated words, ten Arabic speaking students living in Nottingham (5 males and 5 females) participated in the experiment, the reason for choosing Arabic speakers is to ensure correct pronunciation for the words.

The results are shown in table 3.2 below

Word accuracy recognition or word accuracy percentage rates were defined using the formula:

$$\text{Word Accuracy} = \frac{\text{Number of words correctly recognized}}{\text{Total number of words tested}} \times 100$$

Word	Accuracy rate %	Word	Accuracy rate %
Arnab*	90	Dhifdaaa	0
Asad	50	Dha baaab	0
		Dhameer*	20
Batta	20	Taawela	20
Boostan*	90	Taa era	0
Baab	80	Teen*	90
Toofah*	80	The laam*	50
Toot	50	Tharf	10
Tem sah	60		
Thaalab	10	Ayn*	50
Thoor	30	Asal	40
Thoom*	100	Aali	90
Jamal	40	Gazal*	90
Jazar	70	Ghoraab	70
Jowz*	100		
Hemar	10	Feeel	70
Ham mama*	90	Fanoos*	90
Hessan	10		
Khaa roof	20	Galam	30
Khawkh	10	Galb	40
Kho soof*	90	Gassi*	100
Dob	50	Kalb	50
Deek*	100	Korrssay*	90
Thora	20	Laimoon	50
The a bab*	60	Lail	30
		Lee bas*	80
Roomaaan	20	Maawz	30
Reeesh*	100	Madrasa*	80
Zahraa	20	Nasr*	80
Zarafa	10	Naml	40
Zak kaah*	80		
Samaaka	50	Hood hood*	90
Samak	50	Herra	30
Sakan*	80		
Shams*	90	Wa rdda	20
Shabaka	50	Wadi*	90
		Wet waat	10
Sagor	30	Yas meen*	100
Soorah*	90	Yad	40
Sadeeq	40	Yam mama	80

Table 3.2 Recognition rates for candidate words
*=The words with the best recognition rate i.e. chosen words

3.4.1 Refining the selection

From tables 3.2 & 3.3 it can be seen that if the most popular words in each section were to be selected as the vocabulary (if words were listed by popularity and the top words chosen) then the average recognition rate would only be 46%. However, if the vocabulary were to be chosen based on the words with the best recognition rate (indicated by *, and shown in table 3.3) in each section then the average recognition rate would be 85%. The following set therefore shows the set of best recognised words which were selected for further evaluation.

أ Arnab	ح Ham mama	ز Zak kaah	ط Teen	ق Gassi	هـ Hood hood
ب Boostan	خ Kho soof	س Sakan	ظ The laam	ك Korrssay	و Wadi
ت Toofah	د Deek	ش Shams	ع Aali	ل Lee bas	ي Yas meen
ث Thoom	ذ The a bab	ص Soorah	غ Gazal	م Madrasa	
ج Jowz	ر Reeesh	ض Dhameer	ف Fanoos	ن Nasr	

Table 3.3 Set of Chosen words

3.5 Evaluation

A new list that contained only the chosen words was then tested more systematically with a range of different Arabic speakers including the 10 that were part of the first experiment. Of the thirty subjects, 16 were females and 14 males. They included a marketing specialist, 23 students (4 school students and 19 university students), two managers, and 4 teachers participated in the study.

The experiment took place in the gathering room at the University of Bahrain over 2 days, the 30 subjects were dealt with in the same room, each person was asked to come at a different time, 15 minutes between each test and the next. The first 15 subjects conducted the experiment on one day and the other 15 came next day.

The Arabic diacritised and English transliterated versions of the 28 chosen words were presented to the subjects (all users read and spoke English) and each person was asked to read each word clearly using a microphone. The recognised words were saved into a log file and recognition rates were calculated.

Letter being presented	Word representing the letter	No. Times correctly recognized/ 30	% recognition	% recognition in table 3.2	Number completely unrecognized	List of words misrecognised as
أ	Arnab	27	93	90	3	
ب	Boostan	24	82	90	6	
ت	Toofah	24	82	80	6	
ث	Thoom	29	97	100	1	
ج	Jowz	26	87	100	4	
ح	Ham mama	19	64	90	11	
خ	Kho soof	18	61	90	4	Gazal (27%)
د	Deek	21	70	100	0	Teen (30%)
ذ	The a bab	15	50	60	6	Reeesh (33%)
ر	Reeesh	21	71	100	9	
ز	Zak kaah	8	28	80	22	
س	Sakan	15	50	80	15	
ش	Shams	29	97	90	1	
ص	Soorah	27	90	90	3	
ض	Dhameer	4	14	20	26	
ط	Teen	15	53	90	15	
ظ	The laam	5	18	50	25	
ع	Aali	15	50	90	1	Gassi (5%) Wadi (45%)
غ	Gazal	20	68	90	10	
ف	Fanoos	27	93	90	3	
ق	Gassi	26	87	100	4	
ك	Korrssay	21	73	90	2	Gassi (25%)
ل	Lee bas	15	50	80	15	
م	Madrasa	19	64	80	11	
ن	Nasr	21	71	80	9	
ه	Hood hood	17	57	90	13	
و	Wadi	29	97	90	1	
ي	Yas meen	27	90	100	3	

Table 3.4 Evaluation results.

It was noticed that from time to time background noise would increase due to students gathering and talking outside the gathering room and also computers and air conditioning systems were on which added to the background noise.

The overall results gave an average recognition rate of 68%.
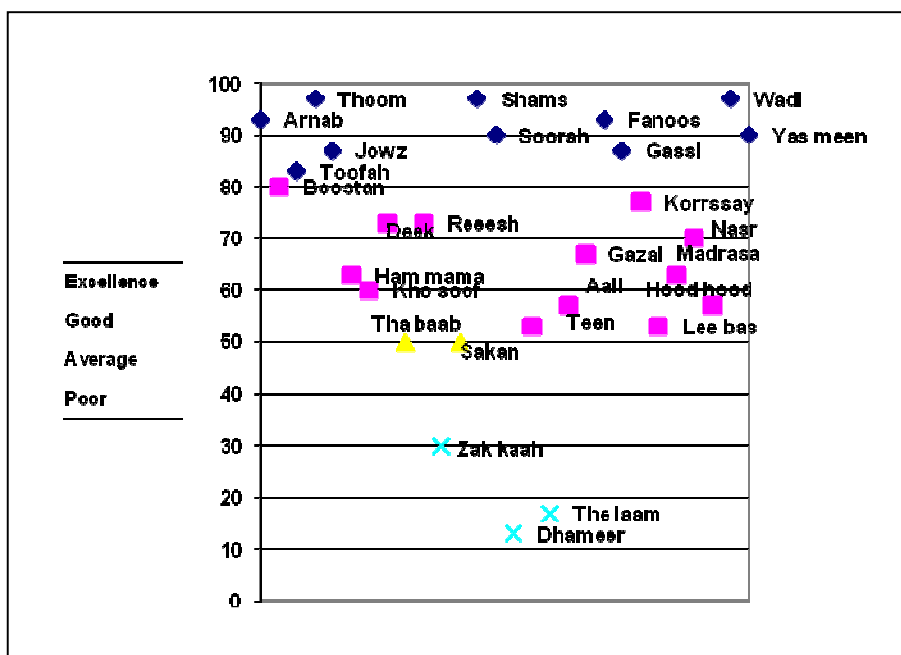
Figure 3.4 The Accuracy Rates of Words Recognition

Figure 3.4 shows that the words Dhameer, The Laam and Zak Kaah had very poor recognition rates, which was due to the application failing to make a match at all when these words were spoken. In other cases words were wrongly identified as shown in Table 3.5.

Word	Misrecognised as	
Kho soof	Gazal (27%)	
Deek	Teen (30%)	
Aaali	Gassi (5%)	Wadi (45%)
Korrssay	Gassi (25%)	
The a bab	Reeesh (33%)	

Table 3.5 Misrecognition of words

A first attempt was conducted to change some words or select different words that can be more easily distinguished by the application to obtain better recognition rates,

an Arabic/Arabic dictionary Almawrid was used to ensure that all the words used are proper Arabic words. The steps for choosing the spellings for the words followed in section 3.2 were conducted again to replace the poorly recognised words.

Some of the misrecognised words shown in table 3.5 will be replaced in the hope that the new words would be recognised more reliably; hence (Kho soof) is changed to (khoorfa kaan). The word kho soof contains the letter (س) seen which can be easily confused as the letter (ص) saad and (ز) zain, therefore the new replaced word contains easy to pronounce letters even for non-native speakers like (ر) raa, (ف) faa, (ك) kaaf and (ن) noon and the word can easily be distinguished.
(The a bab) is changed to (Thee kkraa) for the same reason as kho soof, the letters (ر) raa, and (ك) kaaf in thee kkraa are easy to pronounce and easily distinguished, and (Gassi) to (Ghaa noon). Most of the misrecognised words got recognised as Gassi which makes changing it a must, thus it was changed to ghaa noon.
The option of changing the words to the second highest recognised words in the table 3.2 was considered but by looking at the recognition rates of these words it appeared that these words might not be the best words to replace the current misrecognised words, thus new words were chosen.

The following table shows the new set of chosen words:

أ Arnab	ح Ham mama	ز Zak kaah	ط Teen	ق Ghaa noon	هـ Hood hood
ب Boostan	خ Khoorfa kaan	س Sakan	ظ The laam	ك Korrssay	و Wadi
ت Toofah	د Deek	ش Shams	ع Aaali	ل Lee bas	ي Yas meen
ث Thoom	ذ Thee kkraa	ص Soorah	غ Gazal	م Madrasa	
ج Jowz	ر Reeesh	ض Dhameer	ف Fanoos	ن Nasr	

Table 3.6 New set of Words

This vocabulary was then tested two weeks later on a range of different Arabic speakers (the subjects are a sub-set of the 30 speakers used in the main experiment) in two different environments, a quiet and a noisy environment during the experiment. Of the twenty subjects, 10 were females and 10 males.

Then the recognition rates were calculated. First the experiment was conducted in the same gathering room but it was noticed that the background noise was very high so, all experiments were conducted again in a quiet office near the library. The presentation conditions were exactly the same as in the previous experiment. The subjects were presented with a diacritised Arabic and English transliterated versions of the vocabulary and they all read and spoke English. Each person was asked to read each word clearly using a microphone. The recognised words were saved into a log file and recognition rates were calculated.
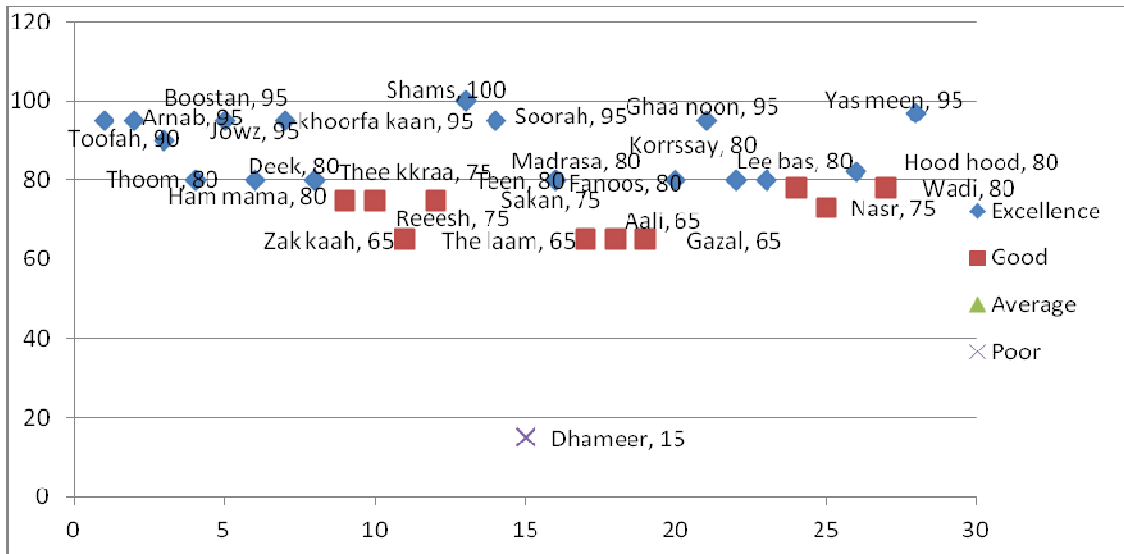
Figure 3.5 The Accuracy Rates of Words Recognition (Test 2)

Environment	Average
Noisy	84%
Quiet	85%

Table 3.7 Noisy vs. quiet environment recognition rates comparison

The accuracy results in a noisy environment gave an average recognition rate of 84% vaguely lower than the rate in a quiet environment 85%.

The overall accuracy results have improved and an average recognition rate of 85% was achieved.

Letter being presented	Word representing the letter	No. Times correctly recognized	% recognition	word used in previous attempt	% recognition in the previous attempt	Number completely unrecognized	List of words misrecognied as
أ	Arnab	19	95		93	1	
ب	Boostan	19	95		82	1	
ت	Toofah	18	90		82	2	
ث	Thoom	16	80		97	1	Lee bas (15%)
ج	Jowz	19	95		87	1	
ح	Ham mama	16	80		64	4	
خ	khoorfa kaan	19	95	Kho soof	61	1	
د	Deek	16	80		70	4	
ذ	Thee kkraa	15	75	The a bab	50	5	
ر	Reeesh	15	75		71	5	
ز	Zak kaah	13	65		28	7	
س	Sakan	15	75		50	5	
ش	Shams	20	100		97	0	
ص	Soorah	19	95		90	1	
ض	Dhameer	3	15		14	27	
ط	Teen	16	80		53	4	
ظ	The laam	13	65		18	7	
ع	Aali	13	65		50	7	
غ	Gazal	13	65		68	1	The laam (30%)
ف	Fanoos	16	80		93	2	Shams (10%)
ق	Ghaa noon	19	95	Gassi	87	1	
ك	Korrssay	16	80		73	4	
ل	Lee bas	16	80		50	4	
م	Madrasa	16	80		64	5	
ن	Nasr	15	75		71	5	
ه	Hood hood	16	80		57	4	
و	Wadi	16	80		97	0	Thoom (20%)
ي	Yas meen	19	95		90	1	

Table 3.8 Evaluation results (Test 2)

However, despite taking care to choose an appropriate set of words, and although speech is assumed to be the most natural input method, the recognition rates are still limited.

Changing and replacing some words helped to increase the recognition rates. A final attempt is conducted to change some words to obtain better recognition rates.

Following the same procedure as changing words in the first test, the misrecognised words will be replaced, so (Gazal) got changed to (Ghaanna) because the letter (ز) zain can be confused with the letters (س) seen or (ص) saad by the application, (Wadi) to (Waseela) because it ends in a unique way and hence can be easily distinguished by the application and (Thoom) to (Thamer) for the same reason as the word waseela.

The only one word that got recognised only five times in both tests above was the word Dhameer, the letter (Thad) is unique to Arabic language and even if other words were chosen the problem will remain, hence a final try to change the spelling of the word to Dhameeer, and if doesn't work another word will be picked to represent the letter (ض) dhaa.

Final set of chosen words is shown in the following table:

أ	ح	ز	ط	ق	هـ
Arnab	Ham mama	Zak kaah	Teen	Ghaa noon	Hood hood
ب	خ	س	ظ	ك	و
Boostan	Khoorfa kaan	Sakan	The laam	Korrssay	Waseela
ت	د	ش	ع	ل	ي
Toofah	Deek	Shams	Aaali	Lee bas	Yas meen
ث	ذ	ص	غ	م	
Thamer	Thee kkraa	Soorah	Ghaanna	Madrasa	
ج	ر	ض	ف	ن	
Jowz	Reeesh	Dhameeer	Fanoos	Nasr	

Table 3.9 Final set of Words

Two weeks later, this final vocabulary was then tested on a sub-set of the 30 subjects used in the main experiment. Of the 10 subjects, 5 were males and 5 females.

This time users were asked to repeat each word with a short pause between every word and the other, the test was conducted in the meeting room at the library to ensure that no background noise would affect the results.

The following table shows the number of times each word was correctly recognised and the recognition rates of each of the chosen words.

Letter being presented	Word representing the letter	No. Times correctly recognized	% recognition	word used in previous attempt	% recognition in the previous attempt
أ	Arnab	10	100		95
ب	Boostan	10	100		95
ت	Toofah	10	100		90
ث	Thamer	10	100	Thoom	80
ج	Jowz	10	100		95
ح	Ham mama	10	100		80
خ	khoorfa kaan	10	100		95
د	Deek	10	100		80
ذ	Thee kkraa	10	100		75
ر	Reeesh	10	100		75
ز	Zak kaah	10	100		65
س	Sakan	10	100		75
ش	Shams	10	100		100
ص	Soorah	10	100		95
ض	Dhameeer	10	100	Dhameer	15
ط	Teen	10	100		80
ظ	The laam	10	100		65
ع	Aali	10	100		65
غ	Ghaanna	10	100	Gazal	65
ف	Fanoos	10	100		80
ق	Ghaa noon	10	100		95
ك	Korrssay	10	100		80
ل	Lee bas	10	100		80
م	Madrasa	10	100		80
ن	Nasr	10	100		75
ه	Hood hood	10	100		80
و	Waseela	10	100	Wadi	80
ي	Yas meen	10	100		95

Table 3.10 Evaluation results 3

A final average recognition rate of 100% was achieved. Every word got recognized every time. This is a significant improvement. But it can be due to the small vocabulary used and the particular transliterations used. A larger vocabulary would definitely decrease the recognition rates.

أ	ح	ز	ط	ق	هـ
Arnab	Ham mama	Zak kaah	Teen	Gassi <mark>Ghaa noon</mark>	Hood hood
ب	خ	س	ظ	ك	و
Boostan	Kho soof <mark>khoorfa kaan</mark>	Sakan	The laam	Korrssay	Wadi <mark>Waseela</mark>
ت	د	ش	ع	ل	ي
Toofah	Deek	Shams	Aali	Lee bas	Yas meen
ث	ذ	ص	غ	م	
Thoom <mark>Thamer</mark>	The a bab <mark>Thee kkraa</mark>	Soorah	Gazal <mark>Ghaanna</mark>	Madrasa	
ج	ر	ض	ف	ن	
Jowz	Reeesh	Dhameer <mark>Dhameeer</mark>	Fanoos	Nasr	

Table 3.11 The words that got changed to create the new table

Highlighted in <mark>yellow</mark> are the words that got changed in the first test.

Highlighted in <mark>green</mark> are the words that got changed in the second test.

3.6 Discussion & Conclusion

The results and the feedback show that it is possible to create an application for the purpose of recognizing Arabic words using a Standard English speech recognition engine.

Achieving 100% recognition rate is possible for a limited vocabulary generated manually, and, although this is entirely appropriate for a single application of 28 words, this could be very limiting for larger and more volatile vocabularies.

There are two factors that could affect the recognition rates: the transliteration of words and the choice of words i.e. choosing a set of words that are different enough from each other makes the speech recognizer's job easier to tell the words apart, but choosing easily distinguishable words can also limit the use of the application to

evaluate transliterations, this is more of an issue if the vocabulary has limited number of words like the 28 previously tested words.

The fact that the recognition results changed as the words were transliterated differently for example the word (dhameeer) transliterated and used in the lexicon of the speech engine, the engine must find a match to the sound that is internally created for this word with the spoken word (ضمير) pronounced in Arabic. If other transliterations for the same word were found, that are very similar in pronunciation but have only slightly difference, each word will be part of the lexicon and also phonetic representations and sounds will be created internally by the engine for every saved version of these words, and when a word is spoken the speech engine selects the closest matching spelling from the list of words that match the speakers pronunciation for this word. The recognized word would be displayed in Arabic and the transliterated version will also be shown in order to distinguish the accuracy rates for the different transliterations for the same word.

There are differences in the recognition rates for the same word transliterated differently. For a word to be recognized, the spoken input is matched to the phonetic representations of the words in the lexicon of the speech recognizer. If the internal representations are good, they will give good matches, so that recognition rate can be an indicator of how good a transliteration is. However, recognition rate is also influenced by other factors (e.g., vocabulary size and content; matching method), and it is not a direct evaluation of the quality of transliteration. The recognition rate for the recognised transliterated word matched with the original spoken Arabic word offers a way to measure the accuracy of transliterations automatically, with some degree of consistency and repeatability. One way of evaluating transliterations is through human expert judges (see chapter 7); but there would be some variability even in this case, since, because there is not a simple one-to-one mapping between Arabic and English orthographies, transliterations will, to some degree, be a matter of opinion.

A short experiment was done to examine which transliteration of the word ضمير (dhameer; transliterated in 3 different ways: dhameer, dhameeer and dhamir) resulted in the better recognition performance. The three versions are pronounced similarly as spoken by the text-to-speech application and judged by the author, there was only slight and not easy to hear difference in the pronunciation. The phonetic representation for these words produced by the ASR will be slightly different. The three words formed the vocabulary for the speech recognizer and the same test was done to measure the accuracy rates by 10 users. The results were as follows:

Word	Recognition rate average %
dameer	30
dhameeer	70
dhamir	0

Table 3.12 The effect of transliterations on recognition rates.

Note that (dhameeer) wasn't considered as one of the transliterations when the best transliteration for each of the 28 words were found by the author and the expert, as adding three Es to a word is not acceptable in English text. This transliteration was proposed at a later stage of the experiments, in an attempt to increase the recognition rates.

Different transliterations for the same word can affect the recognition rates. The word (ضمير) as an example was transliterated and approved by an expert as (dhameer), but the recognition rate for this word was very low which is normal and expected as the letter (ض) is unique to Arabic and finding an exact equivalent is impossible. But it was worth investigating whether an improvement could be made and if finding other transliterations for the word would improve the recognition. For the word (ضمير) changing the transliteration gave better results, and given that this Arabic word contains a letter unique to Arabic, this may be a better transliteration.

The results from table 3.12 suggest that high recognition rates don't always mean that the transliterations used are the best. The transliteration for the word (dhameer) as chosen by both the author and the expert achieved lower recognition rate than (dhameeer) which achieved the highest recognition rates.

Transliteration of Arabic words can often be a matter of judgment, and recognition rate is not a perfect method of judgment of the transliteration since other factors can influence the recognition, such as accent or the way people pronounce different letters, if the way they speak matches the phonetic representation or the sound made internally by the speech engine for this word or letters it will be likely to be recognized correctly.
Also background noise can affect the results slightly, since the recognition results increased when the tests were conducted in a quiet environment.

Using the ASR to evaluate transliterations has the advantage of providing tests that are repeatable and the recognition results will always be consistent, if recorded voices were used. In comparison with if humans were to judge the evaluation, different results might be obtained every time even if the voices used were recorded.

In addition, although some care has been taken to get a range of Arabic speakers, they mainly came from or lived in Bahrain. It is likely that the recognition rates would be lower for the full Arabic speaking population. To overcome these difficulties it was necessary to improve the application by testing it on wider range of people and/or using a bigger vocabulary.

Another problem that occurred while conducting the tests is that it was very difficult to find the same people for each experiment and also external factors can corrupt the results like background noise, which meant that it was necessary to change to recorded voices, to allow the creation of exact replica to ensure repeatability and consistency.

Future work will therefore concentrate on looking at automatic methods of transliteration because transliteration is complex, critical and time consuming, as it was established that in order to get good results it was necessary to select the right transliterations of the words that would form the vocabulary. Also the next stage will look into investigating whether it is possible to computerize the process in order to save time and effort and to examine the transliteration process with the objective of creating a transliteration program that could be used in conjunction with an English speech recognition engine in order to evaluate transliteration tables.

CHAPTER 4
Automatically transliterating and generating words from diacritised Arabic

The design of systematic evaluation of recognition consists of three different stages represented by three separate applications: the transliteration application (refer to section 4.2), the speech recognition application and the analysis process application. Figure 4.1 shows the design of the Evaluation process.
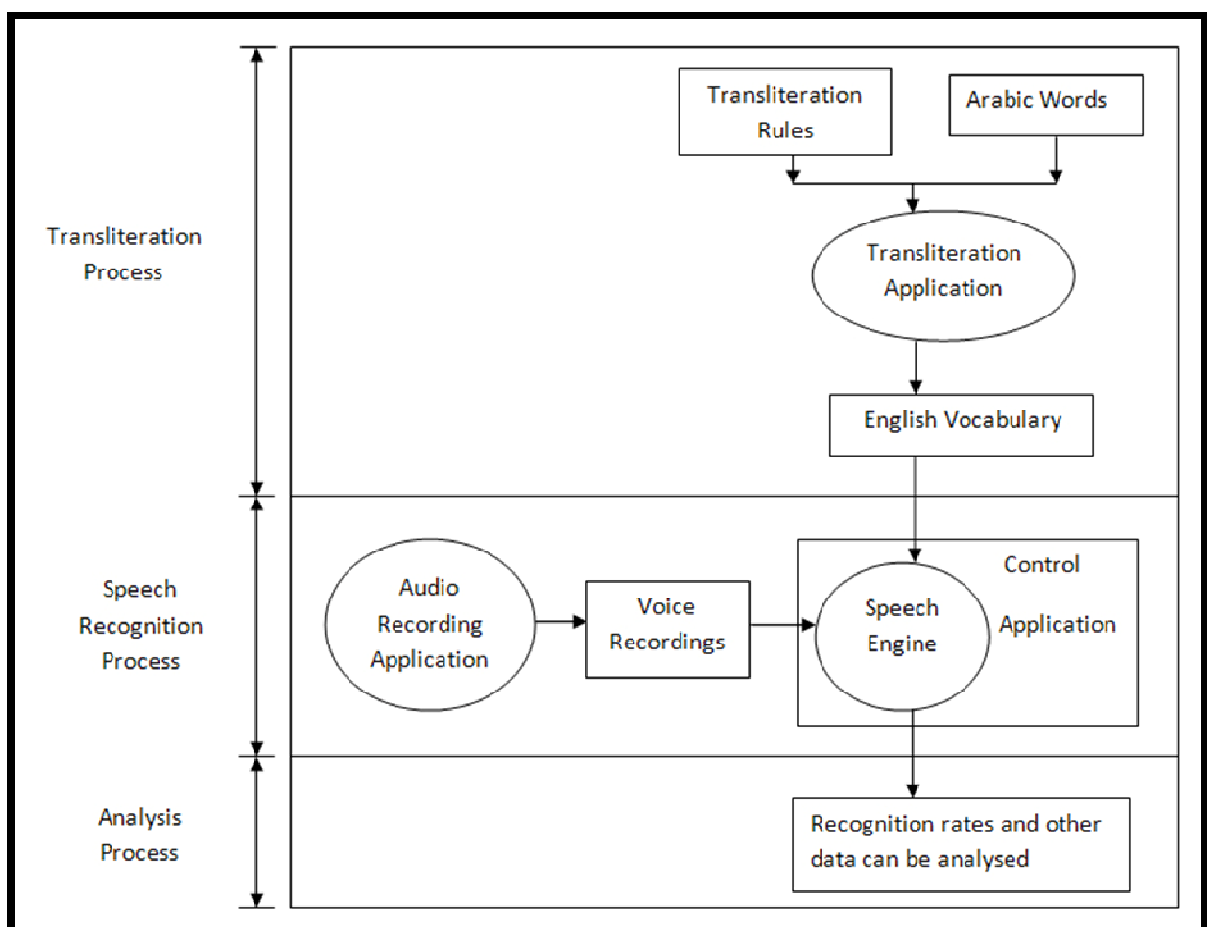
Figure 4.1: The design of evaluating transliteration tables' process

The transliteration stage (see figure 4.1) prepares the list of words for the speech recognition process by transliterating the Arabic words into their English equivalent. After that, the speech recognition engine with the aid of the pre-recorded audio files tries to recognise the words and then sends them to an external file for the last stage

which is analysing the recognition rates and other data like what words got misrecognised as others.

4.1 The use of voice recordings instead of live voices

As discussed in the previous chapter, although manual transliteration and live voices can be successful for a limited vocabulary, the results could be very inconsistent and finding the same people to repeat the experiments would require a lot of time and effort. Thus it is essential to consider the use of recordings to ensure consistency and repeatability of experiments rather than using live voices.

The first problem that was looked into was recording the 28 words shown in table 3.9. Ten Arabic speakers were presented with the 28 Arabic manually diacritised and English transliterated words. Window's sound recorder was used to record one word at a time and store the files separately. The recording of each word is stored in a separate file and is called the (word's name) for example the recording of the word Arnab is called "Arnab", etc. The files are saved as wav files (PCM 44, 100 KHz, and 16 bit sample rate, stereo); with an average file length of 2 seconds. The length of wav files range from 0.8 seconds to 2.4 seconds, Additionally, a 0.2 second silence period is added to the beginning and end of each file. The files are recorded in a quiet environment (air-conditioning systems and computers were off) the sets of recordings are kept in different folders for different speakers each titled with the speaker's first names. The voices of the same 10 subjects used in the main experiment in 3.4 were recorded reading the 28 chosen words.

The quality of the recording was checked by listening to the recording to ensure that they were clear and sounded right. If the quality wasn't good enough, the subjects were asked to repeat the recordings.
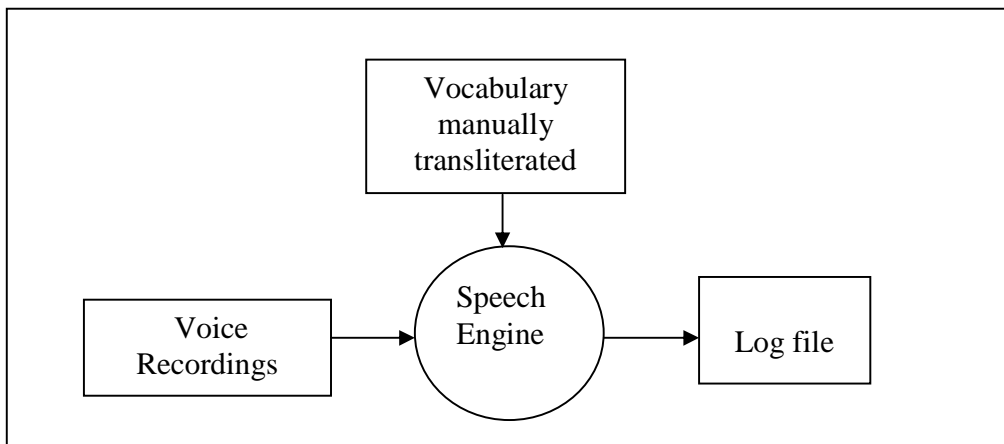
Figure 4.2 Diagram of the use of voice recordings to aid transliteration experiment methodology

The recordings of the list of words shown in table 3.9 were played (the sound data fed directly into the program) and recognition results calculated. The control application creates an interface to the speech engine and also controls feeding the recorded files into the speech engine and logging the results. Each word was played one after the other, the recognised words were displayed on screen in Arabic and English and the results for each user saved into a log file. Then the recognised words saved in the log file were compared with the words the testers read from the list, and recognition rates were calculated.

The results are shown in table 4.1 below

Letter being presented	Word representing the letter	No. Times correctly recognized/10	% recognition
أ	Arnab	10	100
ب	Boostan	10	100
ت	Toofah	10	100
ث	Thamer	9	90
ج	Jowz	10	100
ح	Ham mama	10	100
خ	khoorfa kaan	10	100
د	Deek	10	100
ذ	Thee kkraa	10	100
ر	Reeesh	10	100
ز	Zak kaah	10	100
س	Sakan	10	100
ش	Shams	10	100
ص	Soorah	10	100
ض	Dhameeer	9	90
ط	Teen	10	100
ظ	The laam	10	100
ع	Aali	10	100
غ	Ghaanna	9	90
ف	Fanoos	10	100
ق	Ghaa noon	10	100
ك	Korrssay	10	100
ل	Lee bas	9	90
م	Madrasa	10	100
ن	Nasr	10	100
ه	Hood hood	10	100
و	Waseela	10	100
ي	Yas meen	10	100

Table 4.1 Using recorded voices evaluation results

The experiment was conducted twice to ensure consistency. The results were exactly the same each time.

The overall results gave an average recognition rate of 98.6%. This demonstrates that it is possible to use recorded voices instead of using live voices to ensure consistency, repeatability and to save time and effort.

4.2 Automatic transliteration of diacritised Words

Arabic letters are pronounced differently when diacritised.

There are 6 vowels in Arabic, 3 short and 3 long and there are 2 semi-vowels.

Long vowels written in the middle of a word of unvocalized text are treated like consonants with a *sukūn*. Arabic short vowels are written with diacritics placed above or below the consonant that precedes them, the diacritics are shown in table 4.2.

Short vowels /a/, /u/, /i/		Double case ending		Syllabi-fication marks	
Fatha	◌́	fathatayn	◌		
damma	◌	dammatayn	◌	shadda	◌
kasra	◌	kasratayn	◌	sukuun	◌

Table 4.2 Arabic Diacritics

It is essential to create an application that would allow applying diacritics; hence the next phase is diacritising and testing the words manually then checking whether diacritising affects the results.

So the next phase was then to design and implement an application that would generate an English vocabulary by transliterating each Arabic diacritised word into its English equivalent.

The letters were transliterated according to the United Nations Educational, Scientific and cultural Organization Transliteration table (UNESCO, 2006c). The UNESCO table offered transliterations for the 28 Arabic letters only, but didn't mention the diacritics that's why the diacritics were transliterated to their equivalent according to (IPA) the International Phonetic Alphabet (Alghamdi, 2003).

Arabic letter	UNESCO Transliteration	Arabic letter	UNESCO Transliteration
أ	A	ظ	Z
ب	B	ع	'
ت	T	غ	Gh
ث	Th	ف	F
ج	J	ق	Q
ح	H	ك	K
خ	Kh	ل	L
د	D	م	M
ذ	Dh	ن	N
ر	R	هـ	H
ز	Z	و	W
س	S	ي	Y
ش	Sh		
ص	S		
ض	D		
ط	T		

Table 4.3 United Nations Educational, Scientific and cultural Organization, Transliteration table

IPA	Diacritic
U	ُ
I	ِ
A	َ
xx	ّ
x	ْ
An	ً
Un	ٌ
In	ٍ

Table 4.4 IPA diacritics transliteration table
*xx =duplicated letter, e.g. لّ = LL

The short vowels or diacritics have associated sounds and when placed above or below the consonant that precedes them they add to the vowel sound. For example the Arabic word for "school" without short vowels is مدرسة. If the letters were transliterated we would end up with <mdrsa> – the two short "a"-sounds are merely

implied. The reader needs to know already that the word means "school" and that school is *madrasa* and not something else like *mudarisa* or *midarusa.*

The word (مَدْرَسة) there is a small sign above the first letter from the right م (*meem*) that indicates a short <a>-sound. This little stroke is called "fatHa". The next letter د (*daal*) bears a symbol indicating that there is no short vowel at this point. This small circle is called "sukoon". There is no sign above the fourth letter س (*seen*), because it is followed by a ة (*taa' marbooTa*) that is pronounced as an "a"sound in any case.
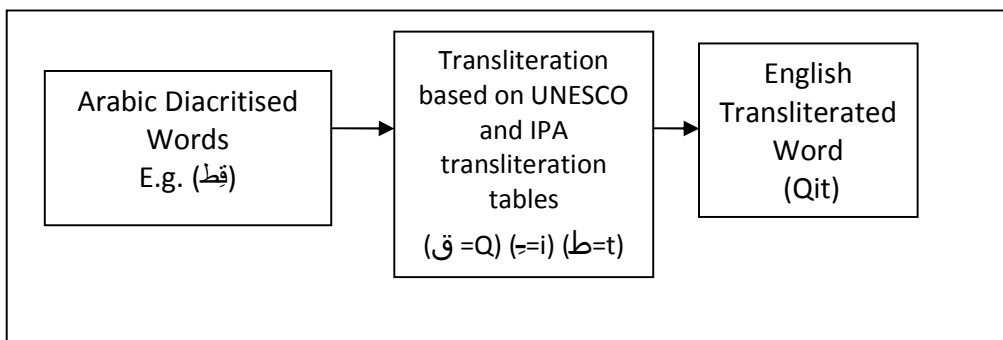
Figure 4.3 Diagram of automatic transliteration methodology

4.3 Testing of words transliterated automatically.

The previous 28 set of words was diacritised and transliterated using the UNESCO and IPA transliteration tables. The application was tested using the same recordings used in section 4.1. The use of the same recordings provides the opportunity to check whether the changes in recognition are due to transliterations or changes in voices.
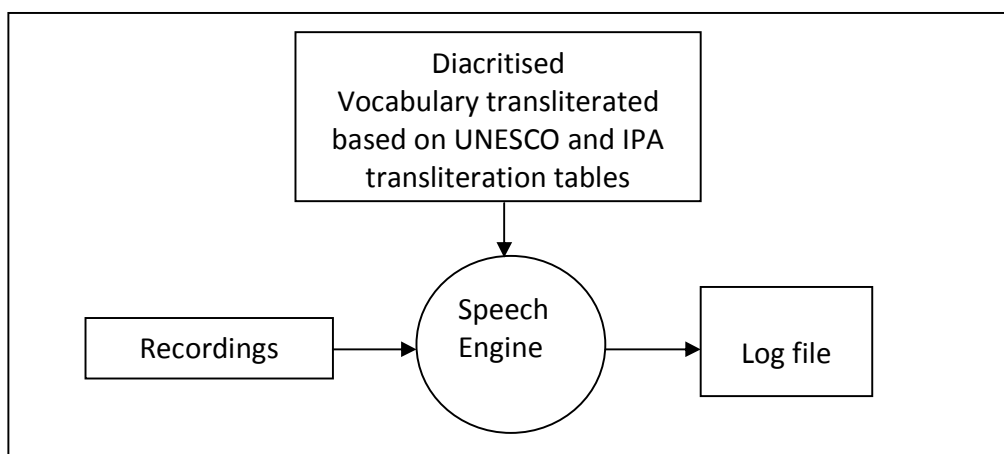
Figure 4.4 Testing methodology using recordings and automatically transliterated vocabulary

This experiment is similar to the previous one. The same set of 28 words was used, but in this case they were diacritised and then transliterated automatically using the UNESCO and IPA transliteration tables. An application was developed to automatically transliterate the 28 diacritised (diacritised by the author) words to form the vocabulary for the speech recognition engine. The application provided a form that contains a text box and a button, the Arabic diacritised word should be entered in the text box field and when the button is clicked the English transliterated version is displayed and it could be used to form the lexicon.

The control application helped to control testing recordings of the same ten Arabic speakers from Bahrain used in the previous experiment (refer to figure 4.2).

The 10 recordings were played one after the other with the control application managing the process and the recognition rates were saved into the log file and calculated. The purpose for automating the transliterations and the use of recordings instead of live voices is to speed up the process and ensure repeatability and consistency.

Arabic Word	English word obtained from manual transliteration	% correct recognition (from table 4.1)	Words misrecognised as	English word obtained from automatic transliteration	% correct recognition	Words misrecognised as
أرنب	Arnab	100	-	arNab	100	
بستان	Boostan	100	-	bustaN	80	tiYN, YasmiYN
تفاح	Toofah	100	-	tuFah	70	Thee kkraa, MadrasaH
ثامر	Thamer	90	-	thaMir	60	Nasr
جوز	Jowz	100	-	jaWz	90	
حمامه	Ham mama	100	-	haMaMaH	70	WasiyLa
خورفكان	Khoorfa kaan	100	-	khurFaKaN	50	bustaN
ديك	Deek	100	-	diYK	40	riYsh
ذكرى	Thee kkraa	100	-	dhiKra	20	thMir
ريش	Reeesh	100	-	riYsh	60	Shams
زكاة	Zak kaah	100	-	zakaH	50	suWraH
سكن	Sakan	100	-	saKaN	100	
شمس	Shams	100	-	Shams	100	
صورة	Soorah	100	-	suWraH	10	tuFah
ضمير	Dhameer	90	-	daMiYr	0	
طين	Teen	100	-	tiYN	20	saKaN
ظلام	The laam	100	-	ZaLaM	0	
عالي	Aaali	100	-	'aLY	0	
غانا	Ghaanna	90	Lee bas	GhaNa	10	dhiKra
فانوس	Fanoos	100	-	FaNuWs	20	Libas, Shams
قانون	Ghaa noon	100	-	QaNuwN	10	saKaN
كرسي	Korrssay	100	-	KursY	20	khurFaKaN
لباس	Lee bas	90	Aali	Libas	50	ZaLaM
مدرسة	Madrasa	100	-	MadrasaH	60	suWraH
نسر	Nasr	100	-	Nasr	90	Shams
هدهد	Hood hood	100	-	HudHud	30	saKaN
وسيلة	Waseela	100	-	WasiyLa	20	dhiKra,
ياسمين	Yas meen	100	-	YasmiYN	70	QaNuwN

Table 4.5 Evaluation results of testing of the 28 words generated automatically using recorded voices.

The accuracy results of testing the 28 words generated automatically using recorded voices gave an average recognition rate of 46.4% noticeably lower than the results of testing the same words generated manually using the same recorded voices, where an accuracy of 98.6% was achieved. Table 4.5 compares the recognition of the 28 words using the same 10 recorded voices. The only difference is that the 28 words were transliterated differently in both experiments, automatically using the UNESCO and IPA transliteration tables and manually. Words that are the same when transliterated using both methods (automatically and manually) like سكن sakan and شمس shams scored the same recognition rate of 100%. Whereas words like فانوس,

which when transliterated manually as *fanoos* scored a recognition rate of 100% and when transliterated automatically as *FaNuWs* the recognition rate was only a disappointing 20%. The reason for the lower recognition rate achieved using the automatically transliterated words is because of the different transliterations produced. For example, when transliterated automatically the letter (و) in the words فانوس is always transliterated as (W) and the (ٌ) diacritic is always transliterated as (U), whereas in manual transliteration the author chooses the appropriate transliteration for each letter, hence the letter (و) is transliterated as (O) in فانوس, but transliterated as (W) in وردة.

4.4 Discussion & Conclusion

Manual transliteration can be very successful for a limited vocabulary; as the number of words increases the need to consider automating the process increases. The use of automatic transliteration to generate a good set of words without manual adjustment also proved to be possible but the recognition rates were not very good compared to the results from manual transliteration which are almost perfect, also the transliterations are not as perfect as manual transliteration and a lot of unreal words are included. When a word is transliterated manually only one equivalent is generated whereas using the transliteration application to transliterate an undiacritised word would generate more than one possibility which increases the vocabulary and hence places more load on the speech engine. According to (AbdulJaleel and Larkey, 2003) even though some transliteration systems are provided with online translation, little is published about them and there is no information concerning how effective they are or how they generate transliterations.

AbdulJaleel and Larkey (2003) established a statistical transliteration technique which uses English to Arabic transliteration model from pairs of names. They used a selected n-gram model which is a two stage training procedure and managed to get accuracy rates of around 50-80%.

These rates are consistent with the 46.4% accuracy rates achieved in section 4.3 using the automatic transliteration with the aid of transliteration rules which involved using words in a list, which makes them more difficult to transliterate than words in context.

Although introducing automatic transliteration can save time and effort, the recognition rates depended heavily on the transliteration rules because of the inaccurate transliterations produced, which did not correspond to the input. If automatic transliteration is to be used in the coming experiments more attention should be paid to the transliteration rules.

The use of voice recordings instead of live voices in both methods ensured consistency and repeatability of the experiments and saved time and effort. Hence it is ideal to use voice recordings in the coming experiments.

The results and the feedback confirm that using an English engine to recognise an Arabic word is very sensitive to the transliteration and hence the efficiency of the recognition can be measured to assess the quality of transliteration. Nevertheless this doesn't exactly give a direct evaluation of the quality of transliteration. The quality of transliteration can be established by recognizing the correct word from the list.
The transliterated word's recognition rate compared with the original word provides an effective way to measure the accuracy of transliteration schemes.

The subsequent chapters will focus on examining the efficiency of this as a measurement.

CHAPTER 5

Transliteration of undiacritised words

As mentioned in chapter 2, almost all of modern Arabic text is written without diacritics. Readers of Arabic normally use the context to work out which of the possible words it actually is. However, if the word was in a list, there is no context, but there is the possibility to generate all the possible words for the undiacritised word. This is considered next.

5.1 Automatic transliteration of undiacritised Words

Diacritising manually is time and effort consuming thus it is essential to improve it by developing a transliteration application that can apply diacritics automatically.

One approach to dealing with the diacritisation problem is for the transliteration program to start by generating all possible diacritised versions of an undiacritised Arabic word and constructing an English vocabulary incorporating all of these automatically. So if the 28 words were to be recognized, the 28 undiacritised words would be diacritised and transliterated, by generating all the possible diacritised versions of each word of the 28 words list. Each one of these possibilities then links back to the original Arabic word, so that if any of them is recognised by the speech engine, the originating word is identified (refer to figure 5.1)
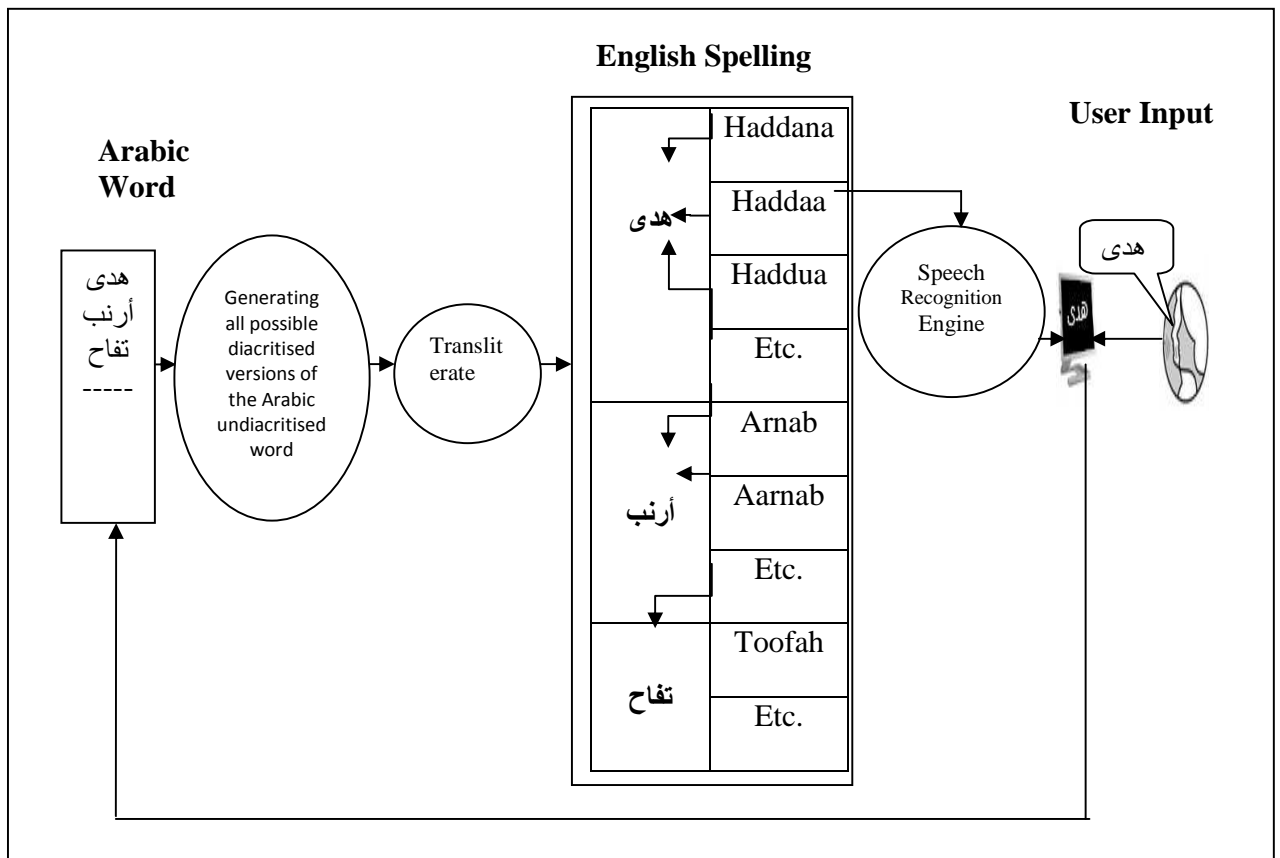
Figure 5.1 Diagram of the process of diacritising and transliterating Arabic undiacritised words and using speech recognition engine to test the accuracy of the transliterations.

When an Arabic diacritised word is transliterated and the transliteration is used in the lexicon of the English speech engine, the engine tries to match the sound that is internally created for this word with the spoken word pronounced in Arabic.

If the sounds match, this could indicate that the transliteration is good, and vice versa poor transliteration could lead to misrecognition of sounds and words.

The engine chooses the closest matching word from a list of words to match the spoken word. If the word is not diacritised, then an application that would apply diacritics to all words and generate all possible transliterations should be implemented and used. The generated diacritised transliterations for a word are then used in the lexicon of the English speech engine, and the engine compares the spoken word with the sound created internally from the text of words stored in the lexicon.

This is not a direct evaluation of the match between written and spoken words. By selecting the correct word from the list, the accuracy of the similarity of the spoken and written word is verified.

A word that consists of 3 letters like (هدى) Huda, can have up to 37 different possibilities when diacritised.

Transliteration	Arabic Diacritised word	Transliteration	Arabic Diacritised Word
Haddana	هَدَّنى	hadana	هَدَّنى
Haddaa	هَدَّى	hadaa	هَدَى
Haddua	هَدَّى	hadua	هَدُى
Haddia	هَدِّى	hadia	هَدِى
huddana	هُدَّنى	huda	هُدى
huddaa	هُدَّى	hudana	هُدَّنى
huddua	هُدُّى	hudaa	هُدَى
huddia	هُدِّى	hudua	هُدُى
hiddana	هِدَّنى	hudia	هُدِى
hiddaa	هِدَّى	hida	هِدى
hiddua	هِدُّى	hidana	هِدَّنى
hiddia	هِدِّى	hidaa	هِدَى
hddana	هدَّنى	hidua	هِدُى
hddaa	هدَّى	hidia	هِدِى
hddua	هدُّى	hdana	هدَّنى
hddia	هدِّى	hdaa	هدَى
hda	هدى	hdua	هدُى
hada	هَدى	hdia	هدِى

Table 5.1: The possibilities of adding diacritics to the word Huda

The word Huda consists of the letters (هـ), (د), and (ى), and by setting the rules so that all diacritics should be applied to every letter in every position, for example the letter (هـ) at the beginning of the word Huda, will have all the diacritics added to it (هَـ), (هِـ),

(ﻫِ), (ﻫَّ), (ﻫَ), etc and the same applies to the other two letters. Only 5 of the 37 possibilities have meaning, whereas most of the possibilities have no meaning.

This process can be successful if the word consists of few letters, and if the word has vowels as the generated possibilities are fewer than words that contain consonants only. If the word consists of 4 or more letters and contains consonants and no vowels, this can be challenging as the number of possibilities rises and some of the possibilities can even be unrealistic and can't be pronounced, which can add load to the transliteration application.

Building a sensible vocabulary by generating all the possibilities for each word can be problematic, as duplication of words might occur (i.e. the possibility of generating the same English words if the words were spelled the same in Arabic when not diacritised for example the word (وَرَق) meaning paper and the word (وَرِق) meaning silver, both of these words would be spelled as (ورق) when not diacritised and in a list, and the generated possibilities would be exactly the same. This might be confusing and can lead to recognizing the wrong word.

Also the generated number of possibilities might be large which means more work for the speech recognition engine to look for the right word from all these possibilities as it would be difficult to filter these words if the number of possibilities was large. This is not a problem for the current 28 chosen words, as every word is unique and the vocabulary doesn't contain two Arabic words spelled the same when not diacritised.

Calculating the number of possibilities for each word can vary. It depends on the number of letters in each word, and if it contains vowels or just consonants.

In order to generate an application that would diacritise and generate all the possibilities of a word, the following application has been developed in Microsoft Visual Basic and uses SQL queries and statements.

The application consists of 3 forms. The first form allows the user to enter a diacritised Arabic word and clicks a button to get the transliterated version of the word displayed. The diacritised words are entered and transliterated automatically

using the UNESCO and IPA transliteration tables. These automatically transliterated diacritised (diacritised by the author) words form the vocabulary for the speech recognition engine. A text to speech facility has been added to this form to aid with the pronunciation of the transliterated words

The second form permits the user to enter an undiacritised word and with a click of a button it displays all the diacritised transliterated possibilities and saves them to an external file. With the aid of a table that has a listing of all the Arabic letters and the diacritics in the 3 positions start, middle and end and states the rules if this letter could accompany a specific diacritic in a specific position. This table allows the user to set some rules for diacritisation. Also using the UNESCO and IPA transliteration tables to transliterate the generated possibilities.

A text to speech facility has been added to this form to aid with the pronunciation of the transliterated diacritised words. Refer to appendix D, for more details about the forms and the code. The generated list contains lots of unreal words. These words could be pointed out and removed with human interference and with providing a vocabulary of acceptable words to check them against.

The third form contains a button that transliterates a list of diacritised names or words in a text file, and generates an xml file, so that it can be used for the speech recognition process. The generated xml file works with the VB code in the speech recognition application. Refer to appendix D, for more details about the third form.
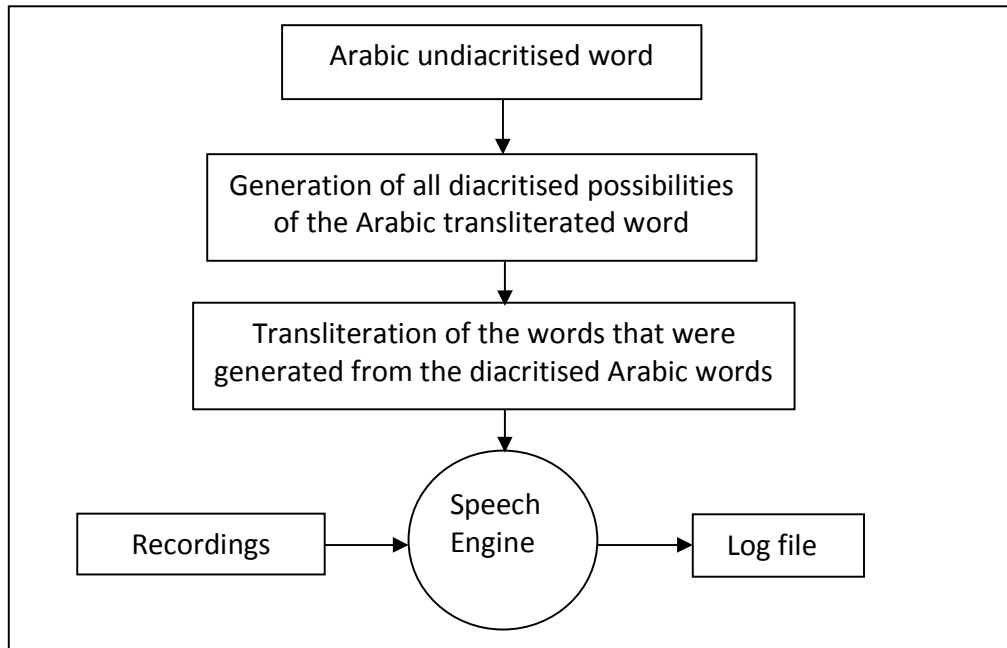
Figure 5.2 Diagram of automatically generating all the transliterated diacritised possibilities of the undiacritised Arabic word experiment methodology.

The development was undertaken in the following steps:

1. A 3 letter diacritisation and transliteration application was developed.

2. The 3 letter diacritisation and transliteration application was upgraded to process 4 letter words.

3. The application was then tested on words spoken by a sample of Arabic speaking population.

4. A 5 letter diacritisation and transliteration application was then developed.

The diacritisation and transliteration application takes the Arabic undiacritised word and generates all diacritised possibilities for this Arabic transliterated word by checking each letter in the word and its position with Table 1 in appendix E, as it has a list of all the letters in every position (beginning, middle and end) and a list of which diacritics it can accommodate at this specific position.

Then all the diacritised words that were generated and transliterated according to the United Nations Educational, Scientific and Cultural Organization Transliteration table (UNESCO, 2006c), and the diacritics to their equivalent according to The International

Phonetic Alphabet (IPA, Alghamdi, 2003) and becomes the vocabulary of the speech recognition engine including the non-words. When the voice recordings are played, the speech recognition engine searches for the closest match and displays the recognised word, this word gets saved in a log file for further analysis.

5.1.1 Three letter words

An initial application was created in visual basic that processed up to 3 letter Arabic words. It generated a vocabulary of English transliterations based on all possible diacritised versions of the original word and then transliterated them according to published rules. Researchers from different academic and research institutions were invited to participate in building a system that would be able to diacritize Arabic text automatically.

The team investigated different approaches for diacritizing Arabic Automatic Diacritizer of Arabic Text Using Hidden Markov Model, Automatic Diacritizer of Arabic Text Using Viterbi and an Independent Diacritizer of Arabic. The Independent System is considered an achievement for several reasons. It is independent and its performance is higher and faster than the other systems (Alghamdi et al, 2006). Hence the diacritical rules used in this system were also used in the 3 letter application.

The 3 letter application was successful, diacritics were applied to words consisting of 3 letters automatically and then transliterations were produced for these diacritised words and words were recognised even though an individual word generated as many as 482 diacritised versions of words. Unreal words are also included as it is difficult to filter these out without human interference.

A recognition test for a subset of the 28 chosen words, words that consist of 3 or less letters were used in the test. All diacritised possibilities for each of these words were transliterated and included as vocabulary (xml file) for the speech recognition engine. After that the recordings of these words were played and results calculated. Each

word was played one after the other, and the recognised words got displayed on screen and the results for each user were saved into the log file (written file).

The overall results gave a recognition rate of 100%. 2 of the 7 words (سكن, شمس) generated 482 diacritised possibilities, and the other 5 (جوز, ديك, ريش, طين, نسر) produced 478.

Three letter words consisting of only consonants generate more possibilities as the rules for adding diacritics to all the letters in all positions (first, middle, last) states that it is possible to add most diacritics to all letters in any position. But when it comes to words that contain vowels, vowels can have specific diacritics in specific position which limit the number of possibilities.

5.1.2 Longer words

The 28 chosen words and Arabic words in general contain words of more than 3 letters hence upgrading the application to diacritise and transliterate words of 4 letters and less is essential. The application works by looking at the letter in the middle and then checks the letters before and after before applying the diacritics and then transliterating the word, so the code was upgraded to work with longer words (more than 3 letters).

Unfortunately a typical 4 letter word would produce about 24000 possibilities, and most of the possibilities would be unrealistic and diacritised wrongly, so the application was then further developed to filter out meaningless words by adding more diacritical rules, (UIUC linguistics, 2007), (Algamdi and Zeeshan, 2007), and (Alghamdi et al., 2006a).

If we assume that a four letter word is represented as, $Z^d \, Y^c \, X^b$, W^a where W is the first letter, X the second letter, Y the third and Z the fourth letter. A, b, c and d are the diacritics following each of the four letters.

Some of the rules include the following:

1. The first diacritic a can only be represented as ◌�
, ◌ُ and ◌ِ

2. b can be represented by all diacritics but ◌ٓ , ◌ُ and ◌ِ

3. c can be represented by all diacritics but ◌ٓ , ◌ُ and ◌ِ

4. d the fourth diacritic can be presented by all diacritics.

5. If one of the three vowel diacritics is doubled, it may only appear at the end of a word.

6. Alif maqsourah is always undiacritised and the madda can only appear on top of an Alif

Refer to Appendix E for a table containing the rules applicable for each letter and its position.

This reduced the number of diacritised version of a word tremendously. For example, the word "نوال" Nawal, consists of 4 letters, and before applying the rules to the transliteration application, it had about 24000 possibilities; however after applying the rules, there were only 200. Most of the possibilities have no meaning (non-words) but filtering them out is difficult without human interference and it is also time consuming (Appendix F shows all the diacritised possibilities of the word Nawal after applying the rules). Table 1 in appendix E presents all the letters in all position (beginning, middle and end) and a list of all the diacritics. So if a letter in a specific position can accommodate a specific diacritic, a tick (✓) is shown in the cell. This table formed the basis to the (tblDiacritics) table, which is the table used by the application to apply diacritics to words automatically.

Refer to (appendix D) for the application code and process diagrams.

5.1.3 Limitation

Unfortunately, even applying these rules, resulted in an unacceptable number of possibilities for 5 letter words. A 5 letter word took about 40 minutes to be processed and produced about 10000 filtered possibilities. In some cases the word could be recognized because these possibilities are linked to the original word, so if

any transliteration possibility is recognized the original word is displayed. In many cases errors occurred.

When diacritised using the transliteration application, each word that represents the letters of the Arabic alphabet used in the previous experiment, had different number of possibilities, this depends on the number of letters, and whether the word contains vowels and consonants or just consonants. The code was upgraded to accommodate words that contain 5 letters, so it starts by looking at the letter in the middle (3rd letter) and then checks the two letters before and after and follows the rules in table 1 appendix E for applying the diacritics and then transliterates the possibilities.

The number of possibilities for each of the 28 words is as follows:
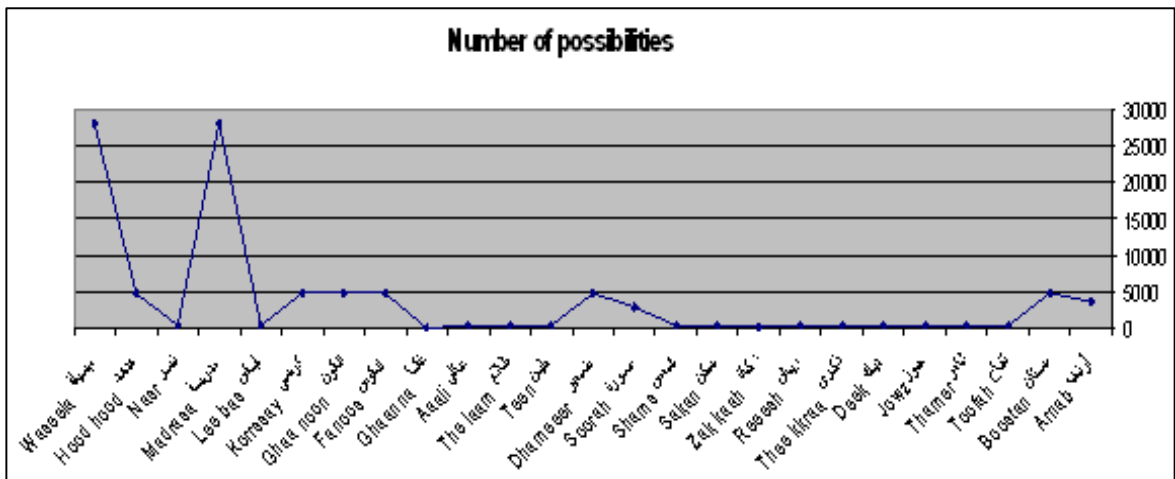
Figure 5.3 Line chart for the number of possibilities for the 28 Arabic alphabet words.

أرنب Arnab	بستان Boostan	تفاح Toofah	ثامر Thamer	جوز Jows	ديك Deek	ذكرى Thee kkraa	ريش Reeesh
3596	4796	482	482	478	478	396	478
زكاة Zak kaah	سكن Sakan	شمس Shams	صورة Soorah	ضمير Dhameeer	طين Teen	ظلام The laam	عالي Aaali
277	482	482	2805	4796	478	480	480
غانا Ghaanna	فانوس Fanoos	قانون Ghaa noon	كرسي Korrssay	لباس Lee bas	مدرسة Madrasa	نسر Nasr	هدهد Hood hood
39	4804	4804	4796	478	27998	478	4804
وسيلة Waseela							
27998							

Table 5.2 Number of possibilities for each of the 28 chosen words.

Consonants generate more possibilities as the rules for adding diacritics to all the letters in all positions (first, middle, last) states that it is possible to add most diacritics to all letters in any position. But when it comes to words that contain vowels, vowels can have specific diacritics in specific positions which limit the number of possibilities. Refer to Table 1 in the appendix E. Words like (ريش) Reeesh and (جوز) Jows both have the same number of possibilities because they both consist of three letters, a vowel in the middle surrounded by two consonants, etc.

5.1.4 Evaluation test

A recognition test for a subset of the 28 chosen words, words that produce 480 or less possibilities were used in the test. All the diacritised possibilities for each of these words were transliterated and included as vocabulary (xml file) for the speech recognition engine. After that the recordings of these words were played and recognition results calculated. Each word was played one after the other, and the recognised words got displayed on screen and the results for each user were saved into the log file (written file). The results are shown in table 5.3.

Arabic word	English Word	Number of possibilities	No. Times correctly recognized	% recognition
تفاح	Toofah	482	10	100
ثامر	Thamer	482	10	100
جوز	Jowz	478	10	100
ديك	Deek	478	10	100
ذكرى	Thee kkraa	396	10	100
ريش	Reeesh	478	10	100
زكاة	Zak kaah	277	10	100
سكن	Sakan	482	10	100
شمس	Shams	482	10	100
طين	Teen	478	10	100
ظلام	The laam	480	9	90
عالي	Aali	480	10	100
غانا	Ghaanna	39	10	100
لباس	Lee bas	478	10	100
نسر	Nasr	478	10	100

Table 5.3 Evaluation results for a subset of the 28 chosen words that produce 480 or less possibilities

The overall results gave an average recognition rate of 99.3%. This demonstrates that including more than one transliterated possibility for each word can be successful and can increase the recognition rate, as recognising any of the possibilities means recognising the original word that is linked to it. This can only be successful if the displayed recognised word was an undiacritised Arabic word. But this is not useful when using the speech engine to evaluate transliteration tables as only one word can be correct to measure the accuracy of the transliteration and there is a need to concentrate on studying specific letters and the accompanying diacritics in specific position increases.

5.2 Discussion & Conclusion

Automatic diacritisation, transliteration and producing all the possibilities for each word can be successful to an extent but the huge number of possibilities means more time for the speech engine to recognize the correct match.

Using undiacritised words for transliteration can only be used in very limited cases (i.e. a few small words in the vocabulary) and there is a potential for automatic transliteration but it critically depends on the transliteration and the rules used.

Also this approach to the first stage in the transliteration process can succeed in specific cases, especially where the vocabulary consists of shorter words. There is also the logical problem that it would be impossible for a speech recognition system to distinguish between two different words that were spelt with identical Arabic letters without some form of additional intervention (diacritics).

It is only sensible to proceed with diacritised words and the experiment is very limited with just 28 words so there is a need to develop a more comprehensive system for testing. Also different and more sophisticated transliteration rules should be explored.

It is clear that the success of the recognition depends on the quality of the transliteration rules that are used for automatic transliteration. Consequently, it means that this system could be used to test transliteration rules. Using the same vocabulary and voice recordings, different transliteration tables can be tested and compared by comparing recognition results recorded.

CHAPTER 6
System for Testing Transliteration Rules

6.1 Introduction

The research so far showed that the English speech engines could be used to recognise Arabic words. As the recognition of lists of words using this method was very sensitive to the transliteration rules used it was a motivation for this work to see whether this method could be used to test and compare transliteration rules.

This chapter describes a proposed novel system for testing transliteration rules.

6.2 Preparation of data

In order to test transliteration rules, it was necessary to choose a suitable vocabulary and record a selection of voices. The same evaluation process introduced in chapter 4 was used to test transliteration rules (refer to figure 4.1) so the transliteration stage prepares the list of words for the speech recognition process by transliterating the Arabic words into their English equivalent and the speech recognition engine with the aid of the pre-recorded audio files tries to recognise the words and then sends the results to an external file for analysis.

There are only 28 letters in the Arabic alphabet, but their sound can depend on their position in a word (start, middle or end) and they can be further altered by the use of additional symbols called diacritics. The term diacritem was defined to mean a particular letter in a particular word position with a particular diacritic. This letter/diacritic relationship can affect transliterations, as breaking the words into letters and diacritics in the three positions and studying the recognition for each to highlight the weaknesses and try to come up with better transliterations sounds promising. Hence the need to have a vocabulary that includes all of these possibilities arises.

The methodology of testing transliteration rules includes comparing different existing transliteration tables, and the attempt to try to find a better transliteration table than the best by changing the letter for letter transliterations, then more sophisticated rules where different transliterations for letters depending on whether they are adjacent to specific diacritics will be studied. Finally different transliterations of letter diacritic pairs should be considered depending on where their position is in the word. This process of improvement is described in Chapter 7.

This research is concerned with speech recognition from lists of words not from written texts; hence large chunks of established text will not be used in testing. The context of words will not be considered that's why attempting to use big texts is not useful for this research. Also any ordinary block of text will, by coincidence, contain lots of examples of some letters, or letter combinations, but very few examples of others, so it is not going to be efficient for testing this idea.

6.2.1 Selection of vocabulary words

The selection of words was made with the aid of Almawrid Arabic/ Arabic dictionary (Ba'Albaki, 1998), and then the list was presented to and approved by three experts in the Arabic linguistics field. The list of chosen words was presented to each expert and was asked to check whether all letters and diacritics were included in the three positions of the chosen words. The need for the experts is because some of the words contain only 1 or 2 shown diacritic, hence there is a need to determine the other diacritic(s).

In fact, no such similar comprehensive vocabulary exists for the Arabic language, so it was essential to carefully choose 3 letter words that would be familiar to Arabic speakers, but would cover all the Arabic 28 letters, in the 3 different positions beginning, middle, and end with the main diacritics ('Fat ha', Dhamma and kasra).

A vocabulary of 499, 3 letter words was found that contained all of the diacritems (in all positions) that are used in Arabic.

For example, the words حلبَ Halaba, قلبِ qalbii, and نابُ naabu, end with the letter "ب" baa, but it is diacritised differently in each word, in the first word, it is diacritised with a 'fat ha', in the second word the diacritem is bi which is a kasra, and the third word ends with damma. The letter "ب" in the end of these three words is pronounced differently.

The vocabulary was constructed with the following principles in mind:

1. All words were selected from Al Mawrid Arabic/Arabic dictionary
2. The vocabulary contains every letter in the Arabic alphabet
3. The vocabulary contains every letter in combination with every possible diacritic
4. Every letter and diacritic should be in the three different positions, start, middle and end. For example
   the letter raa (ر), start (* * رَ), (* * رُ), (* * ر), middle (*ر*), (* رُ *), (* ر *), and end (رَ * *), (رُ * *), (ر * *).

The table in Appendix G shows that all of the diacritems exist in the chosen vocabulary. Along the top, the complete list of possible diacritics is listed, and a list of all alphabet in three positions (start, middle, and End) listed down the side. The 499 words are listed in the correct cells according to their diacritems. Refer to Appendix H for a list of the 499 chosen words.

6.3 Recording of voices

The first problem that was looked into was the quality of the recordings. Because recording 499 words is time consuming, a couple of methods were tried to help in recording one word at a time and storing them in a separate files.

A simple audio recording application was implemented (Gaudio) using visual basic, this application allows the recording of a new set of words semi automatically and manually and then saves them in a predefined location. The application also displays a list of Arabic predefined words. Hence the user reads each Arabic word and they are recorded.

The user is asked to locate a folder or create a new location for saving the audio files. The user has the option of choosing the time interval (2, 3, or 4 seconds) between each word in the semi-automatic recording option.

Four subjects from the main experiment in 3.4 recorded their voices reading the 499 chosen words, one word at a time.
An initial test, playing the files and listening to the words proved that the quality of the recordings made using this application needed to be enhanced as they contained noise and were not clear enough.

Therefore another method was used to record a new set of recordings and thus increase the quality. The speech acquisition tool was provided by the Centre for Innovation and Technology Exploitation (CITE) at Nottingham Trent University. CITE provided a python tool, developed, using the snack library, to record user utterances (snack, 2006).

This application allows the recording of a new set of words manually by displaying a list of predefined words. Hence the user reads the word and records it by clicking the record icon, then clicks on the play icon to hear the word.

The best feature about this application is that it shows the distortion or noise, so that the user can tell if the word or part of it is not clear by looking at the distortion or noise level, if any part of the word was above or below the noise level, and if there

were more red dots than the yellow or blue then this part is distorted or noisy, so they had to re-record.

Two days later the same 4 subjects recorded the same 499 words using this application, it took about an hour to complete each set, roughly the same time it took them to record using Gaudio application with a 10 minutes break every half an hour.

The new recordings of each word are stored in a separate file and are called w1, w2, (word 1, word 2 etc.), the files are saved as wav files (705 kbps, 8bit sample rate, stereo, CCITT U-law). U-law or MU law is used in America and Japan for digital telecommunication.

The file has an average length of 2 seconds. The length of wave files range from 0.8 seconds to 2.4 seconds, Additionally, a 0.2 second silence period is added to the beginning and end of each file. The files are recorded in a quiet environment (ac's and computers off) the sets of recordings are kept in different folders for different speakers each titled with the speaker's first names.

Some of the files have background noise: although this kind of noise was deliberately avoided while recording, some files might have faint noise at the beginning.

The quality of the recording was checked by the author by looking at the waveform, and by listening to the recording to see if they were clear and sounded right.

If the quality wasn't good enough, the subjects were asked to repeat the recordings.

The application was used to record 4 new sets of high quality recordings, and then tested them automatically, using the same method used in chapter 4 where the 28 diacritised words were transliterated automatically using the same method introduced in chapter 4, using the UNESCO and IPA transliteration tables and the voice recordings were played and using the speech recognition engine, results were recorded.

The experiment was conducted using the Gaudio recordings to find the recognition rates and then repeated using the recordings by the application from CITE to compare recognition rates and use the recordings with the highest recognition rates in the upcoming experiments.

Refer to the appendix and attached CD for each set of the recordings (recorded using the CITE application).

Recording	Recordings using Gaudio recognition rate %	Recordings Using application from CITE recognition rate%
1	13	16
2	11	14
3	9	13
4	7.2	12

Table 6.1 Recognition rates of the recordings using the application from CITE compared to the recordings by Gaudio.

The recordings using the application from CITE recognition rate results show an improvement, but are similar to the results by Gaudio, both in the range of 0-20%. Although that the results are low they still serve their purpose by ensuring repeatability and consistency of the experiment results. This makes them more useful than using live voices.

6.4 Selection of transliteration tables

This section starts with a comparison of different transliteration tables' rules, to find the best transliteration table that can be constructed that is basically capable of transliterating one letter at a time.

The next stage is to explore different transliteration tables, hence five commonly used transliteration tables were tested and the results were compared, UN, Qalam, Buckwalter, ArabTeX and a table introduced in a paper by Alghamdi (Alghamdi et al., 2006).

The United Nations recommended romanization system was approved in (1972) (resolution II/8), based on the system adopted by Arabic experts at the conference held at Beirut in 1971 with the practical amendment carried out and agreed upon by the representatives of the Arabic-speaking countries at their conference. (UNESCO, 2006c)

According to Becker, (1987): **Qalam** (1985) is an Arabic-Latin-Arabic transliteration system between the Arabic script and the Latin script embodied in the ASCII (American Standard Code for Information Interchange) character set. The goal of the Qalam system is to transliterate Arabic script for computer communication by those literate in the language, it is a system that focuses upon preserving the spelling, rather than the pronunciation, and uses mixed case.

Buckwalter Transliteration (1990s) was developed at Xerox by Tim Buckwalter; and doesn't require unusual diacritics. (Habash et al., 2007b), (Lagally, 2004), (Micher and Voss, 2008), (Buckwalter, 2004) and (Buckwalter, 2002).

ArabTeX (since 1992) its "native" input is 7-bit ASCII: "has been modelled closely after the transliteration standards ISO/R 233 and DIN 31635"(Lagally, 2004).

Alghamdi's table Alghamdi et al., (2006) introduced a transliteration table that uses only plain Roman alphabets that can be processed and printed easily, so that ordinary people can read the transliterations.

All previous tests depended on UNESCO, (2006). A comparison between transliteration tables is shown in table 6.2

Letter	Letter name	Qalam	Buckwalter	ArabTeX	Alghamdi	UN
أ	Alef	aa	A	a	a	a
ب	Baa	b	B	b	b	b
ت	Taa	t	T	t	t	t
ث	Thaa	th	V	_t	th	th
ج	Jeem	j	J	^g	j	j
ح	Haa	H	H	.h	h	h
خ	Khaa	kh	X	_h	kh	kh
د	Dal	d	D	d	d	d
ذ	Thal	dh	*	_d	th	dh
ر	Raa	r	R	r	r	r
ز	Zain	z	Z	z	z	z
س	Seen	s	S	s	s	s
ش	Sheen	sh	$	^s	sh	sh
ص	Saad	S	S	.s	s	s
ض	Dhad	D	D	.d	dh	d
ط	Ta	T	T	.t	t	t
ظ	THa	Z	Z	.z	th	z
ع	Ain	`	E	`	A	'
غ	Ghain	gh	G	.g	gh	gh
ف	Faa	f	F	f	f	f
ق	Qaaf	q	Q	q	q	q
ك	Kaaf	k	K	k	k	k
ل	Lam	l	L	l	l	l
م	Meem	m	M	m	m	m
ن	Noon	n	N	n	n	n
ه	haa	h	H	h	h	h
و	waw	w	W	w	w	w
ي	Yaa	y	Y	y	y	Y
َ	'Fat ha'	a	A	a	a	a
ُ	dhamma	o	U	U or o	u	u
ِ	kasra	e	I	i or e	i	I

Table 6.2 UN, Qalam, Buckwalter, ArabTeX and Alghamdi's Transliteration Tables

For the transliteration of the rest of the diacritics, which were not available to the author for the above tables, the (IPA) International Phonetic Alphabet diacritics transliteration table will be used (Algamdi, 2003).

IPA	Diacritic
Xx	ّ—
X	ْ—
An	ً—
Un	ٌ—
In	ٍ—

Table 6.3 IPA diacritics transliteration table
*Xx= letter doubled
*X=letter

6.5 Results

To compare and find out the recognition rates for the previously mentioned transliteration tables, the automated application described in chapter 4 was used.

In each test, the transliteration table letters were fed into the application and the results were saved in the database, the application treats upper and lower cases the same.

The recordings by the application from CITE were played and fed into the speech recognition engine and finally the results were recorded.

6.6 Overall recognition rates

The following results were obtained:

Transliteration Table	Recognition Rate
UN	17.9%
Qalam	14.6%
Backwalter	13.2%
Arabtext	10%
Alghamdi	19.1%

Table 6.4 UN, Qalam, Buckwalter, ArabTeX and Alghamdi's transliteration tables' comparison tests results

Recording	Results
1	20.2%
2	23%
3	18.4%
4	14.8%

Table 6.5 Alghamdi's transliteration table recognition results

The results clearly highlight that Alghamdi's transliteration table achieved 19.1% using the previous method. This makes this transliteration table the best compared to the other 4 transliteration tables using this method of testing.
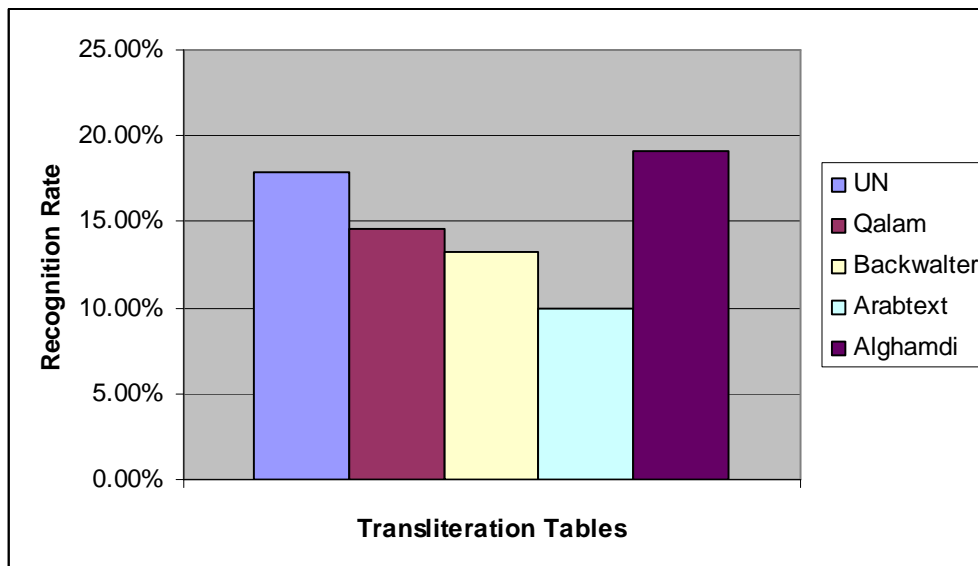
Figure 6.1 UN, Qalam, Buckwalter, ArabTeX and Alghamdi's transliteration table's comparison results

The next stage of the research was to identify if further improvements to Alghamdi's table could be made.

Refer to appendix I for a list of the transliterated words using different transliteration tables.

6.7 Analysis of individual letters

The letters that make up each word of the 499 words were highlighted for example the word (باد) was analysed to (ب), (ا), and (د) and the words were categorised according to the letters that it contained, so the same word (باد) can be in the letters (ب), (ا), and (د) categories . Then the total numbers of words that contain that specific letter were calculated. This provided an opportunity to calculate the recognition rate for each letter and therefore analyse this specific letter.

Using the transliteration table provided by Alghamdi, the recognition rate for each word averaged across all speakers is shown in table 1 (refer to appendix J). A further average was then taken of the recognition rates of all words that contained each letter and diacritic individual of the Arabic alphabet (only the main three diacritics will be studied, 'fat ha', dhamma, and kasra). This is presented in table 6.6

Arabic letter	Name of letter	English letter	TNOW	RR of test 1	RR of test 2	RR of test 3	RR of test 4	ARR
أ	alef	a	19	5.3%	10.5%	5.3%	5.3%	6.6%
ب	baa	b	71	15.5%	18.3%	17%	18.3%	17.3%
ت	taa	t	31	12.9%	12.9%	12.9%	12.9%	12.9%
ث	thaa	th	34	11.8%	23.5%	17.6%	11.8%	16.2%
ج	jeem	j	51	15.7%	23.5%	25.5%	13.7%	19.6%
ح	Haa	h	43	18.6%	34.9%	25.6%	18.6%	24.4%
خ	khaa	kh	28	10.7%	10.7%	7.1%	17.9%	11.6%
د	daal	d	58	17.2%	22.4%	22.4%	19%	20.3%
ذ	thaal	th	33	12.1%	15.2%	24.2%	24.2%	18.9%
ر	raa	r	126	12.7%	15.9%	15.1%	14.3%	14.5%
ز	zain	z	28	32.1%	32.1%	32.1%	32.1%	32.1%
س	seen	s	64	20.3%	23.4%	26.6%	18.8%	22.3%
ش	sheen	sh	34	38.2%	44.1%	50%	47.1%	44.9%
ص	Saad	s	39	20.5%	23.1%	20.5%	23.1%	21.8%
ض	DHad	dh	30	6.7%	6.7%	6.7%	6.7%	6.7%
ط	Ta	T	36	8.3%	11.1%	11.1%	11.1%	10.4%
ظ	THa	th	22	13.6%	13.6%	9.1%	13.6%	12.5%
ع	ain	A	75	12%	16%	18.7%	17.3%	16%
غ	ghain	gh	34	0%	17.6%	17.6%	2.9%	9.5%
ف	faa	f	57	17.5%	33.3%	17.5%	14%	20.6%
ق	qaaf	q	51	11.8%	11.8%	11.8%	13.7%	12.3%
ك	kaaf	k	43	20.9%	20.9%	20.9%	20.9%	20.9%
ل	laam	l	77	6.5%	11.7%	10.4%	11.7%	10.1%
م	meem	m	62	25.8%	27.4%	27.4%	27.4%	27%
ن	noon	n	61	23%	47.5%	37.7%	39.3%	36.9%
ه	haa	h	45	15.6%	24.4%	28.9%	26.7%	23.9%
و	waaw	w	71	18.3%	18.3%	18.3%	18.3%	18.3%
ي	yaa	y	50	22%	22%	22%	22%	22%
ﹷ	'Fat ha'	a	672	15.9%	24.1%	23.4%	22.6%	21.5%
ﹹ	dhamma	u	150	8.7%	22.7%	14.7%	13.3%	14.9%
ﹻ	kasra	i	118	12.7%	25.4%	21.2%	17.8%	19.3%

Table 6.6 Alghamdi's single letter or diacritic recognition rates analysis

*TNOW=Total number of words: Total Number of words that contain that specific letter

* RR of tests 1, 2, 3, and 4=Recognition rate of tests 1, 2, 3, and 4 recognition rate of each test for a specific letter (test 1=recording 1, test 2= recording 2, etc)

*ARR=Average recognition rate: average of the 4 recognition rate for a specific letter

In the diacritics case ('fat ha', dhamma and kasra), they can appear more than once in a word.

6.8 Summary

Table 6.6 presented the recognition rates for each letter and diacritic as transliterated according to Alghamdi's transliteration table. The results are low especially for the letter (Thad), which is expected because it is unique to the Arabic language and there is no equivalent to this letter in any other language; the letter (Thad) achieved a disappointing 6.7%.

A 499 word vocabulary that was designed to cover all common sounds in the Arabic language was found. This vocabulary was used to test currently published transliteration tables. The term Diacritem was defined to mean a combination of a letter and a diacritic at a specific location in a word. This vocabulary contained all possible Arabic diacritems.

Alghamdi's transliteration table achieved better recognition rates than the other transliteration tables. The author met Professor Alghamdi in Riyadh, he explained that his table is newer than the other 4 transliteration tables and when he was trying to come up with this table he studied nearly all of the published transliteration tables and tried to come up with a better table, he looked at their weakness and avoided them.

Hence the next step is trying to analyse Alghamdi's table and come up with a better transliteration table.

CHAPTER 7
Improvements to Alghamdi's transliteration table

7.1 Introduction

The previous chapter described a proposed system for testing transliteration rules, and the methodology for testing transliteration rules which includes comparing different existing transliteration tables was introduced. The results demonstrated that it is possible to test transliteration rules with the aid of speech recognition.

7.2 Finding improvements in the transliteration rules

In this chapter an attempt to try to find a better transliteration table than Alghamdi's table is described. Although Alghamdi's transliteration table achieved better recognition results than the other transliteration systems considered, the results are still somewhat limited and could be improved. Therefore an attempt to improve this table by changing letter for letter transliterations, then more sophisticated rules where different transliterations for letters depending on whether they are adjacent to specific diacritics were studied. Finally different transliterations of letter diacritic pairs were considered depending on where their position is in the word.

7.3 Improvements to single letter transliteration

In this section a method is described for improving the transliteration of each individual letter.

Using the transliteration table provided by Alghamdi, the recognition rate for each letter of the 499 words averaged across the four recordings are shown in table 1 (see appendix J). A further average was then taken of the recognition rates of all words that contained each letter and diacritic individual of the Arabic alphabet (only the

main three diacritics will be studied, 'fat ha', dhamma, and kasra). This is presented in table 6.6 which shows Alghamdi's single letter or diacritic recognition rate analysis.

7.3.1 The method used to identify how to improve the transliteration of single letter

Each letter or diacritic was studied and all words containing that specific letter or diacritic were analysed, and alternatives for each letter or diacritic were presented based on different transliteration tables like ALA-LC/UNGEGN and online transliteration applications and using the experimental method described in chapter 4, where the transliterations of the letters were fed into the application and the results were saved in the database, and the recordings were played and finally the results were recorded and the best results were chosen to form an improved table.

1. All the words that contain that specific letter or diacritic were found.
2. For each letter of the alphabet alternatives i.e. different possibilities were found with the help of different transliteration tables like ALA-LC/UNGEGN and the use of online transliteration applications also based on the author's knowledge of the language.
3. All the words were transliterated
4. The application was run, and the tests were done manually.

For example the letter baa, out of the 499 word vocabulary, the words that contain the letter baa were chosen and saved on a piece of paper. They were written as w1, w154 (word 1, word 154) then three alternatives were found (bb, p, and pp) and the fourth alternative (b) is Alghamdi's choice. (Using a text to speech facility, to make sure that the alternatives sound the same as Arabic). Next, all the words were transliterated so the baa was changed to be represented as bb, p and pp. Finally the application was run. Choosing a specific word number and running this word then

testing it is an option in this application, hence, only the words that contain the baa were chosen and run and the results were recorded. The experimental method is the same as the method described in chapter 4. The new transliterations of the letters were fed into the application to form the speech recognition lexicon but in this test only the recordings of the words containing the evaluated letter were run manually. And the results were saved in the log file

The (bb)'s recognition rate is 15.2 %, the (p) 10.9 % and the (pp) 10.2 %. Whereas Alghamdi's other alternative (b) got 17.3% which is the highest, so in this case the best alternative is chosen which is Alghamdi's (b), after that all the best alternatives were gathered to create a new table.

Another example is the letter (khaa), the (kh) which is Alghamdi's choice recognition rate is 11.6%, the (k) 16.1%, the (kk) 8% and (x) 8% and hence the (kh) will be replaced with (k). For the rest of the letters (see table 1 – appendix k)

7.3.2 The new transliteration rule based on single letter

The transliterations of letters or diacritics that got the best recognition rates and were different from those in Alghamdi's table, and were used to form the improved single letter transliteration table (SLT) are:

Arabic letter	Name of letter	Alghamdi's English letter	SLT
ج	Jeem	J	Jj
خ	Khaa	Kh	K
ر	Raa	R	Rr
ز	Zain	Z	Zz
ط	Ta	T	Tt
غ	Ghain	Gh	G
ق	Qaaf	Q	K
ك	Kaaf	K	Kk
ل	Laam	L	Ll
َ	'Fat ha'	A	Aa

Table 7.1 Differences between Alghamdi's table and the improved transliteration table (SLT)
*SLT= Single letter transliteration

For a comparison between Algahmdi's and the SLT table refer to Appendix S.

The improved (SLT) table is as follows:

Arabic letter	Name of letter	English letter	Arabic letter	Name of letter	English letter
أ	Alef	a	ظ	tha	th
ب	Baa	b	ع	ain	a
ت	Taa	t	غ	ghain	g
ث	Thaa	th	ف	faa	f
ج	Jeem	jj	ق	qaaf	k
ح	Haa	h	ك	kaaf	kk
خ	Khaa	k	ل	laam	ll
د	Daal	d	م	meem	m
ذ	Thaal	dh	ن	noon	n
ر	Raa	rr	ه	haa	h
ز	Zain	zz	و	waaw	w
س	Seen	s	ي	yaa	y
ش	Sheen	sh	َ	'Fat ha'	aa
ص	Saad	s	ُ	dhamma	u
ض	Dhad	dh	ِ	kasra	I
ط	Ta	tt			

Table 7.2 The improved SLT table

Another test was conducted using the improved single letter transliteration table, so these alternatives were fed into the application and the 499 were transliterated according to this new table, and the 4 recordings were run and the results were analysed.

7.3.3 Recognition results using the new single letter transliterations

The SLT table's recognition rate is 34.3%, clearly higher than rates achieved using Alghamdi's table. (See table 6.5).

	Letter overall test				SLT Table Average	Alghamdi's Average
	R 1	R 2	R 3	R 4		
Recognition Rate	39.1%	34.9%	28.9%	34.5%	34.3%	19.1%
Number of recognized words	195	174	144	172	-	
Words Total	499					

Table 7.3 overall test results for the new rule based on single letter transliteration
*R1, R2, R3, R4 = Recording1, Recording 2, Recording 3, and Recording 4.
*SLT Table Average= Single letter transliteration table average

7.4 Improvements using letter diacritic pairs

In this section a method is described for improving the transliteration of diacritic/letter pairs.

The recognition rate for each word averaged across all speakers is shown in table 1 (see appendix L). Using the improved SLT table a further average was then taken of the recognition rates of all words that contained each letter/diacritic pair of the Arabic alphabet. Only the main three diacritics will be studied ('Fat ha', dhamma and kasra). The analysis is presented in table 1– (appendix M).

Table 1 (appendix M) presents the letter and diacritic pair analysis, in alphabetical order according to the overall recognition rates. The table clearly shows that some of the recognition rates were odd for some of the pairs (highlighted). For example: for the letter ta, the overall recognition rate is 43.1% but when it is paired with dhamma diacritic, the recognition rate is only 8.3%, which is very low Also the letter baa, the overall recognition rate is 40.1%, but when it paired with kasra diacritic, the recognition rate is a disappointing 12.5%.

Hence a further analysis of these cases was performed.

Alef dhamma analysis

Refer to table 1–(appendix N) Aukht, baarraaa, and mudhi are words that contain the alef dhamma case. The word *aukht* consists of the letters alef, khaa and taa and the recognition rates for each is (18.4%, 27.7% and 29%). The dhamma diacritic's recognition rate is 33.4%, therefore in this case the low recognition rate cannot be from the letters or the diacritic because their recognition rates are higher than the alef, and thus the transliteration of the alef and dhamma pair might have caused the low recognition. Also the combination of English letters alef (au) and khaa (kh) makes an odd sound aukh. Another thing is that the letter khaa is one of the letters that appears only in the Arabic alphabet but has no equivalent in English.

Baarraau consists of baa (40.1%), raa (28.4%) and alef (18.4%), and the dhamma's recognition is (33.4%), same as above, the misrecognition is from the pair alef and dhamma also from the odd sound (au).

The same applies to the word *mudhi*, as the misrecognition is from the pair alef dhamma.

Baa kasra analysis

Refer to table 2-(appendix N).

The baa kasra case contains 4 words, bishrr, jjubillaa, kaallbi and birraakku.

The word *bishrr* consists of the following letters and diacritic, baa (40.1%), sheen (57.4%), raa (28.4%) and kasra (34.5%). The letter raa's recognition rate is close but lower than the baa, so this is one reason for the misrecognition also the poor transliteration of the pair baa kasra.

In *Jjubillaa* case, the letters jeem and laam recognition rates are lower than the baa, so this lowered the recognition rates, as well as the kasra diacritic which is lower than the baa and the pair baa kasra.

The same applies for the words *kaallbi* and *birraakku*. But in kaalbi's case, the letter gaaf which appears only in Arabic is the reason for the misrecognition.

Taa 'fat ha' Analysis

Refer to table 3-(appendix N). In the Taa 'fat ha' case, the word *thaabaataa* letters and diacritic recognition rates are as follows: thaa 25%, baa 40.1%, taa 29% and the kasra diacritic 37.3%, the misrecognition is caused by the letter thaa which has a lower recognition rate than the taa, in addition to the pair taa 'fat ha'.

Taaht, the reason for the misrecognition is from the transliteration of the taa 'fat ha' pair, as the recognition rates for the other letters and diacritic are higher than the letter taa, also the letter haa, as it is one of the letters that appear only in Arabic and has no equivalent in English, so the combination of the taa and haa, which makes an odd sound in English.

Taathill, the misrecognition is from the letters tha or laam, or the transliteration of the pair taa 'fat ha' and the combination of the taa and tha, (taath), which makes an odd sound in English.

Taaky, the recognition rates for the other letters and diacritics are higher than the letter taa, so the misrecognition is from transliterating the pair taa 'fat ha' as well as the combination of the taa and gaaf, which makes an odd sound in English.

Taamrr, the misrecognition is from the letter raa or the pair taa 'fat ha'.

Kaataallaa, the reason for the misrecognition is the letter laam and the transliteration of the pair taa 'fat ha' also the combination of the gaaf and taa (kaataa) which is an odd sound in English.

Naahaataa, the misrecognition is caused by the transliteration of the pair taa 'fat ha', plus the combination of the letters noon and haa.

Taa dhamma analysis

In table 4-(appendix N) *Twt*, the misrecognition is from transliterating the pair taa dhamma as the recognition rate of the letter waaw and the diacritic dhamma, is higher than the letter taa.

Atumaa, the misrecognition is caused by the transliteration of the letter ain, and transliterating the pair taa dhamma in addition to the combination of the ain and taa, which forms an odd sound in English.

Yumitu, the cause for the misrecognition is the transliteration of the pair taa dhamma.

Thaa dhamma analysis

In table 5-(appendix N) the misrecognition of the word *thullth* is caused by the transliteration of the pair thaa dhamma.

Thullaat, the misrecognition is from transliterating the pair thaa and dhamma.

The misrecognition of the word *jjuthw* is caused by the transliteration of the pair thaa and dhamma.

Baathu is misrecognised, because of the transliteration of the pair thaa and dhamma.

Thulluthin, *thuluthun*, and *thulluthan*, the misrecognition of these words is caused by transliterating the pair thaa dhamma.

Khaa kasra analysis

Refer to table 6-(appendix N). The misrecognition of the words *khidrr* and *mukhi* is caused by the transliteration of the pair khaa kasra, and the combination of khaa and daal in khidrr's case and the combination of meem and khaa in mukhi's case.

Baakhillaa, the misrecognition is caused by the transliteration of the letter laam and the pair khaa kasra as well as the combination of the letter khaa and other letters, like baa and laam, as it is one of the letters that distinguish Arabic language.

Thaal kasra analysis

In table 7–(appendix N) the misrecognition of the words, *dhiib* and *mudhi*, is caused by the transliteration of the letter alef, and the pair thaal kasra.

Kkaadhibaa and dhihni, the misrecognition is from transliterating the pair thaal kasra.

Ta dhamma analysis

Table 8–(Appendix N) the misrecognition of the word *ttaak* is caused by the letter gaaf, and the diacritic dhamma, as well as the transliteration of the pair ta dhamma and the combination of the letter ta and gaaf which is an odd sound to English.

Ottuf, the misrecognition is from the letter ain, and the diacritic dhamma, also the pair ta dhamma and the combination of the letter ain ta, in addition to the ta and faa, both pairs make unusual sounds to English.

The word *kirrttu* is misrecognised because of the letter gaaf, and the letter raa also the diacritic dhamma, besides the pair ta dhamma and the odd sound (kirr) which is a combination of the letters gaaf and raa and the odd sound rrttuu, a combination of raa and taa both strange sounds to English.

Tha dhamma analysis

Table 9-(Appendix N), the misrecognition of the words *kkaathi, thul, naathufaa* and *haathu* is caused by the transliteration of the pair tha dhamma and the combination of the letter tha and other letters (kaa, laam, noon, faa and haa) as the letter tha appears only in Arabic language.

The word *ghaaythu* is misrecognised because of the letter ghain and the pair tha dhamma in addition to the combination of the letters yaa and tha (yth) or letters ghain and yaa (ghaay).

Tha kasra analysis

Table 10-(Appendix N), the misrecognition of the words *taathill, thifrr* and *kaaythi* is caused by the transliteration of the pair tha kasra.

Aaathin is misrecognised because of the letter alef and the pair tha kasra.
The misrecognition of the word *Athimaa* is caused by the letter ain, besides the pair tha kasra.

The misrecognition of all the above words is from the combination of the letter tha and other letters.

Ain dhamma analysis

Table 11-(Appendix N), the misrecognition of the words, *otw, othirraa, orrsan, orrsun, orrsin, ottuf, omrr, naaomaa,* and *olluw* is caused by the transliteration of the pair ain dhamma.

Saao's misrecognition is from the letter alef and the transliteration of the pair ain dhamma. The letter ain is unusual to English and appears only in Arabic so the combination of the letter ain and other letters produce an odd sound in English.

Ghain kasra analysis

Table 12-(Appendix N), the misrecognition of the words ghill, ttaaghiyaa, and saamghi is caused by the misrecognition of the pair ghain kasra plus the combination of the letter ghain and other letters, as the letter ghain, is one of the letters that differentiate Arabic from other languages.

Tables (1 to 12 in appendix N) show that the letters with different diacritics affect the recognition rates, so it is worth investigating whether or not changing the transliteration of the letter/diacritics pair can improve the recognition rates.

7.4.1 The method used to identify how to improve the transliteration of letter and diacritic pair

Every letter/diacritic pair was studied and all words containing that specific letter/diacritic pair were analysed, and alternatives for each pair were presented (using a text to speech facility, to ensure that the alternatives sound the same as the Arabic pair) and using the experimental method described in section 4.3, where the transliterations of the pair were fed into the application and the results were saved in the database, and the recordings were played and finally the results were recorded and the best results were chosen to form a new table.

1. All the words that contain that specific letter/diacritic pair were found.
2. For each pair alternatives were found.
3. All the words were transliterated.
4. Then the application was run, and the tests were done manually.

For example the pair taa 'fat ha', out of the 499 words vocabulary, the words that contain the pair taa 'fat ha' were distinguished and saved on a piece of paper. They were written as w1, w154 (word 1, word 154), in the taa 'fat ha' case, three alternatives were found (tta, ttaa, taa, and ta). Choosing a specific word number and running this word then testing it is an option in this application, hence, only the words that contain the taa 'fat ha' were chosen and run and the results were recorded.

The (ta)'s recognition rate is 7.1 %, the (ttaa) 0 %, the (tta) 3.6 %, and the (taa) got 3.6% , so in this case the best alternative is the (ta), then all the best alternatives were gathered to create a new table. For the rest of the problematic pairs refer to table 1-(appendix O).

The letters or diacritics that got changed from the single letter transliteration table to form the improved transliteration of letter and diacritic pair table are as follows:

		SLT Table			LDPT Table		
Arabic letter	Name of letter	SLT Table English letter	SLT Table English diacritic	SLT Table letter-diacritic pair	LDPT Table English letter	LDPT Table English diacritic	LDPT Table English letter-diacritic pair
بِ	Baa	b	I	Bi	B	e	be
تَ	Taa	t	Aa	Taa	T	a	ta
ثُ	Thaa	th	U	Thu	Th	o	tho
خِ	Khaa	k	I	Ki	Kh	i	khi
طُ	Ta	tt	U	Ttu	T	u	tu
ظُ	Tha	th	U	Thu	Th	o	tho
ظِ	Tha	th	I	Thi	Th	e	the

Table 7.4 Differences between the SLT table and the LDPT table.
*SLT Table = Single letter transliteration table.
*LDPT Table = Letter and diacritic pair table transliteration.

Another test was conducted using the improved LDPT table alternatives, so these alternatives were fed into the application and the 499 were transliterated according to this improved LDPT table, and the 4 recordings were run and the results were analysed (Refer to table 1, appendix P).

The improved LDPT table recognition rates and the improved SLT table recognition rates are both near the middle of the 30-40% range. (See table 7.5).

7.4.2 Recognition results using the new rules based on letter diacritic pair

	Letter/Diacritic overall test				Average	SLT Table Average
	Recording 1	Recording 2	Recording 3	Recording 4		
Recogn. Rate (%)	42.7%	36.7%	31.7%	31.3%	35.6%	34.3%
No. of recog. Words	213	183	158	156	-	
Total words	499					

Table 7.5 The improved LDPT table overall test results
*Recogn. Rate = Recognition Rate
*No. of recog. words= Number of recognised words

7.5 Improvements using diacritems

There is a potential for further analysis of the improved LDPT table to prove that the diacritem (letter/diacritic/position) can affect the recognition rates.

The letter with different diacritics in different positions (start, middle and end) can affect the recognition rates, so it is worth investigating, whether changing diacritem (letter/diacritics/position) can improve the recognition rates.

In this section a method is described for improving the transliteration of diacritem. The method used is basically the same process as in the last section, however in this section diacritem combinations were considered.

The recognition rate for each word averaged across all speakers is shown in table 1 (see appendix L) using the improved SLT table transliteration. A further average was then taken of the recognition rates of all words that contained each letter/diacritic in a specific position (start, middle and end) of the Arabic alphabet. Only the main three diacritics will be studied ('Fat ha', dhamma and kasra). The analysis is presented in tables 1-3 – (appendix Q).

Tables 1, 2 and 3 – (appendix Q) present the diacritem analysis, in alphabetical order according to the overall recognition rates. The table clearly shows that some of the recognition rates were low for some of the diacritems (highlighted in grey). For example: for the letter alef 'fat ha', the overall recognition rate is 29.2% but when it is positioned in the middle, the recognition rate is 0 which is very low.

Hence a further analysis of these cases was performed.

7.5.1 The method used to identify how to improve the transliteration of diacritem

Each letter/diacritic pair in a specific position (start, middle, and end) were studied and all words containing that specific letter/diacritic pair were analysed, and alternatives for each diacritem were presented (using a text to speech facility, to make sure that the alternatives sound the same as Arabic) and using the experimental method described in section 4.3, where the transliterations of the pair were fed into the application and the results were saved in the database, and the recordings were played and finally the results were recorded and the best results were chosen to form a new table.

1. All the words that contain that specific diacritem were found.

2. For each diacritem alternatives were found.

3. All the words were transliterated.

4. The application was run, and the tests were done manually.

For example the pair taa 'fat ha' in the middle, out of the 499 words vocabulary, the words that contain the letter taa 'fat ha' in the middle were distinguished and saved on a piece of paper. They were written as w1, w154 (word 1, word 154), and in the taa 'fat ha' middle case, three alternatives were found (taa, tta, and ttaa). So the taa 'fat ha' in the middle was changed to be represented as taa, tta, and ttaa.

Choosing a specific word number and running this word then testing it is an option in this application, hence, only the words that contain the taa 'fat ha' in the middle were chosen and run and the results were recorded.

The (ta)'s 'fat ha' middle recognition rate is 0 %, the (taa) 0 %, the (tta) 6.3 %, and the (ttaa) got 0% , so in this case the best alternative is the (tta), then all the best alternatives were gathered to create a new table. For the rest of the problematic pairs (see tables 1, 2, and 3- Appendix R).

The diacritems that got changed from the improved LDPT table to form the improved DT (Diacritem transliteration) table are:

'Fat ha'

Arabic letter	Name of letter	Position	From	To
أ	alef	Middle	aaa	aa
ت	taa	Middle	ta	tta
ض	dhad	Middle	dhaa	dha
غ	ghain	End	ghaa	gaa

Table 7.6 'Fat ha' diacritems that got changed

Dhamma

Arabic letter	Name of letter	Position	from	to
خ	khaa	End	khu	khoo
ذ	thal	Start	dhu	thu
ر	raa	Middle	ru	rro
ض	dhad	End	dhu	dho
ط	ta	Start	ttu	tto
ط	ta	End	ttu	ttou
ظ	tha	Start	tho	thu
ظ	tha	End	tho	thu
ع	ain	End	o	au
غ	ghain	Middle	ghu	gu
ل	laam	Middle	llu	lu
ن	noon	Middle	nu	no
ه	haa	Start	hu	ho

Table 7.7 Dhamma diacritems that got changed

Kasra

Arabic letter	Name of letter	position	From	to
ث	thaa	Start	Thi	the
ج	jeem	Start	Jji	jje
د	daal	End	Di	ddi
ر	raa	End	Rri	ri
س	seen	End	Si	ssi
ض	dhad	End	Dhi	dhe
ع	ain	End	Ee	ai
غ	ghain	Middle	Ghi	ghe
ق	qaaf	Start	Ki	kki
م	meem	Middle	mi	me
ه	haa	End	Hi	hhi

Table 7.8 kasra diacritems that got changed

Another test was conducted using tables 7.6, 7.7 and 7.8 alternatives, so these alternatives were fed into the application and the 499 were transliterated according to this new DT table, and the 4 recordings were run and the results were analysed.

7.5.2 Recognition results using the new rules based on Diacritem.

	Diacritem overall test				Average	LDPT Table average	SLT Table average
	R 1	R 2	R 3	R 4			
Recognition Rate	44.9%	40.3%	32.3%	34.3%	37.9%	35.6%	34.3%
Number of recognized words	244	201	161	171	-		
Total words	499						

Table 7.9 The improved DT table overall test results

Although the above method is complex, and time consuming, the new table's recognition rate is 37.9%. This is higher than single letter and letter/diacritic pair methods.
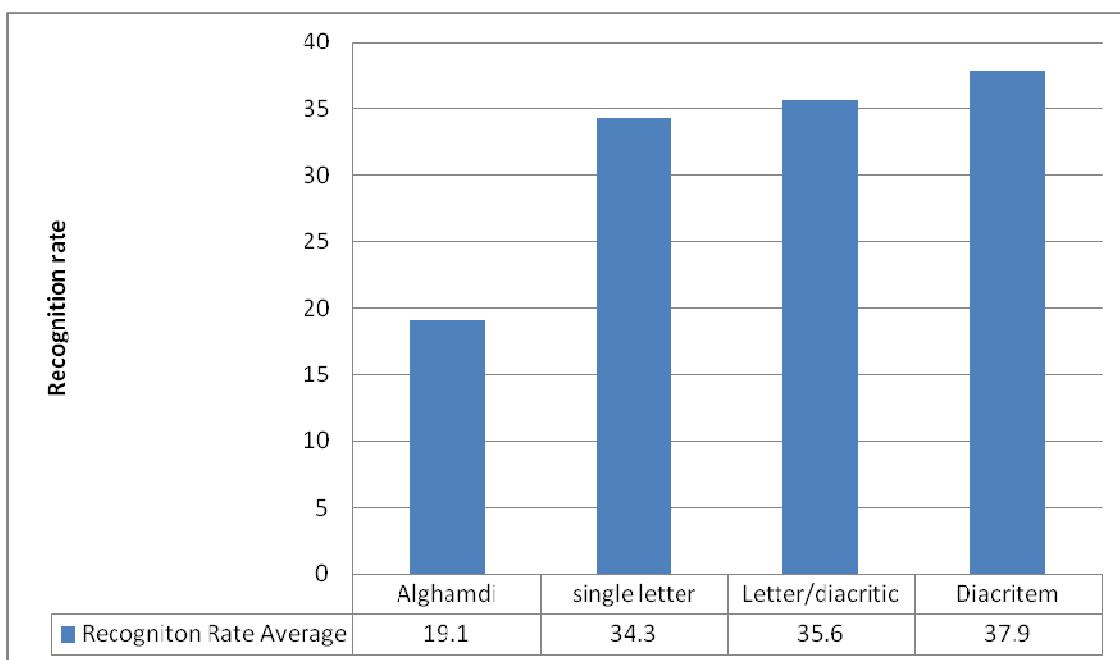
Figure 7.1 comparison of the recognition rates of Alghamdi's, single letter, letter/diacritic pair and diacritem transliteration tables.

7.6 Evaluating the system for testing and improving transliterations

The methods applied above helped to improve recognition rates over those achieved using Alghamdi's transliteration table, but there is still the question whether the improved results are due to this particular set of words or voices. Hence, there is still the need to evaluate whether the system for testing and improving transliteration is really an effective way to do this.

7.6.1 The method used to evaluate the system for improving transliterations

Two sets of words were compared. The two sets of similar words that contain the letter kha in the three positions accompanied with all diacritics, they both have 28 words, and each word consists of three letters.

All the words containing that specific letter from the list of 499 words were chosen, and alternatives were presented (using a text to speech facility, to make sure that the alternatives sound the same as the Arabic pronunciation) and using the experimental method described in section 4.3, where the transliterations of the letter were fed into the application and the results were saved in the database, and the recordings were played and finally the results were recorded. Another similar kha list was found, and the same experimental method applied. Finally the results of both experiments were compared.

1. All the words that contain that specific letter were found.
2. For the specific letter alternatives were found.
3. All the words were transliterated
4. Then the application was run, and the tests were done manually.
5. Another similar list of words was found.
6. For the specific letter alternatives were found.
7. All the words were transliterated
8. Then the application was run, and the tests were done manually.
9. The results of both groups were compared.

For example the letter kha was chosen, out of the 499 words vocabulary, the 28 words that contain the letter kha were distinguished and saved on a piece of paper. They were written as w1, w154 (word 1, word 154), and alternatives were found (kk, q, kkh, kha and kh). Choosing a specific word number and running this word then testing it is an option in this application, hence, only the words that contain the letter kha were chosen and run and the results were recorded.

Another list that consists of 28 words, containing the letter kha in the three positions was found using Almawrid Arabic/Arabic dictionary and the same testing procedure was followed and results compared, refer to table 1- (appendix T) for the new list of words.

The alternatives are as follows:

Group	Alternatives											
	1		2		3		4		5		6	
	Op. 1	RR %	Op. 2	RR %	Op. 3	RR %	Op. 4	RR %	Op. 5	RR %	Op. 6	RR %
1	K	64.2	kh	75	kk	14.2	Khh	53.5	kha	21.4	Q	60.7
2	K	67.8	kh	75	kk	17.8	Khh	42.8	kha	28.5	Q	57.1

Table 7.10 Alternatives for the letter kha comparison
*Op. 1= Option 1, Op. 2 Option 2, etc.
*RR%=Recognition Rate%

The above table shows that the alternative kh should be chosen to represent the transliteration of the letter kha, and the similarity of results between the two groups proved that even if different words in the list were used then the best way of transliterating certain specific letters remains the same. This eliminates the specific set of words factor. There is still the need to establish whether the quality of recordings has any influence on the recognition rates.

Hence, further analysis of the results of the 4 recordings, reading the two sets of words, as the letter kha was represented as 'kh' is necessary to eliminate the quality of recordings factor.

The analysis of the results is as follows:

	Recordings 1		Recordings 2		Recordings 3		Recordings 4	
	G1	G2	G1	G2	G1	G2	G1	G2
No. of recognised words	24	18	20	24	22	22	19	19
Average all	85.7	62.2	71.4	85.7	78.6	78.6	67.8	67.8
Average 'fat ha'	93.3	86.6	66.6	100	86.6	80	60	60
Average 'fat ha' Start	91.6	83.3	75	100	83.3	75	58.3	66.6
Average 'fat ha' Middle	100	100	50	100	100	100	100	50
Average 'fat ha' End	100	100	0	100	100	100	0	0
Average Dhamma	66.66	0	100	66.6	66.6	66.6	33.3	100
Average Dhamma S	100	0	100	100	100	100	0	100
Average Dhamma M	100	0	100	100	100	0	100	100
Average Dhamma E	0	0	100	0	0	100	0	100
Average kasra	100	75	50	75	75	75	100	75
Average kasra S	100	100	100	100	100	100	100	100
Average kasra M	100	50	0	100	50	50	100	50
Average kasra E	100	100	100	0	100	100	100	100

Table 7.11 Analysis of the recognition of the 4 recordings by Groups 1 and 2 reading (kha represented as kh) words.
*G1, G2, G3, and G4=Group1, Group2, Group3, and Group 4.

The previous table demonstrated that the quality of the recordings did not affect the recognition rates as the results of both groups were quite similar across the 4 recordings. This eliminates the recordings quality factor.

7.7 Comparison of Alghamdi's and the improved diacritem transliteration tables

Alghamdi's and the improved DT tables were compared.

The 499 words were transliterated according to the two tables and presented to two Arabic language experts working at the University of Bahrain. The four recordings were also presented to the experts. They were asked to play the recordings and check them against the transliterations.

Both transliteration tables were evaluated based on the phonetic and spelling accuracy and usability, as well as how accurate the word is transliterated to match the recordings.

Accuracy for the purpose of this research is similar to Lawson's (2008) definition which examines how close the pronunciation is to the original Arabic letter, this could be very tricky because there are some letters in Arabic that have no direct English equivalent like the letter ض.

The usability part of the evaluation will investigate each table's adherence to ASCII standards i.e. non use of symbols (non letters).

Each table receives a score from 0 to 100; these points are allotted for spelling and phonetic accuracy for the 4 recordings. They were asked to play the recordings and compare each word with the transliterated version transliterated using both tables and rate them from 0-100.

Usability according to Lawson (2008) measures how accurate is the representation of the transliterated word (spelling).

Accuracy assesses how close the pronunciation of the word is to the original Arabic word and whether it matches the transliteration (phonetics).

The evaluation results were as follows.
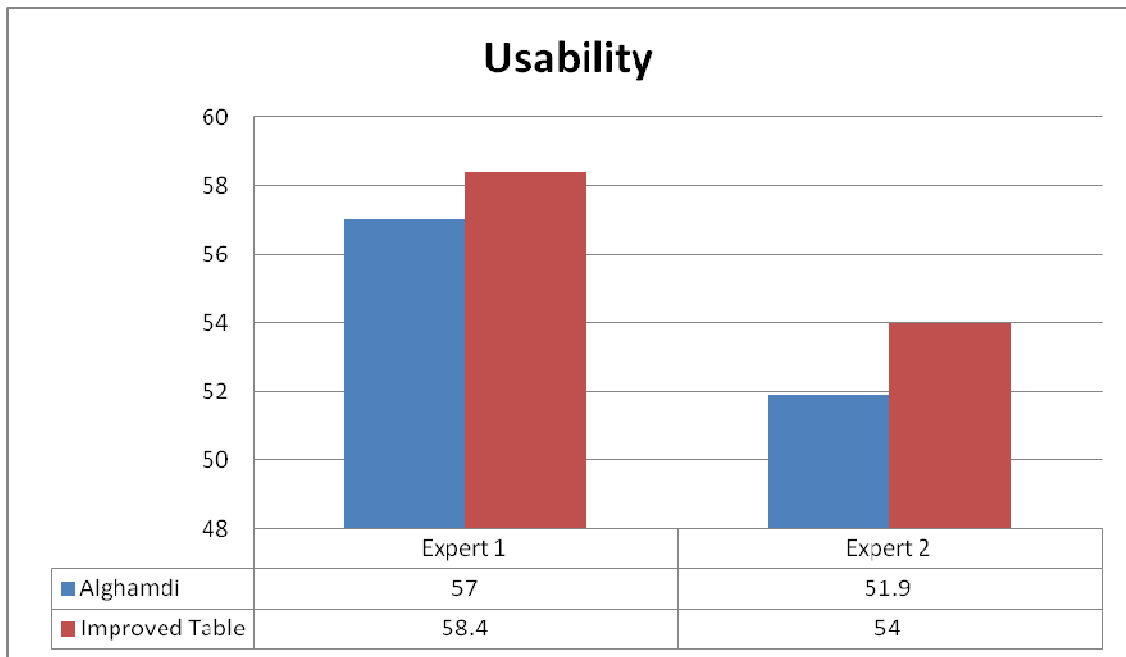
Figure 7.2 Alghamdi and improved DT table comparison usability evaluation

Figure 7.2 clearly shows that both experts thought that the representation of the transliterated words by both table is in the range of 50-60. The improved DT table results were slightly higher.
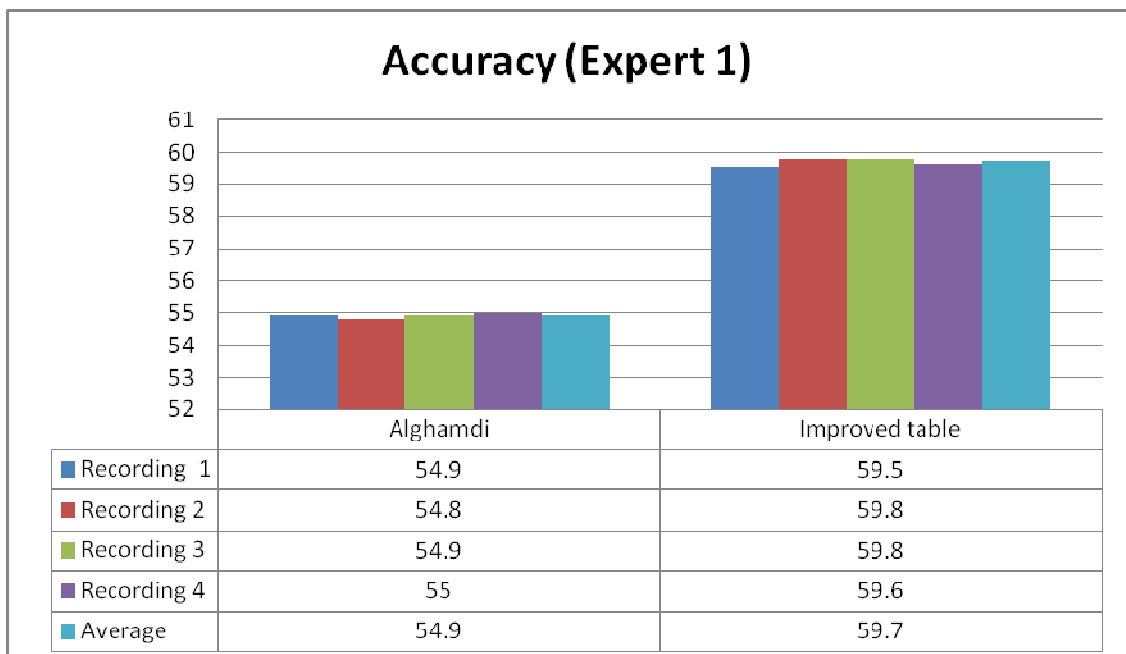
Figure 7.3 Alghamdi and improved DT table comparison Accuracy evaluation by expert 1
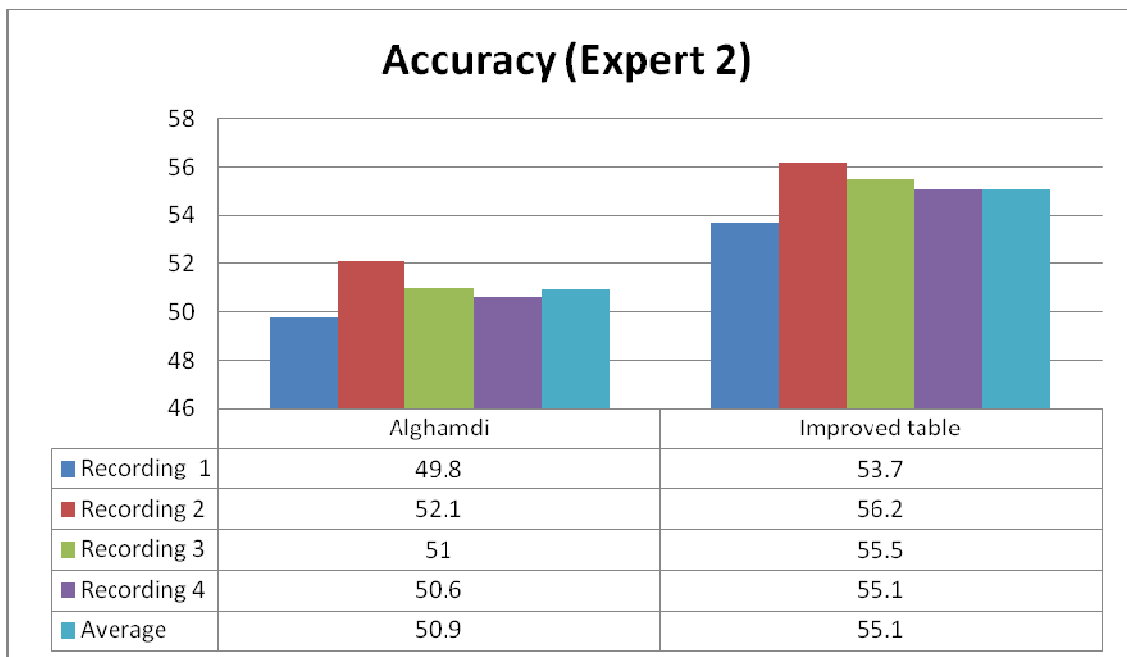
Figure 7.4 Alghamdi and improved DT table comparison Accuracy evaluation by expert 2

Both experts evaluated the accuracy of how close the pronunciation of each word by the four subjects is to the original Arabic word and does it match with the transliteration (phonetics).

Both experts gave both tables an average that ranges between (50-60). But both agreed that the improved DT table's transliterations are closer to the pronunciation of the original word.

Refer to table 1 appendix U for the original evaluation results table by the two experts.

7.8 Conclusion

This section presented an attempt to improve Alghamdi's transliteration table. Methods include changing letter for letter transliterations, then more sophisticated rules where different transliterations for letters depending on whether they are adjacent to specific diacritics were studied. Finally different transliterations of letter diacritic pairs were considered depending on where their position is in the word.

Although the diacritem method is more complex than the other two methods, it achieved the highest recognition rate.

The experts' evaluation results measuring the accuracy and usability for both Alghamdi and the improved diacritem transliteration tables were similar and ranges between (50-60). The original results achieved by automating the transliteration and speech recognition processes were a disappointing 19.1% for Alghamdi's transliteration table and 37.9 % using the diacritem transliteration table.

Both experts agreed that the improved DT table's transliterations are closer to the pronunciation of the original word. In both tests the automated transliteration and recognition and the evaluation by the two experts showed that the improved diacritem transliteration table gave better results than Alghamdi's transliteration table in both cases.

The results demonstrated that the new transliteration rules (diacritem) can improve the recognition rates, but it cannot be established whether these results are statistically significant, or whether they are just randomly based on the chosen words. Also previous chapters showed that high recognition rates do not always indicate that the transliterations used are the best.

These issues will be discussed in the next chapter.

CHAPTER 8 Discussion and Conclusion

This concluding chapter summarizes the results that were presented in this thesis. Furthermore, some open problems are pointed out, giving possible research directions for the future.

This thesis illustrates the development of a novel approach to systematically evaluating transliteration systems with the aid of speech recognition technology. In order to achieve this there were 5 main objectives, first to establish the possibility, for English speech recognition engines to recognize Arabic words with the aid of Arabic transliteration. This research is concerned with speech recognition from lists of words not from written texts; hence large chunks of established text were not used in testing. The context of words will not be considered which is why attempting to use big texts is not useful for this research. Another reason is any ordinary block of text will, by coincidence, contain lots of examples of some letters, or letter combinations, but very few examples of others, so it is not going to be an efficient for testing this idea.

The second objective is to establish the possibility of automatic transliteration of diacritised Arabic words for the purpose of creating a vocabulary for the speech recognition engine, thirdly, to explore the possibility of automatically generating transliterations of non diacritised Arabic words. Also to find the means to construct a general method to test different ways of performing transliteration. Last of all, testing the system and using it to experiment with new transliteration ideas.

Transliteration schemes can be evaluated by humans manually but repeating the results is inconsistent, time consuming and difficult and the results will always be a matter of opinion because there is not a simple one-to-one mapping between Arabic and English orthographies. Hence using a speech recognition engine to test Arabic transliterated words could be an alternative.

Speech recognition technology could be used to evaluate transliterations because both technologies are concerned with the sounds of words.

When the user speaks the transliterated words; the spoken phonemes or sounds should match the phonemes stored internally in the speech recognition system.

The spoken input is matched to the phonetic representations of the words in the lexicon of the speech recognizer in order for a word to be recognised. If the internal versions are good, they will provide good matches, so that recognition rate can be an indicator of how good a transliteration is. Nevertheless, other factors can influence the recognition rate like the vocabulary size and content, and it is not a direct evaluation of the quality of transliteration. The recognition rate for the recognised transliterated word matched with the original spoken Arabic word offers a way to measure the accuracy of transliterations automatically, with some degree of consistency and repeatability especially if recoded voices were used.

The experimental results also propose that high recognition rates don't always signify that the transliterations used are the best.

Transliteration of Arabic words can often be a matter of judgment, and recognition rate is not an ideal technique of judgment of the transliteration since other factors can influence the recognition, such as accents or pronunciation of different letters, if the way they speak matches the phonetic representation or the sound made internally by the speech engine for this word or letters it will be likely to be recognized correctly.

Another finding, is that although transliteration acceptability cannot be measured directly using the speech recognition performance, it does correlate well with human judgment and as well as offering consistency and repeatability.

8.1 Achievements

This section will highlight the achievements of this work and its contribution to transliteration research communities.

8.1.1 Using English speech recognition technology for the recognition of Arabic.

The use of the more developed English speech recognition technology with the aid of transliteration, for the recognition of Arabic vocabularies as an alternative may offer a route to creating practical Arabic language speech engines.

A simple application was created and tested which was to recognise the 28 words of the Arabic Alphabet in a similar way to that used in the Civil Aviation Organisation code to identify letters of the English alphabet (Alpha, Bravo…..Zulu). The next stage looked into investigating whether it is possible to computerize the process in order to save time and effort and to examine the transliteration process with the objective of creating a transliteration program that could be used in conjunction with an English speech recognition engine in order to evaluate transliteration tables. Based on the results and the feedback it is possible to create an application for the purpose of transliterating Arabic into English and then recognizing the words using a Standard English speech recognition engine manually.

In this research, the simple speech recognition application has been developed in Microsoft Visual Basic and uses the Microsoft Speech SDK V5.1 to create an interface to the Microsoft English (U.S.) V6.1 Recognizer speech recognition engine.

Any engine can be chosen as long as the engine has all of the letters to cover the Arabic alphabet. Hence, it is possible to use a UK English Engine, as long as the letter to letter alternatives are altered to match the pronunciation accurately. The UK English is similar to American English and they both have the same alphabet.

Using a Persian engine for example is possible, as the transliterations can be altered so that the pronunciation can be close to Arabic using the Persian letters and phonemes, keeping in mind that any transliteration is considered correct as long as it sounds close Arabic pronunciation because there is no real answer. And any transliteration rules will only be appropriate for a specific accent.

Any transliteration scheme can be thought of as a compromise between the original language and the speaker's accent. At the very best, every transliteration seeks an average accent. For example Modern Standard Arabic is the accent spoken nearly the same by all Arabic speakers, but the transliteration needed for Egyptians is different from the transliteration needed for Bahrainis to pronounce words transliterated in Modern Standard Arabic correctly.

With this finding, it is confirmed that one of the main objectives of this work is achieved.

In addition, although some care has been taken to get a range of Arabic speakers, they mainly came from or lived in Bahrain. It is likely that the recognition rates would be even lower for the full Arabic speaking population.

In the future, further evaluation still needs to be carried out on a larger population of users from other Arabic speaking countries.

More representative groups of voices should be arranged like covering all 6 Gulf countries or all Arabic speaking countries. Evaluating the statistical variation of similar sets of voices can be a solution.

Using different transliteration rules for different geographical areas is another idea, so that a telephone application in Bahrain would have different vocabulary from the one in Morocco. This might be good for telephone systems, but this method is not practical for an agreed global approach to transliteration.

8.1.2 Automatically generating transliterations of diacritised Arabic words

An application that generated an English vocabulary by transliterating each Arabic diacritised word into its English equivalent was designed and implemented.

Automatic diacritisation and transliteration and producing all the possibilities for each word can be successful to an extent but the huge number of possibilities means more time for the speech engine to recognize the correct match.
Automatic transliterations of diacritised words save time and effort also aid to decrease the number of possibilities.

Using undiacritised words for transliteration can only be used in very limited cases (i.e. a few, short words in the vocabulary) and there is a potential for automatic transliteration but it critically depends on the transliteration and the rules used. Trying to automatically diacritise and transliterate undiacritised words means generating a large number of possibilities, and some of these possibilities might be irrelevant and in order to filter these, they should be compared with a list of real words or with human interference.

The findings demonstrated that it is possible to automatically transliterate diacritised words for the purpose of providing vocabulary for the speech recognition engine. This is used for testing transliteration tables.

It is only sensible to proceed with diacritised words and the experiment is very limited with just 28 words so there is a need to develop a more comprehensive system for testing. Also more sophisticated transliteration rules should be explored. This system could be used to test transliteration rules. Using the same vocabulary and voice recordings, different transliteration tables can be tested and compared by comparing recognition results recorded.

In addition to the huge number of possibilities problem, the logical problem exists, as it would be impossible for a speech recognition system to distinguish between two different words that were spelt with identical Arabic letters without some form of additional intervention (diacritics).

A simple test was done to establish whether the quality of recordings has any influence on the recognition rates.

Thus, further analysis of the results of the 4 recordings, reading two different sets of words is necessary to eliminate the quality of recordings factor. And the results proved that the quality of the recordings did not affect the recognition rates as the results of both groups were quite similar across the 4 recordings. This eliminates the recordings quality factor. As time was a limitation for this research, future work should concentrate on conducting further tests to ensure the accuracy of results covering the whole set of 499 words and a different list that consists of similar size of words created using the same original criteria. As testing different sets of voices and words and a statistical analysis of the variation of the results would help to create guidelines about how efficient the testing system is.

8.1.3 Constructing a novel method to test and compare transliteration tables.

The research results proved that the English speech engines could be used to identify Arabic words but there is a need for effective transliteration of the Arabic words in order to create an appropriate English vocabulary (lexicon). As the recognition of lists of words using this method was very sensitive to the transliteration rules used it was a motivation for this work to see whether this method could be used to test and compare transliteration rules.

A novel procedure for systematically testing transliteration rules has been created and software to support this procedure has been produced. The testing system used a

499 word vocabulary that was designed to cover all common sounds in the Arabic language. In order to do this the term Diacritem was defined to mean a combination of a letter and a diacritic at a specific location in a words (start, middle or end). This vocabulary contained all possible Arabic diacritems.

Then different transliteration tables' rules were compared, to find the best transliteration table that can be constructed that is basically capable of transliterating one letter at a time.

The developed system achieved the main objective of the work which makes it an effective transliteration rules testing process.

Although the limitations appear to be that the experiments were only tested on a limited number of voices which only come from one region, the percentages obtained in this work serves as a quantitative indication that the test method and application presented in chapter 4, 5, 6 and 7 is indeed feasible.

A specific vocabulary of 499 words was used to test the transliterations rules. A disadvantage of using this specific vocabulary is that results are bound to be influenced by the chosen 499 words. A simple test was done to check whether the choice of 499 words had a significant effect on the results and a different set of words to test the recognition rate of one particular letter (letter Kha) was chosen. The 28 words, containing the letter kha in the three positions were found using Almawrid Arabic/Arabic dictionary and the same testing procedure was followed and results compared. Choosing a new set of words to represent the same letter did not alter the results; time was a limitation for this research and hence future work should concentrate on conducting further tests to ensure the accuracy of results covering all letters of the alphabet. And a different list that consists of similar size of words and creating using the same original criteria should be found. As the comparison of the two lists and testing complete sets of different voices and words and a statistical

analysis of the variation of the results would provide an opportunity to create guidelines about how effective the testing system is.

For example producing additional complete sets of the 499 words and finding the recognition rates of each of the transliteration tables that were used for each set of words, then analysing the variation of the recognition rate for each table and using a statistical significance test to see if the differences between tables are statistically significant, or whether they are just randomly based on the words that were chosen.

Further testing will determine whether the process used to select the words could have in any way skewed the results. For example, if the words were systematically chosen from the beginning of a dictionary until enough words were found. The chosen words would have been picked from the words that are found early in the alphabet and perhaps a completely different result would be obtained if the chosen words were selected from the end and worked to the beginning.

In addition to doing a statistical analysis on the different voices, this would provide a range of recognition rates for the different voices and there will be a standard deviation between these results. This could be used to evaluate the statistical significance of the overall recognition rates. Obviously, this only measures the statistical variation of the chosen voices and there may of course be other non statistical factors e.g. they all come from Bahrain.

It would also be possible to do a similar sort of test, to see if the decisions about improvements were statistically significant.

Alghamdi's and the improved Diacritem transliteration tables were also manually compared.

The 499 chosen words were transliterated according to the two tables and presented to two Arabic language experts. The four recordings were also presented to the experts. They were asked to play the recordings and check them against the

transliterations. Both transliteration tables were evaluated based on the phonetic and spelling accuracy and usability, as well as how accurate the word is transliterated to match the recordings.

Both experts thought that the improved Diacritem table is more usable and more accurate than Alghamdi's table.

The original results by automating the process also showed that the improved diacritem transliteration table gave better results than Alghamdi's transliteration table.

Both results confirmed that the new transliteration rules (diacritem) can improve the recognition rates, but it cannot be verified whether these results are statistically significant, or whether they are just randomly based on the chosen words and hence the importance of re-testing on a new set of 499 words to determine whether the results were statistically significant arises.

8.1.4 Experimenting with new novel transliterations ideas to find improvements in the transliteration rules

The methodology of testing transliteration rules includes comparing different existing transliteration tables, and the attempt to try to find a better transliteration table than the best by changing the letter for letter transliterations, then more sophisticated rules where different transliterations for letters depending on whether they are adjacent to specific diacritics, and diacritems were studied.

The results demonstrated that the diacritem transliteration achieved higher recognition than the other two methods although it is time and effort consuming. Changing the letter and diacritic pair achieved a similar recognition rate result as changing letter for letter transliterations.

The only limitation is that time was not enough to identify more imaginative ways to perform transliteration. Examples may be to take more account of the anomalies in the way English words are pronounced when certain word combinations are produced. Taking account of the anomaly in English could lead to better transliteration. E.g. 'T' followed by 'H' would not sound 'TH' but (THE).

8.2 Overall contributions of this work

The increased international communication has led to an increased need for transliteration of many things that cannot be translated like many proper names.

Numerous transliteration systems have been developed and used by researchers for many years; however there is still the question of how effective these systems are.
This research demonstrates the development of a novel approach to systematically evaluating transliteration systems.
This testing method could help researchers to compare existing systems and come up with the best transliteration table that could cover all the languages of the world.

Overall the contributions of this work to the related research community are as follows:

1) Identifying that in principle diacritised Arabic words can be identified by English speech engines provided suitable rules for transliteration are available

2) Constructed a novel general method to test different ways of performing transliteration by implementing an application to support this procedure and building an Arabic comprehensive vocabulary as a research infrastructure and have it available for Arabic researchers to stimulate further research in this field and its application.

3) Improvements to currently published transliteration tables were explored and significant improvements were made.

4) The use of more complex rules based on the novel concept of the diacritem and letter and diacritic pair has been explored.

The novel methodology created to test different ways of performing transliteration and to compare existing transliteration rules using novel and more complex rules based on the diacritem concept is not sensitive to the changes of words, i.e. choosing a different set of words of the same creating would achieve the same recognition results as proved by testing the (kha) set of words. However it is not possible generate the perfect rule or the perfect testing method as there is no real answer and any transliteration rule will only be appropriate for a specific accent. Actually the transliteration that would make an American make the same sound as an Egyptian will be different from the transliteration needed to make a Scots man sound like a Bahraini. This is both because the Egyptian and Bahraini will make different sounds when they see the same Arabic word, but also because the Scotsman and the American will make different sounds when they see the same English words. Every transliteration is a compromise. At best, transliteration can aim for a sort of 'average accent'. Although there is no perfect rule or testing method, the methodology found has the advantage over purely subjective methods of being repeatable and consistent and also it is easy to apply that the program is implemented.

Findings from experiments carried out in this work will contribute to the transliteration research community in the area of transliteration rules comparison where much more investigation is necessary as it has not been covered.

8.3 FUTURE WORK

The obtained results from this work have been encouraging and showed many possibilities for future work. The following section will provide some extensions of this work.

Future work aims on developing the testing capabilities for example automating the transliteration testing process and conducting more thorough testing and statistically based analysis of this method, also testing more complex vocabularies like Arabic full names, list of street names, etc and finding a different 499 words list that consists of similar size of words and creating using the same original criteria to compare it with the current 499 words list.

In addition to exploring the possibility of using English speech engines to recognise Arabic words and establishing guidelines about how effective using English based engines to recognise Arabic speech is.

8.3.1 Automating the transliteration testing process

Through the lessons learned from this work it is possible that the testing method could be further improved by automating the whole process. What really happens in the current application is when two tables to be compared are chosen the author inserts the transliterations into the specified tables and then runs the application after choosing the recordings that will test this vocabulary. The same process is repeated for the second table or set of transliteration rules. And finally the two results get compared. This work could be cut short by introducing a method that accepts inserting two tables at the same time and running two applications to conduct the test only once.

8.3.2 Testing the application using more complex Vocabularies and generating guidelines

Testing the application using a more complex vocabulary like full Arabic names is a must because it is more realistic and can be more useful for e.g. like using it by doctors to enter patient names.

After converting the Arabic names into their Roman (English) equivalent using automatic transliteration and diacritisation, the names will form a database or a list of common Arabic names, then they can be tested using the speech recognition application.

Another 499 words list should be found, this list should consist of similar size of words and creating using the same original criteria to ensure the accuracy of results also more recordings should be used. As the comparison of the original and new list would and the recordings provide an opportunity to create guidelines about how effective the testing system is.

Further testing is necessary to establish whether the process used to select the words had any effect on the results.

In addition to conducting more thorough testing and statistically based analysis of this method. After analyzing the results, it will be possible to create guidelines about how effective using English based engines to recognise Arabic speech.

8.3.3 Covering other languages and accents.

Another area of improvement is the scope of the transliteration covered. While Arabic script has been adopted by some other languages, such as Urdu and Persian, this research supports only Arabic.
Characters representing non-Arabic consonants and scheme of transliteration can be considered in future works.

References R
....

References

AbdulJaleel, N., and Larkey, L.S., 2003. *English to Arabic Transliteration for Information Retrieval: A Statistical Approach* [Online], November 3-8, Louisiana: USA. Available at: http://ciir.cs.umass.edu/pubfiles/ir-261.pdf [Accessed: 16 January 2007]

Abdulmun'im, N., 1993. *Sibawayh the Phonologist*: *A Critical Study of the Phonetic and Phonological Theory of Sibawayh As Presented in His Treatise Al-Kitab*, Kegan Paul Intl.

AbuZeina, D., and Elshafei, M., 2012. *Cross-Word Modeling for Arabic Speech Recognition*. New York: Springer.

Agaram, K.K., Keckler, S.W., Burger, D., 2001.*Characterizing the SPHINX Speech Recognition System*, [online] University of Texas at Austin, Technical Report TR2001-18. Available at: http://www.research.ibm.com/acas/projects/00agaram.pdf [Accessed: 16 January 2010]

Ajeeb, 2010. *Tarjim Dictionary*. [Online] Ajeeb. Available at: http://tarjim.ajeeb.com/ajeeb/ [Accessed: 24 July 2007].

Al-bab, 2009. *Arabic words and the Roman alphabet* [Online] Al-bab. Available at: http://www.al-bab.com/arab/language/roman1.htm [Accessed: October 2010].

Al Badrashiny, M., 2009. *Automatic Diacritizer for Arabic Texts* [online]. M.Sc. thesis, Cairo University Available at: http://www.rdi-eg.com/Downloads/ArabicNLP/Mohamed-Badashiny_MSc-Thesis_June2009.pdf [Accessed: October 2010].

Alghamdi, M., Muzaffar, Z., and Alhakami, H., 2010. Automatic Restoration of Arabic Diacritics: A Simple, Purely Statistical Approach. *The Arabian Journal for Science and Engineering*, Volume 35, Number 2C.

Alghamdi, M., 2009. *Romanizing Arabic Proper Names: Saudi Arabia Experience. Symposium Towards a Translation Standard of Arabic: Challenges and Solutions.* Abu Dhabi. 15-16 December 2009.

Alghamdi, M., and Zeeshan, M., 2007. KACST Arabic Diacritizer[Online]. The First International Symposium on Computers and Arabic Language. Available at: http://www.mghamdi.com/KAD.pdf [Accessed: 7 January 2008]

Alghamdi, M., Alsalman, A., Alshamsan, A., Almuhanna, F., Salih, M., Alwayili, M., Alhuqayl, K., Alsubai, S., 2006. *Romanization System for Arabic Names*: *Final Repot* [Online]. Available at: http://www.mghamdi.com/ANRS.pdf [Accessed: 6 June 2007]

Alghamdi, M., Khursheed, M., Elshafei, M., Alhargan, F., Alkanhal, M., Alshamsan, A., Alqahtani, S., Muzaffar, S., Altowim, Y., Yusuf, A., and Al-Muhtaseb, H. 2006 Automatic Arabic Text Diacritizer (Final Report) KACST, 2006. (In Arabic)

Alghamdi, M., 2005. Algorithms for Romanizing Arabic names. *Journal of King Saud University*, Computer Sciences and Information, pp.17, 1-27.

Algamdi, M., 2003. *KACST Arabic Phonetics Database*. The Fifteenth International Congress of Phonetics Science. Barcelona, 3109-3112, 2003.

Al-Onaizan, Y., and Knight, K., 2002. *Machine Transliteration of Names in Arabic Text*. ACL Workshop on Computational Approaches to Semitic Languages. [Online]. Available at: http://acl.ldc.upenn.edu/W/W02/W02-0505.pdf [Accessed: 9 June 2008]

Alotaibi, Y.A., 2003. *High performance Arabic digits recognizer using neural networks*, Proceedings of the International Joint Conference on Neural Networks (24 July), pp. 670-674 vol.1.

Al-Otaibi, F., 2001. *Speaker-Dependant Continuous Arabic Speech Recognition*, M.Sc. Thesis, King Saud University.

Alshamsan, I., 2004. *Arabic Text Diacritization* [online], Diacritization Rules. Available at:
http://www.mghamdi.com/AATD.pdf [Accessed: November 2007]

Alumäe, T. and Võhandu, L., 2004. Limited-Vocabulary Estonian Continuous Speech Recognition System using Hidden Markov Models. *INFORMATICA*, 2004, Vol. 15, No. 3, pp. 303–314.

Ananthakrishnan, S., Narayanan, S. and Bangalore, S., 2005. *Automatic Diacritization of Arabic Transcripts for Automatic Speech Recognition*. In Proceedings of International Conference on Natural Language Processing, Kanpur, India.

Arabic Transliteration Wiki, 2012. *Arabic Transliteration* [website]. Available at: http://arabic-transliteration.software.informer.com/wiki/ [Accessed: February 2012]

Arbabi, M., Fischthal, S., Cheng, V., and Bart, E., 1994. Algorithms for Arabic name transliteration. *IBM Journal of Research and Development*, 38(2):183-193.

Atkielski, A., 2005. *Using Phonetic Transcription in Class* [online]. Available at: http://www.atkielski.com/ESLPublic/Phonetics%20%20Using%20Phonetic%20Transcription%20in%20Class.pdf [Accessed: February 2013]

Bazzi, I. and Glass, J., 2000. *Modeling Out-Of-Vocabulary Words for Robust Speech Recognition. Proceedings of International Conference on Spoken Language Processing (ICSLP 2000)*, Beijing, China, pp.401-404.

Baumgarten, J. A., Barksdale, K., Rutter, M. and Barksdale, K., 2000. *IBM ViaVoice Recognition Software: Quicktorial*. South-Western Educational Publishing.

Ba'Albaki, R., 1998. Al-Mawrid: A Modern Arabic-English Dictionary. Dar Ilm Lil Malayin.

Becker, J., 1987. *Arabic word processing*. Communications of the ACM, 30 (7), PP 600-611.

Ben Sassi, S., Braham, R., Belghith, A., 2001. *Neural speech synthesis system for Arabic language using CELP algorithm*. ACS/IEEE international conference on computer systems and applications (AICCSA'01), pp 119-121, Lebanon.

Bevana, N., Kirakowskib, J., and Maissel, J., 1991. *What is Usability?*. Proceedings of the 4th International Conference on HCI, Stuttgart.

Billa, J., Noamany, M., Srivastava, A., Liu, D., Stone, R., Xu, J., Makhoul, J., Kubala, F., 2002. *Audio indexing of Arabic broadcast news*. Proceedings. (ICASSP '02) IEEE International Conference on Acoustics, Speech, and Signal Processing, pp.I-5 - I-8 vol.1

Buckwalter, T., 2004. *Buckwalter Arabic Morphological Analyzer (BAMA)*, Version 2.0, LDC Catalog number LDC2004L02 [Online], Available at: www.ldc.upenn.edu/Catalog [Accessed: 5 June 2007].

Buckwalter, T., 2002. *Buckwalter Arabic Morphological Analyzer Version 1.0*. Linguistic Data Consortium, University of Pennsylvania, LDC Catalog No.: LDC2002L49.

Cho, P., 2005. Takeluma: An Exploration of Sound, Meaning, and Writing [online]. MFA Thesis, UCLA Available at:
http://www.pcho.net/takeluma/takelumapaper.pdf [Accessed: April 2013]

Citizendium. 2011. American English [website]. Available at:
http://en.citizendium.org/wiki/American_English [Accessed: November 2012]

Collins English Dictionary, 2003. *Accuracy* [online]. Available at:
http://www.thefreedictionary.com/accuracy [Accessed: 24 July 2012].

Cook, S., 2002. *Speech Recognition* [online]. How To. Available at:
http://tldp.org/HOWTO/pdf/Speech-Recognition-HOWTO.pdf [Accessed: 24 July 2008]

Deb, P., Singh, N., Kumar, S., Rai, N., 2010. Offline Navigation System for Mobile Devices [Online]. *International Journal of Software Engineering & Applications (IJSEA)*, Vol.1, No.2, April 2010 . Available at:
http://www.airccse.org/journal/ijsea/papers/0410ijsea3.pdf [Accessed: 23 January 2013]

Dictionary.com, 2010. *Transliteration* [Online]. Dictionary.com. Available at:
http://dictionary.reference.com/browse/transliterate?qsrc=2446 [Accessed 23 July 2007]

Dialect Blog. 2012 *American Accents* [online blog]. Available at:
http://dialectblog.com/northamerican-accents/ [Accessed: December 2012]

Dobrovolsky, M., Katamba, F., 2008. *Phonetics: The Sounds of Language*[online].
Available at:
http://catalogue.pearsoned.co.uk/assets/hip/gb/uploads/Katamba9781405899307_Ch2.pdf
[Accessed: February 2013]

Doe, H., 1998. *Evaluating the Effects of Automatic Speech Recognition Word Accuracy*.
[Online] M.Sc. thesis. Virginia Polytechnic Institute and State University Available at:
http://scholar.lib.vt.edu/theses/available/etd-7598-165040/unrestricted/thesis1.pdf
[Accessed: 21 August 2007]

Du Bois, J., Schuetze-Coburn, S., Cumming, S., Paolino, D., 1993. *Outline of Discourse Transcription* [Online]. Available at: http://anthro.ucsd.edu/~jhaviland/AudVid/AudVidReadings/DuBoisDiscourseTrs.pdf [Accessed: April 2013]

Duchan F., 2006. The Phonetic Notation System of Melville Bell and its Role in the History of Phonetics. *Journal of Speech-Language Pathology and Audiology.* [Online] Spring 2006, Vol. 30, No. 1. . Available at: http://www.caslpa.ca/PDF/monthly_featured_articles/Spring_JSLPA_2006.pdf [Accessed: April 2012]

Dyslexia, 2011. *Dyslexia* [website], dyslexia, Available at: http://www.dyslexia.com [Accessed: 7 June 2007].

Edward, H., Jones, A., Zhang, Q., Rijmen, F., 2007. *Mixed-effects Hidden Markov Model* [Online]. Available at: http://www.phs.wfubmc.edu/public/downloads/MHMM_Ip.pdf [Accessed: 24 July 2008]

El-Imam, Y., 2004. Phonetization of Arabic: Rules and Algorithms. *Computer Speech and Language*, 18(4).

El-Sadany, T. and M. Hashish (1989). An Arabic morphological system. *IBM System Journal*, 28/4.

Elshafei, M., Ali, M., Al-Muhtaseb, H., and Al-Ghamdi, M. 2007. *Automatic segmentation of Arabic speech. Workshop on information technology and Islamic sciences*, Imam Mohammad Ben Saud University, Riyadh, March 2007.

Elshafei , M., Almuhtasib, H., and Alghamdi, M., 2006. *Statistical Methods for Automatic Diacritization of Arabic text*, Proceedings of the 18th National computer Conference NCC'18, Riyadh.

Elshafei, M., Almuhtasib, H., and Alghamdi, M., 2002. *Techniques for High Quality Text-to-speech*, *Information Science*, 140 (3-4) 255-267.

Emam, O. and Volker, F., 2005. Hierarchical Approach for the Statistical Vowelization of Arabic Text. *Technical report*, IBM Corporation Intellectual Property Law, Austin, TX, US.

Frankfurt International School, 2012. *The differences between English and Arabic* [website] FIS. Available at: http://esl.fis.edu/grammar/langdiff/arabic.htm [Accessed: April 2013]

Franz, A., and Milch, B., 2002. *Searching the Web by Voice. Proceedings of 19th International Conference on Computational Linguistics*, Taipei, Taiwan, China, (2), pp.1213-1217.

Glass, R. and Hazen, T J., 1998. *Telephone-Based Conversational Speech Recognition in the JUPITER Domain, In Proceedings of the Fifth International Conference on Spoken Language Processing*, pp. 1327-1330, December 1998.

Grasso, M., 2005. *Speech Input in Multimodal Environments: Effects of Perceptual Structure on Speed, Accuracy, and Acceptance.* PhD Thesis, University of Maryland, Baltimore, USA, Available at: http://ebiquity.umbc.edu/_file_directory_/papers/192.pdf [Accessed 12 November 2006].

Gupta, R., 2005. *Speech Recognition for Hindi.* [online]. M.Tech thesis, The Centre for Development of Advanced Computing, Mumbai Available at: http://www.cdacmumbai.in/design/corporate_site/override/pdfdoc/Speech_Recognition_for_Hindi.pdf [accessed: February 2013]

Habash, N., and Rambow, O., 2007a. *Arabic Diacritization through Full Morphological Tagging.* Proceedings of NAACL HLT 2007, [online]. Companion Volume, pages 53–56, Rochester, NY, April 2007. Available at: http://acl.ldc.upenn.edu/N/N07/N07-2014.pdf [accessed: November 2012]

Habash, N, Soudi, A., and Buckwalter, T., 2007b. *On Arabic Transliteration, In Arabic Computational Morphology: Knowledge-based and Empirical Methods.* Soudi, Abdelhadi; van den Bosch, Antal; Neumann, Günter (Eds.), 2007.

Halpern, J., 2007. The *Challenges and Pitfalls of Arabic Romanization and Arabization.*[Online]. Available at: http://www.cjk.org/cjk/arabic/arannana.pdf [Accessed: 15 June 2008].

Haque, M., Azad, M., Mahabubuzzaman, A., 2010. *Designing and Manufacturing a Voice Control Switching System of Electrical Devices* [Online]. ISSN-1997-2571 J. Innov. Dev.Strategy 4(2):23-27(December2010). Available at: http://ggfagro.com/books/JIDS/JIDS%20Vol4%20Issue2/MIN-173%20%2823-27%29.pdf [Accessed: 12 February 2013].

Hieronymus, J.,1993. *ASCII Phonetic Symbols for the World's Languages* [Online]. Available at: http://www.stanford.edu/class/cs224s/worldbet.pdf [Accessed: May 2010].

Hussein, M., 1998. Arabic string searching in the context of character code standards and orthographic variations. *Computer Standards & Interfaces*, (20,1), 16 November, pp. 31-51

IBM, 2006. *Embedded ViaVoice Multiplatform Edition* [Online]. IBM. Available at: http://www306.ibm.com/software/pervasive/embedded_viavoice_multiplatform/ [Accessed 12 July 2007]

IPA The, 2003. *International Phonetic Association* [website] IPA. Available at: http://www.langsci.ucl.ac.uk/ipa/ [Accessed: May 2008]

Ismail, S., and Ahmad, A., 2004. *Recurrent Neural Network with Back propagation through Time Algorithm for Arabic Recognition.* Proceedings of the 18th European Simulation Multiconference, SCS Europe.

Jurafsky, D., and Martin, J., 2009. *Speech and Language Processing: An Introduction to Natural Language Processing*, Speech Recognition, and Computational Linguistics. 2nd edition. Prentice-Hall.

Karimi, S., Scholer, F., and Turpin, A., 2007. *Collapsed consonant and vowel models: New approaches for English-Persian transliteration and back-transliteration*. In *Proceedings of the Annual Meeting of the Association of Computational Linguistics (ACL2007), pages 648-655.*

Kemble, K. A., 2001. *An Introduction to Speech Recognition*. [Online] *Voice System Middleware Education* - IBM Corporation. Available at: ftp://ftp.software.ibm.com/software/pervasive/info/products/Introduction_to_Speech_Reco gnition.pdf [Accessed: 24 July 2007].

Kirchhoff, K., Vergyri, D., Bilmes, J., Duh, K., and Stolcke, A., 2004. *Morphology-based language modeling for Arabic speech recognition.* Proceedings of ICSLP 2004, Jeju, South Korea.

Kirchhoff, K., Bilmes, J., Das, S., Duta, N., Egan, M., Ji, G., He, F., Henderson, J., Liu, D., Noamany, M., Schoner, P., Schwartz, R., and Vergyri, D., 2003. *Novel Approaches to Arabic Speech Recognition*. Report from the 2002 John-Hopkins Summer Workshop", ICASSP 2003, pp. I-344-I347.

Knight, K. and Graehl, J., 1998. *Machine Transliteration*. Computational Linguistics. 24(4):599–612

Krauss, R., 2002. The Psychology of Verbal Communication [online]. *The International Encyclopedia of the Social and Behavioral*. Available at: http://www.columbia.edu/~rmk7/PDF/IESBS.pdf [Accessed: May 2011].

Lagally, K., 2004. *ArabTEX, Typesetting Arabic and Hebrew*, User Manual Version 4.00. (11 March) Report Nr. 2004/03, [online] University at Stuttgart, Fakult at Informatik. Available at: http://www.scribd.com/doc/47494374/Arabtex-Typesetting-Arabic-and-Hebrew [Accessed April 2007]

Lawson David R., 2008. *An Evaluation of Arabic Transliteration Methods*, [online] M.Sc. thesis, University of North Carolina at Chapel Hill. Available at: http://dc.lib.unc.edu/cdm/singleitem/collection/s_papers/id/1061 [Accessed February 2013].

Lee, J., and Choi, K., 1998. English to Korean Statistical transliteration for information retrieval. *Computer Processing of Oriental Languages* [Online]. 12(1):17-37. Available at: http://dev.swrc.kaist.ac.kr/paper/15.pdf [Accessed: 24 August 2008]

Lee, K., Hon, H., and Reddy, R., 1990. An overview of the sphinx speech recognition system. *IEEE Transactions of Acoustics, Speech, and Signal Processing*, Vol. 38 (1), 35-45.

Lewin, I., Russell, M., Carter, D., Browning, S., Ponting, K., and Pulman, S., 1993. *A speech-based route enquiry system built from general purpose components*, in proceedings of the 3rd European Conference on Speech Communication and Technology, Berlin, Germany, September 1993.

Lin, B., Chen, B., Wang, H., and Lee, L., 2002. *A hierarchical tag-graph search scheme with layered grammar rules for spontaneous speech understanding* [Online]. Pattern Recognition Letters 23 (2002) 819–831. Available at: http://ntur.lib.ntu.edu.tw/bitstream/246246/142095/1/12.pdf [Accessed: 6 June 2007].

Mansikkaniemi, A., 2010. *Acoustic Model and Language Model Adaptation for a Mobile Dictation Service*. [Online] M.Sc. thesis, Available at: http://lib.tkk.fi/Dipl/2010/urn100143.pdf [Accessed: November 2012].

Micher, J., and Voss, C., 2008. *Buckwalter-based Lookup Tool as Language Resource for Arabic Language Learners*, Software Engineering, Testing, and Quality Assurance for Natural Language Processing, Association for Computational Linguistics, pages 66–67. Columbus, Ohio, USA, June 2008.

Mubarak, H., Al Sharqawy, M., Al Masry, E., 2005. *Diacritization and Transliteration of Proper Nouns from Arabic to English*. [Online] Sakhr Software. Cairo, Egypt. Available at: http://www.mt-archive.info/MEDAR-2009-Mubarak.pdf [Accessed: 23 April 2008]

MSDN, 2012. *About Lexicons and Phonetic Alphabets (Microsoft. Speech)* [Online] MSDN. Available at: http://msdn.microsoft.com/en-us/library/hh378451(v=office.14).aspx [Accessed: 23 April 2013].

Natural Reader, 2012. *Natural reader* [website]. Available at: http://www.naturalreaders.com/index.php [Accessed: June 2008].

Nelken, R., and Shieber, S., 2005. *Arabic diacritization using weighted finite-state transducers*. In Proceedings of the 2005 ACL Workshop on Computational Approaches to Semitic Languages, pages 79-86, Ann Arbor, Michigan, June 2005

Niculescu, A., and De Jong, F., 2008. *Development of a Speech Recognition System for Spanish Broadcast News*, CTIT-technical Report, version 1.0, January 2008.

Noeman, S., 2009. *Language Independent Transliteration system using phrase based SMT approach on substrings*. In Proceedings of the 2009 Named Entities Workshop, ACL-IJCNLP 2009, pages 112-115.

Nuance, 2006. *Dragon NaturallySpeaking*. [Online]. Nuance. Available at: http://www.nuance.com/naturallyspeaking/ [Accessed: 7 July 2007].

Peacocke, R., and Graf, D., 1990. An Introduction to Speech and Speaker Recognition. *Journal of IEEE Computer*, 23(8), pp.26-33.

Pedersen, T., 2008. *Transliteration of Arabic* [online]. Available at: http://transliteration.eki.ee/pdf/Arabic_2.2.pdf [Accessed: July 2011].

Petrie, G., 2003. *SPEECH RECOGNITION SOFTWARE: ITS POSSIBLE IMPACT ON THE LANGUAGE LEARNING CLASSROOM* [online]. Teaching English with Technology, vol. 3, no. 3, pp. 40-48. Available at: http://www.iatefl.org.pl/call/callnl.htm [Accessed: February 2012]

Peissner, M., 2002. *What the Relationship between Correct Recognition Rates and Usability Measures Can Tell Us about the Quality of a Speech Application*. Proceedings of 6th International Scientific Conference on Work with Display Units, Berchtesgaden, Germany, PP. 296-298.

Philips, 2005. *Philips Dictation Systems* [Online]. Philips. Available at: http://www.dictation.philips.com/index.php?id=start [Accessed: 12 November 2005].

Phonemic Chart, 2013. *English Club* [website]. Available at: http://www.englishclub.com/pronunciation/phonemic-chart.htm [Accessed: February 2012]

Pouliquen, B., Steinberger, R., Ignat, C., Temnikova, I., Widiger, A., Zaghouani, W., and Žižka, J., 2005. *Multilingual person name recognition and transliteration*.[online] Available at: http://arxiv.org/ftp/cs/papers/0609/0609051.pdf [Accessed: 12 April 2013]

Qiu, L., 2011. *British English vs. American English* [website]. Available at: http://www.scribd.com/doc/95548934/British-and-American-English-Lingyu [Accessed: February 2012]

Rabiner, L. and Juang, B., 1986. *An Introduction to Hidden Markov Models. IEEE ASSP Magazine, pp. 4-16.*

Rambow, O., Chiang, D., Diab, M., Habash, N., Hwa, R., Sima'an, K., Lacey, V., Levy, R., Nichols, C., Shareef, S., 2006. *Parsing Arabic Dialects, Final Report* [online]. Version 1,
January 18, 2006. Available at: http://old-site.clsp.jhu.edu/ws05/groups/arabic/documents/finalreport.pdf [Accessed: July 2008]

Roe, D.B., and Wilpon, J. G., 1993. *Whither Speech Recognition*: The Next 25 Years, *IEEE Communications Magazine*, Nov, pp.54-62.

Sagheer, A., Tsuruta, N., Ichiro, R., & Maeda, S., 2005. *Visual speech features representation for automatic lip reading*. In IEEE International conference on acoustics, speech and signal processing (Vol. 2, pp. 781-784).

Sakhr, 2011, [online] diacritization. Available at: http://demo.sakhr.com/technology_a/diacritization/default.aspx?sec=Technology&item=Di acritization Accessed: January 2011]

Sakti, S., Kelana, E. Riza, H., Sakai, S. Markov, K. and Nakamura, S., 2007. *Development of Indonesian Large Vocabulary Continuous Speech Recognition System within A-STAR Project*.[online] Available at: http://aclweb.org/anthology/I/I08/I08-8004.pdf [Accessed: July 2010]

Saleem, S., Kao, C., Prasad, R., Choi, F., Natarajan, P., Stallard, D., Krstovski, K., Kamali, M., 2008. *Rapid Development of an English/Farsi Speech-to-Speech Translation System.* Proceedings of IWSLT 2008, Hawaii, USA.

Seneff, S., 2002. *Response Planning and Generation in the MERCURY Flight Reservation System*. Computer Speech and Language, (16), pp.283-312.

Shaalan, K., Abo Bakr, I., 2009. *A Hybrid Approach for Building Arabic Diacritizer*. Proceedings of the EACL 2009 Workshop on Computational Approaches to Semitic Languages, pages 27–35, Athens, Greece, 31 March, 2009.

Sherif, T., and Kondrak, G., 2007. *Bootstrapping a Stochastic Transducer for Arabic-English Transliteration Extraction* [Online]. Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics, pages 864–871, Prague, Czech Republic, June 2007. Available at: http://www.aclweb.org/anthology-new/P/P07/P07-1109.pdf [Accessed: 23 April 2008]

Shishtla, P., Ganesh, S., Subramaniam, S., Varma, V., 2009. *A Language-Independent Transliteration Schema Using Character Aligned Models* [Online]. NEWS 2009. Available at: http://www.mt-archive.info/NEWS-2009-Shishtla.pdf [Accessed: 23 July 2010]

Slavic information literacy, 2012. *Transliteration history* [Online], UOAL, Available at: http://intranet.library.arizona.edu/users/brewerm/sil/lib/transhist.html [Accessed: 6 December 2013].

Snack, 2006. *The Snack Sound Toolkit* [Online], KTH, Available at: http://www.speech.kth.se/snack/ [Accessed: 6 December 2005].

Srinivasamurthy, N., and Narayanan, S., 2003. *Language adaptive Persian speech recognition*. [online]In proceedings of Eurospeech 2003. Available at: http://iranianlinguistics.org/papers/LAAM.pdf [Accessed: July 2008]

Sugumaran, K., 2013. Speech recognition systems [online]. Available at: http://www.doc.ic.ac.uk/~nd/surprise_95/journal/vol1/ks4/article1.html [Accessed: April 2013]

Tebelskis, J., 1995. *Speech Recognition using Neural Networks* [online]. Ph.D. Thesis, Carnegie Mellon University. Available at: http://isl.anthropomatik.kit.edu/cmu-kit/english/2168_2309.php [Accessed: 21 April 2013]

The British Academy, 1917. *The British Academy Transliteration Of Arabic And Persian*. From the Proceedings of the British Academy, Vol, VIII] London.

Tomokiyo, M., Black, A., and Lenzo, K., 2003. *Arabic in my Hand: Small-footprint Synthesis of Egyptian Arabic*. Eurospeech 2003, Geneva, Switzerland.

Trost, H., 1991. Recognition and generation of word forms for natural language understanding systems. Integrating two-level morphology and feature unification, *Applied Artificial Intelligence*, v 5, n 4, October, pp. 411-457.

UIUC linguistics, 2007. Arabic Online: Diacritics [Online], UIUC, Available at: http://www.linguistics.uiuc.edu/ngurevic/ciber/samples/diacritics/index.html [Accessed: 6 June 2007].

UNESCO, 2006a. *Serbian- transliteration table* [Online], UNESCO, Available at: http://portal.unesco.org/culture/en/files/32321/11625496373serbian_en.pdf/serbian_en.pdf [Accessed: 6 July 2007]

UNESCO, 2006b. *Russian - transliteration table* [Online], UNESCO, Available at: http://portal.unesco.org/culture/en/files/32320/11625495633russian_en.pdf/russian_en.pdf [Accessed: 6 July 2007]

UNESCO, 2006c. *Arabic - transliteration table* [Online], UNESCO, Available at: http://portal.unesco.org/culture/en/ev.php-URL_ID=32265& URL_DO=DO _TOPIC& URL_SECTION=201.html [Accessed: 6 July 2007]

UOAL, 2010. *Transliteration history* [Online], UOAL, Available at: http://intranet.library.arizona.edu/users/brewerm/sil/lib/transhist.html [Accessed: 6 July 2007].

Viet-Bac, L., Besacier, L., Seng, S., Bigi, B., Do, T., 2007. *Recent Advances in Automatic Speech Recognition for Vietnamese, [Online].* Grenoble Cedex 9, FRANCE. Available at: http://www-clips.imag.fr/geod/User/laurent.besacier/Publis/sltu08-vn.pdf [Accessed: 24 June 2007]

Vimala, C., and Radha, V., 2012. A Review on Speech Recognition Challenges and Approaches. *World of Computer Science and Information Technology Journal (WCSIT)* ISSN: 2221-0741 Vol. 2, No. 1, 1-7, 2012.

Vollmann, R., Deutsch, W., Koechert, A., Moosmuller, S., Noll, A., Pribbenow, S., Schalhofer, J., *Some Aspects of Annotation of Sound Data within the Framework of the "Multimedia Language Documentation and Language Research Laboratory (MLL)"* [online]. Available at: http://www.mpi.nl/ISLE/documents/papers/Vollmann_paper.pdf [Accessed: 12 April 2013]

Wadhwani, O., KolheSanjay, A., Dekate, S., 2011. Recognition of Vernacular Language Speech for Discrete Words using Linear Predictive Coding Technique. *International Journal of Soft Computing and Engineering (IJSCE)* ISSN: 2231-2307, Volume-1, Issue-5, November 2011. [Online], Available at: http://ijsce.org/attachments/File/Vol-1_Issue-5/E0187091511.pdf [Accessed: 2 February 2013].

Walker, W., Lamere, P., Kwok, P., Raj, B., Singh, R., Gouvea, E., Wolf, P.,and Woelfel, J., 2004. *Sphinx-4: A Flexible Open Source Framework for Speech Recognition.* Technical report, Sun Microsystems Inc, Technical Report TR-2004-139

Whitaker, B., 2008. *Arabic words and the Roman alphabet.* [Online], al-bab, Available at: http://www.al bab.com/arab/language/roman1.htm [Accessed: 7 May 2007]

Whitaker, B., 2002. *Lost in translation* [Online] The guardian, Available at: http://www.guardian.co.uk/world/2002/jun/10/israel1[Accessed: 6 December 2012].

Wikipedia, 2010e. *Romanization* [Online] Wikipedia. Available at: http://en.wikipedia.org/wiki/Romanization [Accessed: 7 October 2010]

Zhang, Z., and Li, L., 2012. *Effects of Cultural Differences on Advertising Translation.* 2012 International Conference on Education Technology and Management Engineering Lecture Notes in Information Technology, Vols.16-17.

Zitouni, I., Sorensen, J., and Sarikaya, R., 2006. *Maximum Entropy Based Restoration of Arabic Diacritics* [online]. *Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the ACL*, pages 577–584, Sydney, July 2006. Available at: http://www.aclweb.org/anthology-new/P/P06/P06-1073.pdf [Accessed: June 2012]

Appendices

Appendix A

The International Phonetic Alphabet Chart

THE INTERNATIONAL PHONETIC ALPHABET (revised to 2005)

CONSONANTS (PULMONIC)   © 2005 IPA

	Bilabial	Labiodental	Dental	Alveolar	Post alveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b			t d		ʈ ɖ	c ɟ	k ɡ	q ɢ		ʔ
Nasal	m	ɱ		n		ɳ	ɲ	ŋ	N		
Trill	ʙ			r					R		
Tap or Flap		ⱱ		ɾ		ɽ					
Fricative	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Lateral fricative				ɬ ɮ							
Approximant		ʋ		ɹ		ɻ	j	ɰ			
Lateral approximant				l		ɭ	ʎ	L			

Where symbols appear in pairs, the one to the right represents a voiced consonant. Shaded areas denote articulations judged impossible.

CONSONANTS (NON-PULMONIC)

Clicks		Voiced implosives		Ejectives	
ʘ	Bilabial	ɓ	Bilabial	ʼ	Examples:
ǀ	Dental	ɗ	Dental/alveolar	pʼ	Bilabial
ǃ	(Post)alveolar	ʄ	Palatal	tʼ	Dental/alveolar
ǂ	Palatoalveolar	ɠ	Velar	kʼ	Velar
ǁ	Alveolar lateral	ʛ	Uvular	sʼ	Alveolar fricative

OTHER SYMBOLS

ʍ Voiceless labial-velar fricative   ɕ ʑ Alveolo-palatal fricatives
w Voiced labial-velar approximant   ɺ Voiced alveolar lateral flap
ɥ Voiced labial-palatal approximant   ɧ Simultaneous ʃ and x
ʜ Voiceless epiglottal fricative
ʢ Voiced epiglottal fricative
ʡ Epiglottal plosive

Affricates and double articulations can be represented by two symbols joined by a tie bar if necessary. k͡p t͡s

VOWELS

	Front	Central	Back
Close	i y	ɨ ʉ	ɯ u
	ɪ ʏ		ʊ
Close-mid	e ø	ɘ ɵ	ɤ o
		ə	
Open-mid	ɛ œ	ɜ ɞ	ʌ ɔ
	æ	ɐ	
Open	a ɶ		ɑ ɒ
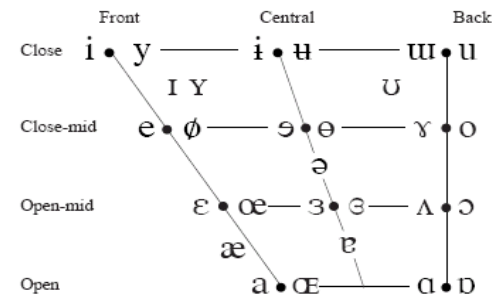
Where symbols appear in pairs, the one to the right represents a rounded vowel.

SUPRASEGMENTALS

ˈ Primary stress
ˌ Secondary stress   ˌfoʊnəˈtɪʃən
ː Long   eː
ˑ Half-long   eˑ
˘ Extra-short   ĕ
| Minor (foot) group
‖ Major (intonation) group
. Syllable break   ɹi.ækt
‿ Linking (absence of a break)

TONES AND WORD ACCENTS

LEVEL			CONTOUR		
e̋ or ˥	Extra high		ě or ˩˥	Rising	
é ˦	High		ê ˥˩	Falling	
ē ˧	Mid		e᷄ ˦˥	High rising	
è ˨	Low		e᷅ ˩˨	Low rising	
ȅ ˩	Extra low		e᷈	Rising-falling	
↓ Downstep			↗ Global rise		
↑ Upstep			↘ Global fall		

DIACRITICS

Diacritics may be placed above a symbol with a descender, e.g. ŋ̊

̥	Voiceless	n̥ d̥		̤	Breathy voiced	b̤ a̤		̪	Dental	t̪ d̪
̬	Voiced	s̬ t̬		̰	Creaky voiced	b̰ a̰		̺	Apical	t̺ d̺
ʰ	Aspirated	tʰ dʰ		̼	Linguolabial	t̼ d̼		̻	Laminal	t̻ d̻
̹	More rounded	ɔ̹		ʷ	Labialized	tʷ dʷ		̃	Nasalized	ẽ
̜	Less rounded	ɔ̜		ʲ	Palatalized	tʲ dʲ		ⁿ	Nasal release	dⁿ
̟	Advanced	u̟		ˠ	Velarized	tˠ dˠ		ˡ	Lateral release	dˡ
̠	Retracted	e̠		ˤ	Pharyngealized	tˤ dˤ		̚	No audible release	d̚
̈	Centralized	ë		̴	Velarized or pharyngealized	ɫ				
̽	Mid-centralized	e̽		̝	Raised	e̝	(ɹ̝ = voiced alveolar fricative)			
̩	Syllabic	n̩		̞	Lowered	e̞	(β̞ = voiced bilabial approximant)			
̯	Non-syllabic	e̯		̘	Advanced Tongue Root	e̘				
˞	Rhoticity	ə˞ a˞		̙	Retracted Tongue Root	e̙				

Figure 1 The International Phonetic Alphabet (Adapted from the official IPA chart) (IPA, 2003)

monophthongs				diphthongs			
iː sheep	ɪ ship	ʊ good	uː shoot	ɪə here	eɪ wait		
e bed	ə teacher	ɜː bird	ɔː door	ʊə tourist	ɔɪ boy	əʊ show	
æ cat	ʌ up	ɑː far	ɒ on	eə hair	aɪ my	aʊ cow	
p pea	b boat	t tea	d dog	tʃ cheese	dʒ June	k car	g go
f fly	v video	θ think	ð this	s see	z zoo	ʃ shall	ʒ television
m man	n now	ŋ sing	h hat	l love	r red	w wet	j yes

Phonemic Chart
voiced
unvoiced

Figure 2 Phonemic Chart (English Club, 2013)

Appendix B

....

Survey on developing an Arabic voice spelling alphabet

SURVEY

From:

Ghadeer Khalil (N0007771@ntu.ac.uk)

Sent: 20 November 2005 21:21:45

To:

I am conducting a survey on "Developing an Arabic voice spelling alphabet"

I would really appreciate it if you would complete this very simple and short questionnaire which would contribute tremendously to my research.
You can find the questionnaire by following this link:

http://FreeOnlineSurveys.com/rendersurvey.asp?id=129635

Much Appreciated

Ghadeer Khalil

Sample of the survey results

Simply type the first word that comes into your mind, which starts with the following letters.

No.	Response ID	أ	ب	ت	ث	ج	ح
1	814784						
2	814791	اسد	بطة	تفاحة	ثعلب	جمل	حصان
3	814948						
4	814954	أسد	بطة	تفاحة	ثعلب	جمل	حصان
5	814986	أرقام	بدر	تفاح	ثمرة	جاسم	حمد
6	814991	أحمد	بطة	تفاحة	ثعلب	جمل	حصان
7	815106	احمد	بدر	تامر	ثعلب	جمل	حسام
8	815147	ارنب	برتقال	تمساح	ثعلب	جمل	حمامة
9	815926	أرنب	بطة	تفاحة	ثور	جمل	حصان
10	816047		بطة				
11	816155	الله	بطة	تفاحة	ثمرة	جمل	حب
12	817131	أرنب	بطة	تفاحة	ثور	جمل	حمار
13	817852	أرنب	بطة	تفاحة	ثعلب	جمل	حمار
14	826145	ـأسد	بطة	تمساح	ثعلب	جدار	حمار
15	828725		بطة				
16	830307	أرنب	بطة	تمر	ثلج	جبل	حليب
17	830399	اسنان	بطة	تفاحه	ثمود	جزر	حبل
18	830414	أسد	بطة	تفاحه	ثعلب	جمل	حلاوة
19	830425	الوان	بطة	تلفون	ثلاجه	جبس	حرام
20	834302	ارنب	بطة	تفاحة	ثعلب	جام	حمد
21	834305	ارنب	بطة	تاج	ثلج	جمل	حليب
22	834306	ارنب	بطة	تمر	ثعلب	جمل	حمار
23	834309	ارنب	بطة	تفاحة	ثوب	جبن	حوت
24	834310	احمد	بطة	تفاحة	ثعلب	جوز	حمد
25	834315	احمد	بطة	توت	ثعلب	جمل	حوت
26	834320	ازهار	بطة	تفاحة	ثعلب	جمل	حورية
27	834322	امل	بطة	تمر	ثريا	جميل	حيدر
28	834326	ارنب	بطة	تلفون	ثوب	جمل	حلم
29	834337	احمد	بطة	تي شيرت	ثوب	جمل	حمار
30	834340	ارنب	بطة	توت	ثوب	جليد	حداد
31	836746	أرنب	بطة	تفاحة	ثلاجة	جزرة	حوت
32	838093	أنف	بطة	توت	ثعلب	جزر	حوت
33	838232	ارنب	بطة	تاريخ	ثور	جمل	حمار
34	842583	أرنب	بطة	توت	ثوب	جمل	حليب
35	843142	اسد	بطة	تفاحة	ثعلب	جمل	حامل

خ	د	ذ	ر	ز	س	ش	ص
خروف	دب	ذرة	رمان	زهرة	سمكه	شبكة	صقر
خروف	دب	ذرة	رمانه	زهرة	ساعة	شمس	صقر
خيال	داهية	ذمة	رمية	زهرة	ساعة	شراب	صيف
خبز	دب	ذئب	رمان	زرافة	سيارة	شبل	صبي
خلود	دودة	ذروة	رمان	زهرة	سمكة	شريف	صلى
خروف	دمعه	ذئب	ريشة	زرافة	سمك	شنطة	صديق
خروف	دلة	ذرة	رمان	زهرة	سمكه	شمس	صرصور
خائن	درس	ذهب	رحلة	زهرة	سفر	شمال	صديق
خاروف	دال	ذبابة	رمان	زهرة	سمك	شمس	صنوبر
خبز	دب	ذبانة	روبة	زبالة	سمجة	شوربة	صلعة
خنزير	دب	ذبابه	رمل	زوراق	ساعه	شاشه	صندوق
خروف	دندله	ذره	رمل	زيتون	سنافر	شمس	صرصور
خوخ	دكان	ذره	رسمه	زعفران	سم سم	شجره	صاروخ
خوخ	دلفين	ذيل	روب	زرافة	سمك	شامبانزي	صرصور
خلف	داله	ذره	ريش	زهرة	سهم	شعر	صدف
خوخة	دونت	ذبابة	رز	زمبابوي	سارة	شمس	صالح
خوخ	دم	ذيل	ربيع	زر	سمكة	شمس	صدفة
خل	دجاجة	ذرة	راديو	زهرة	سيارة	شيطان	صالون
خروف	دبوس	ذيل	رمال	زرافة	سحب	شرار	صالون
خروف	ديك	ذرة	رزان	زرافة	سنجاب	شجرة	صياد
خوخ	دلفين	ذرة	رمان	زيد	سرير	شمس	صوم
خديجة	دلال	ذيب	رزان	زينب	سالم	شيرين	صدقة
خليل	دمعة	ذيب	راشد	زهرة	سارة	شجون	صفاء
خليل	دلال	ذيب	رنا	زهاري	سيما	شذى	صدى
خنفساء	دب	ذيل	ريم	زهرة	سيارة	شمام	صديق
خالد	دعاء	ذهول	راس	زاد	سؤال	شمس	صباح
خرتيت	دمام	ذبابة	ريش	زرافة	سلمان	شرير	صنوبر
خالد	دائرة	ذيل	رأس	زرافة	سراب	شتاء	صياد
خيرية	دال	ذيب	روب	زرافة	سلطة	شوربة	صاروخ
خوخ	درج	ذرة	رياض	زهور	سمكة	شراع	صوت
خوخة	دودة	ذرة	رسالة	زرافة	سيارة	شبح	صرصور

ض	ط	ظ	ع	غ	ف	ق	ك
ضفدع	طائرة	ظرف	عين	غراب	فيل	قط	كلب
ضفدع	طائرة	ظرف	عنب	غزال	فيل	قلم	كتاب
ضفدع	طالب	ظلام	عالم	غابة	فاتورة	قلم	كف
ضفدع	طبل	ظرف	عين	غدير	فيل	قمر	كلب
ضرس	طاولة	ظب	عمر	غبي	فأس	قملة	كلب
ضبع	طائرة	ظبي	عين	غزال	فلفل	قلب	كلب
ضفدع	طائرة	ظرف	عنب	غزال	فيل	قلم	كلب
ضعيف	طيب	ظبي	عقل	غيم	فرح	قديس	كون
ضفدع	طاولة	ظهر	عسل	غزال	فرس	قلم	كرسي
ضفدع	طبل	ظب	عسل	غراب	فراولة	قنبلة	كلب
ضباب	طاوله		عين	غزال	فراش	قنبله	كف
ضفدع	طابور	ظرف	عيون	غدير	فطر	قلم	كلبه
ضفدع	طاوس	ظفر	عنب	غيمه	فلفل	قفل	كرت
ضفدع	طاولة		عيون	غراب	فهد	قطه	كراميل
ضفدع	طاوله	ظفر	عين	غزال	فانوس	قلم	كرتون
ضب	طلال	ظلام	عبدالله	غنم	فؤاد	قمر	كلب
ضفدع	طبل	ظلام	عنب	غراب	فاطمة	قلب	كرسي
ضاري	طيارة	ظبية	عنب	غباء	فراولة	قمر	كرسي
ضب	طاولة	ظفر	عسل	غنم	فراولة	قوس	كلاب
ضفدع	طراد	ظفيرة	علم	غزال	فهد	قطار	كهف
ضفدع	طاولة	ظرف	عيون	غزال	فراولة	قصيدة	كلب
ضروري	طلال	ظبي	علي	غراب	فاطمة	قارورة	كلب
ضفدع	طاهرة	ظابط	عيون	غيم	فاطمة	قمر	كنافة
ضب	طلب	ظمأ	علياء	غالية	فرس	قهر	كنافة
ضب	طماطم	ظفر	عجوز	غنم	فأر	قمر	كويت
ضد	طلب	ظهور	عكس	غامق	فاضي	قمر	كائن
ضربة	طبل	ظهر	عنبر	غدير	فراولة	قنبلة	كنغر
ضباب	طيارة	ظل	عمارة	غاز	فيل	قاموس	كنز
ضرس	طيارة	ظرف	عين	غدير	فراولة	قسم	كلب
ضياء	طاقة	ظرف	عنب	غزال	فهد	قلم	كتاب
ضفدع	طاووس	ظابط	عروس	غراب	فيل	قطار	كلب

ل	م	ن	هـ	و	ي
ليل	موز	نسر	هدهد	وردة	ياسمين
ليمون	موز	نسر	هدهد	وردة	يد
لعبة	محمد	نمل	هادئ	وادي	يكتب
ليمون	مدرسة	نمر	هدهد	وقت	يمامة
ليمون	منذر	نمر	همزة	واحد	يأكل
ليمون	موز	نور	هدهد	وزة	يد
لمبه	موز	نسر	هدهد	ورده	يد
لون	منزل	نجاح	هوس	وحدة	يتيم
ليمون	مسرح	نسر	هدهد	وزة	يد
ليت	منقة	نمر	هدهد	وردة	يرادة
ليمون	مدرسه	نحله	هواء	ورده	ياسمين
لومينا	منقا	نسوان	هدوء	ورده	ياسمين
ليمون	منجا	نور الدنيا	هدهد	ورل	يدرس
لوح	مها	نورس	هدهد	وطواط	ياربي
ليت	محايه	نعال	هريس	ورده	يحه
ليمون	مريم	نهر	همبرغر	وليد	ياسمين
ليمون	مريم	نار	هواء	وادي	يمامة
لولوة	مريم	نعامة	هدهد	وردة	يويو
ليمون	مجنون	نورس	هدهد	وسواس	يمامة
لاما	منصور	نور	هيثم	وداد	ياسمين
ليمون	محمد	نايف	هواء	وردة	ياسمين
لميس	مريم	نور	هاشم	وردة	يارا
لوسي	مها	نسر	هدى	وداد	ياسمين
لمى	موز	ندى	همس	وزن	ياسر
لولوة	موز	نرجس	هوز	وردة	ياسمين
لولوة	مزاج	نمر	هدى	ولد	ياسر
ليمون	مكة	نادي	هندي	وردة	ياسمين
ليل	منزل	نبات	هواء	وكر	يمين
ليمون	مشروم	نورس	هيا	وطواط	يمامة
لؤلؤ	مركب	نجمة	هدهد	ورد	ياسمين
لعبة	موزة	نمل	هدهد	وطواط	يمامة

Table 1- Sample of survey results

Appendix C

Voice speller application code

Voice speller application code
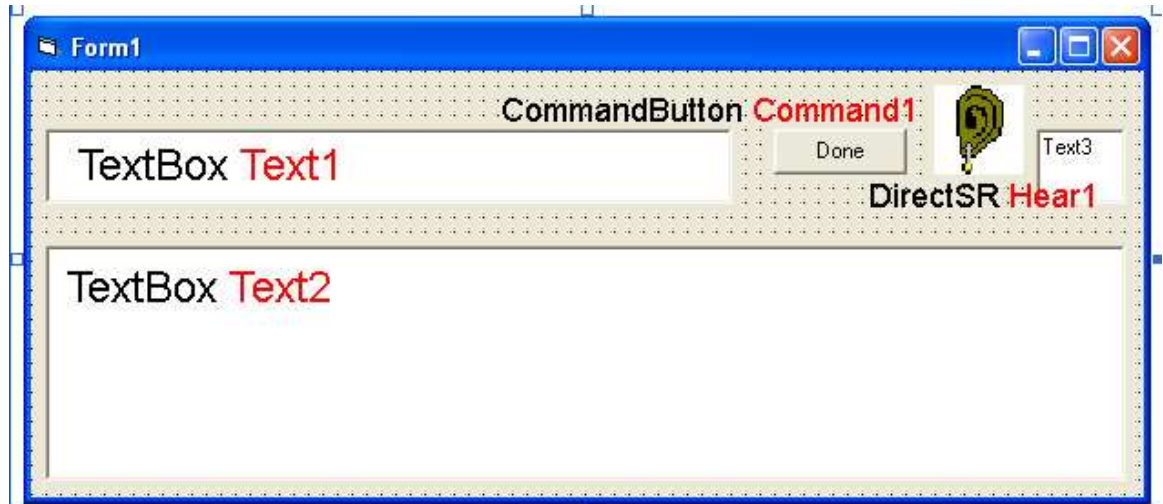
Figure 1-Snapshot of the Development application.

The voice speller application code is as follows:

```
Private Sub Command1_Click()
Dim DesktopDrive As String
Dim HomePath As String
Dim DesktopPath As String

' /* ***************************************** */
' /*     Get The Path To The Desktop          */
' /* ***************************************** */

' /* First, Get The Drive That The Profile Is On */
DesktopDrive = Environ("HOMEDRIVE")

' /* Then, Get The Path To Where The Profile Is Saved At */
HomePath = Environ("HOMEPATH")

' /* Now Make The DesktopPath Variable Equal The Drive, The Path To The Profile, And
\Desktop */
DesktopPath = DesktopDrive & HomePath & "\Desktop"


' /* *********************************** */
' /*       Write The Log File            */
' /* *********************************** */

' /* First, Open The File For APPEND (Add Stuff To It, Don't Over-write It) */
Open DesktopPath & "\logfile.txt" For Append As #1

 ' /* Then, Actually Write The Stuff In The Textbox, To The File */
```

```vb
    Print #1, Date & Time & Text2.Text

    '/* when command1 is clicked write to log file

    ' /* Anything That Is Opened, Must Be Closed */
    Close #1

    ' /* *********************************** */
    ' /*      Close The Program           */
    ' /* *********************************** */

    ' /* For All The Forms That Are In Our Project */
    For Each XFrm In Forms

        ' /* Unload The Form */
        Unload XFrm

    Next XFrm

    ' /* Code Should Never Reach Here, But Just In Case */
    End

End Sub

Private Sub Data1_Validate(Action As Integer, Save As Integer)
End Sub


' /* When the form is loaded

Private Sub Form_Load()

st$ = "[Grammar]" + vbNewLine
st$ = st$ + "type=cfg" + vbNewLine
st$ = st$ + "[<start>]" + vbNewLine
st$ = st$ + "<start>=arrnab" + vbNewLine
st$ = st$ + "<start>=boo staan" + vbNewLine
st$ = st$ + "<start>=tofah" + vbNewLine
st$ = st$ + "<start>=thoom" + vbNewLine
st$ = st$ + "<start>=jowz" + vbNewLine
st$ = st$ + "<start>=ham mama" + vbNewLine
st$ = st$ + "<start>=khaadim" + vbNewLine
st$ = st$ + "<start>=deek" + vbNewLine
st$ = st$ + "<start>=thee kkraa" + vbNewLine
st$ = st$ + "<start>=reeesh" + vbNewLine
st$ = st$ + "<start>=Zak kaah" + vbNewLine
st$ = st$ + "<start>=sakan" + vbNewLine
st$ = st$ + "<start>=shams" + vbNewLine
st$ = st$ + "<start>=soora" + vbNewLine
st$ = st$ + "<start>=dhameeer" + vbNewLine
```

```vb
st$ = st$ + "<start>=teen" + vbNewLine
st$ = st$ + "<start>=The laam" + vbNewLine
st$ = st$ + "<start>=aaali" + vbNewLine
st$ = st$ + "<start>=ghazal" + vbNewLine
st$ = st$ + "<start>=fa noos" + vbNewLine
st$ = st$ + "<start>=ghaa noon" + vbNewLine
st$ = st$ + "<start>=korrssay" + vbNewLine
st$ = st$ + "<start>=lee bas" + vbNewLine
st$ = st$ + "<start>=madrasa" + vbNewLine
st$ = st$ + "<start>=nasr" + vbNewLine
st$ = st$ + "<start>=hood hood" + vbNewLine
st$ = st$ + "<start>=waseela" + vbNewLine
st$ = st$ + "<start>=yas meen" + vbNewLine
st$ = st$ + "<start>=hamza" + vbNewLine
st$ = st$ + "<start>=space" + vbNewLine
st$ = st$ + "<start>=back space" + vbNewLine

' /* Hear1 is basically DirectSR which is represented by the ear icon.
' /* This is the procedure that will be called by DirectSR when it has finish processing a '/*
voice command.

' /* this is the parameter string that will contain the recognized word that was processed
'/* by the DirectSR  engine.


hear1.GrammarFromString st$
hear1.Activate
End Sub

Public Function GetSupportedThresholdValues() As Single()
GetSupportedThresholdValues = threshvalues
Text3.Text = threshvalues
End Function




Private Sub Hear1_PhraseFinish(ByVal flags As Long, ByVal beginhi As Long, ByVal beginlo
As Long, ByVal endhi As Long, ByVal endlo As Long, ByVal Phrase As String, ByVal parsed
As String, ByVal results As Long)

'/* One of the bad things about DirectSR is when the DirectSR recognizes a sound; it will

 '/* process the sound into a word that may closely match the one you provided.

 '/* If it matches, the Phrase variable will contain the matched word.

'/* Select Case code is used just for the Phrase variable for the words we fed in the engine in the load

 '/* function.

Select Case Phrase
   Case "arrnab"
```

```vb
        Text1.Text = Text1.Text & Chr$(199)
    Case "boo staan"
        Text1.Text = Text1.Text & Chr$(200)
    Case "tofah"
        Text1.Text = Text1.Text & Chr$(202)
    Case "thoom"
        Text1.Text = Text1.Text & Chr$(203)
    Case "jowz"
        Text1.Text = Text1.Text & Chr$(204)
    Case "ham mama"
        Text1.Text = Text1.Text & Chr$(205)
Case "khaadim"
        Text1.Text = Text1.Text & Chr$(206)
    Case "deek"
        Text1.Text = Text1.Text & Chr$(207)
    Case "thee kkraa"
        Text1.Text = Text1.Text & Chr$(208)
Case "reeesh"
        Text1.Text = Text1.Text & Chr$(209)
    Case "Zak kaah"
        Text1.Text = Text1.Text & Chr$(210)
    Case "sakan"
        Text1.Text = Text1.Text & Chr$(211)
Case "shams"
        Text1.Text = Text1.Text & Chr$(212)
    Case "soora"
        Text1.Text = Text1.Text & Chr$(213)
    Case "dhameeer"
        Text1.Text = Text1.Text & Chr$(214)
    Case "teen"
        Text1.Text = Text1.Text & Chr$(216)
    Case "The laam"
        Text1.Text = Text1.Text & Chr$(217)
     Case "aaali"
        Text1.Text = Text1.Text & Chr$(218)
    Case "ghazal"
        Text1.Text = Text1.Text & Chr$(219)
    Case "fa noos"
        Text1.Text = Text1.Text & Chr$(221)
     Case "ghaa noon"
        Text1.Text = Text1.Text & Chr$(222)
Case "korrssay"
        Text1.Text = Text1.Text & Chr$(223)
    Case "lee bas"
        Text1.Text = Text1.Text & Chr$(225)
    Case "madrasa"
        Text1.Text = Text1.Text & Chr$(227)
    Case "nasr"
        Text1.Text = Text1.Text & Chr$(228)
    Case "hood hood"
```

```vbnet
         Text1.Text = Text1.Text & Chr$(229)
      Case "waseela"
         Text1.Text = Text1.Text & Chr$(230)
      Case "yas meen"
         Text1.Text = Text1.Text & Chr$(237)
      Case "hamza"
         Text1.Text = Text1.Text & Chr$(198)
      Case "space"
         Text1.Text = Text1.Text & Chr$(32)
      Case "back space"
         Text1.Text = Delete
         Text2.Text = Chr$(8)
      End Select


If Phrase <> "" Then
  Text2.Text = Text2.Text & Phrase

Else
Text2.Text = Text2.Text & "No word matched"

End If
End Sub
```

Appendix D

Transliteration application code and process diagrams

Transliteration application code and process diagrams

Transliteration process

The transliteration application is saved in a folder that includes an .mdb file (Microsoft Access databasefile), a text file and an .xml file.

The .mdb file contains 2 tables and 3 forms available to the user, (table 1, tbDiaciritcs, form 1, form 2 and form 33).
The three forms are three alternative ways to transliterate Arabic words. Form 33 is the form used to prepare the information in the final experiments.
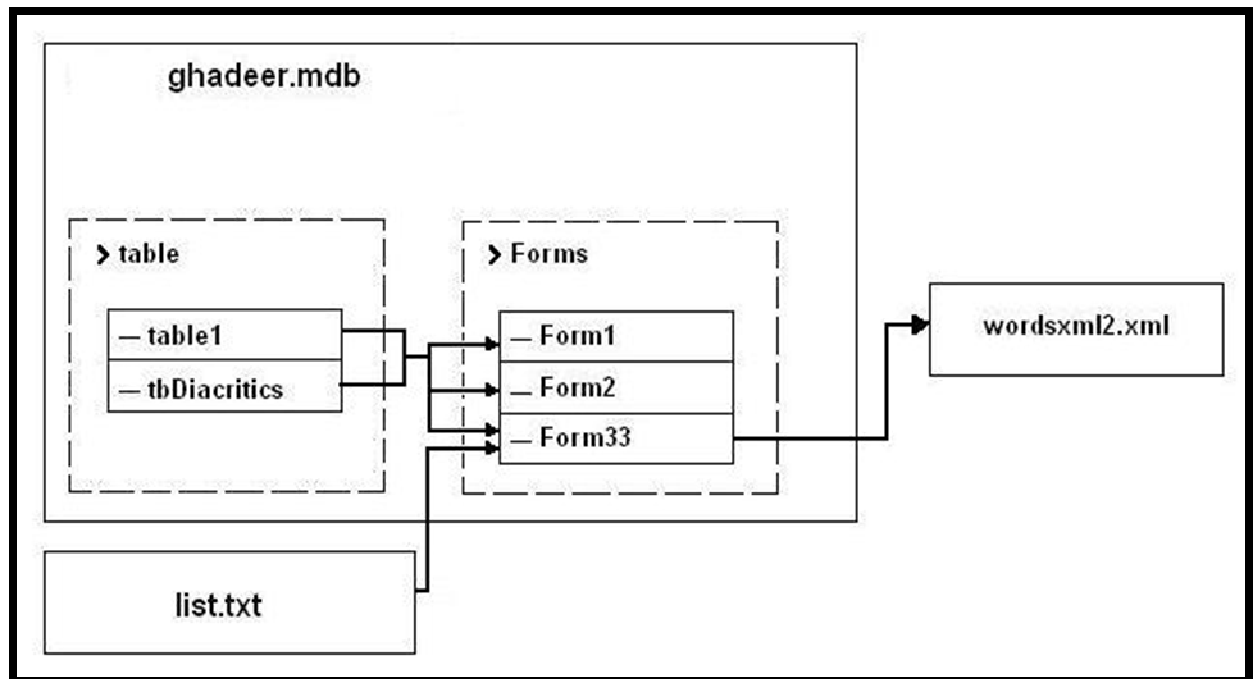
Figure 1- Files used in transliteration process

Figure 2-Sanpshot of Table 1

Table 1: contains the main transliteration table, the 2 columns in the middle represents Arabic letters and their transliterations according to any transliteration table, this can be changed to test different transliteration tables.
The second column from the right (Arabic column), has a listing of undiacritsed Arabic letters and each letter with diacritics, also transliteration of the diacritics separately are provided.

Figure 3-Snapshot of tblDiacritics

TblDiacritcs: Has a listing if all the Arabic letters and the diacritics in the 3 positions start, middle and end. This table allows the user to set some rules for the transliteration, for example it is not possible for baa to have tanween kasr in the start or middle.

After setting these rules, the user can enter the words or transliterate the text file using the following forms, according to the rules he/she already set using the previous tables.

Figure 4-Snapshot of form 1 (Basic form)

Form 1 (Basic form): Allows the user to type in a diacritised name, and shows the transliterated version of the name, moreover it has a text to speech facility that reads the transliterated word, so that the user can check whether the pronunciation is close to Arabic.
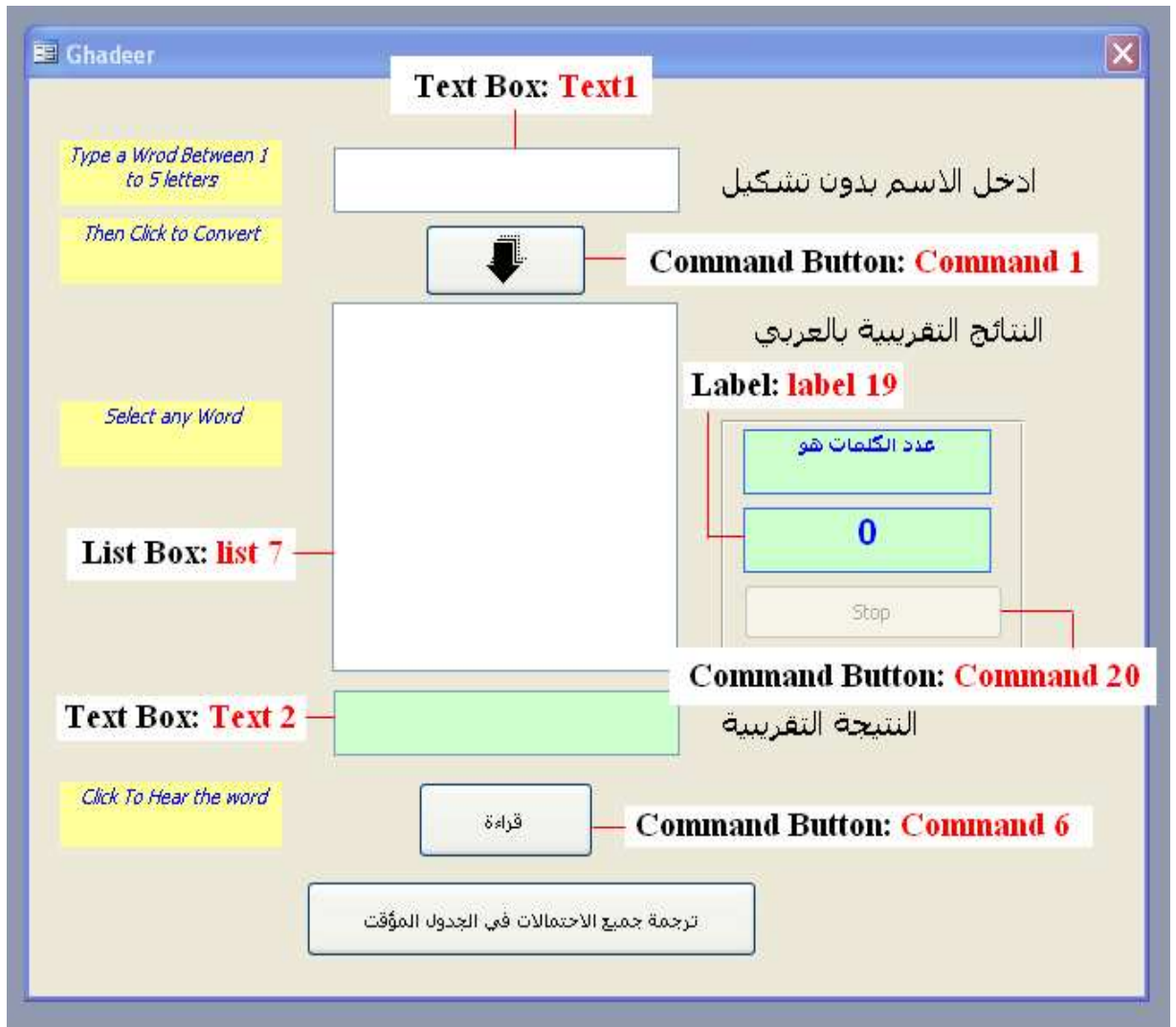
Figure 5-Snapshot of Form 2

Form 2: Is the upgraded form, it allows the user to enter any undiacritised word, and shows all the possible diacritised versions of this word. The TTS feature is also available.
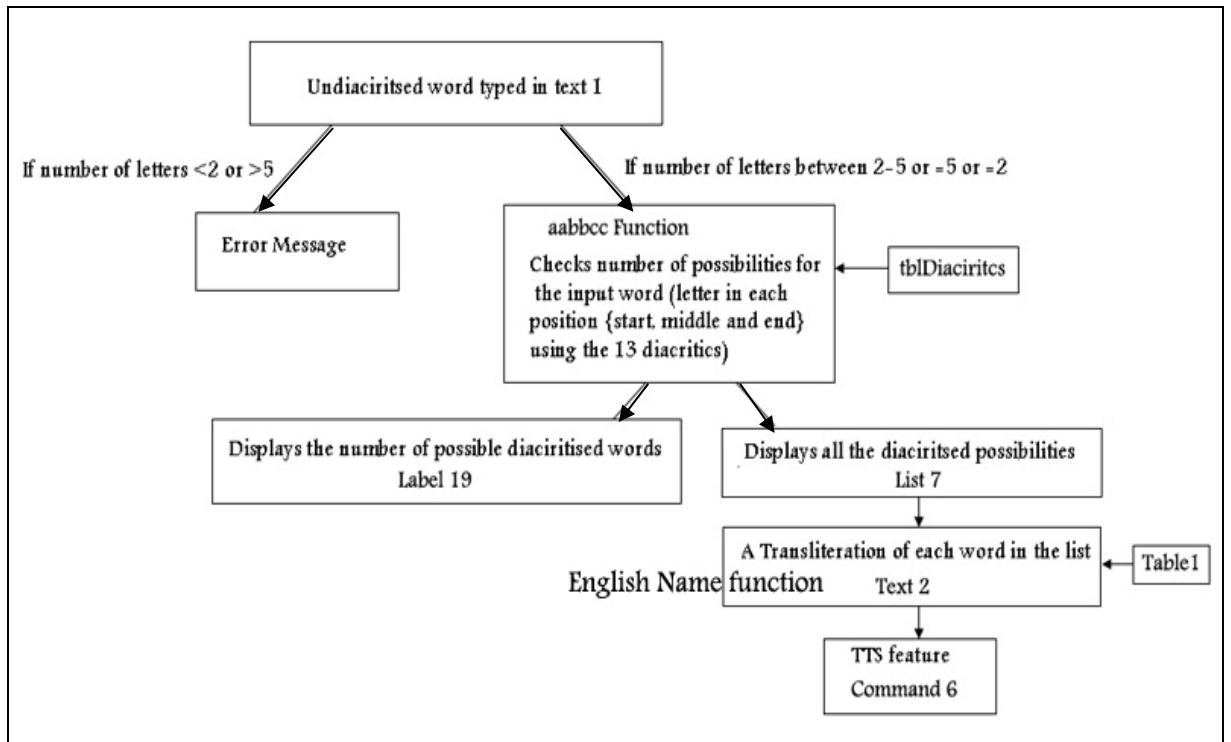
Figure 6-Transliteration process diagram

Figure 7-Snapshot of form 33

Form 33: The button in this form transliterates all the names or words in text file, and prepares an xml file, so that it can be used for the speech recognition process.
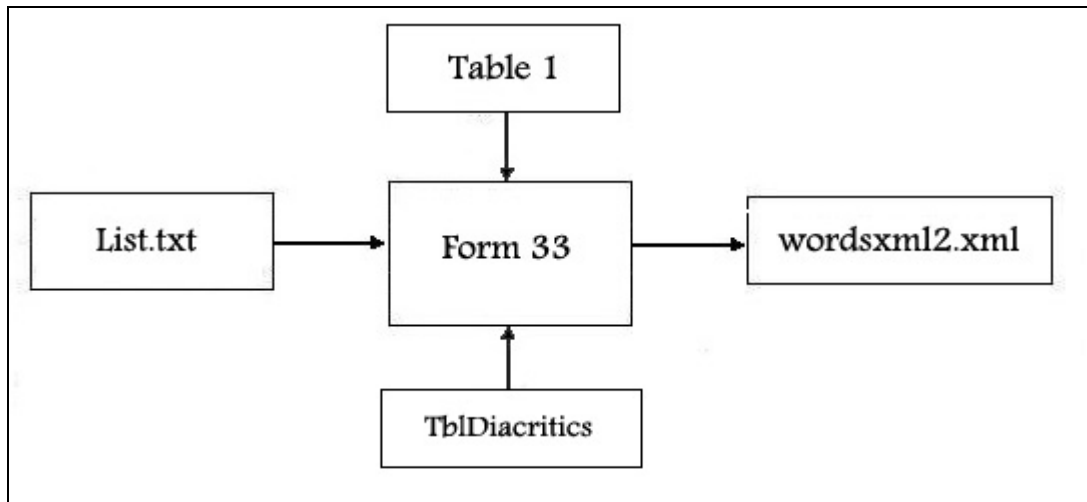
Figure 8- Diagram of the process of transliterating a list of words from a txt file to xml file
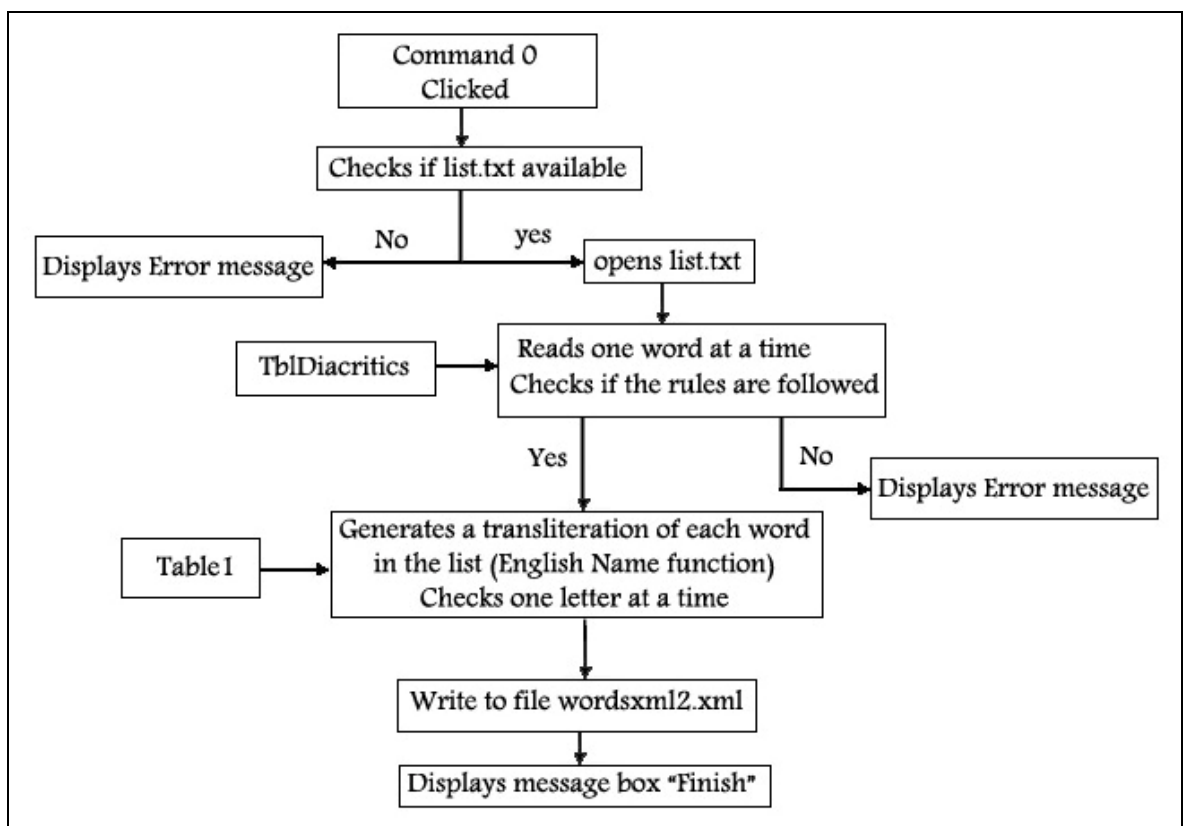
Figure 9-Transliteration list of words process Diagram

Figure 10-Snapshot of the Text file (list.txt)

This file contains a list of all the words that needs to be transliterated.

```
- <GRAMMAR LANGID="409">
  - <DEFINE>
      <ID NAME="RID_Rank" VAL="103" />
      <ID NAME="RID_Ranko" VAL="104" />
      <ID NAME="RID_Rankom" VAL="105" />
      <ID NAME="RID_Rankon" VAL="106" />
    </DEFINE>
  - <RULE NAME="rank" ID="RID_Rank" TOPLEVEL="ACTIVE">
    - <P>
      - <L>
          <P>DHaaAa</P>
          <P>ATHin</P>
          <P>Saagha</P>
          <P>zaaar</P>
          <P>qaas</P>
          <P>Aamal</P>
          <P>jathaa</P>
          <P>shaaaH</P>
          <P>Taaaf</P>
          <P>hayaAA</P>
          <P>kaAs</P>
          <P>Aukht</P>
          <P>baaada</P>
          <P>Aaw</P>
          <P>Aakala</P>
          <P>edhaa</P>
          <P>saAala</P>
          <P>DHuUul</P>
          <P>baisa</P>
          <P>baraAa</P>
          <P>suwuai</P>
```

Figure 11- Snapshot of The XML file (words2xml.xml)

The generated XML file works with the VB file to recognise the speech.

When the button is clicked in *form33*, the applications opens the text file *list.txt*, and checks one word at a time, to see if it follows the rules, that are ready set in *TblDiaciritcs*, for example it is not possible to add fat ha to the letter Fa in the middle, so it checks the words in the list, letters, diacritics, and positions (diacritems) to determine whether the rules are followed correctly, if not error message will be displayed.

Then the actual transliteration happens with the aid of *table1*, the application looks for the English equivalent for each character (letter or diacritic) from table 1.

Finally it writes the transliteration into *words2xml.xml*.

The code of the forms is as follow:

```vba
Option Compare Database
Option Base 1 ' to set the array base to be 1 not zero
Public arr1() As String
Public db As Database
Public rsTemp As Recordset

Sub aabbcc(xStart As String, xMiddle As String, xEnd As String, strWord As String)
On Error GoTo errnames1
    intLength = Len(strWord) ' integer value to store the length
    Dim rsMiddleOther() As Recordset ' this is an array recordset for saving the result of the
xmiddle string
    Dim rsStart As Recordset
    Dim rsMiddle As Recordset
    Dim rsEnd As Recordset
    Dim rs2 As Recordset
    xx = "Select * from tblDiacritics "
    xStart = xx & xStart & " Order by ID"
    xMiddle = xx & xMiddle & " Order by ID"
    xEnd = xx & xEnd & " Order by ID"
    'xx = xx & "(Letter_Asc_Code=223 AND Letter_Pos='Start') OR
(Letter_Asc_Code=202 AND Letter_Pos='Middle') OR (Letter_Asc_Code=200 AND
Letter_Pos='End')"
    Set db = CurrentDb() ' current database
    Set rs2 = db.OpenRecordset("select * from tblDiacritics where Letter_Asc_Code=999",
dbOpenSnapshot)
    Set rsStart = db.OpenRecordset(xStart, dbOpenSnapshot)
    Set rsMiddle = db.OpenRecordset(xMiddle, dbOpenSnapshot)
    Set rsEnd = db.OpenRecordset(xEnd, dbOpenSnapshot)
    Set rsTemp = db.OpenRecordset("tmpKeyWord")
    rsMiddle.MoveLast
    rsMiddle.MoveFirst
    rs2.MoveLast
    rs2.MoveFirst
If intLength > 3 Then
    ReDim rsMiddleOther(intLength - 3) ' redim array for new size
    For i = 1 To rsMiddle.RecordCount - 1 ' this loop to know all record in table for middle
character
        Set rsMiddleOther(i) = db.OpenRecordset(xMiddle, dbOpenSnapshot)
        'rsMiddleOther(i).Move i
    Next
    rsMiddle.MoveFirst
    Do While Not rsMiddle.EOF
        If rsMiddle("Letter_Name") = Mid(strWord, 2, 1) Then Exit Do ' if the character in
the table is found then exit
        rsMiddle.MoveNext
    Loop
    Do While Not rsMiddleOther(1).EOF
        If rsMiddleOther(1)("Letter_Name") = Mid(strWord, 3, 1) Then Exit Do
        rsMiddleOther(1).MoveNext
    Loop
If intLength = 5 Then
```

```
        Do While Not rsMiddleOther(2).EOF
           If rsMiddleOther(2)("Letter_Name") = Mid(strWord, 4, 1) Then Exit Do
           rsMiddleOther(2).MoveNext
        Loop
End If

Select Case intLength
Case 4
   For iStart = 1 To 13
      For iMiddle = 1 To 13
         For iMiddle2 = 1 To 13
            For iEnd = 1 To 13
               cntr = cntr + 1
               ReDim Preserve arr1(cntr) ' we save all the result in array
               arr1(cntr) = rsStart("Letter_Name") & IIf(rsStart("f" & iStart) = "Yes", rs2("f"
& iStart), "") & rsMiddle("Letter_Name") & IIf(rsMiddle("f" & iMiddle) = "Yes", rs2("f"
& iMiddle), "") & rsMiddleOther(1)("Letter_Name") & IIf(rsMiddleOther(1)("f" &
iMiddle2) = "Yes", rs2("f" & iMiddle2), "") & rsEnd("Letter_Name") & IIf(rsEnd("f" &
iEnd) = "Yes", rs2("f" & iEnd), "")
                  DoEvents
                  Form_Form2.Label19.Caption = Format((cntr / 28561), "Percent")
               Next iEnd
            Next iMiddle2
         Next iMiddle
      Next iStart
Case 5
   For iStart = 1 To 13
      For iMiddle = 1 To 13
         For iMiddle2 = 1 To 13
            For iMiddle3 = 1 To 13
               For iEnd = 1 To 13
                  cntr = cntr + 1
                  ReDim Preserve arr1(cntr)
                  arr1(cntr) = rsStart("Letter_Name") & IIf(rsStart("f" & iStart) = "Yes",
rs2("f" & iStart), "") & rsMiddle("Letter_Name") & IIf(rsMiddle("f" & iMiddle) = "Yes",
rs2("f" & iMiddle), "") & rsMiddleOther(1)("Letter_Name") & IIf(rsMiddleOther(1)("f" &
iMiddle2) = "Yes", rs2("f" & iMiddle2), "") & rsMiddleOther(2)("Letter_Name") &
IIf(rsMiddleOther(2)("f" & iMiddle3) = "Yes", rs2("f" & iMiddle3), "") &
rsEnd("Letter_Name") & IIf(rsEnd("f" & iEnd) = "Yes", rs2("f" & iEnd), "")
                     DoEvents
                     Form_Form2.Label19.Caption = Format((cntr / 371293), "Percent")

                  Next iEnd
               Next iMiddle3
            Next iMiddle2
         Next iMiddle
      Next iStart
End Select
Else
   If xMiddle = "Select * from tblDiacritics  Order by ID" Then
      For iStart = 1 To 13
         For iEnd = 1 To 13
               cntr = cntr + 1
```

```
            ReDim Preserve arr1(cntr)
            arr1(cntr) = rsStart("Letter_Name") & IIf(rsStart("f" & iStart) = "Yes", rs2("f"
& iStart), "") & rsEnd("Letter_Name") & IIf(rsEnd("f" & iEnd) = "Yes", rs2("f" & iEnd),
"")
            DoEvents
            Form_Form2.Label19.Caption = Format((cntr / 169), "Percent")
        Next iEnd
    Next iStart

    Else
    For iStart = 1 To 13
      For iMiddle = 1 To 13
        For iEnd = 1 To 13
            cntr = cntr + 1
            ReDim Preserve arr1(cntr)
            arr1(cntr) = rsStart("Letter_Name") & IIf(rsStart("f" & iStart) = "Yes", rs2("f" &
iStart), "") & rsMiddle("Letter_Name") & IIf(rsMiddle("f" & iMiddle) = "Yes", rs2("f" &
iMiddle), "") & rsEnd("Letter_Name") & IIf(rsEnd("f" & iEnd) = "Yes", rs2("f" & iEnd),
"")
          DoEvents
          Form_Form2.Label19.Caption = Format((cntr / 2197), "Percent")
        Next iEnd
      Next iMiddle
    Next iStart
    End If
End If
db.Execute "delete * from tmpKeyWord"
Form_Form2.Command20.Enabled = True
Form_Form2.StopLoop = False

For i = 1 To UBound(arr1)
   If Form_Form2.StopLoop Then Exit For
   rsTemp.AddNew
   rsTemp("KeyWord") = arr1(i) ' we show the result in list box in the form
   rsTemp.Update
   DoEvents
   Form_Form2.Label19.Caption = "Wait ... " & UBound(arr1) - i
Next i
   Form_Form2.Command20.Enabled = True
   Form_Form2.StopLoop = False
   MsgBox "Finish Function"
Exit Sub
errnames1:
If Err.Number = 9 Then
   ReDim arr1(cntr)
   Resume Next
ElseIf Err.Number = 3022 Then
   Resume Next
Else
   MsgBox "حصل خطأ ارجو الكتابة بطريقة اخرى"
End If
End Sub
--
```

```
Function GetA(strChar As String, strPos As String) As Boolean
On Error GoTo errnames1
'Dim arr1() As String
Dim db As Database
Dim rs As Recordset
Dim rs2 As Recordset
Set db = CurrentDb()
Set rs2 = db.OpenRecordset("select * from tblDiacritics where Letter_Asc_Code=999",
dbOpenSnapshot)
Set rs = db.OpenRecordset("select * from tblDiacritics where Letter_Pos='" & strPos & "'
and Letter_Asc_Code=" & Asc(strChar), dbOpenSnapshot)
rs.MoveLast
rs.MoveFirst
rs2.MoveLast
rs2.MoveFirst
cntr = 0
GetA = False
For Field_no = 1 To 13
   If rs("f" & Field_no) = "Yes" Then
      cntr = cntr + 1
      ReDim Preserve arr1(cntr)
      arr1(cntr) = strChar & rs2("f" & Field_no)
   End If
Next Field_no

If UBound(arr1) > 1 Then GetA = True

Exit Function
errnames1:
MsgBox "حصل خطأ ارجو الكتابة بطريقة اخرى"
'MsgBox Err.Description, , Err.Number
End Function
--

Function EnglishName(ArabicName As String, Optional HideError As Boolean) As String
On Error GoTo errnames
Dim db As Database
Dim rs As Recordset
Dim strTemp As String
Dim str1 As String
Dim str2 As String
Dim ifound As Boolean
Set db = CurrentDb()
Set rs = db.OpenRecordset("table1")

For i = 1 To Len(ArabicName)
   str1 = Mid(ArabicName, i, 1)
   If str1 = " " Then
      str2 = str1
      GoTo loop1
   End If
   GoTo loop2
```

```vba
loop1:
    If ifound Then
        strTemp = strTemp & str2
    End If

Next i
EnglishName = strTemp
Exit Function

loop2:
rs.MoveFirst
Do While Not rs.EOF
    If rs("Arabic") = str1 Then
        ifound = True
        If rs("english") = "xxx" Then
            str2 = str2
        ElseIf rs("english") = "xx" Then
            str2 = str2 & "n"
        Else
            str2 = rs("english")
        End If
        GoTo loop1
    End If
    rs.MoveNext
Loop
Exit Function
errnames:
If Not HideError Then MsgBox "حصل خطأ ارجو الكتابة بطريقة اخرى"
End Function
--
Option Compare Database

Sub WriteFile()
On Error Resume Next
Dim RSlist As Recordset
Dim db As Database
Set db = CurrentDb() ' current database
Set RSlist = db.OpenRecordset("select * from List", dbOpenSnapshot)

FileName = "wordsxml2.xml"
Open FileName For Output As #1

Print #1, "<GRAMMAR LANGID='409'>"

RSlist.MoveFirst
Do While Not RSlist.EOF' this part to know the length of the word max 5 letter
    Print #1, "<p>" & EnglishName(RSlist(0)) & "</p>"
    RSlist.MoveNext
Loop

Print #1, "</GRAMMAR>"

Close #1
```

```vba
MsgBox "Finish"
End Sub
--

Option Compare Database
Public StopLoop As Boolean
Private Sub Command11_Click()
On Error Resume Next
Dim xxStart As String 'this is a variable to save the first part of the SQL query
Dim xxMiddle As String 'this is a variable to save the middle part of the SQL query
Dim xxEnd As String 'this is a variable to save the last part of the SQL query
Command20.Enabled = False
StopLoop = False 'Boolian variable to exit from the loop
If Len(Text1) > 5 Or Len(Text1) < 1 Then
    MsgBox "Must the length between 1 and 5, in VER1"
    Exit Sub
End If
If Len(Text1) = 1 Then
    GetA Text1, "Start" 'if the length of the word is 1 the we will call the function GetA
Else
    'Text1 = "ßÊÈ"
    xxMiddle = "" ' reset the variables
    xxStart = ""
    xxEnd = ""
    strPos = "Start"
    xxStart = "(Letter_Asc_Code=" & Asc(Left(Text1, 1)) & " AND Letter_Pos='" & strPos
& "')"
    strPos = "End"
    xxEnd = "(Letter_Asc_Code=" & Asc(Right(Text1, 1)) & " AND Letter_Pos='" &
strPos & "')"
    If Len(Text1) > 2 Then
        For j = 2 To Len(Text1) - 1 ' this loop to take all the letters in word and make an
xxmiddle variable and save it as SQL query
            strPos = "Middle"
            If j <> 2 Then xx = " OR "
            xxMiddle = xxMiddle & xx & "(Letter_Asc_Code=" & Asc(Mid(Text1, j, 1)) & "
AND Letter_Pos='" & strPos & "')"
        Next j
         aabbcc "Where " & xxStart, "Where " & xxMiddle, "Where " & xxEnd, (Text1)
    Else ' if the length of the word = 2 then we will call the function with two variable
xxstart and xxend.
        aabbcc "Where " & xxStart, "", "Where " & xxEnd, (Text1)
    End If
End If
' this part is used after calling the function aabbcc or geta
Command20.Enabled = False
For i = 0 To List7.ListCount - 1
    DoEvents
    Label19.Caption = "Delete..." & i ' delete old data in the list
    List7.RemoveItem 0
Next i
StopLoop = False
```

```vb
Command20.Enabled = True
rsTemp.MoveFirst
cnt = 1
Do While Not rsTemp.EOF 'add new data in the list
   If StopLoop Then Exit Do
   DoEvents
   List7.AddItem rsTemp("KeyWord")
   Label19.Caption = "ADD..." & cnt
   rsTemp.MoveNext
   cnt = cnt + 1
Loop
StopLoop = False
Command20.Enabled = False
'For i = 1 To UBound(arr1)
 ReDim arr1(0) ' reset the array
End Sub


--
Private Sub Command20_Click()
StopLoop = True
End Sub
--
Private Sub Command25_Click()
DoCmd.SetWarnings False
DoCmd.OpenQuery "Qry_TransAll_Temp_word"
DoCmd.SetWarnings True
MsgBox "finish"
End Sub
--
Private Sub Command6_Click()
   Dim x As SpeechLib.SpVoice
   Set x = New SpeechLib.SpVoice
   x.Speak Text2
   Set x = Nothing
End Sub
--
Private Sub List7_Click()
   Text2 = EnglishName(List7)

End Sub
--
Option Compare Database

Private Sub Command4_Enter()
   Text2 = EnglishName(Text0)
End Sub
--
Private Sub Command6_Click()
   Dim x As SpeechLib.SpVoice
   Set x = New SpeechLib.SpVoice
   x.Speak Text2
   Set x = Nothing
End Sub
```

```
--
Option Compare Database

Sub WriteFile()
On Error Resume Next
Dim RSlist As Recordset
Dim db As Database
Set db = CurrentDb() ' current database
Set RSlist = db.OpenRecordset("select * from List", dbOpenSnapshot)

FileName = "wordsxml2.xml"
Open FileName For Output As #1

Print #1, "<GRAMMAR LANGID='409'>"

RSlist.MoveFirst
Do While Not RSlist.EOF
    Print #1, "<p>" & EnglishName(RSlist(0)) & "</p>"
    RSlist.MoveNext
Loop

Print #1, "</GRAMMAR>"

Close #1

MsgBox "Finish"
End Sub
--
```

Speech recognition process

An application has been developed in Microsoft Visual Basic and uses the Microsoft Speech SDK 5.1 to create an interface to the Microsoft English (U.S.) V6.1 Recognizer speech engine.
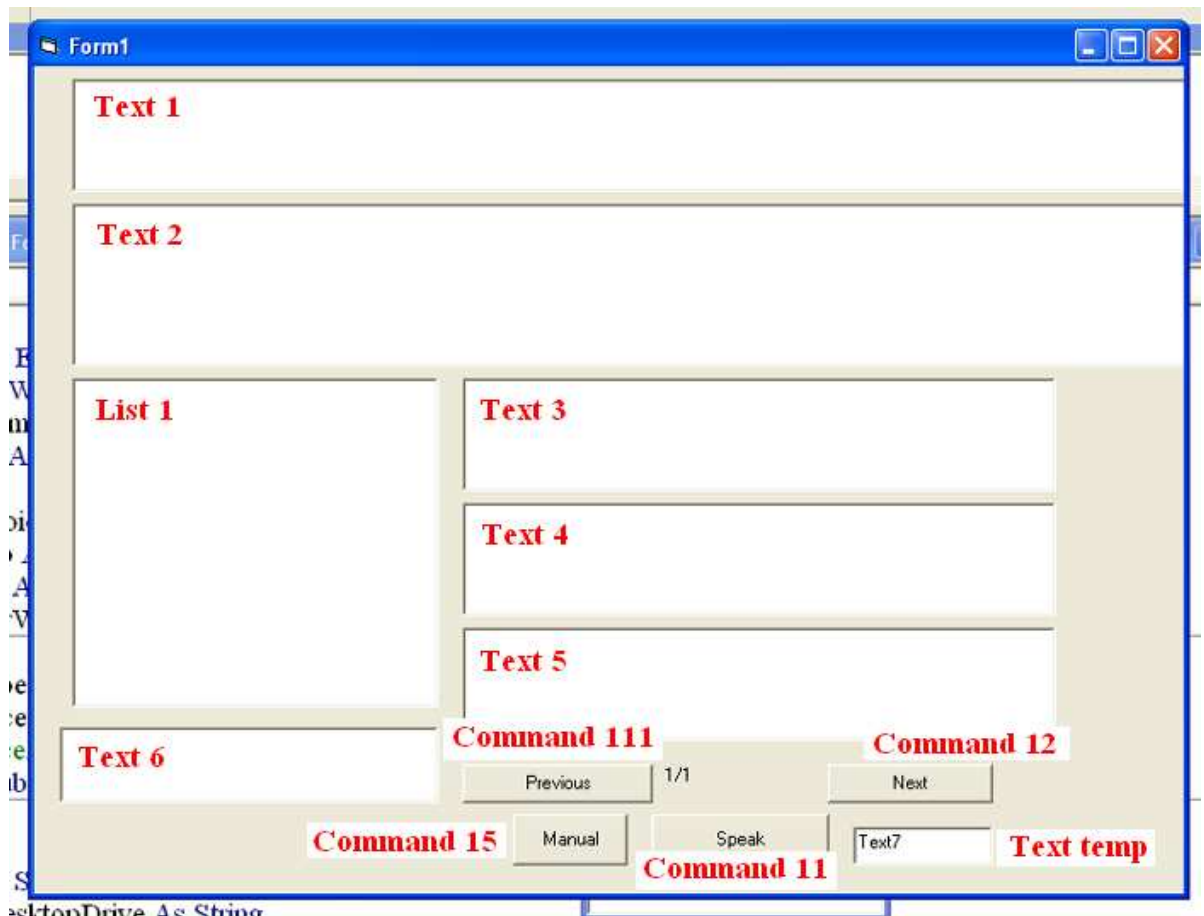
Figure 12- Snapshot of Form 1

This form, allows the user to speak into a microphone or play a set of recordings and the recognised words are recorded.
A manual and auto feature has been added to enable users to choose whether they want to play a set of recordings or choose a specific file or word to play.
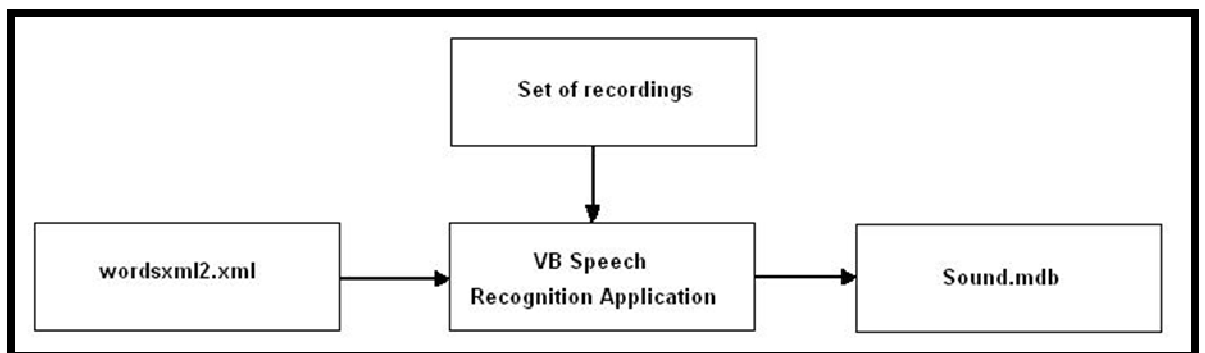
Figure 13-The speech recognition process

The speech recognition application takes the transliterated words saved in the xml file from the previous process and uses these with the set of recordings. The recognised words are saved in an .mdb file.

Code:

```
Option Explicit
Public WithEvents RC As SpSharedRecoContext
Public myGrammar As ISpeechRecoGrammar
Dim E As SpeechLib.ISpeechPhraseElement

Dim Voice As SpVoice
Dim db As Database
Dim rs As Recordset
Dim strWavPath As String

Sub SpeakVoice(filename As String)

   Voice.Speak filename, 15
 ' Voice.WaitUntilDone 1000000
End Sub


Private Sub Command1_Click()
Dim DesktopDrive As String
Dim HomePath As String
Dim DesktopPath As String


' /* ***************************************** */
' /*     Get The Path To The Desktop          */
' /* ***************************************** */

' /* First, Get The Drive That The Profile Is On */
DesktopDrive = Environ("HOMEDRIVE")

' /* Then, Get The Path To Where The Profile Is Saved At */
HomePath = Environ("HOMEPATH")

' /* Now Make The DesktopPath Variable Equal The Drive, The Path To The Profile, And
\Desktop */
DesktopPath = DesktopDrive & HomePath & "\Desktop"



' /* ************************************** */
' /*      Write The Log File               */
' /* ************************************** */

' /* First, Open The File For APPEND (Add Stuff To It, Don't Over-write It)

Open DesktopPath & "\logfile2.txt" For Append As #1

   ' /* Then, Actually Write The Stuff In The Textbox, To The File */
   Print #1, Date & Time & Text2.Text

' /* Anything That Is Opened, Must Be Closed */
```

```vb
        Close #1


' /* *************************************** */
' /*       Close The Program            */
' /* *************************************** */

' /* For All The Forms That Are In Our Project */
Dim XFrm As Form
For Each XFrm In Forms

    ' /* Unload The Form */
    Unload XFrm

Next XFrm

' /* Code Should Never Reach Here, But Just In Case */
End


End Sub

Private Sub Command11_Click()

SpeakVoice strWavPath & rs("Code") & ".wav"
End Sub

Private Sub Command111_Click()
If rs.BOF Then rs.MoveFirst
    rs.MovePrevious
Label1.Caption = rs.AbsolutePosition + 1 & "/" & rs.RecordCount


'       "E:\bashar\jeem\w102.wav"
End Sub

Private Sub Command12_Click()

    rs.MoveNext
    Label1.Caption = rs.AbsolutePosition + 1 & "/" & rs.RecordCount

End Sub

Private Sub Command15_Click()
Timer1.Enabled = Not Timer1.Enabled
If Command15.Caption = "Auto" Then
Command15.Caption = "Manual"
ElseIf Command15.Caption = "Manual" Then
Command15.Caption = "Auto"
End If


End Sub
```

```vb
Private Sub Form_Load()
'-------- read from database
Set Voice = New SpVoice
Set db = OpenDatabase("C:\Documents and Settings\Dell\Desktop\SoundProject\sounds")
Set rs = db.OpenRecordset("tblSounds", dbOpenDynaset)
strWavPath = "D:\rashid\"

rs.MoveLast
rs.MoveFirst
Label1.Caption = rs.AbsolutePosition + 1 & "/" & rs.RecordCount
'---- end

   Set RC = New SpSharedRecoContext
   Set myGrammar = RC.CreateGrammar
      myGrammar.CmdLoadFromFile "C:\Program Files\Microsoft Speech SDK
5.1\Samples\Common\wordsxml2.xml", SLODynamic
   myGrammar.CmdSetRuleIdState 0, SGDSActive
End Sub




Private Sub Label2_Click()

End Sub

Private Sub RC_FalseRecognition(ByVal StreamNumber As Long, ByVal StreamPosition
As Variant, ByVal Result As SpeechLib.ISpeechRecoResult)
   Beep
   Text2.Text = Text2.Text & "no recognition"
   txtTemp = Result.PhraseInfo.GetText
End Sub

Private Sub RC_Hypothesis(ByVal StreamNumber As Long, ByVal StreamPosition As
Variant, ByVal Result As SpeechLib.ISpeechRecoResult)
Text6.Text = Text6.Text & Result.PhraseInfo.GetText
End Sub

Dim X As String
   Dim ii As Integer
   Dim PR As ISpeechPhraseRule
   Dim PRs As ISpeechPhraseRules

   ii = 0
   For Each E In Result.PhraseInfo.Elements
      X = "element" & Str(ii) & ": " & E.DisplayText
      List1.AddItem X
      ii = ii + 1
   Next

   'This is the rule that recognition was based on
   Set PR = Result.PhraseInfo.Rule
```

```
    List1.AddItem ""
    List1.AddItem "Id:            " & PR.Id
    List1.AddItem "Rule:          " & PR.Name
    List1.AddItem "NumberOfElements: " & PR.NumberOfElements
    List1.AddItem "FirstElement:    " & PR.FirstElement
    List1.AddItem "EngineConfidence: " & PR.EngineConfidence
    List1.AddItem "Confidence:      " & PR.Confidence
    List1.AddItem ""

If PR.Confidence = 1 Then
Text3.Text = Text3.Text & Result.PhraseInfo.GetText
End If
If PR.Confidence = 0 Then
Text4.Text = Text4.Text & Result.PhraseInfo.GetText
End If
If PR.Confidence = -1 Then
Text5.Text = Text5.Text & Result.PhraseInfo.GetText
End If

End Sub




Private Sub Timer1_Timer()
Command12_Click
Command11_Click

End Sub

Private Sub txtTemp_Change()
rs.Edit
rs("EnglishDesc2") = txtTemp
rs.Update

rs.Edit
If rs("EnglishDesc") = rs("EnglishDesc2") Then
   rs("Match") = "Yes"
Else
   rs("Match") = "No"
End If
rs.Update

End Sub
```

Analysis process

Two files are required for this process, and excel and an access file.
The recognised words get recorded in an Access file.

D	C	B	A	
	DHaaAa	ضاع	w1	1
	ATHin	أظن	w2	2
	Saagha	صاغ	w3	3
	edhaa	اذا	w4	4
	zaaar	زار	w5	5
	qaas	قاس	w6	6
	Aamal	أمل	w7	7
	jathaa	جثى	w8	8
	shaaaH	شاح	w9	9
	Taaaf	طاف	w10	10
	hayaAA	هيأ	w11	11
	kaAs	كأس	w12	12
	Aukht	أخت	w13	13
	baaada	باد	w14	14
	Aaw	او	w15	15
	Aakala	أكل	w16	16
	saAala	سأل	w17	17
	DHuUul	ضؤول	w18	18
	baisa	بئس	w19	19
	baraAa	برأ	w20	20
	suwuai	سوع	w21	21
	daaaan	داء	w22	22
	daaaun	داء	w23	23
	daaain	داء	w24	24
	THaby	ظبي	w25	25
	DHahAa	هبأ	w26	26

\ Sheet3 \ Sheet2 \ Sheet1 / ⊮ ◂ ▸ ⊩

NUM

Figure 14-Snapshot of the excel file

This excel file identifies the list of words in the database, so word 1 (W1) is basically the word DHaaAa and the sound file should be recognised as DHaaAa.

ID	Code	ArabicDesc	EnglishDesc	EnglishDesc2	Match
1	w1	ضاغ	DHaaAa	tyn	No
2	w2	انطن	ATHin	ttaabllin	No
3	w3	صاغ	Saagha	zzaaamaa	No
4	w4	اذا	edhaa	yin	No
5	w5	زان	zaaar	zzaaamaa	No
6	w6	قانس	qaas	shaams	No
7	w7	أمل	Aamal	kaarrnin	No
8	w8	جثى	jathaa	dhaaba	No
9	w9	شاح	shaaaH	shaah	No
10	w10	طانف	Taaaf	ghaat	No
11	w11	هيا	hayaAA	haayuUa	No
12	w12	كفس	kaAs	ghaat	No
13	w13	أخت	Aukht	ghill	No
14	w14	بلا	baaada	ghaadha	No
15	w15	أو	Aaw	ghill	No
16	w16	اكل	Aakala	maakkaanaa	No
17	w17	سق	saAala	thaawbanaa	No
18	w18	منول	DHuUul	daaaun	No
19	w19	بثس	baisa	ghaaythu	No
20	w20	برأ	baraAa	baadaa	No
21	w21	سوع	suwuai	saawyi	No
22	w22	داء	daaaan	daahrran	No
23	w23	داء	daaaun	daaaun	Yes
24	w24	داع	daaain	daaaan	No
25	w25	ظبي	THaby	maahddi	No
26	w26	ضبح	DHabAa	dhaab	No
27	w27	بزغ	bazagha	baaiisaa	No
28	w28	بصل	baSal	baasaattaa	No
29	w29	بهق	bahaq	sum	No

Figure 15- Snapshot of the Access file

The Access file records the recognised words and states if they got recognised or not to allow the user to calculate the recognition rates and analyse the list.

Appendix E

Diacritical Rules

Diacritical Rules

الحركة / Diacritic															الحرف Letter	الموقع Start, middle, End
ُّ	ِّ	َّ	ُ	ِّ	َّ	ْ	ُ	ِ	َ	ُ	ِ	َ	•	⊗		
-	-	-	-	-	-	-	-	-	-	-	-	-	-	✗	ء	البداية start
✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗	✗	✓	✓	✓	ء	الوسط middle
✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	ء	النهاية End
-	-	-	-	-	-	-	-	-	-	-	-	-	-	✗	ؤ	البداية start
✗	✗	✗	✓	✗	✓	✓	✗	✗	✓	✓	✗	✓	✓	✓	ؤ	الوسط middle
✗	✗	✗	✗	✗	✗	✓	✓	✓	✗	✓	✗	✓	✓	✓	ؤ	النهاية End
-	-	-	-	-	-	-	-	-	-	-	-	-	-	✗	ئ	البداية start

ــّة	ــِّ	ــًّ	ـّه	ـِّ	ـَّ	ه	ـٌ	ــ	ــ	ـُ	ــ	ــ	●	⊗	Letter	
✘	✘	✘	✘	✔	✘	✔	✘	✘	✔	✔	✔	✔	✔	✔	ئ	الوسط middle
✘	✘	✘	✘	✘	✘	✔	✔	✔	✘	✔	✔	✔	✔	✔	ئ	النهاية End
✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✔	✔	✔	✘	✔	ا	البداية start
✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✔	ا	الوسط middle
✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✔	ا	النهاية End
✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✔	✔	آ	البداية start
✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✔	✔	آ	الوسط middle
✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✔	✔	آ	النهاية End
✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✔	✘	✔	✔	✔	أ	البداية start

													●	⊗	Letter	
✗	✗	✗	✗	✗	✓	✓	✗	✗	✗	✗	✗	✓	✓	✓	أ	الوسط middle
✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	أ	النهاية End
✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗	✓	✓	إ	البداية start
✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗	✓	✓	إ	الوسط middle
✗	✗	✗	✗	✗	✗	✗	✗	✓	✗	✗	✓	✗	✓	✓	إ	النهاية End
✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	ب	البداية start
✗	✗	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓	ب	الوسط middle
✓	✓	✗	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	ب	النهاية End
✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	ت	البداية start

ﺔّ	ﻤّ	ﻼّ	ﺸ	ﻴّ	ﻼ	ه	ﻮّ	ﻌّ	ﻌ	ﻮ	ﻌ	ﻌ	●	⊗	Letter	
✗	✗	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓	ت	الوسط middle
✓	✓	✗	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	ت	النهاية End
-	-	-	-	-	-	-	-	-	-	-	-	-	-	✗	ة	البداية start
-	-	-	-	-	-	-	-	-	-	-	-	-	-	✗	ة	الوسط middle
✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	ة	النهاية End
✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	ث	البداية start
✗	✗	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓	ث	الوسط middle
✓	✓	✗	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	ث	النهاية End
✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	ج	البداية start

ـٌّ	ـِّ	ـَّ	ـُ	ـِ	ـَ	ـْ	ـٌ	ـٍ	ـً	ـُ	ـِ	ـَ	●	⊗	Letter	
✗	✗	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓	ج	الوسط middle
✓	✓	✗	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	ج	النهاية End
✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	ح	البداية start
✗	✗	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓	ح	الوسط middle
✓	✓	✗	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	ح	النهاية End
✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	خ	البداية start
✗	✗	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓	خ	الوسط middle
✓	✓	✗	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	خ	النهاية End
✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	د	البداية start

													●	⊗	Letter	
✗	✗	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓	د	الوسط middle
✓	✓	✗	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	د	النهاية End
✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	ذ	البداية start
✗	✗	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓	ذ	الوسط middle
✓	✓	✗	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	ذ	النهاية End
✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	ر	البداية start
✗	✗	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓	ر	الوسط middle
✓	✓	✗	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	ر	النهاية End
✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	ز	البداية start

ـّ	ـّ	ـّ	ـ	ـ	ـ	ه	ـ	ـ	ـ	ـ	ـ	ـ	●	⊗	Letter	
✗	✗	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓	ز	الوسط middle
✓	✓	✗	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	ز	النهاية End
✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	س	البداية start
✗	✗	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓	س	الوسط middle
✓	✓	✗	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	س	النهاية End
✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	ش	البداية start
✗	✗	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓	ش	الوسط middle
✓	✓	✗	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	ش	النهاية End
✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	ص	البداية start

ـّ	ـّ	ـّ	ـّ	ـّ	ـّ	ه	ـّ	ـّ	ـّ	ـّ	ـّ	ـّ	●	⊗	Letter	
✗	✗	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓	ص	الوسط middle
✓	✓	✗	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	ص	النهاية End
✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	ض	البداية start
✗	✗	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓	ض	الوسط middle
✓	✓	✗	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	ض	النهاية End
✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	ط	البداية start
✗	✗	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓	ط	الوسط middle
✓	✓	✗	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	ط	النهاية End
✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	ظ	البداية start

ـّٔ	ـّٕ	ـّ	ـٔ	ـٕ	ـ	ه	ـٌ	ـٍ	ـً	ء	ـ	ـ	●	⊗	Letter	
✗	✗	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓	ظ	الوسط middle
✓	✓	✗	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	ظ	النهاية End
✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	ع	البداية start
✗	✗	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓	ع	الوسط middle
✓	✓	✗	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	ع	النهاية End
✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	غ	البداية start
✗	✗	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓	غ	الوسط middle
✓	✓	✗	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	غ	النهاية End
✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	ف	البداية start

ـّۢ	ـۙ	ـّۤ	ـۀ	ـۚ	ـۜ	ـ۟	ـّٗ	ـۣ	ـۗ	ـٔ	ـۖ	ـَ	●	⊗	Letter	
✗	✗	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓	ف	الوسط middle
✓	✓	✗	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	ف	النهاية End
✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	ق	البداية start
✗	✗	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓	ق	الوسط middle
✓	✓	✗	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	ق	النهاية End
✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	ك	البداية start
✗	✗	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓	ك	الوسط middle
✓	✓	✗	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	ك	النهاية End
✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	ل	البداية start

						°							●	⊗	Letter	
✗	✗	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓	ل	الوسط middle
✓	✓	✗	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	ل	النهاية End
✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	م	البداية start
✗	✗	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓	م	الوسط middle
✓	✓	✗	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	م	النهاية End
✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	ن	البداية start
✗	✗	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓	ن	الوسط middle
✓	✓	✗	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	ن	النهاية End
✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	ه	البداية start

						ه							●	⊗	Letter	
✗	✗	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓	ه	الوسط middle
✓	✓	✗	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	ه	النهاية End
✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	و	البداية start
✗	✗	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✗	✓	و	الوسط middle
✓	✓	✗	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✗	✓	و	النهاية End
-	-	-	-	-	-	-	-	-	-	-	-	-	-	✗	ى	البداية start
-	-	-	-	-	-	-	-	-	-	-	-	-	-	✗	ى	الوسط middle
✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	ى	النهاية End
✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗	✓	✓	✓	ي	البداية start

													•	⊗	Letter	
✗	✗	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✗	✓	ي	الوسط middle
✓	✓	✗	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✗	✓	ي	النهاية End

Table 1- Diacritical rules (Alghamdi et al., 2006)

⊗ The letter can occur at this position of the word.

• The letter can occur without diacritics.

This table provides the rules for each letter in the three positions (start, middle and end), with a list of all diacritic.

Appendix F

The possibilities of the word (Nawal) after applying diacritical rules to the transliteration application.

The possibilities for the word Nawal after applying diacritical rules to the transliteration application.

Arabic	Translit	Arabic	Translit	Arabic	Translit	Arabic	Translit
نوال	nwaal	نَوَالٌ	nwaaalun	نُوَالُ	nuwuaalu	نُوَالٍّ	nuwanaallin
نَوال	nawaal	نَوَالٍ	nwaaalin	نُوَالِ	nuwuaali	نُوَالَّ	nuwanaalla
نَوَال	nawanaal	نَوَالَ	nwaaala	نُوَال	nuwiaal	نُوَالُّ	nuwanaallu
نَوَالٌ	nawanaalun	نَوَالُ	nwaaalu	نُوَالٌ	nuwiaalun	نُوَالٍّ	nuwanaalli
نَوَالٍ	nawanaalin	نَوَالِ	nwaaali	نُوَالٍ	nuwiaalin	نُوَالٌّ	nuwaaallun
نَوَالَ	nawanaala	نُوَال	nwuaal	نُوَالَ	nuwiaala	نُوَالٍّ	nuwaaallin
نَوَالُ	nawanaalu	نُوَالٌ	nwuaalun	نُوَالُ	nuwiaalu	نُوَالَّ	nuwaaalla
نَوَالِ	nawanaali	نُوَالٍ	nwuaalin	نُوَالِ	nuwiaali	نُوَالُّ	nuwaaallu
نَوَال	nawaaal	نُوَالَ	nwuaala	نُوَالٌ	nuwaalun	نُوَالٌّ	nuwaaallu
نَوَالٌ	nawaaalun	نُوَالُ	nwuaalu	نُوَالٍ	nuwaalin	نُوَالٌّ	nuwuaallun
نَوَالٍ	nawaaalin	نُوَالِ	nwuaali	نُوَالَ	nuwaala	نُوَالٍّ	nuwuaallin
نَوَالَ	nawaaala	نوال	nwiaal	نُوَالُ	nuwaalu	نُوَالَّ	nuwuaalla
نَوَالُ	nawaaalu	نوالٌ	nwiaalun	نُوَالِ	nuwaali	نُوَالُّ	nuwuaallu
نَوَالِ	nawaaali	نوالٍ	nwiaalin	نِوال	niwaal	نُوَالٍّ	nuwuaalli
نَوَال	nawuaal	نوالَ	nwiaala	نِوَال	niwanaal	نُوَالٌّ	nuwiaallun
نَوَالٌ	nawuaalun	نوالُ	nwiaalu	نِوَالٌ	niwanaalun	نُوَالٍّ	nuwiaallin
نَوَالٍ	nawuaalin	نوالِ	nwiaali	نِوَالٍ	niwanaalin	نُوَالَّ	nuwiaalla
نَوَالَ	nawuaala	نوالٌ	nwaalun	نِوَالَ	niwanaala	نُوَالُّ	nuwiaallu
نَوَالُ	nawuaalu	نوالٍ	nwaalin	نِوَالُ	niwanaalu	نُوَالٍّ	nuwiaalli
نَوَالِ	nawuaali	نوالَ	nwaala	نِوَالِ	niwanaali	نُوَالٌّ	nuwaallun
نَوَال	nawiaal	نوالُ	nwaalu	نِوَال	niwaaal	نُوَالٍّ	nuwaallin
نَوَالٌ	nawiaalun	نوالِ	nwaali	نِوَالٌ	niwaaalun	نُوَالَّ	nuwaalla
نَوَالٍ	nawiaalin	نَوَالٌّ	nawanaallun	نِوَالٍ	niwaaalin	نُوَالُّ	nuwaallu
نَوَالَ	nawiaala	نَوَالٍّ	nawanaallin	نِوَالَ	niwaaala	نُوَالٍّ	nuwaalli
نَوَالُ	nawiaalu	نَوَالَّ	nawanaalla	نِوَالُ	niwaaalu	نِوَالٌّ	niwanaallun
نَوَالِ	nawiaali	نَوَالٌّ	nawanaallu	نِوَالِ	niwaaali	نِوَالٌّ	niwanaallin

نَوَالٌ	nawaalun	نَوَالِّ	nawanaalli	نِوُال	niwuaal	نِوَالَ	niwanaalla
نَوَالٍ	nawaalin	نَوَالٌ	nawaaallun	نِوُالٌ	niwuaalun	نِوَالُ	niwanaallu
نَوَالَ	nawaala	نَوَالٍ	nawaaallin	نِوُالٍ	niwuaalin	نِوَالِ	niwanaalli
نَوَالُ	nawaalu	نَوَالَ	nawaaalla	نِوُالَ	niwuaala	نِوَالٌ	niwaaallun
نَوَالِ	nawaali	نَوَالُ	nawaaallu	نِوُالُ	niwuaalu	نِوَالٍ	niwaaallin
نُوَال	nuwaal	نَوَالِ	nawaaalli	نِوُالِ	niwuaali	نِوَالَ	niwaaalla
نُوَال	nuwanaal	نَوُالٌ	nawuaallun	نِوِال	niwiaal	نِوَالُ	niwaaallu
نُوَالٌ	nuwanaalun	نَوُالٍ	nawuaallin	نِوِالٌ	niwiaalun	نِوَالِ	niwaaalli
نُوَالٍ	nuwanaalin	نَوُالَ	nawuaalla	نِوِالٍ	niwiaalin	نِوُالٌ	niwuaallun
نُوَالَ	nuwanaala	نَوُالُ	nawuaallu	نِوِالَ	niwiaala	نِوُالٍ	niwuaallin
نُوَالُ	nuwanaalu	نَوُالِ	nawuaalli	نِوِالُ	niwiaalu	نِوُالَ	niwuaalla
نُوَالِ	nuwanaali	نَوِالٌ	nawiaallun	نِوِالِ	niwiaali	نِوُالُ	niwuaallu
نُوَال	nuwaaal	نَوِالٍ	nawiaallin	نِوَالٌ	niwaalun	نِوُالِ	niwuaalli
نُوَالٌ	nuwaaalun	نَوِالَ	nawiaalla	نِوَالِ	niwaalin	نِوِالٌ	niwiaallun
نُوَالٍ	nuwaaalin	نَوِالُ	nawiaallu	نِوَالَ	niwaala	نِوِالٍ	niwiaallin
نُوَالَ	nuwaaala	نَوِالِ	nawiaalli	نِوَالُ	niwaalu	نِوِالَ	niwiaalla
نُوَالُ	nuwaaalu	نَوَالٌ	nawaallun	نِوَالِ	niwaali	نِوِالُ	niwiaallu
نُوَالِ	nuwaaali	نَوَالٍ	nawaallin	نَوَال	nwanaal	نِوِالِ	niwiaalli
نُوُال	nuwuaal	نَوَالَ	nawaalla	نَوَالٌ	nwanaalun	نِوَالٌ	niwaallun
نُوُالٌ	nuwuaalun	نَوَالُ	nawaallu	نَوَالِ	nwanaalin	نِوَالٍ	niwaallin
نُوُالٍ	nuwuaalin	نَوَالِ	nawaalli	نَوَالَ	nwanaala	نِوَالَ	niwaalla
نُوُالَ	nuwuaala	نُوَالٌ	nuwanaallun	نَوَالُ	nwanaalu	نِوَالُ	niwaallu
نَوَالَ	nwanaalla	نَوَالٍ	nwanaallin	نَوَالِ	nwanaali	نِوَالِ	niwaalli
				نَوَال	nwaaal	نَوَالٌ	nwanaallun

Table 1-Nawal diacritised possibilities

Appendix G

•••••••••••••••••••••••••••••••••••

The 499 words analysis

ﹻ	ﹾ	ﹼ	ﹷ	ﹾ	ﹷ	Diacritics Position	Arabic letters
			اذا	أخت	أَظِنُ - أمل - أو - أَكَلَ - أجد - أَذُن - أَشُدُ - أرضٍ - أَمِل - أودُ	S	أ
			بئس	ضوّل	سأل	M	
داءٍ	داءٌ	داءً	سوءٍ	برأ	هيأ	E	
			بشر - برَكُ	بُرج	باد - بئس - برأ - بزغ - بصل - بهق - بَشِم - بَطَلَ - بثه - بَثُ - بَخَسَ - بَخِلَ - بَخٌ - بذخ - بذخًا - بَسَط - بَرقًا - بَرقٌ - بَرق	S	ب
			جُبلَ - طُبعَ	سُبل	خبط - ثبت - ربط - ثبط - حَبَسَ - سَبَقَ - عَبق	M	
ثوبٍ	ثوبٌ	ثوبًا	قلبٍ - ذئب	فرَبٌ - نابُ	حلب - كبّ - صب - وجب - وَثَبَ - نَدَبَ - كذِب - رُبًّ - سَكَبَ - ضَرَبَ - رَكِبَ - رَهِبَ	E	
			تين	توت	تحت - تظل - تقي - تمر	S	ت
			سُترَ - عَتيًا - عَتيٌ - عَتي	عتو - عَثُمَ	قتَلَ	M	
صمتٍ	صمتٌ	صمتًا	يختِ	يُمتُ	ثبت - نحتَ	E	

≈	ۏ	≈	ـَ	و	ـَ	Diacritics Position	Arabic letters
			ثِي	ثلث ـ ثلة ـ ثلثٌ ـ ثلثٍ ـ ثلثًا	ثبت ـ ثوبًا ـ ثوبٌ ـ ثوبٍ ـ ثنت ـ ثقف ـ ثوي ـ ثخن ـ ثبط ـ ثَجَ ـ ثَقُلَ ـ ثه	S	ث
			عُثِرَ	جثو	جثى ـ بثه ـ وَثَبَ	M	
ثلثٍ	ثلثٌ	ثلثًا	رَثِ	بَثُ	مَكَثَ ـ عاثَ	E	
			جدُّ	جُبلَ ـ جُثو ـ جذم ـ جُهد ـ جُحر ـ جُزُر	جثى ـ جفت ـ جرك ـ جص ـ جحظ ـ جدي ـ جز ـ جوق ـ جمل ـ جوزُ ـ جرك ـ جَلِي ـ جَرواً ـ جَروٌ ـ جَروِ	S	ج
			وجب ـ أجد ـ وُجدَ	حُجُبْ ـ سجق	لجج ـ خجل ـ طجن ـ عجز ـ سجع ـ وَجَدَ ـ شَجر ـ غجر	M	
عوج	عوجٌ	عوجاً	وَهَج	سَرجُ	ثجَ ـ ضَجَ ـ شج ـ هج ـ دَرَجَ	E	
			حِبر	حُجُبْ ـ حُسن ـ حُدِّدَ ـ حُنو	حلب ـ حدث ـ حضر ـ حذو ـ حطم ـ حسن ـ حك ـ حله ـ حي ـ حَمَلَ ـ حَرَمَ ـ حَبَسَ ـ حَرَسُ ـ حَفظ ـ حَظ ـ حَرَّكَ	S	ح
			يَحِلُ	صحف ـ سُحُب ـ جُحر	نحتَ ـ جحظ ـ طحَنَ ـ شَحَذُ ـ زُحل ـ ضُحى	M	
قزح	قزحٌ	قزحً	فُرَح	مَرَحُ	مَسَحَ	E	

	ۇو	=		و	ـَ	Diacritics Position	Arabic letters
=			ـَ				
			خِدر	خُمس	خبط - خجل - خدك - خث - خشع - خص - خزق - خسف - خمن - خوي - خس - خَلفُ	S	خ
			بَخِلَ	رَخُصَ	ضخه - ذخر - بَخَسَ	M	
بذخ	بذخٌ	بذخًا	مُخ	سَلخُ	صَرَخَ	E	
			دِيك	دُب	دَرَجَ - دس - دغل - دعه - دف - دمع - دهرِّ - دهر - دهراً - دَفِرَ - دقك - دَغْلُ	S	د
			جدي - حُدِّدَ	خدك - مدُن	شدت - حدث - قدم - نَدَبَ - غَدَرَ	M	
وعدٍ	وعدٌ	وعداً	مهدِ	جدُّ - أجد - عهدُ - أشُدُ - يَصُدُ - ضِدِ - أودُ	باد - وَجَد - رصد - حُدِّدَ - سَدَّ - رُشِدَ - قدَّ - فقِدَ - هُود - وُجِدَ	E	
			ذَنب - ذِهن	ذُل	ذنب - ذرت - ذخر - ذكي - ذم - ذل - ذود - ذنب - ذاب - ذُوُ	S	ذ
			كذِب	حذو - أُذن	بذخٌ - بذخ - بذخًا - قذع - عذَرَ	M	
فذٍ	فذٌ	فذًا	موْذِ	منذُ	شذ - شَحَذُ	E	

ّ	ّ ُ	ّ	َ	ُ	َ	Diacritics Position	Arabic letters
			رمشُ - ريش	رُبَّ - رُزِق - رُسُل - رُشِدَ	ربط - رَثِ - رَخُصَ - رصد - ركل - رغو - رد - رَشَّ - رَصَّ - رَضِيَ - رَطِب - رَغْدَ - رَفعَ - ركض - رَكِبَ - رَكَعَ - رَهِبَ - رَهُفَ	S	ر
			سري - وَرَعاً - وَرِعّ - وَرَع	سُرر	برأ - فرَبُّ - ذرت - شرث - جرك - دَرَجَ - فَرَح - مَرَحُ - صَرَخَ - حَرَمَ - زَرَعَ - حَرَسُ - فَرَس - ضَرَبَ - مَرَضُ - عَرَضَ - غرق - مَرَغَ - شَرَفَ - طرفِ - جرك - حَرَّكَ - برَكُ - غره	M	
دهرِ	دهرٌ	دهراً	جُحر	فِكرُ	سُترَ - عُثِرَ - حضر - ذخر - عذَرَ - ظهر - ضر - فطرَ - عَصَرَ - نُصِرَ - نظَرَ - غذَرَ - صَغُرَ - دَفِرَ - قهَرَ - سَيَّرَ	E	
			زر	زُحل	زار - زند - زف - زعم - زَرَعَ	S	ز
			رُزق	جُزر	بزغ - هزت - قزحَ - قزحاً - قزح - قزحّ - خزق - عزَفَ - زكي	M	
فِلزِ	فِلزٌ	فِلزاً	كنز	جوزُ	عجز - جز - فازَ	E	
			سر - سري - سِحر - سِلكاً - سِلكٌ - سِلكِ	سوءٍ - سُبل - سُترَ - سُحُب - سُرر - سُم - سجق	سأل - سطت - سجع - سَرجُ - سَلَخُ - سَدَّ - سهو - سكَبَ - سكَبَ - سعف - سَعِر - سَقط - سَبَقَ - سَمَكِ - سَني - سَهوَ - سَيَّرَ - سوي	S	س
			نَسِيَ - وَسِعَ	رُسُل	نسب - حسن - مَسَحَ - خسف - غسل - عَسَل - وَسَطِ - بَسَطَ - غَسَقُ	M	
عُرسِ	عُرسٌ	عُرساً	فُرَس	حَرَسُ	بئس - بَخَسَ - دس - حَبَسَ - نَعَسَ	E	

ــٍ	ــٌ	ــاً	ــَ	ــُ	ــَ	Diacritics Position	Arabic letters
			شِبل	شغل شُكر	شاح - شدت - شرث - شج - شذ - شَحَذْ - شمس - شذو - شص - شظف - شط - شك - شهي - شَجر - شَرَّفَ	S	ث
			بَشِم - نشز - رُشِدَ	أشُدُ	خشع	M	
كبشٍ	كبشٌ	كبشاً	ريش	رمشُ	رَشَّ - غش	E	
			صِهر	صحف - صم -صُوصُ	صاغ - صب - صكت - صمتٌ - صمتَ - صمتٍ - صَرَخَ - صنع - صه - صَيد - صَغُرَ - صَدغْ - صَمغ - صَوغاً - صَوغٌ - صَوغ - صَنَم	S	ص
			وصي - نُصِرَ	يَصُدُ - غُصن	بصل - رصد عَصَرَ	M	
فصٍ	فصٌ	فصاً	قرص	صُوصُ	خص - رَخُصَ - قص - رَصَّ	E	
			ضد - ضرس - ضيق	ضَوْل - ضُحى	ضاعَ - ضبع - ضمت - ضَج - ضخه - ضر - ضغث - ضن - ضل - ضفر - ضَرَبَ	S	ض
			رَضِيَ	عَضُد	حضر - وضع - فضَّلَ	M	
قرضٍ	قرضٌ	قرضاً	أرض	مَرَضُ - غض	عَرَضَ - ركض	E	

⸗	ٌ	⸗	ـَ	و	ـَ	Diacritics Position	Arabic letters
			طِب	طبعَ	طاف - طجن - طحَنَ - طود - طير - طق - طمَعَ - طغِيَ - طرفٍ - طبلاً - طبلٌ - طبلِ - طيُّ	S	ط
			رَطب	عُطف	بطلَ - سطت - حطم - فطرَ - هطل - وطن - عطش	M	
ثُقطِ	ثُقطُ	ثُقطاً	وَسطِ	قرطُ	ربط - ثبط - بَسَط - غط - سَقط	E	
			ظِفر	ظل	ظبي - ظهر - ظرف	S	ظ
			أظنُ - تظل - عَظِمَ	نظفَ	شظف - وظف - نظرَ	M	
وَعظِ	وَعظٌ	وَعظاً	قيظِ	كظ - حَظُ - غيظ	جحظ - حَفظ	E	
			عوجاً - عوجّ - عوج - عجل	عتو - عُثْرَ - عُطف - عُرساً - عُرسٌ - عُرسٍ - عُمر - عُلو	عَثُمَ - عث - عجز - عهدُ - عدَرَ - عزَفَ - عَسَل - عَصَرَ - عَضُد - عَرَضَ - عَظِمَ - عذْل - عطش - عكس - عقر - عَين - عَفوّ - عَلَفاً - عَلَفٌ - عَلَفِ - عَبَق - عَمَلِ - عَلَماً - عَلَمّ - عَلَمِ - عَيِيَ - عَتياً - عَتيّ - عَتي	S	ع
			سَعِر	نَعُمَ	زعم - سعف - نَعَسَ - فَعَلَ	M	
وَرَع	وَرعٌ	وَرعاً	قاع	صاعُ	ضاعَ - خشع - زَرَعَ - وضع - طمَعَ - طبعَ - وَسِعَ - رَفَعَ - رَكَعَ	E	

≈	ورو	≈	ـَ	ورو	ـَ	Diacritics Position	Arabic letters
			غِل	غُصن	غت - غث - غسل - غجر - غذى - غش - غض - غفي - غرق - غيظ - غط - غَدَرَ - غَسَقُ - غَلى - غَنَمُ - غره	S	غ
			طَغِيَ	رغو - صَغُرَ	ضغث - رَغَدَ	M	
صَوغُ	صَوغٌ	صَوغاً	صَمغ	صَدغ	صاغ - بزغ - مَرَغْ	E	
			فكرُ - فلزاً - فلزٌ - فلزِ - فِجل - فَيهِ	فُرن - فُقِدَ	فرَبٌ - فرَح - فذ - فذّا - فذّ - فذٍ - فطرَ - فرَس - فُصٌّ - فُصاً - فص - فضَّلَ - فك - فن - فعَلَ - فهَمَ - فنَر	S	ف
			غفي - وفي - دَفِرَ	عَفْوٌ	جفت - ضفر - حَفظ - رَفعَ	M	
عَلَفٍ	عَلَفٌ	عَلَفاً	طرفِ	خَلفُ	خسف - زف - عزَفَ - وظف - نَظفَ - شَرَفَ - رَهُفَ	E	
			قِرطُ - قِدِر	قرص - قَدَّ	قاس - قلبِ - قتَلَ - قزح - قزحً - قزح - قزحّ - قدم - قذع - قرن - قش - قص - قرضاً - قرضٌ - قرضٍ - قيظِ - قلَم - قرنأ - قرنّ - قرنِ - قهَرَ	S	ق
			تقي - فُقِدَ	ثَقُلَ	ثقف - نُقطاً - نُقطٌ - نُقطٍ - عقر - دقك - سَقط	M	
بَرق	بَرقٌ	بَرقاً	عَبَق	غَسَقُ	خزق - رُزق - غرق - سَبَقَ	E	

ـٍ	ـٌ	ـً	ـَ	ـُ	ـَ	Diacritics Position	Arabic letters
			كيس	كُوع	كأس ـ كبّ ـ كذِب ـ كنز ـ كبشاً ـ كبشٌ ـ كبشٍ ـ كظ ـ كوى ـ كهن ـ كَلب	S	ك
			ذكي ـ زكي ـ رَكِبَ	مَكُنَ	أكَلَ ـ صكت ـ مَكَثَ ـ جرك ـ ركل ـ سَكَبَ ـ عكس ـ ركض ـ رَكَعَ ـ مَكَّنَ	M	
سِلكٍ	سِلكٌ	سِلكاً	سَمَكِ	برَكُ	حك ـ خدك ـ فك ـ دقك ـ جرك ـ حَرَّكَ	E	
			لين	لجج ـ لُمَّة	قتَلَ ـ بَخِلَ ـ غسل ـ لیث ـ لَهو	S	ل
			فِلزاً ـ فِلزٌ ـ فِلزٍ ـ جَلي	ثلثٌ ـ ثلثٍ ـ ثلثاً ـ قلو ـ عُلو	حلب ـ ثُلة ـ حله ـ ملف ـ عَلفاً ـ عَلفٌ ـ عَلفٍ ـ قلم ـ عَلماً ـ عَلمٌ ـ عَلمٍ ـ ثله	M	
طبلٍ	طبلٌ	طبلاً	عَمَل	يَحلُ ـ دَغلُ	أكَلَ ـ سأل ـ بَطلَ ـ جُبلَ ـ حَمَلَ ـ ذل ـ ركل ـ ضل ـ فضَّلَ ـ هطل ـ فعَلَ ـ ثَقلَ ـ عَمَلَ	E	
			مِن ـ مِنهُ	مُخ ـ مدُن ـ منذُ ـ موذِ	مَكثَ ـ مَسَحَ ـ مَرَحُ ـ مهدِ ـ مَرَضُ ـ مَرَّعَ ـ ملف ـ مَكُن ـ مَوز ـ مَكَّنَ ـ مَحوُ	S	م
			يُمتُ ـ أمِل	ثمُو	أمل ـ ضمت ـ جمل ـ حَمَلَ ـ من ـ طمَعَ ـ سَمَكِ ـ لُمَّة ـ عَمَل	M	
علمٍ	علمٌ	علماً	صَنَم	غنَمُ	بَشِمٍ ـ عَثَمَ ـ حطم ـ ذم ـ حَرَمَ ـ زعم ـ عَظِمَ ـ نَعِمٍ ـ فهَم	E	

ِّ	ُّ	ـً ّ	ِ	و ُ	َ	Diacritics Position	Arabic letters
			نِبر	نُصِرَ ـ نُقطَاً ـ نُقطُ ـ نُقطٍ ـ نُمُو ـ نُور	نسب ـ نحتَ ـ نَدَبَ ـ نَسِيَ ـ نشز ـ نظرَ ـ نظفَ ـ نَعَسَ ـ نَعُمَ ـ نَهر ـ نَحنُ ـ نَوَّهَ	S	ن
			سَنِي	حُنُو	ثنت ـ صنع ـ غَنَمُ ـ صَنَم ـ فَنَر	M	
قرنِ	قرنٌ	قرناً	ذِهن	أَظنُ ـ نَحنُ	طَحَنَ ـ خمن ـ أَدْن ـ ضن ـ مَكَّنَ ـ ون ـ مَكَّنَ	E	
			هِر	هُود	هيأ ـ هزت ـ هج ـ هطل ـ هف ـ هم ـ هَوَس ـ هَيُوَ	S	هـ
			شهي ـ رَهِبَ	رَهفَ ـ سَهوَ	بهق ـ وَهَج ـ ظهر ـ كهن ـ فُهَمَ ـ نَهر ـ قَهَرَ	M	
جاهِ	جاهٌ	جاهاً	فيهِ	بثه ـ حله ـ ضخه ـ دعه ـ غره ـ ثله ـ مِنهُ	نَوَّهَ	E	
			وتر	وُجِدَ	وجب ـ وَثَبَ ـ وَجَد ـ وَهَج ـ وعداً ـ وعدٌ ـ وعدٍ ـ وَشَم ـ وصي ـ وضع ـ وَطن ـ وَسَطِ ـ وظف ـ وَعظاً ـ وَعظٌ ـ وَعظِ ـ وَسِعَ ـ وَرعاً ـ وَرعٌ ـ وَرَعَ ـ وفي ـ ون	S	و
			ثوي ـ خوي ـ هُود ـ أودُ ـ سوي	ذُوُو	عوجاً ـ عوجٌ ـ عوج ـ هَوَس ـ نَوَّهَ	M	
جَروِ	جَروٌ	جَرواً	لهو	رغو ـ سهو ـ عَفُوُ ـ قلو ـ حُنُو ـ ذُوُو ـ مَحوُ	سَهوَ	E	

	و	=		و	◌َ	Diacritics Position	Arabic letters
=			◌َ				
			ين	يُمتُ - يُسر	يختِ - يَحِلُ - يَصدُ -يوم - يَد	S	
			غيظ - عَيِيَ	هَيُوَ	هيأ - سَيَّرَ	M	ي
عَتِي	عَتيٌّ	عَتياً	جَلِي - سوي	طيُ	نَسِيَ - رَضِيَ - غَفِي - طغِيَ - سَني - عَيِيَ	E	

Table 1 The 499 words transliteration analysis

Appendix H

The 499 chosen words

The 499 chosen words

w1	ضاع	DHaaAa		**w180**	دِيك	ddiyik		**w359**	وَرَع	waraAin
w2	أظنُ	AThin		**w181**	نَدَبَ	nadaba		**w360**	غجر	ghajar
w3	صاغ	Saagha		**w182**	حُدِّدَ	Huuddidda		**w361**	غذى	ghadhdhaaa
w4	اذا	edhaa		**w183**	مدُن	mudun		**w362**	غش	ghashsha
w5	زار	zaaar		**w184**	سَدَّ	sadda		**w363**	غض	ghaDHDHu
w6	قاس	qaas		**w185**	عهدُ	Aahdu		**w364**	غفي	ghafya
w7	أمل	Aamal		**w186**	مهدِ	mahhdi		**w365**	غرق	gharqa
w8	جثى	jathaa		**w187**	وعداً	waAdaaan		**w366**	غيظ	ghayiTHu
w9	شاح	shaaaH		**w188**	وعدٌ	waAdun		**w367**	غط	ghaTTaa
w10	طاف	Taaaf		**w189**	وعدٍ	waAdin		**w368**	غَدَرَ	ghadara
w11	هيأ	hayaAA		**w190**	ذكي	dhaakiy		**w369**	غصن	ghuSun
w12	كأس	kaAs		**w191**	ذم	dhama		**w370**	غِل	ghil
w13	أخت	Aukht		**w192**	ذل	dhaalla		**w371**	صَغُرَ	Saghura
w14	باد	baaada		**w193**	فذ	fadh		**w372**	رَغَدَ	raghad
w15	أو	Aaw		**w194**	قذع	qadhaA		**w373**	طغِيَ	Taghya
w16	أكلَ	Aakala		**w195**	شذ	shadhdha		**w374**	مَرَّغَ	marragha
w17	سأل	saAala		**w196**	ذود	dhaawuud		**w375**	صَدغ	Sadaghu
w18	ضؤل	DHuUul		**w197**	ذنب	dhaiib		**w376**	صَمغ	Samghi
w19	بنس	baisa		**w198**	ذاب	dhaaab		**w377**	صَوغاً	Sawghaaan
w20	برأ	baraAo		**w199**	ذُل	dhul		**w378**	صَوغ	Sawghun
w21	سوء	suwuai		**w200**	كذِب	kadhiba		**w379**	صَوغ	Sawghin
w22	داءً	daaaan		**w201**	عذَر	Aadhara		**w380**	هف	haf
w23	داءٌ	daaaun		**w202**	أذُن	Aadhuna		**w381**	وفي	wafy
w24	داءٍ	daaain		**w203**	شَحَذْ	shaHadha		**w382**	ملف	malaf
w25	ظبي	THaby		**w204**	منذُ	mundhu		**w383**	فك	fakka
w26	ضبع	DHabAa		**w205**	مؤذِ	muuUdhi		**w384**	فن	fan
w27	بزغ	bazagha		**w206**	فذًا	fadhdhaaan		**w385**	فجل	fijl
w28	بصل	baSal		**w207**	فذٌ	fadhdhun		**w386**	فرن	furn
w29	بهق	bahaq		**w208**	فذٍ	fadhdhin		**w387**	فَعَلَ	faAala

w30	خبط	khabaT		w209	ظهر	THahara		w388	رَفَعَ	rafaAa
w31	كبّ	kabba		w210	قرن	qaarn		w389	دَفِرَ	dafira
w32	ذنب	dhanb		w211	ركل	rakala		w390	عَفوُ	Aafwu
w33	بَشِم	bashima		w212	ضر	DHaarra		w391	شَرَفَ	sharfa
w34	صب	Sabba		w213	رغو	raghwu		w392	طرفِ	Tarafi
w35	فرَبَّ	faraabbu		w214	طير	Taayr		w393	خَلفُ	khalfu
w36	نسب	nasab		w215	سر	sir		w394	عَلفًا	Aalafaaan
w37	وجب	wajiba		w216	رد	rad		w395	عَلفٌ	Aalafun
w38	ثبت	thabataa		w217	رُبَّ	rubba		w396	عَلفٍ	Aalafin
w39	بَطَلَ	baTala		w218	سُرر	suurur		w397	سجق	sujuq
w40	بشر	bishr		w219	حَرَمَ	Haarama		w398	قلو	qluwu
w41	بُرج	burj		w220	سري	siry		w399	دقك	daqqaka
w42	جُبِلَ	jubiila		w221	فطرَ	faTara		w400	قلَم	qalam
w43	ربط	rabaTaa		w222	جُحر	juHurri		w401	قِدر	qidr
w44	سُبل	subul		w223	فِكرُ	fikru		w402	قَدَّ	qudda
w45	حلب	Halaba		w224	دهرٌ	dahiruun		w403	سَقط	saqaTa
w46	قلبِي	qalbii		w225	دهرٍ	dahirin		w404	فُقدَ	fuuqida
w47	نابُ	naabu		w226	دهرًا	dahiraaan		w405	ثَقُلَ	thaqula
w48	ثوبًا	thawbaaan		w227	زف	zaffa		w406	سَبَقَ	sabaqa
w49	ثوبٌ	thawbun		w228	زعم	zaAama		w407	عَبَق	Aabaqe
w50	ثوبٍ	thawbin		w229	زكي	zaky		w408	غَسَقُ	ghasaquu
w51	تحت	taaHt		w230	زُحل	zuHal		w409	بَرقًا	barqaaan
w52	ضمت	DHaammat		w231	زَرَعَ	zaraAa		w410	بَرقٌ	barquun
w53	تظل	taTHil		w232	زر	zir		w411	بَرقٍ	barqin
w54	سطت	saTaat		w233	رُزق	ruziqa		w412	ركض	rakaDHa
w55	صكت	Saakkat		w234	عزَفَ	Aazafa		w413	جرك	jarraka
w56	ذرت	dhaarat		w235	جُزُر	juzur		w414	كوى	kaawaaa
w57	هزت	hazaat		w236	فازَ	faaaza		w415	كهِن	kahan
w58	شدت	shaaddat		w237	جوزُ	jawzuu		w416	كَلب	kalb
w59	ثنت	thanat		w238	كنز	kanzi		w417	كِيس	kiyis
w60	جفت	jaffat		w239	فِلزًا	filizaaan		w418	كُوع	kwuA

w61	عتو	Autuw		**w240**	فِلِزٌّ	fillizzun		**w419**	رَكِبَ	rakiba
w62	غت	ghaat		**w241**	فِلِز	fillizzin		**w420**	رَكَعَ	rakaAa
w63	تقي	taqiy		**w242**	شمس	shams		**w421**	مَكُنَ	makuuna
w64	تمر	taamr		**w243**	غسَل	ghasala		**w422**	حَرَّكَ	Haarraka
w65	تين	tiyn		**w244**	سهو	saahwu		**w423**	برَكُ	biraaku
w66	توت	tuwut		**w245**	كيس	kaiys		**w424**	سَمَكِ	samaki
w67	قَتَلَ	qattala		**w246**	ضرس	DHirs		**w425**	سِلكاً	silkaaan
w68	سُتِرَ	sutiira		**w247**	سُم	suum		**w426**	سِلكٌّ	silkun
w69	عَتُّمَ	Aatuuma		**w248**	سَكَبَ	sakaba		**w427**	سِلكٍ	silkin
w70	يُمتُ	yumitu		**w249**	سِحر	siHur		**w428**	لَيث	layth
w71	يخت	yakhti		**w250**	رُسُل	rusul		**w429**	لِين	liyn
w72	نحتَ	naHata		**w251**	عَسَل	Aasal		**w430**	لُمَّة	luumaat
w73	صمتٌ	Samtun		**w252**	نَسِيَ	naaisiya		**w431**	عُلو	Auuluwu
w74	صمتاً	Samtaaan		**w253**	حَبَسَ	Habasa		**w432**	غَلى	ghalaa
w75	صمتٍ	Samtin		**w254**	حَرَسَ	Haarasa		**w433**	جَلِي	jaaliyi
w76	ثلث	thulth		**w255**	فَرَس	faarasi		**w434**	دَغْلُ	daghlu
w77	ثقف	thaqaf		**w256**	عُرساً	Aursaaan		**w435**	عَمَل	Aamali
w78	مَكَثَ	makathaa		**w257**	عُرسٌّ	Aursuun		**w436**	طبلاً	Tablaaan
w79	غث	ghath		**w258**	عُرس	Aursin		**w437**	طبلٌ	Tablun
w80	حدث	Haadath		**w259**	شذو	shadhw		**w438**	طبلٍ	Tablin
w81	شرث	sharrath		**w260**	شص	shaS		**w439**	هم	haam
w82	عث	Aath		**w261**	شظف	shaTHaf		**w440**	يوم	yawm
w83	ثوي	thawiy		**w262**	شط	shat		**w441**	مَوز	mawz
w84	ثخن	thakhn		**w263**	شغل	shughl		**w442**	مِن	min
w85	بثه	baththahu		**w264**	قش	qash		**w443**	أمِل	Aamil
w86	ثبط	thabaTaa		**w265**	شك	shak		**w444**	عَمَلَ	Aamala
w87	ثَجَ	ththajja		**w266**	نشز	nashiz		**w445**	نُمُو	numuw
w88	ثِنْي	thiny		**w267**	شهي	shahy		**w446**	فَهَمَ	fahama
w89	ثلة	thuullah		**w268**	شَجر	shajar		**w447**	غَنَمُ	ghanamu
w90	وَثَبَ	wathabaa		**w269**	شِبل	shibl		**w448**	صَنَم	Sanami
w91	عَثِرَ	Authiira		**w270**	شُكر	shukr		**w449**	عَلَماً	Aaalamaaan

w92	جُثُو	juthuw		**w271**	وَشَم	washm		**w450**	عَلَمٌ	Aaalamun
w93	عاثَ	Aaatha		**w272**	رُشِدَ	rushida		**w451**	عَلَم	Aaalamin
w94	رَثِ	raathy		**w273**	أَشُدُ	Ashudu		**w452**	ون	wanna
w95	بَثُ	bathu		**w274**	رَشَّ	rasha		**w453**	نَهر	nahar
w96	ثلثٌ	thuulthun		**w275**	رمشُ	rimshu		**w454**	نِبر	nibr
w97	ثلثٍ	thulthin		**w276**	ريش	riyshi		**w455**	نُور	nuuwr
w98	ثلثًا	thulthaan		**w277**	كبشاً	kabshaaan		**w456**	فَنر	fanar
w99	لجج	lujaj		**w278**	كبشٌ	kabshun		**w457**	سَنِي	saanya
w100	جرك	jarrraka		**w279**	كبش	kabshin		**w458**	حُنُو	Huunuuw
w101	ضَجَ	DHaajjaaa		**w280**	قص	qaSSa		**w459**	مَكَّنَ	maakkana
w102	جص	jaaS		**w281**	صم	Sum		**w460**	نحنُ	naHnu
w103	خجل	khaajaal		**w282**	صنع	SanaAa		**w461**	ذِهن	dhihni
w104	جحظ	jaHaTHaa		**w283**	صه	Sah		**w462**	قرنًا	qarnaaan
w105	طجن	Taajjan		**w284**	وصي	waSy		**w463**	قرنٌ	qarnun
w106	شج	shajja		**w285**	صُوصُ	SuwuSo		**w464**	قرنِ	qarnin
w107	عجز	Ajjaaza		**w286**	صَيد	Sayd		**w465**	غره	ghaarraahu
w108	سجع	saajA		**w287**	صِهر	Sihr		**w466**	ثله	thaallaahu
w109	جذم	judhm		**w288**	عَصَرَ	AaSara		**w467**	هِر	hir
w110	جدي	jaddy		**w289**	نُصِرَ	nuSira		**w468**	هوَس	hawaaas
w111	جز	jazzaaa		**w290**	يَصُدُ	yaSudu		**w469**	هُود	huuwida
w112	هج	hajaa		**w291**	رَصَّ	raSSa		**w470**	رَهِبَ	rahiba
w113	جوق	jaawwq		**w292**	قُرص	qurSi		**w471**	رَهُفَ	rahuufa
w114	جمل	jamaal		**w293**	فصٌّ	faSSun		**w472**	قَهَرَ	qahara
w115	جُهد	juhhd		**w294**	فصاً	faSSan		**w473**	نوَّهَ	nawwaha
w116	جِدَّ	jiiddu		**w295**	فصٍ	faSSin		**w474**	مِنهُ	minhu
w117	وَجَد	wajada		**w296**	ضغث	DHaghath		**w475**	فيهِ	fiyhi
w118	أجد	Ajidu		**w297**	وضع	waDHaAa		**w476**	جاهاً	jaaahaaan
w119	حُجُبْ	Hujuub		**w298**	ضن	DHanna		**w477**	جاةٌ	jaaahun
w120	دَرَجَ	daraja		**w299**	ضل	DHalla		**w478**	جاهٍ	jaaahin
w121	سَرجُ	sarju		**w300**	ضيق	DHayq		**w479**	وتر	witr
w122	وَهَج	wahaji		**w301**	ضفر	DHafar		**w480**	وُجَد	wujida

w123	عوجاً	Aiiwajaaan		w302	ضَرَبَ	DHaraba		w481	أودُ	Aawiddu
w124	عوجٌ	Aiiwajun		w303	ضُحى	DHuHa		w482	دُوُو	dhawuwu
w125	عوج	Aiiwajin		w304	ضِدُ	DHiddu		w483	مَحوُ	maHwu
w126	حضر	HaaDHara		w305	رَضِيَ	raDHiya		w484	لَهو	lahuwi
w127	قزح	qazaH		w306	عَضُد	AaDHud		w485	سَهوَ	sahuwaa
w128	صحف	SuHuf		w307	فضّلَ	faDHDHala		w486	جَروأً	jarwaaan
w129	حذو	Hadhw		w308	مَرَضُ	maraDHa		w487	جَروٌ	jarwuun
w130	حطم	HaaTTama		w309	عَرَضَ	AaraDHa		w488	جَرو	jarwin
w131	حسن	Hasan		w310	أرض	AarDHi		w489	يَد	yad
w132	حك	Hakkaa		w311	قرضاً	qarDHaaan		w490	يُسر	yusr
w133	حله	Hallahu		w312	قرضٌ	qarDHun		w491	ين	yin
w134	حي	Haay		w313	قرض	qarDHin		w492	سَيَّرَ	sayyaara
w135	حَمَلَ	Haamala		w314	طق	Taq		w493	عِيَ	Aayiya
w136	حِبر	Hibr		w315	هطل	haTala		w494	سوي	saawayaii
w137	حُسن	Husn		w316	طمَعَ	TamaAa		w495	طيُ	Taayauu
w138	طحَنَ	TaHana		w317	طِب	Tib		w496	هَيُوَ	hayuUa
w139	سُحُب	suHub		w318	طبعَ	TubiiAa		w497	عَتِياً	Aatiyaaan
w140	يَحِلُ	yaHiilu		w319	وَطن	waTan		w498	عَتِيٌ	Aatiyun
w141	مَسَحَ	masaHa		w320	رَطِب	raTib		w499	عَتِي	Aatiyin
w142	فرَح	faraHi		w321	عُطف	AuTuf				
w143	مَرَحُ	maraaHu		w322	قِرطُ	qirTu				
w144	قزحً	qazaHan		w323	وَسَطِ	wasaTi				
w145	قزح	qazaHin		w324	بَسَط	basaTa				
w146	قزحٌ	qazaHun		w325	نُقطأً	nuqaTaaan				
w147	ضخه	DHakhahu		w326	نُقطَ	nuqaTun				
w148	خدك	khadduk		w327	نُقطِ	nuqaTin				
w149	خث	khath		w328	ظهر	THahar				
w150	خشع	khashaAa		w329	كظ	kaTHu				
w151	خص	khaSsa		w330	وظف	waTHafa				
w152	ذخر	dhakhkhara		w331	ظرف	THarf				
w153	خزق	khazaqa		w332	ظِفر	THifr				

w154	خسف	khasafa		**w333**	ظل	THul
w155	خمن	khamana		**w334**	نظرَ	naTHara
w156	خوي	khawy		**w335**	نظفَ	naTHufa
w157	خس	khas		**w336**	عَظِمَ	AaTHima
w158	خِدر	khidr		**w337**	حفظ	HafaTHa
w159	خُمس	khuums		**w338**	قيظِ	qayTHi
w160	بَخَسَ	bakhasa		**w339**	حَظ	HaTHTHu
w161	بَخِلَ	bakhila		**w340**	وَعظاً	waATHaaan
w162	رَخُصَ	rakhuSa		**w341**	وَعظ	waATHun
w163	صَرَخَ	Saarakha		**w342**	وَعظٍ	waATHin
w164	مُخ	mukhkhi		**w343**	عذل	Aadhl
w165	سَلَخُ	salkhu		**w344**	سعف	saAf
w166	بذخٌ	badhakhun		**w345**	عطش	AaTash
w167	بذخ	badhakhin		**w346**	عكس	Aakas
w168	بذخًا	badhakhaaan		**w347**	عقر	Aaqar
w169	ضد	DHid		**w348**	عَين	Aayn
w170	زند	zand		**w349**	عِجل	Aijl
w171	رصد	raSSada		**w350**	عُمر	Aumr
w172	قدم	qadam		**w351**	سَعِر	saAir
w173	طود	Tawwd		**w352**	نَعَسَ	naAasa
w174	دس	dassa		**w353**	نْعُمَ	naAuma
w175	دغل	daghghl		**w354**	وَسِعَ	wasiAa
w176	دعه	daAAhu		**w355**	قاع	qaaaAi
w177	دف	daf		**w356**	صاعُ	SaaaAu
w178	دمع	damA		**w357**	وَرعاً	wariAaaan
w179	دُب	duub		**w358**	وَرعٌ	wariAun

Table 1- The 499 chosen words

Appendix I

The transliterations of the 499 words using the Buckwalter, Arabtex, Alghamdi, Qalam, United Nations, and the two improved tables (SLT & LDPT).

arabtex words	Buckwalter words	Alghamdi words	Qalam words	UN words	Improved SLT table	Improved LDPT table
.da`a	DAEa	dhaaa	Daa'a	daa'a	DHaaAa	Dhaaa
Aa.zen	AaZen	aathin	AaZen	aathin	AaTHen	Aaathen
.sa.ga	SAga	saagha	Saagha	saagha	Saaga	Saaghaa
e_da	e*A	ethaa	edhaa	edhaa	edhaa	Edhaa
zar	zAr	zaar	zaar	zaar	zzaarr	Zzaarr
qas	qAs	qaas	qaas	qaas	kaas	Kaas
Aamal	Aamal	aamal	Aamal	aamal	Aamall	Aaamaall
^ga_t_A	javY	jatha	jathae	jatha	jjatha	jjaatha
^sa.h	$AH	shaah	shaaH	shaah	shaaH	shaah
.taf	TAf	taaf	Taaf	taaf	TTaaf	ttaaf
hayaAa	hayaAa	hayaaa	hayaAa	hayaaa	hayaAa	haayaaaaa
kaAs	kaAs	kaas	kaAs	kaas	kkaAs	kkaaasu
Ao_ht	Aoxt	aukht	Aokht	aokht	Aokt	aukht
bada	bAda	baada	baada	baada	baada	baadaa
Aaw	Aaw	aaw	Aaw	aaw	Aaw	aaaw
Aakala	Aakala	aakala	Aakala	aakala	Aakkalla	aaakkaallaa
saAal	saAal	saaal	saAal	saaal	saAall	saaaaall
.doUl	Doaul	dhuUl	DoUl	doUl	DhoUll	dhuUll
baiusa	baausa	baiisa	baiusa	baiusa	baiesa	baaiisaa
baraA	baraA	baraa	baraA	baraa	barraA	baarraau
sw'a	sw'i	swai	sw'a	sw'ai	swae	swai
da'a	dA'an	daaaan	daa'a	daa'aan	daaaan	daaaan
da'on	dA'on	daaaun	daa'on	daa'aon	daaaon	daaaun
da'en	dA'en	daaain	daa'en	daa'ain	daaaen	Daaain

arabtex words	Buckwalter words	Alghamdi words	Qalam words	UN words	Improved SLT table	Improved LDPT table
.zaby	Zaby	thaby	Zabye	thaby	THaby	thaaby
.dab`	DabE	dhaba	Dab'	dab'	DhabA	dhaaba
baza.ga	bazaga	bazagha	bazagha	bazagha	bazzaga	baazzaaghaa
ba.sal	baSal	basal	baSal	basal	baSall	baasaall
bahaq	bahaq	bahaq	bahaq	bahaq	bahak	baahaak
_haba.t	xabaT	khabat	khabaT	khabat	kabaTT	khaabaatt
kaba	kaba	kaba	kaba	kaba	kkaba	kkaabaa
_danb	*anb	thanb	dhanb	dhanb	dhanb	dhaanb
ba^sema	ba$ema	bashima	bashema	bashima	bashema	baashimaa
.saba	Saba	saba	Saba	saba	Saba	saabaa
farabo	farabo	farabu	farabo	farabo	farrabo	faarraabu
nasab	nasab	nasab	nasab	nasab	nasab	naasaab
wa^geba	wajeba	wajiba	wajeba	wajiba	wajjeba	waajjibaa
_tabata	vabata	thabata	thabata	thabata	thabata	thaabaata
ba.tala	baTala	batala	baTala	batala	baTTalla	baattaallaa
be^sr	be$r	bishr	beshr	bishr	beshrr	beshrr
bor^g	borj	burj	borj	borj	borrjj	burrjj
^gobela	jobela	jubila	jobela	jobila	jjobella	jjubellaa
raba.ta	rabaTa	rabata	rabaTa	rabata	rrabaTTa	rraabaattaa
sobol	sobol	subul	sobol	sobol	soboll	subull
.halaba	Halaba	halaba	Halaba	halaba	Hallaba	haallaabaa
qalbe	qalbe	qalbi	qalbe	qalbi	kallbe	kaallbe
nabo	nAbo	naabu	naabo	naabo	naabo	naabu

arabtex words	Buckwalter words	Alghamdi words	Qalam words	UN words	Improved SLT table	Improved LDPT table
_tawbana	vawbanA	thawbanaa	thawbanaa	thawbanaa	thawbanaa	thaawbanaa
_tawbon	vawbon	thawbun	thawbon	thawbon	thawbon	thaawbun
_tawben	vawben	thawbin	thawben	thawbin	thawben	thaawbin
ta.ht	taHt	taht	taHt	taht	taHt	taht
.damat	Damat	dhamat	Damat	damat	Dhamat	dhaamaat
ta.zel	taZel	tathil	taZel	tathil	taTHell	tathell
sa.tat	saTat	satat	saTat	satat	saTTat	saattaat
.sakat	Sakat	sakat	Sakat	sakat	Sakkat	saakkaat
_darat	*arat	tharat	dharat	dharat	dharrat	dhaarraat
hazat	hazat	hazat	hazat	hazat	hazzat	haazzaat
^sadat	$adat	shadat	shadat	shadat	shadat	shaadaat
_tanat	vanat	thanat	thanat	thanat	thanat	thaanaat
^gafat	jafat	jafat	jafat	jafat	jjafat	jjaafaat
`otw	Eotw	otw	'otw	'otw	Aotw	otw
.gat	gat	ghat	ghat	ghat	gat	ghaat
taqy	taqy	taqy	taqye	taqy	taky	taky
tamr	tamr	tamr	tamr	tamr	tamrr	tamrr
tyn	tyn	tyn	tyen	tyn	tyn	tyn
twt	twt	twt	twt	twt	twt	twt
qatala	qatala	qatala	qatala	qatala	katalla	kaatallaa
sotera	sotera	sutira	sotera	sotira	soterra	sutirraa
`atoma	Eatoma	atuma	'atoma	'atoma	Aatoma	atumaa
yometo	yometo	yumitu	yometo	yomito	yometo	yumitu
ya_hte	yaxte	yakhti	yakhte	yakhti	yakte	Yaakhti

arabtex words	Buckwalter words	Alghamdi words	Qalam words	UN words	Improved SLT table	Improved LDPT table
na.hata	naHata	nahata	naHata	nahata	naHata	naahaata
.samton	Samton	samtun	Samton	samton	Samton	saamtun
.samtan	Samtan	samtan	Samtan	samt'an	Samtan	saamtan
.samten	Samten	samtin	Samten	samtin	Samten	saamtin
_tol_t	volv	thulth	tholth	tholth	thollth	thollth
_taqaf	vaqaf	thaqaf	thaqaf	thaqaf	thakaf	thaakaaf
maka_ta	makava	makatha	makatha	makatha	makkatha	maakkaathaa
.ga_t	gav	ghath	ghath	ghath	gath	ghaath
.hada_t	Hadav	hadath	Hadath	hadath	Hadath	haadaath
^sara_t	$arav	sharath	sharath	sharath	sharrath	shaarraath
`a_t	Eav	ath	'ath	'ath	Aath	ath
_tawy	vawy	thawy	thawye	thawy	thawy	thaawy
_ta_hn	vaxn	thakhn	thakhn	thakhn	thakn	thaakhn
ba_taho	bavaho	bathahu	bathaho	bathaho	bathaho	baathaahu
_taba.ta	vabaTa	thabata	thabaTa	thabata	thabaTTa	thaabaattaa
_ta^ga	vaja	thaja	thaja	thaja	thajja	thaajjaa
_teny	veny	thiny	thenye	thiny	theny	thiny
_tolat	volat	thulat	tholat	tholat	thollat	thollaat
wa_taba	wavaba	wathaba	wathaba	wathaba	wathaba	waathaabaa
`o_tera	Eovera	othira	'othera	'othira	Aotherra	othirraa
^go_tw	jovw	juthw	jothw	jothw	jjothw	jjuthw
`a_ta	EAva	aaatha	'aatha	'aatha	Aaatha	aaathaa
ra_te	rave	rathi	rathe	rathi	rrathe	Rraathi

arabtex words	Buckwalter words	Alghamdi words	Qalam words	UN words	Improved SLT table	Improved LDPT table
ba_to	bavo	bathu	batho	batho	batho	baatho
_tolo_ten	voloven	thuluthin	tholothen	tholothin	thollothen	tholluthin
_tolo_ton	volovon	thuluthun	tholothon	tholothon	thollothon	tholluthun
_tolo_tan	volovan	thuluthan	tholothan	tholoth'an	thollothan	tholluthan
lo^ga^g	lojaj	lujaj	lojaj	lojaj	llojjajj	llujjaajj
^garaka	jaraka	jaraka	jaraka	jaraka	jjarrakka	jjaarraakkaa
.da^ga	Daja	dhaja	Daja	daja	Dhajja	dhaajjaa
^ga.s	jaS	jas	jaS	jas	jjaS	jjaas
_ha^gal	xajal	khajal	khajal	khajal	kajjall	khaajjaall
^ga.ha.za	jaHaZa	jahatha	jaHaZa	jahatha	jjaHaTHa	jjaahaathaa
.ta^gan	Tajan	tajan	Tajan	tajan	TTajjan	ttaajjaan
^sa^ga	$aja	shaja	shaja	shaja	shajja	shaajjaa
`a^gaza	Eajaza	ajaza	'ajaza	'ajaza	Aajjazza	ajjaazzaa
sa^ga`	sajaE	sajaa	saja'	saja'	sajjaA	saajjaaa
^go_dm	jo*m	juthm	jodhm	jodhm	jjodhm	jjudhm
^gady	jady	jady	jadye	jady	jjady	jjaady
^gaza	jaza	jaza	jaza	jaza	jjazza	jjaazzaa
ha^ga	haja	haja	haja	haja	hajja	haajjaa
^gawq	jawq	jawq	jawq	jawq	jjawk	jjaawk
^gamal	jamal	jamal	jamal	jamal	jjamall	jjaamaall
^gohd	johd	juhd	johd	johd	jjohd	jjuhd
^gedo	jedo	jidu	jedo	jido	jjedo	jjidu
wa^gada	wajada	wajada	wajada	wajada	wajjada	waajjaadaa
Aa^gedo	Aajedo	aajidu	Aajedo	aajido	Aajjedo	Aaajjidu

arabtex words	Buckwalter words	Alghamdi words	Qalam words	UN words	Improved SLT table	Improved LDPT table
.ho^gob	Hojob	hujub	Hojob	hojob	Hojjob	hujjub
dara^ga	daraja	daraja	daraja	daraja	darrajja	daarraajjaa
sar^go	sarjo	sarju	sarjo	sarjo	sarrjjo	saarrjju
waha^ge	wahaje	wahaji	wahaje	wahaji	wahajje	waahaajji
`ewa^gan	Eewajan	ewajan	'ewajan	'iwaj'an	Aewajjan	eewaajjan
`ewa^gon	Eewajon	ewajun	'ewajon	'iwajon	Aewajjon	eewaajjun
`ewa^gen	Eewajen	ewajin	'ewajen	'iwajin	Aewajjen	eewaajjin
.ha.dara	HaDara	hadhara	HaDara	hadara	HaDharra	haadhaarraa
qaza.h	qazaH	qazah	qazaH	qazah	kazzaH	kaazzaah
.so.hof	SoHof	suhuf	SoHof	sohof	SoHof	suhuf
.ha_dw	Ha*w	hathw	Hadhw	hadhw	Hadhw	haadhw
.ha.tama	HaTama	hatama	HaTama	hatama	HaTTama	haattaamaa
.hasan	Hasan	hasan	Hasan	hasan	Hasan	haasaan
.haka	Haka	haka	Haka	haka	Hakka	haakkaa
.halaho	Halaho	halahu	Halaho	halaho	Hallaho	haallaahu
.hay	Hay	hay	Haye	hay	Hay	haay
.hamala	Hamala	hamala	Hamala	hamala	Hamalla	haamaallaa
.hebr	Hebr	hibr	Hebr	hibr	Hebrr	hibrr
.hosn	Hosn	husn	Hosn	hosn	Hosn	husn
.ta.hana	TaHana	tahana	TaHana	tahana	TTaHana	ttaahaanaa
so.hob	soHob	suhub	soHob	sohob	soHob	suhub
ya.helo	yaHelo	yahilu	yaHelo	yahilo	yaHello	yaahillu
masa.ha	masaHa	masaha	masaHa	masaha	masaHa	maasaahaa
fara.he	faraHe	farahi	faraHe	farahi	farraHe	Faarraahi

arabtex words	Buckwalter words	Alghamdi words	Qalam words	UN words	Improved SLT table	Improved LDPT table
mara.ho	maraHo	marahu	maraHo	maraho	marraHo	maarraahu
qaza.hana	qazaHanA	qazahanaa	qazaHanaa	qazahanaa	kazzaHanaa	kaazzaahanaa
qaza.hen	qazaHen	qazahin	qazaHen	qazahin	kazzaHen	kaazzaahin
qaza.hon	qazaHon	qazahun	qazaHon	qazahon	kazzaHon	kaazzaahun
.da_haho	Daxaho	dhakhahu	Dakhaho	dakhaho	Dhakaho	dhaakhaahu
_hadok	xadok	khaduk	khadok	khadok	kadokk	khaadukk
_ha_t	xav	khath	khath	khath	kath	khaath
_ha^sa`a	xa$aEa	khashaa	khasha'a	khasha'a	kashaAa	khaashaaa
_ha.sa	xaSa	khasa	khaSa	khasa	kaSa	khaasaa
_da_hara	*axara	thakhara	dhakhara	dhakhara	dhakarra	dhaakhaarraa
_hazaqa	xazaqa	khazaqa	khazaqa	khazaqa	kazzaka	khaazzaakaa
_hasafa	xasafa	khasafa	khasafa	khasafa	kasafa	khaasaafaa
_hamana	xamana	khamana	khamana	khamana	kamana	khaamaanaa
_hawy	xawy	khawy	khawye	khawy	kawy	khaawy
_has	xas	khas	khas	khas	kas	khaas
_hedr	xedr	khidr	khedr	khidr	kedrr	khidrr
_homs	xoms	khums	khoms	khoms	koms	khums
ba_hasa	baxasa	bakhasa	bakhasa	bakhasa	bakasa	baakhaasaa
ba_hela	baxela	bakhila	bakhela	bakhila	bakella	baakhillaa
ra_ho.sa	raxoSa	rakhusa	rakhoSa	rakhosa	rrakoSa	Rraakhusaa

arabtex words	Buckwalter words	Alghamdi words	Qalam words	UN words	Improved SLT table	Improved LDPT table
.sara_ha	Saraxa	sarakha	Sarakha	sarakha	Sarraka	saarraakhaa
mo_he	moxe	mukhi	mokhe	mokhi	moke	mukhi
sal_ho	salxo	salkhu	salkho	salkho	sallko	saallkhu
ba_da_hon	ba*axon	bathakhun	badhakhon	badhakhon	badhakon	baadhaakhun
ba_da_hen	ba*axen	bathakhin	badhakhen	badhakhin	badhaken	baadhaakhin
ba_da_hana	ba*axanA	bathakhanaa	badhakhanaa	badhakhanaa	badhakanaa	baadhaakhanaa
.ded	Ded	dhid	Ded	did	Dhed	dhid
zand	zand	zand	zand	zand	zzand	zzaand
ra.sada	raSada	rasada	raSada	rasada	rraSada	rraasaadaa
qadam	qadam	qadam	qadam	qadam	kadam	kaadaam
.tawd	Tawd	tawd	Tawd	tawd	TTawd	ttaawd
dasa	dasa	dasa	dasa	dasa	dasa	daasaa
da.gl	dagl	dagl	daghl	daghl	dagll	daagll
da`ho	daEho	daahu	da'ho	da'ho	daAho	daaahu
daf	daf	daf	daf	daf	daf	daaf
dam`	damE	dama	dam'	dam'	damA	daama
dob	dob	dub	dob	dob	dob	dub
deyk	deyk	diyk	deyek	diyk	deykk	diykk
nadaba	nadaba	nadaba	nadaba	nadaba	nadaba	naadaabaa
.hodeda	Hodeda	hudida	Hodeda	hodida	Hodeda	hudidaa
modon	modon	mudun	modon	modon	modon	Mudun

arabtex words	Buckwalter words	Alghamdi words	Qalam words	UN words	Improved SLT table	Improved LDPT table
sada	sada	sada	sada	sada	sada	saadaa
`ahdo	Eahdo	ahdu	'ahdo	'ahdo	Aahdo	ahdu
mahde	mahde	mahdi	mahde	mahdi	mahde	maahdi
wa`dan	waEdan	waadan	wa'dan	wa'd'an	waAdan	waaadan
wa`don	waEdon	waadun	wa'don	wa'don	waAdon	waaadun
wa`den	waEden	waadin	wa'den	wa'din	waAden	waaadin
_daky	*aky	thaky	dhakye	dhaky	dhakky	dhaakky
_dama	*ama	thama	dhama	dhama	dhama	dhaamaa
_dala	*ala	thala	dhala	dhala	dhalla	dhaallaa
fa_d	fa*	fath	fadh	fadh	fadh	faadh
qa_da`	qa*aE	qathaa	qadha'	qadha'	kadhaA	kaadhaaa
^sa_da	$a*a	shatha	shadha	shadha	shadha	shaadhaa
_dawd	*awd	thawd	dhawd	dhawd	dhawd	dhaawd
_deib	*eab	thiib	dheib	dhiib	dheib	dhiib
_dab	*Ab	thaab	dhaab	dhaab	dhaab	dhaab
_dol	*ol	thul	dhol	dhol	dholl	dhull
ka_deba	ka*eba	kathiba	kadheba	kadhiba	kkadheba	kkaadhibaa
`a_dara	Ea*ara	athara	'adhara	'adhara	Aadharra	adhaarraa
Aa_dona	Aa*ona	aathuna	Aadhona	aadhona	Aadhona	aaadhunaa
^sa.ha_da	$aHa*a	shahatha	shaHadha	shahadha	shaHadha	shaahaadhaa
mon_do	mon*o	munthu	mondho	mondho	mondho	mundhu
mU_de	mau*e	mUthi	mUdhe	mUdhi	mUdhe	mUdhi
fa_dana	fa*anA	fathanaa	fadhanaa	fadhanaa	fadhanaa	Faadhanaa

arabtex words	Buckwalter words	Alghamdi words	Qalam words	UN words	Improved SLT table	Improved LDPT table
fa_don	fa*on	fathun	fadhon	fadhon	fadhon	faadhun
fa_den	fa*en	fathin	fadhen	fadhin	fadhen	faadhin
.zahara	Zahara	thahara	Zahara	thahara	THaharra	thaahaarraa
qarn	qarn	qarn	qarn	qarn	karrn	kaarrn
rakala	rakala	rakala	rakala	rakala	rrakkalla	rraakkaallaa
.dara	Dara	dhara	Dara	dara	Dharra	dhaarraa
ra.gw	ragw	ragw	raghw	raghw	rragw	rraagw
.tayr	Tayr	tayr	Tayer	tayr	TTayrr	ttaayrr
ser	ser	sir	ser	sir	serr	sirr
rad	rad	rad	rad	rad	rrad	rraad
roba	roba	ruba	roba	roba	rroba	rrubaa
soror	soror	surur	soror	soror	sorrorr	surrurr
.harama	Harama	harama	Harama	harama	Harrama	haarraamaa
sery	sery	siry	serye	siry	serry	sirry
fa.tara	faTara	fatara	faTara	fatara	faTTarra	faattaarraa
^go.hre	joHre	juhri	joHre	johri	jjoHrre	jjuhrri
fekre	fekre	fikri	fekre	fikri	fekkrre	fikkrri
dahron	dahron	dahrun	dahron	dahron	dahrron	daahrrun
dahren	dahren	dahrin	dahren	dahrin	dahrren	daahrrin
dahran	dahran	dahran	dahran	dahr'an	dahrran	daahrran
zafa	zafa	zafa	zafa	zafa	zzafa	zzaafaa
za`ama	zaEama	zaama	za'ama	za'ama	zzaAama	zzaaamaa
zaky	zaky	zaky	zakye	zaky	zzakky	zzaakky
zo.hal	zoHal	zuhal	zoHal	zohal	zzoHall	Zzuhaall

arabtex words	Buckwalter words	Alghamdi words	Qalam words	UN words	Improved SLT table	Improved LDPT table
zara`a	zaraEa	zaraa	zara'a	zara'a	zzarraAa	zzaarraaa
zer	zer	zir	zer	zir	zzerr	zzirr
rozeq	rozeq	ruziq	rozeq	roziq	rrozzek	rruzzik
`azafa	Eazafa	azafa	'azafa	'azafa	Aazzafa	azzaafaa
^gozor	jozor	juzur	jozor	jozor	jjozzorr	jjuzzurr
faza	fAza	faaza	faaza	faaza	faazza	faazzaa
^gawzo	jawzo	jawzu	jawzo	jawzo	jjawzzo	jjaawzzu
kanze	kanze	kanzi	kanze	kanzi	kkanzze	kkaanzzi
felezan	felezan	filizan	felezan	filiz'an	fellezzan	fillizzan
felezon	felezon	filizun	felezon	filizon	fellezzon	fillizzun
felezen	felezen	filizin	felezen	filizin	fellezzen	fillizzin
^sams	$ams	shams	shams	shams	shams	shaams
.gasala	gasala	ghasala	ghasala	ghasala	gasalla	ghaasaallaa
sahw	sahw	sahw	sahw	sahw	sahw	saahw
kys	kys	kys	kyes	kys	kkys	kkys
.ders	Ders	dhirs	Ders	dirs	Dherrs	dhirrs
som	som	sum	som	som	som	sum
sakaba	sakaba	sakaba	sakaba	sakaba	sakkaba	saakkaabaa
se.hr	seHr	sihr	seHr	sihr	seHrr	sihrr
rosol	rosol	rusul	rosol	rosol	rrosoll	rrusull
`asal	Easal	asal	'asal	'asal	Aasall	asaall
naseya	naseya	nasiya	naseya	nasiya	naseya	naasiyaa
.habasa	Habasa	habasa	Habasa	habasa	Habasa	haabaasaa
.harasa	Harasa	harasa	Harasa	harasa	Harrasa	Haarraasaa

arabtex words	Buckwalter words	Alghamdi words	Qalam words	UN words	Improved SLT table	Improved LDPT table
farase	farase	farasi	farase	farasi	farrase	faarraasi
`orsan	Eorsan	orsan	'orsan	'ors'an	Aorrsan	orrsan
`orson	Eorson	orsun	'orson	'orson	Aorrson	orrsun
`orsen	Eorsen	orsin	'orsen	'orsin	Aorrsen	orrsin
^sa_dw	$a*w	shathw	shadhw	shadhw	shadhw	shaadhw
^sa.s	$aS	shas	shaS	shas	shaS	shaas
^sa.zaf	$aZaf	shathaf	shaZaf	shathaf	shaTHaf	shaathaaf
^sa.t	$aT	shat	shaT	shat	shaTT	shaatt
^so.gl	$ogl	shugl	shoghl	shoghl	shogll	shugll
qa^s	qa$	qash	qash	qash	kash	kaash
^sak	$ak	shak	shak	shak	shakk	shaakk
na^sez	na$ez	nashiz	nashez	nashiz	nashezz	naashizz
^sahy	$ahy	shahy	shahye	shahy	shahy	shaahy
^sa^gar	$ajar	shajar	shajar	shajar	shajjarr	shaajjaarr
^sebl	$ebl	shibl	shebl	shibl	shebll	shibll
^sokr	$okr	shukr	shokr	shokr	shokkrr	shukkrr
wa^sm	wa$m	washm	washm	washm	washm	waashm
ro^seda	ro$eda	rushida	rosheda	roshida	rrosheda	rrushidaa
Aa^sodo	Aa$odo	aashudu	Aashodo	aashodo	Aashodo	aaashudu
ra^sa	ra$a	rasha	rasha	rasha	rrasha	rraashaa
rem^so	rem$o	rimshu	remsho	rimsho	rremsho	rrimshu
ry^se	ry$e	ryshi	ryeshe	ryshi	rryshe	rryshi
kab^san	kab$an	kabshan	kabshan	kabsh'an	kkabshan	kkaabshan
kab^son	kab$on	kabshun	kabshon	kabshon	kkabshon	Kkaabshun

arabtex words	Buckwalter words	Alghamdi words	Qalam words	UN words	Improved SLT table	Improved LDPT table
kab^sen	kab$en	kabshin	kabshen	kabshin	kkabshen	kkaabshin
qa.sa	qaSa	qasa	qaSa	qasa	kaSa	kaasaa
.som	Som	sum	Som	som	Som	sum
.sana`a	SanaEa	sanaa	Sana'a	sana'a	SanaAa	saanaaa
.sah	Sah	sah	Sah	sah	Sah	saah
wa.sy	waSy	wasy	waSye	wasy	waSy	waasy
.sow.s	SowS	suws	SowS	sows	SowS	suws
.sayd	Sayd	sayd	Sayed	sayd	Sayd	saayd
.sehr	Sehr	sihr	Sehr	sihr	Sehrr	sihrr
`a.sara	EaSara	asara	'aSara	'asara	AaSarra	asaarraa
no.sera	noSera	nusira	noSera	nosira	noSerra	nusirraa
ya.sodo	yaSodo	yasudu	yaSodo	yasodo	yaSodo	yaasudu
ra.sa	raSa	rasa	raSa	rasa	rraSa	rraasaa
qor.se	qorSe	qursi	qorSe	qorsi	korrSe	kurrsi
fa.son	faSon	fasun	faSon	fason	faSon	faasun
fa.sa	faSa	fasa	faSa	fasa	faSa	faasaa
fa.sen	faSen	fasin	faSen	fasin	faSen	faasin
.da.ga_t	Dagav	dhaghath	Daghath	daghath	Dhagath	dhaaghaath
wa.da`a	waDaEa	wadhaa	waDa'a	wada'a	waDhaAa	waadhaaa
.dana	Dana	dhana	Dana	dana	Dhana	dhaanaa
.dala	Dala	dhala	Dala	dala	Dhalla	dhaallaa
.dyq	Dyq	dhyq	Dyeq	dyq	DHyk	dhyk
.dafar	Dafar	dhafar	Dafar	dafar	Dhafarr	Dhaafaarr

arabtex words	Buckwalter words	Alghamdi words	Qalam words	UN words	Improved SLT table	Improved LDPT table
.daraba	Daraba	dharaba	Daraba	daraba	Dharraba	dhaarraabaa
.do.h_A	DoHY	dhuha	DoHae	doha	DhoHa	dhuha
.dedo	Dedo	dhidu	Dedo	dido	Dhedo	dhidu
ra.deya	raDeya	radhiya	raDeya	radiya	rraDheya	rraadhiyaa
`a.dod	EaDod	adhud	'aDod	'adod	AaDhod	adhud
fa.dala	faDala	fadhala	faDala	fadala	faDHalla	faadhaallaa
mara.da	maraDa	maradha	maraDa	marada	marraDha	maarraadhaa
`ara.da	EaraDa	aradha	'araDa	'arada	AarraDha	arraadhaa
Aar.de	AarDe	aardhi	AarDe	aardi	AarrDhe	aaarrdhi
qar.dan	qarDan	qardhan	qarDan	qard'an	karrDHan	kaarrdhan
qar.don	qarDon	qardhun	qarDon	qardon	karrDhon	kaarrdhun
qar.den	qarDen	qardhin	qarDen	qardin	karrDhen	kaarrdhin
.taq	Taq	taq	Taq	taq	TTak	ttuk
ha.tala	haTala	hatala	haTala	hatala	haTTalla	haattaallaa
.tama`a	TamaEa	tamaa	Tama'a	tama'a	TTamaAa	ttaamaaa
.teb	Teb	tib	Teb	tib	TTeb	ttib
.taba`a	TabaEa	tabaa	Taba'a	taba'a	TTabaAa	ttaabaaa
wa.tan	waTan	watan	waTan	watan	waTTan	waattaan
ra.teb	raTeb	ratib	raTeb	ratib	rraTTeb	rraattib
`o.tof	EoTof	otuf	'oTof	'otof	AoTTof	otuf
qer.to	qerTo	qirtu	qerTo	qirto	kerrTTo	kirrtu
wasa.te	wasaTe	wasati	wasaTe	wasati	wasaTTe	Waasaatti

arabtex words	Buckwalter words	Alghamdi words	Qalam words	UN words	Improved SLT table	Improved LDPT table
basa.ta	basaTa	basata	basaTa	basata	basaTTa	baasaattaa
noqa.tan	noqaTan	nuqatan	noqaTan	noqat'an	nokaTTan	nukaattan
noqa.ton	noqaTon	nuqatun	noqaTon	noqaton	nokaTTon	nukaattun
noqa.ten	noqaTen	nuqatin	noqaTen	noqatin	nokaTTen	nukaattin
.zahar	Zahar	thahar	Zahar	thahar	THaharr	thaahaarr
ka.zo	kaZo	kathu	kaZo	katho	kkaTHo	kkaatho
wa.zafa	waZafa	wathafa	waZafa	wathafa	waTHafa	waathaafaa
.zarf	Zarf	tharf	Zarf	tharf	THarrf	thaarrf
.zefr	Zefr	thifr	Zefr	thifr	THefrr	thefrr
.zol	Zol	thul	Zol	thol	THoll	tholl
na.zara	naZara	nathara	naZara	nathara	naTHarra	naathaarraa
na.zofa	naZofa	nathufa	naZofa	nathofa	naTHofa	naathofaa
`a.zema	EaZema	athima	'aZema	'athima	AaTHema	athemaa
.hafa.za	HafaZa	hafatha	HafaZa	hafatha	HafaTHa	haafaathaa
qay.ze	qayZe	qaythi	qayeZe	qaythi	kayTHe	kaaythe
.ha.zo	HaZo	hathu	HaZo	hatho	HaTHo	haatho
wa`.zan	waEZan	waathan	wa'Zan	wa'th'an	waATHan	waaathan
wa`.zon	waEZon	waathun	wa'Zon	wa'thon	waATHon	waaathun
wa`.zen	waEZen	waathin	wa'Zen	wa'thin	waATHen	waaathin
`a_dal	Ea*al	athal	'adhal	'adhal	Aadhall	adhaall
sa`af	saEaf	saaf	sa'af	sa'af	saAaf	saaaf
`a.ta^s	EaTa$	atash	'aTash	'atash	AaTTash	attaash
`aks	Eaks	aks	'aks	'aks	Aakks	akks
`aqr	Eaqr	aqr	'aqr	'aqr	Aakrr	Akrr

arabtex words	Buckwalter words	Alghamdi words	Qalam words	UN words	Improved SLT table	Improved LDPT table
`ayn	Eayn	ayn	'ayen	'ayn	Aayn	ayn
`e^gl	Eejl	ejl	'ejl	'ijl	Aejjll	eejjll
`omr	Eomr	omr	'omr	'omr	Aomrr	omrr
sa`er	saEer	saer	sa'er	sa'ir	saAerr	saaeerr
na`asa	naEasa	naasa	na'asa	na'asa	naAasa	naaasaa
na`oma	naEoma	naoma	na'oma	na'oma	naAoma	naaomaa
wase`a	waseEa	wasia	wase'a	wasi'a	waseAa	waasia
qa`e	qAEe	qaae	qaa'e	qaa'i	kaaAe	kaaee
.sa`o	SAEo	saao	Saa'o	saa'o	SaaAo	saao
ware`an	wareEan	wariaan	ware'an	wari''an	warreAan	waarriaan
ware`on	wareEon	wariaun	ware'on	wari'on	warreAon	waarriaun
ware`en	wareEen	wariain	ware'en	wari'in	warreAen	waarriain
.ga^gar	gajar	ghajar	ghajar	ghajar	gajjarr	ghaajjaarr
.ga_d_A	ga*Y	ghatha	ghadhae	ghadha	gadha	ghaadha
.ga^sa	ga$a	ghasha	ghasha	ghasha	gasha	ghaashaa
.ga.d	gaD	ghadh	ghaDH	ghad	gaDH	ghaadhu
.gafeya	gafeya	ghafiya	ghafeya	ghafiya	gafeya	ghaafiyaa
.garaqa	garaqa	gharaqa	gharaqa	gharaqa	garraka	ghaarraakaa
.gay.zo	gayZo	ghaythu	ghayeZo	ghaytho	gayTHo	ghaaytho
.ga.ta	gaTa	ghata	ghaTa	ghata	gaTTa	ghaattaa
.gadara	gadara	ghadara	ghadara	ghadara	gadarra	ghaadaarraa
.go.sn	goSn	ghusn	ghoSn	ghosn	goSn	ghusn
.gel	gel	ghil	ghel	ghil	gell	Ghill

arabtex words	Buckwalter words	Alghamdi words	Qalam words	UN words	Improved SLT table	Improved LDPT table
.sa.gora	Sagora	saghura	Saghora	saghora	Sagorra	saaghurraa
ra.gad	ragad	raghad	raghad	raghad	rragad	rraaghaad
.ta.geya	Tageya	taghiya	Tagheya	taghiya	TTageya	ttaaghiyaa
mara.ga	maraga	maragha	maragha	maragha	marraga	maarraaghaa
.sad.go	Sadgo	sadghu	Sadgho	sadgho	Sadgo	saadghu
.sam.ge	Samge	samghi	Samghe	samghi	Samge	saamghi
.saw.gan	Sawgan	sawgan	Sawghan	sawgh'an	Sawgan	saawgan
.saw.gon	Sawgon	sawghun	Sawghon	sawghon	Sawgon	saawghun
.saw.gen	Sawgen	sawghin	Sawghen	sawghin	Sawgen	saawghin
haf	haf	haf	haf	haf	haf	haaf
wafy	wafy	wafy	wafye	wafy	wafy	waafy
malaf	malaf	malaf	malaf	malaf	mallaf	maallaaf
faka	faka	faka	faka	faka	fakka	faakkaa
fan	fan	fan	fan	fan	fan	faan
fe^gl	fejl	fijl	fejl	fijl	fejjll	fijjll
forn	forn	furn	forn	forn	forrn	furrn
fa`ala	faEala	faala	fa'ala	fa'ala	faAalla	faaallaa
rafa`a	rafaEa	rafaa	rafa'a	rafa'a	rrafaAa	rraafaaa
dafera	dafera	dafira	dafera	dafira	daferra	daafirraa
`afwo	Eafwo	afwu	'afwo	'afwo	Aafwo	afwu
^sarafa	$arafa	sharafa	sharafa	sharafa	sharrafa	shaarraafaa
.tarafe	Tarafe	tarafi	Tarafe	tarafi	TTarrafe	ttaarraafi
_halfo	xalfo	khalfu	khalfo	khalfo	kallfo	Khaallfu

arabtex words	Buckwalter words	Alghamdi words	Qalam words	UN words	Improved SLT table	Improved LDPT table
`alafan	Ealafan	alafan	'alafan	'alaf'an	Aallafan	allaafan
`alafon	Ealafon	alafun	'alafon	'alafon	Aallafon	allaafun
`alafen	Ealafen	alafin	'alafen	'alafin	Aallafen	allaafin
so^goq	sojoq	sujuq	sojoq	sojoq	sojjok	sujjuk
qolw	qolw	qulw	qolw	qolw	kollw	kullw
daqaka	daqaka	daqaka	daqaka	daqaka	dakakka	daakaakkaa
qalam	qalam	qalam	qalam	qalam	kallam	kaallaam
qedr	qedr	qidr	qedr	qidr	kedrr	kidrr
qoda	qoda	quda	qoda	qoda	koda	kudaa
saqa.ta	saqaTa	saqata	saqaTa	saqata	sakaTTa	saakaattaa
foqeda	foqeda	fuqida	foqeda	foqida	fokeda	fukidaa
_taqola	vaqola	thaqula	thaqola	thaqola	thakolla	thaakullaa
sabaqa	sabaqa	sabaqa	sabaqa	sabaqa	sabaka	saabaakaa
`abaqa	Eabaqa	abaqa	'abaqa	'abaqa	Aabaka	abaakaa
.gasaqo	gasaqo	ghasaqu	ghasaqo	ghasaqo	gasako	ghaasaaku
barqan	barqan	barqan	barqan	barq'an	barrkan	baarrkan
barqon	barqon	barqun	barqon	barqon	barrkon	baarrkun
barqen	barqen	barqin	barqen	barqin	barrken	baarrkin
raka.da	rakaDa	rakadha	rakaDa	rakada	rrakkaDha	rraakkaadhaa
^garaka	jaraka	jaraka	jaraka	jaraka	jjarrakka	jjaarraakkaa
kaw_A	kawY	kawa	kawae	kawa	kkawa	kkaawa
kahan	kahan	kahan	kahan	kahan	kkahan	kkaahaan
kalb	kalb	kalb	kalb	kalb	kkallb	Kkaallb

arabtex words	Buckwalter words	Alghamdi words	Qalam words	UN words	Improved SLT table	Improved LDPT table
keys	keys	kiys	keyes	kiys	kkeys	kkiys
kow`	kowE	kuwa	kow'	kow'	kkowA	kkuwa
rakeba	rakeba	rakiba	rakeba	rakiba	rrakkeba	rraakkibaa
raka`a	rakaEa	rakaa	raka'a	raka'a	rrakkaAa	rraakkaaa
makona	makona	makuna	makona	makona	makkona	maakkunaa
.haraka	Haraka	haraka	Haraka	haraka	Harrakka	haarraakkaa
berako	berako	biraku	berako	birako	berrakko	berraakku
samake	samake	samaki	samake	samaki	samakke	saamaakki
selkan	selkan	silkan	selkan	silk'an	sellkkan	sillkkan
selkon	selkon	silkun	selkon	silkon	sellkkon	sillkkun
selken	selken	silkin	selken	silkin	sellkken	sillkkin
lay_t	layv	layth	layeth	layth	llayth	llaayth
leyn	leyn	liyn	leyen	liyn	lleyn	lliyn
lomat	lomat	lumat	lomat	lomat	llomat	llumaat
`olow	Eolow	oluw	'olow	'olow	Aollow	olluw
.gala_A	galaY	ghalaa	ghalaae	ghalaa	gallaa	ghaallaaa
^galey	jaley	jaliy	jaleye	jaliy	jjalley	jjaalliy
da.glo	daglo	daglu	daghlo	daghlo	dagllo	daagllu
`amale	Eamale	amali	'amale	'amali	Aamalle	amaalli
.tablan	Tablan	tablan	Tablan	tabl'an	TTabllan	ttaabllan
.tablon	Tablon	tablun	Tablon	tablon	TTabllon	ttaabllun
.tablen	Tablen	tablin	Tablen	tablin	TTabllen	ttaabllin
ham	ham	ham	ham	ham	ham	haam
yawm	yawm	yawm	yawm	yawm	yawm	Yaawm

arabtex words	Buckwalter words	Alghamdi words	Qalam words	UN words	Improved SLT table	Improved LDPT table
mawz	mawz	mawz	mawz	mawz	mawzz	maawzz
men	men	min	men	min	men	min
Aamel	Aamel	aamil	Aamel	aamil	Aamell	aaamill
`amala	Eamala	amala	'amala	'amala	Aamalla	amaallaa
nomow	nomow	numuw	nomow	nomow	nomow	numuw
fahama	fahama	fahama	fahama	fahama	fahama	faahaamaa
.ganamo	ganamo	ghanamu	ghanamo	ghanamo	ganamo	ghaanaamu
.saname	Saname	sanami	Saname	sanami	Saname	saanaami
`alaman	Ealaman	alaman	'alaman	'alam'an	Aallaman	allaaman
`alamon	Ealamon	alamun	'alamon	'alamon	Aallamon	allaamun
`alamen	Ealamen	alamin	'alamen	'alamin	Aallamen	allaamin
wana	wana	wana	wana	wana	wana	waanaa
nahr	nahr	nahr	nahr	nahr	nahrr	naahrr
nebr	nebr	nibr	nebr	nibr	nebrr	nibrr
nowr	nowr	nuwr	nowr	nowr	nowrr	nuwrr
fanar	fanar	fanar	fanar	fanar	fanarr	faanaarr
saneya	saneya	saniya	saneya	saniya	saneya	saaniyaa
.honow	Honow	hunuw	Honow	honow	Honow	hunuw
makana	makana	makana	makana	makana	makkana	maakkaanaa
na.hno	naHno	nahnu	naHno	nahno	naHno	naahnu
_dehne	*ehne	thihni	dhehne	dhihni	dhehne	dhihni
qarnan	qarnan	qarnan	qarnan	qarn'an	karrnan	kaarrnan
qarnon	qarnon	qarnun	qarnon	qarnon	karrnon	Kaarrnun

arabtex words	Buckwalter words	Alghamdi words	Qalam words	UN words	Improved SLT table	Improved LDPT table
qarnen	qarnen	qarnin	qarnen	qarnin	karrnen	kaarrnin
.garaho	garaho	gharahu	gharaho	gharaho	garraho	ghaarraahu
_talaho	valaho	thalahu	thalaho	thalaho	thallaho	thaallaahu
her	her	hir	her	hir	herr	hirr
hawas	hawas	hawas	hawas	hawas	hawas	haawaas
howed	howed	huwid	howed	howid	howed	huwid
raheba	raheba	rahiba	raheba	rahiba	rraheba	rraahibaa
rahofa	rahofa	rahufa	rahofa	rahofa	rrahofa	rraahufaa
qahara	qahara	qahara	qahara	qahara	kaharra	kaahaarraa
nawaha	nawaha	nawaha	nawaha	nawaha	nawaha	naawaahaa
menho	menho	minhu	menho	minho	menho	minhu
fyhe	fyhe	fyhi	fyehe	fyhi	fyhe	fyhi
^gahan	jAhan	jaahan	jaahan	jaah'an	jjaahan	jjaahan
^gahon	jAhon	jaahun	jaahon	jaahon	jjaahon	jjaahun
^gahen	jAhen	jaahin	jaahen	jaahin	jjaahen	jjaahin
wetr	wetr	witr	wetr	witr	wetrr	witrr
wo^geda	wojeda	wujida	wojeda	wojida	wojjeda	wujjidaa
Aawedo	Aawedo	aawidu	Aawedo	aawido	Aawedo	aaawidu
_dawow	*awow	thawuw	dhawow	dhawow	dhawow	dhaawuw
ma.hwo	maHwo	mahwu	maHwo	mahwo	maHwo	maahwu
lahwe	lahwe	lahwi	lahwe	lahwi	llahwe	llaahwi
sahwa	sahwa	sahwa	sahwa	sahwa	sahwa	saahwaa
^garwan	jarwan	jarwan	jarwan	jarw'an	jjarrwan	jjaarrwan
^garwon	jarwon	jarwun	jarwon	jarwon	jjarrwon	Jjaarrwun

arabtex words	Buckwalter words	Alghamdi words	Qalam words	UN words	Improved SLT table	Improved LDPT table
^garwen	jarwen	jarwin	jarwen	jarwin	jjarrwen	jjaarrwin
yad	yad	yad	yad	yad	yad	yaad
yosr	yosr	yusr	yosr	yosr	yosrr	yusrr
yen	yen	yin	yen	yin	yen	yin
sayara	sayara	sayara	sayeara	sayara	sayarra	saayaarraa
`ayeya	Eayeya	ayiya	'ayeya	'ayiya	Aayeya	ayiyaa
sawye	sawye	sawyi	sawye	sawyi	sawye	saawyi
.tayo	Tayo	tayu	Tayo	tayo	TTayo	ttaayu
hayoUa	hayooa	hayuUa	hayoUa	hayoUa	hayoUa	haayuUa
`atyan	Eatyan	atyan	'atyean	'aty'an	Aatyan	atyan
`atyon	Eatyon	atyun	'atyon	'atyon	Aatyon	atyun
`atyen	Eatyen	atyin	'atyen	'atyin	Aatyen	atyin

Table 1-Comparison of the transliterations of the 499 words.

Appendix J

Alghamdi's recognition analysis

Alghamdi's transliteration analysis for the four recordings.

Recording 4				Recording 3				Recording 2				Recording 1		
499 words	Recognised word	Match		499 words	Recognised word	Match		499 words	Recognised word	Match		499 words	Recognised word	Match
dhaaaa	thala	No		dhaaaa	tyn	No		dhaaaa	thala	No		dhaaaa	tyn	No
athin	thawbin	No		athin	fathanaa	No		athin	thawbun	No		athin	athin	Yes
saagha	sanami	No		saagha	thahara	No		saagha	haka	No		saagha	zaama	No
ithaa	dhidu	No		ithaa	sum	No		ithaa	sum	No		ithaa	dhidu	No
zaar	zaama	No		zaar	zafa	No		zaar	fahama	No		zaar	sanami	No
qaas	haf	No		qaas	qasa	No		qaas	haf	No		qaas	haf	No
aamal	qarnin	No		aamal	amala	No		aamal	aamal	Yes		aamal	qarnan	No
jatha	ghafiya	No		jatha	jatha	Yes		jatha	jatha	Yes		jatha	jatha	Yes
shaah	shaah	Yes		shaah	shahatha	No		shaah	shaah	Yes		shaah	fahama	No
taaf	yin	No		taaf	thahar	No		taaf	fath	No		taaf	ham	No
hayaaa	shaja	No		hayaaa	sanami	No		hayaaa	sum	No		hayaaa	daaain	No
kaas	qash	No		kaas	qaas	No		kaas	qas	No		kaas	khath	No
aukht	ath	No		aukht	faaza	No		aukht	ath	No		aukht	ath	No
baada	ttaghiiya	No		baada	saagha	No		baada	dhaghath	No		baada	ttaghiiya	No
aaw	daahu	No		aaw	qarn	No		aaw	aaw	Yes		aaw	aaw	Yes
aakala	dhana	No		aakala	makana	No		aakala	sum	No		aakala	sayd	No
saaal	fahama	No		saaal	saqata	No		saaal	zaama	No		saaal	fan	No
dhuUl	sum	No		dhuUl	dahrun	No		dhuUl	sum	No		dhuUl	watan	No
baiusa	bathu	No		baiusa	naasa	No		baiusa	ghaythu	No		baiusa	kiys	No
baraa	dama	No		baraa	bathakhun	No		baraa	dama	No		baraa	dama	No

Recording 4				Recording 3				Recording 2				Recording 1		
499 words	**Recognised word**	**Match**		**499 words**	**Recognised word**	**Match**		**499 words**	**Recognised word**	**Match**		**499 words**	**Recognised word**	**Match**
swai	samghi	No		swai	thala	No		swai	fijl	No		swai	sayd	No
daaaan	daaain	No		daaaan	daaaan	Yes		daaaan	daaaun	No		daaaan	tyn	No
daaaun	daaaun	Yes		daaaun	hunuw	No		daaaun	daaun	No		daaaun	daaaun	Yes
daaain	ayn	No		daaain	daaain	Yes		daaain	daaain	Yes		daaain	ayn	No
thaby	sum	No		thaby	naabu	No		thaby	sum	No		thaby	tawd	No
dhaba	naabu	No		dhaba	dama	No		dhaba	naabu	No		dhaba	naabu	No
bazagha	ghasala	No		bazagha	basata	No		bazagha	ghasala	No		bazagha	bathu	No
basal	fasun	No		basal	dasa	No		basal	wathafa	No		basal	hasan	No
bahaq	baraa	No		bahaq	jahatha	No		bahaq	nahata	No		bahaq	ghadh	No
khabat	qathaa	No		khabat	habasa	No		khabat	sum	No		khabat	qathaa	No
kaba	khaath	No		kaba	sakaba	No		kaba	khath	No		kaba	thaab	No
thanb	thanb	Yes		thanb	daaaan	No		thanb	fan	No		thanb	thanb	Yes
bashima	bashima	Yes		bashima	bashima	Yes		bashima	bashima	Yes		bashima	bashima	Yes
saba	hhir	No		saba	kahan	No		saba	fahama	No		saba	fath	No
farabu	naabu	No		farabu	sum	No		farabu	fath	No		farabu	naabu	No
nasab	nasiya	No		nasab	dasa	No		nasab	naasa	No		nasab	naasa	No
wajiba	hudida	No		wajiba	hudida	No		wajiba	wasia	No		wajiba	mUthi	No
thabata	abaqa	No		thabata	sabaqa	No		thabata	abaqa	No		thabata	dama	No
batala	makana	No		batala	dhafar	No		batala	dhafar	No		batala	hafatha	No
bishr	ghasha	No		bishr	dasa	No		bishr	dhirs	No		bishr	layth	No
burj	daghl	No		burj	dahrun	No		burj	mawz	No		burj	liyn	No
jubila	wujida	No		jubila	sum	No		jubila	thabata	No		jubila	sayd	No
rabata	abaqa	No		rabata	habasa	No		rabata	nadaba	No		rabata	abaqa	No
subul	thawbin	No		subul	sawghun	No		subul	saba	No		subul	thawbin	No

Recording 4				Recording 3				Recording 2				Recording 1		
499 words	**Recognised word**	**Match**		**499 words**	**Recognised word**	**Match**		**499 words**	**Recognised word**	**Match**		**499 words**	**Recognised word**	**Match**
halaba	nadaba	No		halaba	hudida	No		halaba	sum	No		halaba	nadaba	No
qalbi	samghi	No		qalbi	samtun	No		qalbi	sum	No		qalbi	samghi	No
naabu	naabu	Yes		naabu	naabu	Yes		naabu	naabu	Yes		naabu	naabu	Yes
thawban	thawbin	No		thawbana	thawbun	No		thawbana	halaba	No		thawbanaa	aaw	No
thawbun	sum	No		thawbun	thul	No		thawbun	sum	No		thawbun	ghil	No
thawbin	alafin	No		thawbin	samtun	No		thawbin	alafin	No		thawbin	tyn	No
taht	sum	No		taht	thahar	No		taht	qash	No		taht	sum	No
dhamat	dhamat	Yes		dhamat	dama	No		dhamat	lumat	No		dhamat	fahama	No
tathil	tablun	No		tathil	qarnan	No		tathil	sum	No		tathil	tyn	No
satat	sakat	No		satat	sakat	No		satat	sum	No		satat	samtan	No
sakat	fath	No		sakat	thaab	No		sakat	ghadh	No		sakat	fath	No
tharat	thaarrat	No		tharat	dhana	No		tharat	dhara	No		tharat	dama	No
hazat	hazat	Yes		hazat	qazah	No		hazat	hazat	Yes		hazat	hazat	Yes
shadat	sayd	No		shadat	sum	No		shadat	shajar	No		shadat	sayd	No
thanat	thanat	Yes		thanat	dhana	No		thanat	thanat	Yes		thanat	fahama	No
jafat	jafat	Yes		jafat	jatha	No		jafat	jafat	Yes		jafat	shat	No
otw	samtun	No		otw	hathu	No		otw	aakala	No		otw	hathu	No
ghat	ttaq	No		ghat	thanat	No		ghat	taq	No		ghat	ttaq	No
taqy	rakaa	No		taqy	sum	No		taqy	thiny	No		taqy	zaky	No
tamr	fahama	No		tamr	fahama	No		tamr	samghi	No		tamr	thul	No
tyn	kiys	No		tyn	sawghan	No		tyn	ayn	No		tyn	liyn	No
twt	saaal	No		twt	tawd	No		twt	thala	No		twt	sayd	No
qatala	makana	No		qatala	ghasala	No		qatala	rakala	No		qatala	makana	No
sutira	sutira	Yes		sutira	sutira	Yes		sutira	sum	No		sutira	sutira	Yes

Recording 4			Recording 3			Recording 2			Recording 1		
499 words	**Recognised word**	**Match**	**499 words**	**Recognised word**	**Match**	**499 words**	**Recognised word**	**Match**	**499 words**	**Recognised word**	**Match**
atuma	thala	No	atuma	aakala	No	atuma	hatala	No	atuma	sayd	No
yumitu	ghil	No	yumitu	thul	No	yumitu	lumat	No	yumitu	sayd	No
yakhti	liyn	No	yakhti	ayn	No	yakhti	thaky	No	yakhti	yin	No
nahata	jahatha	No	nahata	jahatha	No	nahata	jahatha	No	nahata	jahatha	No
samtun	samtun	Yes	samtun	samtan	No	samtun	samtan	No	samtun	sawghan	No
samtan	samtan	Yes	samtan	samtun	No	samtan	samtun	No	samtan	fan	No
samtin	samtun	No	samtin	ayn	No	samtin	samtdin	No	samtin	ayn	No
thulth	layth	No	thulth	dhana	No	thulth	tharf	No	thulth	sawghan	No
thaqaf	Dafar	No	thaqaf	sum	No	thaqaf	dhaghath	No	thaqaf	Dafar	No
makatha	makaathaa	No	makatha	makatha	Yes	makatha	makatha	Yes	makatha	makaathaa	No
ghath	ghath	Yes	ghath	nahr	No	ghath	ghath	Yes	ghath	ghath	Yes
hadath	hadath	Yes	hadath	hadath	Yes	hadath	hadath	Yes	hadath	hadath	Yes
sharath	shathaf	No	sharath	shatha	No	sharath	shat	No	sharath	fahama	No
ath	haf	No	ath	ath	Yes	ath	haf	No	ath	haf	No
thawy	thawy	Yes	thawy	sum	No	thawy	jaliy	No	thawy	tyn	No
thakhn	kaahan	No	thakhn	kahan	No	thakhn	kahan	No	thakhn	fan	No
bathahu	ghil	No	bathahu	bathahu	Yes	bathahu	ghasala	No	bathahu	bathu	No
thabata	thabbata	No	thabata	nadaba	No	thabata	nadaba	No	thabata	thabbata	No
thaja	sum	No	thaja	shat	No	thaja	sum	No	thaja	sum	No
thiny	samghi	No	thiny	thiny	Yes	thiny	thiny	Yes	thiny	thiny	Yes
thulat	wana	No	thulat	fahama	No	thulat	faala	No	thulat	fahama	No
wathaba	wathaba	Yes	wathaba	sakaba	No	wathaba	wathafa	No	wathaba	wathafa	No
othira	wasia	No	othira	ghafiya	No	othira	sum	No	othira	wasia	No
juthw	ghaythu	No	juthw	ghaythu	No	juthw	jawzu	No	juthw	tathil	No

Recording 4				Recording 3				Recording 2				Recording 1		
499 words	**Recognised word**	**Match**		**499 words**	**Recognised word**	**Match**		**499 words**	**Recognised word**	**Match**		**499 words**	**Recognised word**	**Match**
aaatha	hasan	No		aaatha	qasa	No		aaatha	hasan	No		aaatha	hasan	No
rathi	liyn	No		rathi	bathahu	No		rathi	wafy	No		rathi	tyn	No
bathu	ghil	No		bathu	dafira	No		bathu	ghasala	No		bathu	layth	No
thuluthin	alafan	No		thuluthin	samtun	No		thuluthin	sawghun	No		thuluthin	alafan	No
thuluthu	samtun	No		thuluthun	thuluthun	Yes		thuluthun	samtun	No		thuluthun	tyn	No
thulutha	sawghun	No		thuluthan	samtun	No		thuluthan	sawghun	No		thuluthan	fan	No
lujaj	raghad	No		lujaj	raghad	No		lujaj	nadaba	No		lujaj	nusira	No
jaraka	jaraka	Yes		jaraka	sum	No		jaraka	jaraka	Yes		jaraka	jaraka	Yes
dhaja	fath	No		dhaja	mahdi	No		dhaja	ghajar	No		dhaja	sayd	No
jas	ghath	No		jas	jas	Yes		jas	jas	Yes		jas	jas	Yes
khajal	fathin	No		khajal	fathanaa	No		khajal	ghajar	No		khajal	fathin	No
jahatha	dhidu	No		jahatha	jahatha	Yes		jahatha	jaahun	No		jahatha	thahar	No
tajan	fathin	No		tajan	fathanaa	No		tajan	sum	No		tajan	fathun	No
shaja	sayd	No		shaja	nahr	No		shaja	sum	No		shaja	sayd	No
ajaza	jaza	No		ajaza	ajaza	Yes		ajaza	ajaza	Yes		ajaza	ajaza	Yes
sajaa	sum	No		sajaa	sada	No		sajaa	sadghu	No		sajaa	sayd	No
juthm	juhd	No		juthm	ghatha	No		juthm	juhd	No		juthm	thulth	No
jady	jady	Yes		jady	thaky	No		jady	jady	Yes		jady	taqy	No
jaza	ghadh	No		jaza	jaza	Yes		jaza	jaza	Yes		jaza	juzur	No
haja	ghajar	No		haja	kabshan	No		haja	sum	No		haja	sihr	No
jawq	ghil	No		jawq	sum	No		jawq	jamal	No		jawq	jaliy	No
jamal	daaain	No		jamal	dhana	No		jamal	daain	No		jamal	sayd	No
juhd	ghil	No		juhd	sum	No		juhd	jafat	No		juhd	thala	No
jidu	ghil	No		jidu	naabu	No		jidu	sum	No		jidu	tyn	No

Recording 4				Recording 3				Recording 2				Recording 1		
499 words	Recognised word	Match		499 words	Recognised word	Match		499 words	Recognised word	Match		499 words	Recognised word	Match
wajada	waadan	No		wajada	nadaba	No		wajada	sum	No		wajada	rakadha	No
aajidu	dhidu	No		aajidu	ghadara	No		aajidu	sum	No		aajidu	taqy	No
hujub	shibl	No		hujub	hadath	No		hujub	sum	No		hujub	samtun	No
daraja	ghajar	No		daraja	ghatha	No		daraja	ghajar	No		daraja	ghajar	No
sarju	fan	No		sarju	samtun	No		sarju	sum	No		sarju	samtun	No
wahaji	wahaji	Yes		wahaji	nahata	No		wahaji	wajada	No		wahaji	taqy	No
ewajan	daraja	No		ewajan	ewajan	Yes		ewajan	ewajan	Yes		ewajan	ewajun	No
ewajun	alafan	No		ewajun	ewajan	No		ewajun	aaatdyun	No		ewajun	alafan	No
ewajin	iwajin	No		ewajin	sum	No		ewajin	aaatdyin	No		ewajin	ewajun	No
hadhara	hamalla	No		hadhara	hadhara	Yes		hadhara	hamala	No		hadhara	hamalla	No
qazah	qazah	Yes		qazah	qazah	Yes		qazah	hazat	No		qazah	faaza	No
suhuf	suhub	No		suhuf	sawghun	No		suhuf	thahar	No		suhuf	sawghun	No
hathw	hathw	Yes		hathw	ghadara	No		hathw	hatala	No		hathw	samtun	No
hatama	dama	No		hatama	fahama	No		hatama	hafatha	No		hatama	thama	No
hasan	hasan	Yes		hasan	hasan	Yes		hasan	hasan	Yes		hasan	hasan	Yes
haka	haka	Yes		haka	thul	No		haka	haka	Yes		haka	haka	Yes
halahu	hunuw	No		halahu	halahu	Yes		halahu	halahu	Yes		halahu	qarnan	No
hay	rad	No		hay	hatala	No		hay	ayn	No		hay	ayn	No
hamala	khamana	No		hamala	hamala	Yes		hamala	amala	No		hamala	khamana	No
hibr	tib	No		hibr	sum	No		hibr	nabu	No		hibr	tib	No
husn	mawz	No		husn	hasan	No		husn	fasun	No		husn	husn	Yes
tahana	dhana	No		tahana	nahata	No		tahana	thahara	No		tahana	wajada	No
suhub	suhub	Yes		suhub	suhuf	No		suhub	sawghun	No		suhub	sawghun	No
yahilu	dhidu	No		yahilu	naoma	No		yahilu	dhuUl	No		yahilu	tyn	No

Recording 4				Recording 3				Recording 2				Recording 1		
499 words	Recognised word	Match		499 words	Recognised word	Match		499 words	Recognised word	Match		499 words	Recognised word	Match
masaha	masaha	Yes		masaha	masaha	Yes		masaha	masaha	Yes		masaha	masaha	Yes
farahi	shahy	No		farahi	thahar	No		farahi	farahi	Yes		farahi	liyn	No
marahu	marahu	Yes		marahu	maragha	No		marahu	sum	No		marahu	gharahu	No
qazahan	qazahun	No		qazahana	qazahun	No		qazahana	qazahun	No		qazahanaa	qazahun	No
qazahin	wadhaa	No		qazahin	wadhaa	No		qazahin	waDthaaaa	No		qazahin	wadhaa	No
qazahun	qazahun	Yes		qazahun	thakhn	No		qazahun	kthazahhun	No		qazahun	qazahun	Yes
dhakhah	bathahu	No		dhakhahu	dhakhahu	Yes		dhakhahu	thahar	No		dhakhahu	aaw	No
khaduk	fath	No		khaduk	hudida	No		khaduk	fath	No		khaduk	thulth	No
khath	haf	No		khath	haf	No		khath	haf	No		khath	haf	No
khashaa	rasha	No		khashaa	qazah	No		khashaa	rasha	No		khashaa	bashima	No
khasa	fasa	No		khasa	qasa	No		khasa	fasun	No		khasa	hasan	No
thakhara	nahata	No		thakhara	thakhara	Yes		thakhara	nahata	No		thakhara	hafatha	No
khazaqa	nadaba	No		khazaqa	khazaqa	Yes		khazaqa	qazahun	No		khazaqa	fasun	No
khasafa	basata	No		khasafa	ath	No		khasafa	yasudu	No		khasafa	basata	No
khaman	khamana	Yes		khamana	hamala	No		khamana	hamala	No		khamana	fathanaa	No
khawy	thawy	No		khawy	hunuw	No		khawy	aawidu	No		khawy	liyn	No
khas	haf	No		khas	qaas	No		khas	haf	No		khas	tharf	No
khidr	nibr	No		khidr	fathanaa	No		khidr	sayd	No		khidr	sayd	No
khums	shams	No		khums	khums	Yes		khums	kahan	No		khums	khums	Yes
bakhasa	nasiya	No		bakhasa	jahatha	No		bakhasa	jahatha	No		bakhasa	bakhasa	Yes
bakhila	bashima	No		bakhila	nahata	No		bakhila	bashima	No		bakhila	rahiba	No
rakhusa	nasiya	No		rakhusa	rakhusa	Yes		rakhusa	rahufa	No		rakhusa	naasa	No
sarakha	sharafa	No		sarakha	thahar	No		sarakha	thahar	No		sarakha	fahama	No
mukhi	wafy	No		mukhi	nahata	No		mukhi	sum	No		mukhi	liyn	No

Recording 4			Recording 3			Recording 2			Recording 1		
499 words	**Recognised word**	**Match**	**499 words**	**Recognised word**	**Match**	**499 words**	**Recognised word**	**Match**	**499 words**	**Recognised word**	**Match**
salkhu	samtun	No	salkhu	samtan	No	salkhu	samtun	No	salkhu	samtun	No
bathakh	bathakhin	No	bathakhu	bathakhun	Yes	bathakhu	bathakhun	Yes	bathakhun	bathakhun	Yes
bathakhi	bbathaakhin	No	bathakhin	bathakhin	Yes	bathakhin	bathakhin	Yes	bathakhin	hay	No
bathakh	bathakhun	No	bathakha	bathakhun	No	bathakha	bathakhun	No	bathakhana	bathakhun	No
dhid	layth	No	dhid	sum	No	dhid	sum	No	dhid	layth	No
zand	zand	Yes	zand	zand	Yes	zand	zand	Yes	zand	fan	No
rasada	yasudu	No	rasada	basata	No	rasada	rasada	Yes	rasada	rasada	Yes
qadam	fathin	No	qadam	mudun	No	qadam	athin	No	qadam	thawbun	No
tawd	fathun	No	tawd	faala	No	tawd	aaw	No	tawd	fathun	No
dasa	dasa	Yes	dasa	dasa	Yes	dasa	dasa	Yes	dasa	fasun	No
daghl	tablun	No	daghl	dama	No	daghl	ghanamu	No	daghl	tablun	No
daahu	aaw	No	daahu	dafira	No	daahu	daaahu	No	daahu	aaw	No
daf	ghath	No	daf	daf	Yes	daf	ghath	No	daf	ghath	No
dama	thanb	No	dama	dhana	No	dama	ghanamu	No	dama	thanb	No
dub	ghiil	No	dub	thaab	No	dub	ghil	No	dub	ghil	No
diyk	kiys	No	diyk	dhaaaa	No	diyk	thiny	No	diyk	liyn	No
nadaba	naddaba	No	nadaba	nadaba	Yes	nadaba	nadaba	Yes	nadaba	nadaba	Yes
hudida	wujida	No	hudida	sum	No	hudida	tyn	No	hudida	hibr	No
mudun	mudun	Yes	mudun	mudun	Yes	mudun	mudun	Yes	mudun	mudun	Yes
sada	sayd	No	sada	saba	No	sada	sayd	No	sada	sayd	No
ahdu	ghil	No	ahdu	sum	No	ahdu	sum	No	ahdu	tyn	No
mahdi	hay	No	mahdi	nahata	No	mahdi	naabu	No	mahdi	liyn	No
waadan	waadan	Yes	waadan	watan	No	waadan	waadan	Yes	waadan	tyn	No
waadun	wathun	No	waadun	waadan	No	waadun	daaaun	No	waadun	wathun	No

Recording 4				Recording 3				Recording 2				Recording 1		
499 words	Recognised word	Match		499 words	Recognised word	Match		499 words	Recognised word	Match		499 words	Recognised word	Match
waadin	daaaan	No		waadin	watan	No		waadin	daain	No		waadin	daaaan	No
thaky	jady	No		thaky	yakhti	No		thaky	thaky	Yes		thaky	liyn	No
thama	dama	No		thama	dama	No		thama	dama	No		thama	dama	No
thala	basata	No		thala	ghadara	No		thala	Ddthanna	No		thala	sayd	No
fath	sum	No		fath	fan	No		fath	sum	No		fath	fan	No
qathaa	qathaa	Yes		qathaa	dhana	No		qathaa	qathaa	Yes		qathaa	fath	No
shatha	shathaf	No		shatha	shathaf	No		shatha	saba	No		shatha	tyn	No
thawd	ghil	No		thawd	thaab	No		thawd	sahw	No		thawd	thul	No
thiib	daaaan	No		thiib	nathufa	No		thiib	daaaan	No		thiib	diyk	No
thaab	thaab	Yes		thaab	ghatha	No		thaab	ghadh	No		thaab	tyn	No
thul	daaaan	No		thul	dhana	No		thul	dahrun	No		thul	liyn	No
kathiba	hudida	No		kathiba	kathiba	Yes		kathiba	hudida	No		kathiba	kathiba	Yes
athara	hadhara	No		athara	ghatha	No		athara	athin	No		athara	ahdu	No
athuna	hatala	No		athuna	qatala	No		athuna	hatala	No		athuna	athin	No
shahath	shahatha	Yes		shahatha	shahatha	Yes		shahatha	shahatha	Yes		shahatha	shahatha	Yes
munthu	munthu	Yes		munthu	munthu	Yes		munthu	munthu	Yes		munthu	min	No
mUthi	liyn	No		mUthi	ayn	No		mUthi	sum	No		mUthi	liyn	No
fathanaa	thawbin	No		fathanaa	qarnan	No		fathanaa	fan	No		fathanaa	yin	No
fathun	ghadh	No		fathun	tablun	No		fathun	fath	No		fathun	rad	No
fathin	daaaan	No		fathin	ayn	No		fathin	sayd	No		fathin	tyn	No
thahara	nahata	No		thahara	nahata	No		thahara	nahata	No		thahara	nahata	No
qarn	athin	No		qarn	qarn	Yes		qarn	fan	No		qarn	fan	No
rakala	makana	No		rakala	makana	No		rakala	rakala	Yes		rakala	rakala	Yes
dhara	ghadara	No		dhara	ghadara	No		dhara	ghadara	No		dhara	ghadara	No

Recording 4				Recording 3				Recording 2				Recording 1		
499 words	**Recognised word**	**Match**		**499 words**	**Recognised word**	**Match**		**499 words**	**Recognised word**	**Match**		**499 words**	**Recognised word**	**Match**
raghw	tablun	No		raghw	sum	No		raghw	naabu	No		raghw	tablun	No
tayr	hudida	No		tayr	ghanamu	No		tayr	faala	No		tayr	hudida	No
sir	fahama	No		sir	fan	No		sir	fahama	No		sir	sayd	No
rad	rad	Yes		rad	sum	No		rad	rad	Yes		rad	rad	Yes
ruba	rakiba	No		ruba	ath	No		ruba	sum	No		ruba	ghil	No
surur	shibl	No		surur	sanami	No		surur	farabu	No		surur	sayd	No
harama	harrama	No		harama	hamala	No		harama	hatama	No		harama	dama	No
siry	zaraa	No		siry	sanami	No		siry	sanami	No		siry	sayd	No
fatara	wathafa	No		fatara	sakaba	No		fatara	hafatha	No		fatara	wathafa	No
juhri	mUthi	No		juhri	dhidu	No		juhri	jaahun	No		juhri	liyn	No
fikri	taqy	No		fikri	saqata	No		fikri	taqy	No		fikri	liyn	No
dahrun	daaaun	No		dahrun	daaaan	No		dahrun	jaahun	No		dahrun	samtun	No
dahrin	dhaaaa	No		dahrin	bathakhin	No		dahrin	daaaan	No		dahrin	tyn	No
dahran	jaahun	No		dahran	bathakhun	No		dahran	jaahun	No		dahran	jaahun	No
zafa	zafa	Yes		zafa	zafa	Yes		zafa	zafa	Yes		zafa	zand	No
zaama	fahama	No		zaama	dhamat	No		zaama	thama	No		zaama	zaama	Yes
zaky	jady	No		zaky	khazaqa	No		zaky	thaky	No		zaky	zaky	Yes
zuhal	sawghun	No		zuhal	makana	No		zuhal	jaahun	No		zuhal	sawghun	No
zaraa	fahama	No		zaraa	zaky	No		zaraa	dama	No		zaraa	zaama	No
zir	Did	No		zir	zand	No		zir	sum	No		zir	zir	Yes
ruziq	wasati	No		ruziq	sakat	No		ruziq	wasia	No		ruziq	wasati	No
azafa	hazat	No		azafa	qazahun	No		azafa	azafa	Yes		azafa	azafa	Yes
juzur	ghasala	No		juzur	jaza	No		juzur	juzur	Yes		juzur	juzur	Yes
faaza	ahdu	No		faaza	faaza	Yes		faaza	fasun	No		faaza	zir	No

Recording 4			Recording 3			Recording 2			Recording 1		
499 words	**Recognised word**	**Match**	**499 words**	**Recognised word**	**Match**	**499 words**	**Recognised word**	**Match**	**499 words**	**Recognised word**	**Match**
jawzu	jawzu	Yes	jawzu	jawzu	Yes	jawzu	jawzu	Yes	jawzu	shams	No
kanzi	fasin	No	kanzi	thanb	No	kanzi	kanzi	Yes	kanzi	tyn	No
filizan	filizin	No	filizan	filizan	Yes	filizan	filizin	No	filizan	filizin	No
filizun	kahan	No	filizun	asal	No	filizun	naasa	No	filizun	asal	No
filizin	wasia	No	filizin	ayn	No	filizin	wasiaa	No	filizin	filizin	Yes
shams	shams	Yes	shams	shams	Yes	shams	shams	Yes	shams	shams	Yes
ghasala	basata	No	ghasala	ghasala	Yes	ghasala	rasada	No	ghasala	fasun	No
sahw	saaal	No	sahw	samtun	No	sahw	fijl	No	sahw	sayd	No
kys	kiys	No	kys	samaki	No	kys	kiys	No	kys	kiys	No
dhirs	dhaghath	No	dhirs	sum	No	dhirs	layth	No	dhirs	layth	No
sum	sawghin	No	sum	sahwa	No	sum	fan	No	sum	sawghan	No
sakaba	sakaba	Yes	sakaba	sakaba	Yes	sakaba	sakaba	Yes	sakaba	sayd	No
sihr	sihr	Yes	sihr	thahar	No	sihr	fijl	No	sihr	sihr	Yes
rusul	rusul	Yes	rusul	rasada	No	rusul	hasan	No	rusul	rusul	Yes
asal	hasan	No	asal	basata	No	asal	fasun	No	asal	hasan	No
nasiya	nasiya	Yes	nasiya	nasiya	Yes	nasiya	nasiya	Yes	nasiya	wasy	No
habasa	dasa	No	habasa	habasa	Yes	habasa	habasa	Yes	habasa	fasin	No
harasa	rasa	No	harasa	hasan	No	harasa	hasan	No	harasa	fasun	No
farasi	hasan	No	farasi	basata	No	farasi	fasun	No	farasi	wasy	No
orsan	rushida	No	orsan	qazahin	No	orsan	hasan	No	orsan	fan	No
orsun	rusul	No	orsun	did	No	orsun	rusul	No	orsun	orsin	No
orsin	orsin	Yes	orsin	fasin	No	orsin	aaurrsun	No	orsin	hasan	No
shathw	shibl	No	shathw	shathw	Yes	shathw	shathw	Yes	shathw	sayd	No
shas	shams	No	shas	shams	No	shas	shas	Yes	shas	tharf	No

Recording 4			Recording 3			Recording 2			Recording 1		
499 words	**Recognised word**	**Match**	**499 words**	**Recognised word**	**Match**	**499 words**	**Recognised word**	**Match**	**499 words**	**Recognised word**	**Match**
shathaf	shadat	No	shathaf	shatha	No	shathaf	shathaf	Yes	shathaf	shathaf	Yes
shat	shat	Yes	shat	shat	Yes	shat	shat	Yes	shat	shat	Yes
shughl	fan	No	shughl	shatha	No	shughl	shathw	No	shughl	sayd	No
qash	qash	Yes	qash	jas	No	qash	rasha	No	qash	haf	No
shak	shak	Yes	shak	shat	No	shak	shak	Yes	shak	fan	No
nashiz	nashiz	Yes	nashiz	basata	No	nashiz	nashiz	Yes	nashiz	nashiz	Yes
shahy	mUthi	No	shahy	shatha	No	shahy	shathw	No	shahy	shahy	Yes
shajar	shajar	Yes	shajar	sajjaa	No	shajar	shajar	Yes	shajar	shajar	Yes
shibl	shibl	Yes	shibl	shibl	Yes	shibl	shibl	Yes	shibl	sayd	No
shukr	shat	No	shukr	shathw	No	shukr	shak	No	shukr	shughl	No
washm	washm	Yes	washm	ghasha	No	washm	wasy	No	washm	washm	Yes
rushida	rushida	Yes	rushida	rushida	Yes	rushida	rushida	Yes	rushida	wafy	No
aashudu	ghil	No	aashudu	yasudu	No	aashudu	ayn	No	aashudu	ryshi	No
rasha	rasha	Yes	rasha	rasha	Yes	rasha	rasha	Yes	rasha	rasha	Yes
rimshu	filizin	No	rimshu	rimshu	Yes	rimshu	sum	No	rimshu	min	No
ryshi	ryshi	Yes	ryshi	ryshi	Yes	ryshi	naasa	No	ryshi	filizin	No
kabshan	shaah	No	kabshan	kabshan	Yes	kabshan	kabshan	Yes	kabshan	kabshan	Yes
kabshun	shughl	No	kabshun	kabshun	Yes	kabshun	kabshin	No	kabshun	shughl	No
kabshin	shahy	No	kabshin	kabshin	Yes	kabshin	shahy	No	kabshin	fan	No
qasa	fasa	No	qasa	kaas	No	qasa	khasa	No	qasa	hasan	No
sum	sawghin	No	sum	samghi	No	sum	fan	No	sum	fan	No
sanaa	fadhala	No	sanaa	sanaa	Yes	sanaa	sum	No	sanaa	fahama	No
sah	saaf	No	sah	sah	Yes	sah	saaf	No	sah	sah	Yes
wasy	wasy	Yes	wasy	wasy	Yes	wasy	wafy	No	wasy	wasy	Yes

Recording 4				Recording 3				Recording 2				Recording 1			
499 words	**Recognised word**	**Match**		**499 words**	**Recognised word**	**Match**		**499 words**	**Recognised word**	**Match**		**499 words**	**Recognised word**	**Match**	
suws	shams	No		suws	saaf	No		suws	malaf	No		suws	shams	No	
sayd	shaja	No		sayd	sayd	Yes		sayd	sum	No		sayd	thiny	No	
sihr	sharath	No		sihr	sahwa	No		sihr	fahama	No		sihr	sayd	No	
asara	qazahin	No		asara	ghasala	No		asara	asara	Yes		asara	hasan	No	
nusira	nasiya	No		nusira	nusira	Yes		nusira	nusira	Yes		nusira	nusira	Yes	
yasudu	aaw	No		yasudu	yasudu	Yes		yasudu	yasudu	Yes		yasudu	basal	No	
rasa	rasa	Yes		rasa	rasa	Yes		rasa	rasa	Yes		rasa	rasa	Yes	
qursi	wasati	No		qursi	basata	No		qursi	wasy	No		qursi	wasy	No	
fasun	fasun	Yes		fasun	fasun	Yes		fasun	fasun	Yes		fasun	hasan	No	
fasa	dasa	No		fasa	sawghan	No		fasa	qazahun	No		fasa	fasun	No	
fasin	rafaa	No		fasin	liyn	No		fasin	rafaa	No		fasin	hasan	No	
dhaghat	sum	No		dhaghath	thama	No		dhaghath	thahar	No		dhaghath	sum	No	
wadhaa	watan	No		wadhaa	sum	No		wadhaa	waathun	No		wadhaa	watan	No	
dhana	dhana	Yes		dhana	daaaan	No		dhana	dhana	Yes		dhana	kahan	No	
dhala	hudida	No		dhala	dama	No		dhala	dama	No		dhala	hudida	No	
dhyq	daaaan	No		dhyq	daaain	No		dhyq	sum	No		dhyq	liyn	No	
dhafar	Dafar	No		dhafar	nahata	No		dhafar	dhafar	Yes		dhafar	Dafar	No	
dharaba	ddama	No		dharaba	nadaba	No		dharaba	nadaba	No		dharaba	fahama	No	
dhuha	wahaji	No		dhuha	nawaha	No		dhuha	wahaji	No		dhuha	mukhi	No	
dhidu	ghil	No		dhidu	sum	No		dhidu	sum	No		dhidu	tyn	No	
radhiya	hudida	No		radhiya	hatala	No		radhiya	radhiya	Yes		radhiya	liyn	No	
adhud	adhud	Yes		adhud	hudida	No		adhud	athima	No		adhud	haf	No	
fadhala	sum	No		fadhala	fahama	No		fadhala	sum	No		fadhala	fath	No	
maradha	nadaba	No		maradha	nadaba	No		maradha	nadaba	No		maradha	mahwu	No	

Recording 4				Recording 3				Recording 2				Recording 1		
499 words	**Recognised word**	**Match**		**499 words**	**Recognised word**	**Match**		**499 words**	**Recognised word**	**Match**		**499 words**	**Recognised word**	**Match**
aradha	athara	No		aradha	thama	No		aradha	atuma	No		aradha	hafatha	No
aardhi	thaby	No		aardhi	qarnan	No		aardhi	sayara	No		aardhi	liyn	No
qardhan	tablun	No		qardhan	karnan	No		qardhan	samtan	No		qardhan	kahan	No
qardhun	thawbin	No		qardhun	mudun	No		qardhun	aaalamin	No		qardhun	fath	No
qardhin	samtun	No		qardhin	qarnan	No		qardhin	liyn	No		qardhin	liyn	No
taq	fath	No		taq	thaab	No		taq	fath	No		taq	fath	No
hatala	makana	No		hatala	sakaba	No		hatala	sum	No		hatala	fathun	No
tamaa	fahama	No		tamaa	thama	No		tamaa	thama	No		tamaa	thama	No
tib	fath	No		tib	tib	Yes		tib	fath	No		tib	tyn	No
tabaa	sabaqa	No		tabaa	thawbanaa	No		tabaa	faala	No		tabaa	fahama	No
watan	watan	Yes		watan	waathun	No		watan	waathun	No		watan	kahan	No
ratib	rafaa	No		ratib	samtun	No		ratib	rakadha	No		ratib	rafaa	No
otuf	aakala	No		otuf	aaatha	No		otuf	thulth	No		otuf	aakala	No
qirtu	tayu	No		qirtu	tathil	No		qirtu	sum	No		qirtu	tyn	No
wasati	wathafa	No		wasati	wathafa	No		wasati	wathafa	No		wasati	rafaa	No
basata	fasa	No		basata	ath	No		basata	bbadvakhana	No		basata	fasa	No
nuqatan	nuqatun	No		nuqatan	nuqatun	No		nuqatan	nuqatun	No		nuqatan	nuqatun	No
nuqatun	nuqatun	Yes		nuqatun	nuqatun	Yes		nuqatun	min	No		nuqatun	nuqatun	Yes
nuqatin	nuqatun	No		nuqatin	nukatun	No		nuqatin	nuqatun	No		nuqatin	liyn	No
thahar	fath	No		thahar	nahata	No		thahar	thahar	Yes		thahar	fath	No
kathu	naabu	No		kathu	fahama	No		kathu	naabu	No		kathu	naabu	No
wathafa	wathaba	No		wathafa	wathafa	Yes		wathafa	wathafa	Yes		wathafa	wathaba	No
tharf	dhana	No		tharf	thama	No		tharf	ghath	No		tharf	fath	No
thifr	layth	No		thifr	nathara	No		thifr	wasy	No		thifr	layth	No

Recording 4				Recording 3				Recording 2				Recording 1		
499 words	**Recognised word**	**Match**		**499 words**	**Recognised word**	**Match**		**499 words**	**Recognised word**	**Match**		**499 words**	**Recognised word**	**Match**
thul	daaaan	No		thul	dama	No		thul	daaaun	No		thul	tyn	No
nathara	nadaba	No		nathara	nadaba	No		nathara	nadaba	No		nathara	hafatha	No
nathufa	naabu	No		nathufa	naabu	No		nathufa	nnaththufa	No		nathufa	naabu	No
athima	hudida	No		athima	ghatha	No		athima	hudida	No		athima	atyan	No
hafatha	hafatha	Yes		hafatha	hafatha	Yes		hafatha	hafatha	Yes		hafatha	hafatha	Yes
qaythi	sum	No		qaythi	sum	No		qaythi	sum	No		qaythi	qaae	No
hathu	ghil	No		hathu	habasa	No		hathu	fath	No		hathu	ghil	No
waathan	waadan	No		waathan	waathun	No		waathan	waathun	No		waathan	fan	No
waathun	thihni	No		waathun	watan	No		waathun	daaun	No		waathun	tyn	No
waathin	liyn	No		waathin	watan	No		waathin	waadan	No		waathin	liyn	No
athal	fathin	No		athal	athal	Yes		athal	athin	No		athal	ath	No
saaf	fahama	No		saaf	thahar	No		saaf	ghath	No		saaf	sihr	No
atash	atash	Yes		atash	atash	Yes		atash	atash	Yes		atash	atash	Yes
aks	haka	No		aks	ath	No		aks	haka	No		aks	kanzi	No
aqr	sum	No		aqr	aqr	Yes		aqr	sum	No		aqr	saer	No
ayn	ayn	Yes		ayn	ayn	Yes		ayn	ayn	Yes		ayn	ayn	Yes
ejl	wafy	No		ejl	sum	No		ejl	sum	No		ejl	wafy	No
omr	ayn	No		omr	fahama	No		omr	aamal	No		omr	kahan	No
saer	sanaa	No		saer	saaal	No		saer	sayara	No		saer	sayd	No
naasa	qasa	No		naasa	naasa	Yes		naasa	naasa	Yes		naasa	hasan	No
naoma	naoma	Yes		naoma	dama	No		naoma	naoma	Yes		naoma	yad	No
wasia	wasia	Yes		wasia	nasiya	No		wasia	wasia	Yes		wasia	wasia	Yes
qaae	zaky	No		qaae	saaal	No		qaae	saaal	No		qaae	ayn	No
saao	saaal	No		saao	sahwa	No		saao	saaal	No		saao	sawghan	No

Recording 4			Recording 3			Recording 2			Recording 1		
499 words	**Recognised word**	**Match**	**499 words**	**Recognised word**	**Match**	**499 words**	**Recognised word**	**Match**	**499 words**	**Recognised word**	**Match**
wariaan	waathun	No	wariaan	waathun	No	wariaan	wariaan	Yes	wariaan	atyan	No
wariaun	wariain	No	wariaun	wariaan	No	wariaun	atyun	No	wariaun	tyn	No
wariain	dhaaaa	No	wariain	waadan	No	wariain	waDthaaaa	No	wariain	watan	No
ghajar	haja	No	ghajar	ghatha	No	ghajar	ghajar	Yes	ghajar	taghiya	No
ghatha	rad	No	ghatha	sanami	No	ghatha	ghadh	No	ghatha	ghadh	No
ghasha	rasha	No	ghasha	nashiz	No	ghasha	rasha	No	ghasha	layth	No
ghadh	rathi	No	ghadh	naabu	No	ghadh	rad	No	ghadh	rathi	No
ghafiya	ayiya	No	ghafiya	rafaa	No	ghafiya	wasia	No	ghafiya	tyn	No
gharaqa	alafun	No	gharaqa	ghata	No	gharaqa	alafan	No	gharaqa	alafun	No
ghaythu	hudida	No	ghaythu	sum	No	ghaythu	ghanamu	No	ghaythu	thawy	No
ghata	sum	No	ghata	sah	No	ghata	dhafar	No	ghata	thul	No
ghadara	ghaddara	No	ghadara	nadaba	No	ghadara	nadaba	No	ghadara	ghaddara	No
ghusn	nashiz	No	ghusn	rusul	No	ghusn	hasan	No	ghusn	orsin	No
ghil	min	No	ghil	liyn	No	ghil	tyn	No	ghil	tyn	No
saghura	sahwa	No	saghura	sahwa	No	saghura	fadhala	No	saghura	fath	No
raghad	raghaad	No	raghad	nadaba	No	raghad	raghad	Yes	raghad	fahama	No
taghiya	hudida	No	taghiya	fadhala	No	taghiya	hudida	No	taghiya	thaby	No
maragha	mahdi	No	maragha	mahdi	No	maragha	nadaba	No	maragha	dama	No
sadghu	naabu	No	sadghu	fadhala	No	sadghu	fath	No	sadghu	fath	No
samghi	samghi	Yes	samghi	samghi	Yes	samghi	samghi	Yes	samghi	samghi	Yes
sawgha	sawghin	No	sawghan	samtun	No	sawghan	thawbun	No	sawghan	hay	No
sawghu	samtun	No	sawghun	samtan	No	sawghun	ayn	No	sawghun	ayn	No
sawghin	fathin	No	sawghin	samtun	No	sawghin	thawbun	No	sawghin	sahw	No
haf	haf	Yes	haf	rathi	No	haf	haf	Yes	haf	haf	Yes

Recording 4			Recording 3			Recording 2			Recording 1		
499 words	**Recognised word**	**Match**	**499 words**	**Recognised word**	**Match**	**499 words**	**Recognised word**	**Match**	**499 words**	**Recognised word**	**Match**
wafy	wafy	Yes	wafy	rafaa	No	wafy	wafy	Yes	wafy	liyn	No
malaf	malaf	Yes	malaf	sum	No	malaf	malaf	Yes	malaf	malaf	Yes
faka	tyn	No	faka	ham	No	faka	tyn	No	faka	wafy	No
fan	min	No	fan	fan	Yes	fan	fan	Yes	fan	min	No
fijl	sum	No	fijl	fathun	No	fijl	sum	No	fijl	layth	No
furn	watan	No	furn	fahama	No	furn	watan	No	furn	yin	No
faala	fahama	No	faala	dhana	No	faala	fahama	No	faala	fahama	No
rafaa	rasa	No	rafaa	rahufa	No	rafaa	rasa	No	rafaa	rasa	No
dafira	biraku	No	dafira	ghafiya	No	dafira	ghasala	No	dafira	biraku	No
afwu	min	No	afwu	hasan	No	afwu	afwu	Yes	afwu	min	No
sharafa	sharafa	Yes	sharafa	sharafa	Yes	sharafa	sharafa	Yes	sharafa	sharafa	Yes
tarafi	liyn	No	tarafi	alafan	No	tarafi	fasun	No	tarafi	liyn	No
khalfu	samtun	No	khalfu	samtun	No	khalfu	samtun	No	khalfu	fan	No
alafan	alafin	No	alafan	alafan	Yes	alafan	alafan	Yes	alafan	alafin	No
alafun	sum	No	alafun	allafun	No	alafun	aaalafun	No	alafun	sum	No
alafin	liyn	No	alafin	allafan	No	alafin	aaalafan	No	alafin	liyn	No
sujuq	shibl	No	sujuq	sada	No	sujuq	khajal	No	sujuq	wafy	No
qulw	wariain	No	qulw	aamil	No	qulw	aamil	No	qulw	sayd	No
daqaka	haka	No	daqaka	daqaka	Yes	daqaka	haka	No	daqaka	rathi	No
qalam	qaarnan	No	qalam	qarnan	No	qalam	qarnan	No	qalam	qaarnan	No
qidr	sum	No	qidr	sum	No	qidr	sayd	No	qidr	sayd	No
quda	sum	No	quda	rakiba	No	quda	sum	No	quda	sum	No
saqata	saqata	Yes	saqata	sabaqa	No	saqata	sakaba	No	saqata	saqata	Yes
fuqida	wujida	No	fuqida	thabata	No	fuqida	sum	No	fuqida	hibr	No

	Recording 4				Recording 3				Recording 2				Recording 1		
499 words	Recognised word	Match		499 words	Recognised word	Match		499 words	Recognised word	Match		499 words	Recognised word	Match	
thaqula	makuna	No		thaqula	makuna	No		thaqula	rakala	No		thaqula	makuna	No	
sabaqa	sabaqa	Yes		sabaqa	sabaqa	Yes		sabaqa	sabaqa	Yes		sabaqa	sabaqa	Yes	
abaqa	athima	No		abaqa	habasa	No		abaqa	abaqa	Yes		abaqa	athima	No	
ghasaqu	masaha	No		ghasaqu	sum	No		ghasaqu	qazahun	No		ghasaqu	rasa	No	
barqan	bbathaakhun	No		barqan	ayn	No		barqan	bathakhin	No		barqan	bbathaakhun	No	
barqun	yin	No		barqun	alafin	No		barqun	bathakhun	No		barqun	yin	No	
barqin	liyn	No		barqin	bathakhin	No		barqin	bathakhin	No		barqin	tyn	No	
rakadha	rakala	No		rakadha	sakaba	No		rakadha	sakaba	No		rakadha	rakala	No	
jaraka	jawzu	No		jaraka	sum	No		jaraka	jaraka	Yes		jaraka	fahama	No	
kawa	thala	No		kawa	fahama	No		kawa	qalam	No		kawa	thala	No	
kahan	daaain	No		kahan	kahan	Yes		kahan	kahan	Yes		kahan	daaain	No	
kalb	dhaaaa	No		kalb	qarnan	No		kalb	qalam	No		kalb	dhaaaa	No	
kiys	kiys	Yes		kiys	saaf	No		kiys	sum	No		kiys	mawz	No	
kuwa	thala	No		kuwa	qalam	No		kuwa	qalam	No		kuwa	mahwu	No	
rakiba	dhidu	No		rakiba	rakadha	No		rakiba	rakadha	No		rakiba	hibr	No	
rakaa	haka	No		rakaa	makana	No		rakaa	rakala	No		rakaa	haka	No	
makuna	wana	No		makuna	makuna	Yes		makuna	makuna	Yes		makuna	wana	No	
haraka	harrasa	No		haraka	haka	No		haraka	fahama	No		haraka	fahama	No	
biraku	biraku	Yes		biraku	bathakhin	No		biraku	biraku	Yes		biraku	tyn	No	
samaki	sanaa	No		samaki	samtun	No		samaki	samaki	Yes		samaki	samaki	Yes	
silkan	samtun	No		silkan	samtun	No		silkan	samtun	No		silkan	sayd	No	
silkun	tablin	No		silkun	samtan	No		silkun	samtan	No		silkun	tablin	No	
silkin	samghi	No		silkin	samtun	No		silkin	samtun	No		silkin	samghi	No	
layth	dhaaaa	No		layth	sum	No		layth	sum	No		layth	jady	No	

Recording 4				Recording 3				Recording 2				Recording 1		
499 words	**Recognised word**	**Match**		**499 words**	**Recognised word**	**Match**		**499 words**	**Recognised word**	**Match**		**499 words**	**Recognised word**	**Match**
liyn	naasa	No		liyn	daaain	No		liyn	ayn	No		liyn	liyn	Yes
lumat	numuw	No		lumat	naoma	No		lumat	naoma	No		lumat	naoma	No
oluw	wariaun	No		oluw	hunuw	No		oluw	hunuw	No		oluw	hunuw	No
ghalaa	dhana	No		ghalaa	ghatha	No		ghalaa	tahana	No		ghalaa	dhana	No
jaliy	jady	No		jaliy	sum	No		jaliy	jaliy	Yes		jaliy	sayd	No
daghlu	ghanamu	No		daghlu	daaaan	No		daghlu	ghanamu	No		daghlu	ghanamu	No
amali	qarnin	No		amali	alafan	No		amali	aamal	No		amali	qarnin	No
tablan	min	No		tablan	tablun	No		tablan	athin	No		tablan	fath	No
tablun	ghil	No		tablun	qazahun	No		tablun	thawbun	No		tablun	ghil	No
tablin	liyn	No		tablin	fathun	No		tablin	fukthida	No		tablin	liyn	No
ham	min	No		ham	sum	No		ham	fan	No		ham	min	No
yawm	ghil	No		yawm	sanami	No		yawm	saaal	No		yawm	ghil	No
mawz	mahwu	No		mawz	mahwu	No		mawz	mahwu	No		mawz	mahwu	No
min	min	Yes		min	liyn	No		min	liyn	No		min	min	Yes
aamil	qarnin	No		aamil	aamal	No		aamil	qarnan	No		aamil	sayd	No
amala	khamana	No		amala	amala	Yes		amala	amala	Yes		amala	ham	No
numuw	numuw	Yes		numuw	numuw	Yes		numuw	numuw	Yes		numuw	numuw	Yes
fahama	fahama	Yes		fahama	fahama	Yes		fahama	fahama	Yes		fahama	fahama	Yes
ghanam	ghanamu	Yes		ghanamu	ghanamu	Yes		ghanamu	ghanamu	Yes		ghanamu	qarnin	No
sanami	fathin	No		sanami	sanami	Yes		sanami	samtun	No		sanami	fathun	No
alaman	alaman	Yes		alaman	alamun	No		alaman	alafan	No		alaman	alaman	Yes
alamun	tablun	No		alamun	alaman	No		alamun	alaman	No		alamun	qarnin	No
alamin	qarnin	No		alamin	alamun	No		alamin	alamin	Yes		alamin	tyn	No
wana	wana	Yes		wana	wana	Yes		wana	wana	Yes		wana	wana	Yes

Recording 4				Recording 3				Recording 2				Recording 1		
499 words	**Recognised word**	**Match**		**499 words**	**Recognised word**	**Match**		**499 words**	**Recognised word**	**Match**		**499 words**	**Recognised word**	**Match**
nahr	min	No		nahr	nahata	No		nahr	nahata	No		nahr	min	No
nibr	nibr	Yes		nibr	nathara	No		nibr	nibr	Yes		nibr	ghaythu	No
nuwr	mahwu	No		nuwr	naoma	No		nuwr	mahwu	No		nuwr	mahwu	No
fanar	fahama	No		fanar	fahama	No		fanar	fahama	No		fanar	fahama	No
saniya	saniya	Yes		saniya	sanaa	No		saniya	sanniya	No		saniya	sayd	No
hunuw	qarnin	No		hunuw	aamil	No		hunuw	hunuw	Yes		hunuw	aamil	No
makana	min	No		makana	thama	No		makana	fahama	No		makana	min	No
nahnu	numuw	No		nahnu	sum	No		nahnu	nahata	No		nahnu	numuw	No
thihni	daaaan	No		thihni	ayn	No		thihni	daaain	No		thihni	wafy	No
qarnan	qarnin	No		qarnan	qarnan	Yes		qarnan	qarnan	Yes		qarnan	qarnin	No
qarnun	karnan	No		qarnun	karnan	No		qarnun	faan	No		qarnun	sayd	No
qarnin	min	No		qarnin	karnun	No		qarnin	daain	No		qarnin	fathun	No
gharahu	sahw	No		gharahu	ghanamu	No		gharahu	ghanamu	No		gharahu	sahw	No
thalahu	ghanamu	No		thalahu	faala	No		thalahu	gghannamu	No		thalahu	ghanamu	No
hir	sayara	No		hir	tyn	No		hir	sayara	No		hir	tyn	No
hawas	malaf	No		hawas	haf	No		hawas	malaf	No		hawas	hawas	Yes
huwid	wujida	No		huwid	wujida	No		huwid	wasia	No		huwid	mahwu	No
rahiba	hudida	No		rahiba	rahiba	Yes		rahiba	zaama	No		rahiba	rad	No
rahufa	sahw	No		rahufa	rahufa	Yes		rahufa	rahufa	Yes		rahufa	mahwu	No
qahara	ghata	No		qahara	qahara	Yes		qahara	qahara	Yes		qahara	fath	No
nawaha	nawaha	Yes		nawaha	nawaha	Yes		nawaha	nawaha	Yes		nawaha	mahwu	No
minhu	minhu	Yes		minhu	munthu	No		minhu	sum	No		minhu	min	No
fyhi	kiys	No		fyhi	sum	No		fyhi	ayn	No		fyhi	tyn	No
jaahan	jaahin	No		jaahan	jaahun	No		jaahan	jaahin	No		jaahan	jaahin	No

Recording 4				Recording 3				Recording 2				Recording 1			
499 words	**Recognised word**	**Match**		**499 words**	**Recognised word**	**Match**		**499 words**	**Recognised word**	**Match**		**499 words**	**Recognised word**	**Match**	
jaahun	tablun	No		jaahun	jaahun	Yes		jaahun	jaahun	Yes		jaahun	tablun	No	
jaahin	daaain	No		jaahin	jaahin	Yes		jaahin	jaahin	Yes		jaahin	daaain	No	
witr	wafy	No		witr	wafy	No		witr	wafy	No		witr	wafy	No	
wujida	wujida	Yes		wujida	wujida	Yes		wujida	wasia	No		wujida	shahy	No	
aawidu	ghil	No		aawidu	sum	No		aawidu	ghil	No		aawidu	aawidu	Yes	
thawuw	DuUl	No		thawuw	daahu	No		thawuw	dahrrun	No		thawuw	ghil	No	
mahwu	marahu	No		mahwu	nahata	No		mahwu	nahata	No		mahwu	marahu	No	
lahwi	lahwi	Yes		lahwi	lahwi	Yes		lahwi	lahwi	Yes		lahwi	ayn	No	
sahwa	sahwa	Yes		sahwa	sahwa	Yes		sahwa	sahwa	Yes		sahwa	fijl	No	
jarwan	tablun	No		jarwan	daaaun	No		jarwan	jaahun	No		jarwan	juhd	No	
jarwun	jarwun	Yes		jarwun	daaaan	No		jarwun	daaun	No		jarwun	sayd	No	
jarwin	jaliy	No		jarwin	sum	No		jarwin	jjaliy	No		jarwin	jaliy	No	
yad	yad	Yes		yad	yad	Yes		yad	yad	Yes		yad	yin	No	
yusr	mawz	No		yusr	yasudu	No		yusr	sum	No		yusr	mawz	No	
yin	min	No		yin	liyn	No		yin	yin	Yes		yin	yin	Yes	
sayara	shaja	No		sayara	faala	No		sayara	sayara	Yes		sayara	fahama	No	
ayiya	sum	No		ayiya	ayiya	Yes		ayiya	atyan	No		ayiya	ayiya	Yes	
sawyi	thawy	No		sawyi	saaal	No		sawyi	jarwin	No		sawyi	sayd	No	
tayu	saaal	No		tayu	ayiya	No		tayu	sum	No		tayu	liyn	No	

Recording 4				Recording 3				Recording 2				Recording 1			
499 words	Recognised word	Match		499 words	Recognised word	Match		499 words	Recognised word	Match		499 words	Recognised word	Match	
hayuUa	hayuUa	Yes		hayuUa	sum	No		hayuUa	sum	No		hayuUa	hayuUa	Yes	
atyan	atyan	Yes		atyan	ham	No		atyan	atyan	Yes		atyan	sayd	No	
atyun	samtun	No		atyun	atyan	No		atyun	atyun	Yes		atyun	samtun	No	
atyin	atyun	No		atyin	afwu	No		atyin	taaqy	No		atyin	atyun	No	
	No. of recognised words	93			No. of recognised words	101			No. of recognised words	115			No. of recognised words	74	
	average	18.637			average	20.240			average	23.0460			average	14.829	
													Average all	19.188	

Table 1- Alghamdi's transliteration recogntion analysis

Appendix K

Letter or diacritic alternatives to create an improvement to Alghamdi's transliterations.

Arabic letter	Name of letter	Total words	Alternatives								Modified English letter
			Alghamdi's choice		Alternative 1		Alternative 2		Alternative 3		
			English letter	Recog. rate (%)	English letter	Recog. rate (%)	English letter	Recog. rate (%)	English letter	Recog. rate (%)	
أ	alef	19	a	6.6	aa	5.3					a
ب	baa	71	b	17.3	bb	15.1	p	10.9	pp	10.2	b
ت	taa	31	t	12.9	tt	11.3	td	9.7			t
ث	thaa	34	th	16.2	tth	8.8	dth	12.5	v	8.1	th
ج	jeem	51	j	19.6	jj	21	g	10.8			jj
ح	haa	43	h	24.4	hh	22					h
خ	khaa	28	kh	11.6	k	16.1	kk	8	x	8	k
د	daal	58	d	20.3	dd	15.1					d
ذ	thaal	33	th	18.9	dh	12.1	dhh	14	dv	9.8	th
ر	raa	126	r	14.5	rr	16					rr

Arabic letter	Name of letter	Total words	Alternatives								Modified English letter
			Alghamdi's choice		Alternative 1		Alternative 2		Alternative 3		
			English letter	Recog. rate (%)	English letter	Recog. rate (%)	English letter	Recog. rate (%)	English letter	Recog. rate (%)	
ز	zain	28	z	32.1	zz	36	s	31.3			zz
س	seen	64	s	22.3	ss	18					s
ش	sheen	34	sh	44.9	ssh	35	ch	17.6			sh
ص	saad	39	s	21.8	ss	19.9	sf	12.2			s
ض	dhad	30	dh	6.7	ddh	3.3	dhv	2.5	th	3.3	dh
ط	ta	36	t	10.4	tt	12					tt
ظ	tha	22	th	12.5	tth	10.2	dh	9.1	z	4.5	th
ع	ain	75	a	16	aa	14	e	13			a
غ	ghain	34	gh	9.5	g	10.3	q	4.4			g
ف	faa	57	f	20.6	ff	20					f

Arabic letter	Name of letter	Total words	Alternatives								Modified English letter
			Alghamdi's choice		Alternative 1		Alternative 2		Alternative 3		
			English letter	Recog. rate (%)	English letter	Recog. rate (%)	English letter	Recog. rate (%)	English letter	Recog. rate (%)	
ق	qaaf	51	q	12.3	k	13	kk	11.3			k
ك	kaaf	43	k	20.9	kk	24.4	q	15.1			kk
ل	laam	77	l	10.1	ll	12					ll
م	meem	62	m	27	mm	10					m
ن	noon	61	n	36.9	nn	30					n
ه	haa	45	h	23.9	hh	23					h
و	waaw	71	w	18.3	ww	13	o	13	oo	11.3	w
ي	yaa	50	y	22	yy	21					Y

Arabic letter	Name of letter	Total words	Alternatives								Modified English letter
			Alghamdi's choice		Alternative 1		Alternative 2		Alternative 3		
			English letter	Recog. rate (%)	English letter	Recog. rate (%)	English letter	Recog. rate (%)	English letter	Recog. rate (%)	
ˈ	**Fat ha**	672	**a**	21.5	**aa**	23.4					**aa**
ˈ	**dhamma**	150	**u**	22.7	**o**	14.8	**oo**	10.7	**ou**	8	**u**
ˌ	**kasra**	118	**i**	31.9	**e**	19.3	**ie**	8.5	**ee**	10.2	**i**

Table 1-Letter or diacritic alternatives to create an improved table to Alghamdi's transliterations.

Appendix L

Improved SLT recognition analysis

Recording 1			Recording 2			Recording 3			Recording 4		
Match	**Recognised words**	**499 words**	**Match**	**Recognised words**	**499 words**	**Match**	**Recognised words**	**499 words**	**Match**	**Recognised words**	**499 words**
No	tyn	dhaaa	No	tyn	dhaaa	No	tyn	dhaaa	No	tyn	dhaaa
No	ttaabllin	aaathin	No	ttaabllin	aaathin	No	ttaabllin	aaathin	No	kkaathu	aaathin
No	saah	saaghaa	No	faakkaa	saaghaa	No	saah	saaghaa	No	saawghun	Saaghaa
No	taathill	edhaa	No	sum	edhaa	No	kudaa	edhaa	Yes	edhaa	edhaa
No	zzaaamaa	zzaarr	No	dhaab	zzaarr	No	zzaaamaa	zzaarr	No	zzaafaa	zzaarr
Yes	kaas	kaas	No	omrr	kaas	No	shaams	kaas	No	faazzaa	kaas
No	kaarrnin	aaamaall	No	amaalli	aaamaall	No	kaarrnin	aaamaall	No	amaallaa	Aaamaall
No	sum	jjaatha	Yes	jjaatha	jjaatha	No	dhaaba	jjaatha	No	daasaa	jjaatha
No	shugll	shaah	No	shaarraath	shaah	Yes	shaah	shaah	No	shaahaadhaa	shaah
No	thullth	ttaaf	No	haaf	ttaaf	No	yin	ttaaf	No	zzaafaa	ttaaf
No	ayiyaa	haayaaaaa	No	sum	haayaaaaa	No	haayuUa	haayaaaaa	No	ath	Haayaaaaa
No	khaas	kkaaas	No	khaath	kkaaas	No	ghaat	kkaaas	No	khaas	kkaaas
No	faadh	aukht	No	thullth	aukht	No	ghill	aukht	No	faazzaa	aukht
No	baarrkin	baadaa	No	dhaaghaath	baadaa	No	ghaadha	baadaa	Yes	baadaa	baadaa
No	ghill	aaaw	No	thull	aaaw	No	ghill	aaaw	No	kaarrn	aaaw
No	sum	aaakkaallaa	Yes	aaakkaallaa	aaakkaallaa	No	sum	aaakkaallaa	No	maakkunaa	aaakkaallaa
No	saanaaa	saaaaall	No	sum	saaaaall	No	thaawbanaa	saaaaall	No	thaakullaa	Saaaaall
No	daaaun	dhuUll	No	sum	dhuUll	No	daaaun	dhuUll	No	daahrrun	dhuUll
No	ghaaythu	baaiisaa	No	ghaaythu	baaiisaa	No	ghaaythu	baaiisaa	No	sum	baaiisaa
No	daama	baarraau	No	badhakon	baarraau	No	dhaaba	baarraau	No	dhaaba	baarraau
No	sillkkin	swai	No	saawyi	swai	No	saawyi	swai	No	suhub	swai
No	jjaahan	daaaan	No	daaaun	daaaan	No	daahrran	daaaan	Yes	daaaan	daaaan
No	jjaahin	daaaun	No	Honow	daaaun	Yes	daaaun	daaaun	Yes	daaaun	daaaun
No	daaaan	daaain	No	daaaan	daaain	No	daaaan	daaain	No	daaaan	daaain
No	waafy	thaaby	No	mUdhi	thaaby	No	maahdi	thaaby	No	sum	thaaby
No	dub	dhaaba	No	naabu	dhaaba	No	dhaab	dhaaba	Yes	dhaaba	dhaaba
No	baaiisaa	baazzaaghaa	No	faasin	baazzaaghaa	No	baaiisaa	baazzaaghaa	No	baaiisaa	baazzaaghaa

	Recording 1			Recording 2			Recording 3			Recording 4	
Match	Recognised words	499 words	Match	Recognised words	499 words	Match	Recognised words	499 words	Match	Recognised words	499 words
Yes	baasaall	baasaall	Yes	baasaall	baasaall	No	naasaab	baasaall	Yes	baasaall	baasaall
No	daaahu	baahaak	No	sum	baahaak	No	sum	baahaak	No	daafirraa	baahaak
Yes	khaabaatt	khaabaatt	Yes	khaabaatt	khaabaatt	No	kaadhaaa	khaabaatt	No	haabaasaa	khaabaatt
No	ghaadh	kkaabaa	No	khaath	kkaabaa	Yes	kkaabaa	kkaabaa	No	dhaaba	kkaabaa
No	daama	dhaanb	No	zzaand	dhaanb	Yes	dhaanb	dhaanb	No	daaaun	dhaanb
Yes	baashimaa	baashimaa	No	sum	baashimaa	Yes	baashimaa	baashimaa	Yes	baashimaa	baashimaa
No	subull	saabaa	No	sum	saabaa	Yes	saabaa	saabaa	No	saah	saabaa
Yes	faarraabu	faarraabu	Yes	faarraabu	faarraabu	Yes	faarraabu	faarraabu	No	min	faarraabu
Yes	naasaab	naasaab	Yes	naasaab	naasaab	Yes	naasaab	naasaab	Yes	naasaab	naasaab
Yes	waajjibaa	waajjibaa	Yes	waajjibaa	waajjibaa	Yes	waajjibaa	waajjibaa	Yes	waajjibaa	waajjibaa
No	sum	thaabaataa	No	ttaabllin	thaabaataa	No	rraakkaaa	thaabaataa	No	khaasaafaa	thaabaataa
Yes	baattaallaa	baattaallaa	No	ghaafiyaa	baattaallaa	No	maakkaanaa	baattaallaa	No	ghaasaallaa	baattaallaa
No	rryshi	bishrr	No	dhirrs	bishrr	No	ghaashaa	bishrr	No	naashizz	bishrr
No	min	burrjj	No	sum	burrjj	No	daaaun	burrjj	Yes	burrjj	burrjj
Yes	jjubillaa	jjubillaa	No	sum	jjubillaa	No	rraadhiyaa	jjubillaa	Yes	jjubillaa	jjubillaa
Yes	rraabaattaa	rraabaattaa	Yes	rraabaattaa	rraabaattaa	Yes	rraabaattaa	rraabaattaa	No	waathaafaa	rraabaattaa
Yes	subull	subull	No	sillkkun	subull	No	saamtun	subull	No	saamtun	subull
Yes	haallaabaa	haallaabaa	Yes	haallaabaa	haallaabaa	No	naadaabaa	haallaabaa	Yes	haallaabaa	haallaabaa
No	kaaythi	kaallbi	No	thiny	kaallbi	No	kaaythi	kaallbi	No	kkaanzzi	kaallbi
Yes	naabu	naabu	Yes	naabu	naabu	Yes	naabu	naabu	Yes	naabu	naabu
No	saawghun	thaawbanaa	No	saawghun	thaawbanaa	No	ttaabllin	thaawbanaa	No	thaawbun	thaawbanaa
No	ayn	thaawbun	No	thulluthun	thaawbun	No	thaawbin	thaawbun	No	thaawbin	thaawbun
No	saawgan	thaawbin	No	allaafin	thaawbin	No	saadaa	thaawbin	Yes	thaawbin	thaawbin
No	thaakhn	taaht	No	kaash	taaht	No	sum	taaht	No	saahw	taaht

Match	Recognised words	499 words	Match	Recognised words	499 words	Match	Recognised words	499 words	Match	Recognised words	499 words
	Recording 1			**Recording 2**			**Recording 3**			**Recording 4**	
Yes	dhaamaat	dhaamaat	Yes	dhaamaat	dhaamaat	No	dhaamaa	dhaamaat	Yes	dhaamaat	dhaamaat
No	ttaabllin	taathill	No	ttaabllin	taathill	No	ttaabllin	taathill	No	thaawbun	taathill
Yes	saattaat	saattaat	Yes	saattaat	saattaat	No	saakkaat	saattaat	Yes	saattaat	saattaat
No	ghaat	saakkaat	No	ghaat	saakkaat	No	ghaat	saakkaat	No	khaath	saakkaat
Yes	dhaarraat	dhaarraat	Yes	dhaarraat	dhaarraat	Yes	dhaarraat	dhaarraat	No	daama	dhaarraat
Yes	haazzaat	haazzaat	Yes	haazzaat	haazzaat	Yes	haazzaat	haazzaat	No	faazzaa	haazzaat
No	shaajjaa	shaadaat	No	shaajjaa	shaadaat	Yes	shaadaat	shaadaat	No	yin	shaadaat
Yes	thaanaat	thaanaat	Yes	thaanaat	thaanaat	Yes	thaanaat	thaanaat	No	ghaadha	thaanaat
No	ghaat	jjaafaat	Yes	jjaafaat	jjaafaat	Yes	jjaafaat	jjaafaat	No	dhaafaarr	jjaafaat
No	taathill	otw	No	taathill	otw	No	saamtan	otw	No	taathill	otw
Yes	ghaat	ghaat	No	sum	ghaat	Yes	ghaat	ghaat	No	faanaarr	ghaat
No	dhaakky	taaky	No	dhaakky	taaky	No	rraakkaaa	taaky	No	sum	taaky
No	thull	taamrr	No	saamtan	taamrr	No	dhaamaat	taamrr	No	saamghi	taamrr
Yes	tyn	tyn	No	lliyn	tyn	No	kkiys	tyn	No	mudun	tyn
No	dhuUll	twt	No	thullaat	twt	No	sum	twt	No	ttaaf	twt
No	haattaamaa	kaataallaa	No	maakkaanaa	kaataallaa	No	haattaamaa	kaataallaa	No	aaadhunaa	kaataallaa
No	sillkkin	sutirraa	Yes	sutirraa	sutirraa	Yes	sutirraa	sutirraa	Yes	sutirraa	sutirraa
No	haattaamaa	atumaa	Yes	atumaa	atumaa	No	aaadhunaa	atumaa	No	sum	atumaa
No	sillkkan	yumitu	No	sum	yumitu	No	thull	yumitu	No	thull	yumitu
No	dhaakky	yaakhti	No	dhaakky	yaakhti	No	dhaakky	yaakhti	No	llaayth	yaakhti
Yes	naahaataa	naahaataa	No	saamtan	naahaataa	Yes	naahaataa	naahaataa	No	maakkaathaa	naahaataa
Yes	saamtun	saamtun	Yes	saamtun	saamtun	No	saamtan	saamtun	Yes	saamtun	saamtun
No	Samton	saamtan	No	Samton	saamtan	No	saamtun	saamtan	Yes	saamtan	saamtan
Yes	saamtin	saamtin	No	saamtan	saamtin	No	saamtan	saamtin	No	saamtan	saamtin
No	saaaf	thullth	No	thaarrf	thullth	No	daaf	thullth	No	dhaanaa	thullth

	Recording 1			Recording 2			Recording 3			Recording 4	
Match	Recognised words	499 words	Match	Recognised words	499 words	Match	Recognised words	499 words	Match	Recognised words	499 words
No	ottuf	thaakaaf	No	dhaaghaath	thaakaaf	No	dhaaghaath	thaakaaf	Yes	thaakaaf	thaakaaf
No	dhaaghaath	maakkaathaa	Yes	maakkaathaa	maakkaathaa	No	maakkaanaa	maakkaathaa	No	naathufaa	maakkaathaa
Yes	ghaath	ghaath	Yes	ghaath	ghaath	No	rraad	ghaath	Yes	ghaath	ghaath
Yes	haadaath	haadaath	No	sum	haadaath	Yes	haadaath	haadaath	Yes	haadaath	haadaath
Yes	shaarraath	shaarraath	No	shaadaat	shaarraath	No	shaathaaf	shaarraath	No	shaah	shaarraath
No	haaf	ath	No	Aath	ath	No	haaf	ath	Yes	ath	ath
No	saawyi	thaawy	No	saawyi	thaawy	No	saadaa	thaawy	No	sum	thaawy
Yes	thaakhn	thaakhn	No	faadhun	thaakhn	Yes	thaakhn	thaakhn	No	kkaahaan	thaakhn
No	taathill	baathaahu	No	ghaasaallaa	baathaahu	No	dhaawuw	baathaahu	No	dhaakhaahu	baathaahu
Yes	thaabaattaa	thaabaattaa	Yes	thaabaattaa	thaabaattaa	Yes	thaabaattaa	thaabaattaa	Yes	thaabaattaa	thaabaattaa
Yes	thaajjaa	thaajjaa	No	sum	thaajjaa	No	dhaajjaa	thaajjaa	No	shaatt	thaajjaa
No	tyn	thiny	No	saaniyaa	thiny	No	shaahy	thiny	No	sum	thiny
No	faaallaa	thullaat	No	faaallaa	thullaat	No	waanaa	thullaat	No	faaallaa	thullaat
No	waasia	waathaabaa	Yes	waathaabaa	waathaabaa	Yes	waathaabaa	waathaabaa	Yes	waathaabaa	waathaabaa
No	fillizzin	othirraa	No	fillizzin	othirraa	No	waasia	othirraa	No	sum	othirraa
No	taathill	jjuthw	No	sum	jjuthw	No	taathill	jjuthw	No	dhaakhaahu	jjuthw
No	haathu	aaathaa	No	haathu	aaathaa	No	haakkaa	aaathaa	No	kaasaa	aaathaa
Yes	rraathi	rraathi	Yes	rraathi	rraathi	No	waafy	rraathi	No	llaayth	rraathi
Yes	baathu	baathu	No	ghaaythu	baathu	No	baasaall	baathu	Yes	baathu	baathu
No	fillizzin	thulluthin	No	kaarrdhun	thulluthin	No	fillizzan	thulluthin	No	saamtan	thulluthin
No	thulluthin	thulluthun	No	fillizzan	thulluthun	No	fillizzin	thulluthun	No	fillizzun	thulluthun
No	fillizzin	thulluthan	No	haasaan	thulluthan	No	fillizzan	thulluthan	No	saamtan	thulluthan
Yes	llujjaajj	llujjaajj	Yes	llujjaajj	llujjaajj	Yes	llujjaajj	llujjaajj	Yes	llujjaajj	llujjaajj
Yes	jjaarraakkaa	jjaarraakkaa	Yes	jjaarraakkaa	jjaarraakkaa	Yes	jjaarraakkaa	jjaarraakkaa	Yes	jjaarraakkaa	jjaarraakkaa
No	baattaallaa	dhaajjaa	No	sum	dhaajjaa	Yes	dhaajjaa	dhaajjaa	No	dhaaba	dhaajjaa

	Recording 1			Recording 2			Recording 3			Recording 4	
Match	Recognised words	499 words	Match	Recognised words	499 words	Match	Recognised words	499 words	Match	Recognised words	499 words
Yes	jjaas	jjaas	Yes	jjaas	jjaas	No	ghaath	jjaas	Yes	jjaas	jjaas
No	ttaajjaan	khaajjaall	No	faadhin	khaajjaall	No	faadhin	khaajjaall	No	faadhanaa	khaajjaall
No	shaahaadhaa	jjaahaathaa	No	jjaahun	jjaahaathaa	No	dhaakhaahu	jjaahaathaa	Yes	jjaahaathaa	jjaahaathaa
No	ewaajjin	ttaajjaan	Yes	ttaajjaan	ttaajjaan	No	faadhin	ttaajjaan	No	faadhanaa	ttaajjaan
Yes	shaajjaa	shaajjaa	Yes	shaajjaa	shaajjaa	Yes	shaajjaa	shaajjaa	No	daaahu	shaajjaa
Yes	ajjaazzaa	ajjaazzaa	Yes	ajjaazzaa	ajjaazzaa	No	faadhanaa	ajjaazzaa	Yes	ajjaazzaa	ajjaazzaa
No	sujjuk	saajjaaa	No	sum	saajjaaa	No	thaajjaa	saajjaaa	No	thaajjaa	saajjaaa
No	jjuthw	jjudhm	No	jjuhd	jjudhm	No	daagll	jjudhm	No	sum	jjudhm
No	dhaakky	jjaady	No	dhakky	jjaady	No	sum	jjaady	No	dhaakky	jjaady
No	jjuzzurr	jjaazzaa	Yes	jjaazzaa	jjaazzaa	No	kaazzaah	jjaazzaa	Yes	jjaazzaa	jjaazzaa
No	tyn	haajjaa	Yes	haajjaa	haajjaa	No	ttaajjaan	haajjaa	No	kkaabshan	haajjaa
Yes	jjaawk	jjaawk	Yes	jjaawk	jjaawk	No	ghill	jjaawk	No	sum	jjaawk
No	sum	jjaamaall	Yes	jjaamaall	jjaamaall	No	daahrrun	jjaamaall	No	ghaanaamu	jjaamaall
Yes	jjuhd	jjuhd	No	jjaafaat	jjuhd	No	dub	jjuhd	No	ghill	jjuhd
No	ghill	jjidu	No	sum	jjidu	No	ghaadh	jjidu	No	naabu	jjidu
No	wujjidaa	waajjaadaa	No	waajjibaa	waajjaadaa	Yes	waajjaadaa	waajjaadaa	No	naadaabaa	waajjaadaa
Yes	aaajjidu	aaajjidu	Yes	aaajjidu	aaajjidu	Yes	aaajjidu	aaajjidu	Yes	aaajjidu	aaajjidu
Yes	hujjub	hujjub	No	sum	hujjub	No	sum	hujjub	Yes	hujjub	hujjub
No	dhaajjaa	daarraajjaa	No	ghaajjaarr	daarraajjaa	No	dhaajjaa	daarraajjaa	No	ghaadha	daarraajjaa
Yes	saarrju	saarrju	No	khaajjaall	saarrju	No	saamtan	saarrju	No	saamtan	saarrju
Yes	waahaajji	waahaajji	Yes	waahaajji	waahaajji	Yes	waahaajji	waahaajji	Yes	waahaajji	waahaajji
No	ewaajjin	ewaajjan	No	aaawidu	ewaajjan	No	arraadhaa	ewaajjan	No	waaadun	ewaajjan
Yes	ewaajjun	ewaajjun	No	sum	ewaajjun	No	aaadhunaa	ewaajjun	No	waaadan	ewaajjun
Yes	ewaajjin	ewaajjin	Yes	ewaajjin	ewaajjin	No	aaajjidu	ewaajjin	No	sum	ewaajjin
No	thaabaattaa	haadhaarraa	No	haamaallaa	haadhaarraa	No	haabaasaa	haadhaarraa	No	haattaamaa	haadhaarraa

Match	Recognised words	499 words	Match	Recognised words	499 words	Match	Recognised words	499 words	Match	Recognised words	499 words
No	sum	kaazzaah	No	haazzaat	kaazzaah	No	faazzaa	kaazzaah	No	sum	kaazzaah
Yes	suhuf	suhuf	Yes	suhuf	suhuf	No	suhub	suhuf	Yes	suhuf	suhuf
No	haathu	haadhw	No	haathu	haadhw	Yes	haadhw	haadhw	No	khaadukk	haadhw
Yes	haattaamaa	haattaamaa	Yes	haattaamaa	haattaamaa	No	daama	haattaamaa	No	daama	haattaamaa
Yes	haasaan	haasaan	No	faasin	haasaan	Yes	haasaan	haasaan	Yes	haasaan	haasaan
No	sum	haakkaa	Yes	haakkaa	haakkaa	No	dhaanaa	haakkaa	No	khaath	haakkaa
No	ghaanaamu	haallaahu	Yes	haallaahu	haallaahu	No	sum	haallaahu	Yes	haallaahu	haallaahu
No	fyhi	haay	No	thiny	haay	Yes	haay	haay	No	haadaath	haay
No	faadhanaa	haamaallaa	Yes	haamaallaa	haamaallaa	No	khaamaanaa	haamaallaa	No	khaamaanaa	haamaallaa
Yes	hibrr	hibrr	Yes	hibrr	hibrr	No	naabu	hibrr	Yes	hibrr	hibrr
Yes	husn	husn	No	faasun	husn	No	faasin	husn	No	haasaan	husn
Yes	ttaahaanaa	ttaahaanaa	No	kaahaarraa	ttaahaanaa	Yes	ttaahaanaa	ttaahaanaa	Yes	ttaahaanaa	ttaahaanaa
Yes	suhub	suhub	Yes	suhub	suhub	Yes	suhub	suhub	Yes	suhub	suhub
Yes	yaahillu	yaahillu	No	sum	yaahillu	No	dhiib	yaahillu	No	faahaamaa	yaahillu
Yes	maasaahaa	maasaahaa	Yes	maasaahaa	maasaahaa	Yes	maasaahaa	maasaahaa	Yes	maasaahaa	maasaahaa
No	ttaarraafi	faarraahi	No	sum	faarraahi	No	shaahy	faarraahi	No	thaahaarraa	faarraahi
Yes	maarraahu	maarraahu	Yes	maarraahu	maarraahu	Yes	maarraahu	maarraahu	No	yaahillu	maarraahu
No	kaazzaahun	kaazzaahanaa	No	kaazzaahun	kaazzaahanaa	No	kaazzaahun	kaazzaahanaa	No	kaazzaahun	kaazzaahanaa
Yes	kaazzaahin	kaazzaahin	No	kazzaHen	kaazzaahin	No	waadhaaa	kaazzaahin	No	waadhaaa	kaazzaahin
Yes	kaazzaahun	kaazzaahun	No	kazzaHon	kaazzaahun	Yes	kaazzaahun	kaazzaahun	Yes	kaazzaahun	kaazzaahun
Yes	dhaakhaahu	dhaakhaahu	Yes	dhaakhaahu	dhaakhaahu	Yes	dhaakhaahu	dhaakhaahu	Yes	dhaakhaahu	dhaakhaahu
No	faadh	khaadukk	No	faadh	khaadukk	No	faadh	khaadukk	No	sum	khaadukk
No	haaf	khaath	No	sum	khaath	No	haaf	khaath	No	ghaath	khaath
Yes	khaashaaa	khaashaaa	No	sum	khaashaaa	No	rraashaa	khaashaaa	No	haazzaat	khaashaaa
No	faasun	khaasaa	No	faasun	khaasaa	No	faasaa	khaasaa	No	kaasaa	khaasaa

	Recording 1			Recording 2			Recording 3			Recording 4	
Match	Recognised words	499 words	Match	Recognised words	499 words	Match	Recognised words	499 words	Match	Recognised words	499 words
No	daafirraa	dhaakhaarraa	No	naahaataa	dhaakhaarraa	No	naahaataa	dhaakhaarraa	No	daakaakkaa	dhaakhaarraa
No	khaasaafaa	khaazzaakaa	No	kaazzaahun	khaazzaakaa	Yes	khaazzaakaa	khaazzaakaa	Yes	khaazzaakaa	khaazzaakaa
Yes	khaasaafaa	khaasaafaa	Yes	khaasaafaa	khaasaafaa	Yes	khaasaafaa	khaasaafaa	Yes	khaasaafaa	khaasaafaa
No	faadhanaa	khaamaanaa	No	haamaallaa	khaamaanaa	No	dhaamaa	khaamaanaa	Yes	khaamaanaa	khaamaanaa
No	saawyi	khaawy	No	saawyi	khaawy	No	sum	khaawy	No	sum	khaawy
Yes	khaas	khaas	No	haaf	khaas	No	haaf	khaas	Yes	khaas	khaas
No	hibrr	khidrr	No	hibrr	khidrr	No	hibrr	khidrr	No	faadhanaa	khidrr
No	shaams	khums	Yes	khums	khums	No	shaams	khums	No	shaams	khums
Yes	baakhaasaa	baakhaasaa	No	naathufaa	baakhaasaa	Yes	baakhaasaa	baakhaasaa	Yes	baakhaasaa	baakhaasaa
No	naathufaa	baakhillaa	No	naasiyaa	baakhillaa	No	baashimaa	baakhillaa	Yes	baakhillaa	baakhillaa
No	orrsun	rraakhusaa	No	rraahufaa	rraakhusaa	Yes	rraakhusaa	rraakhusaa	Yes	rraakhusaa	rraakhusaa
No	saattaat	saarraakhaa	No	waathaafaa	saarraakhaa	No	shaarraafaa	saarraakhaa	Yes	saarraakhaa	saarraakhaa
No	dhaakky	mukhi	No	sum	mukhi	No	rraathi	mukhi	No	mUdhi	mukhi
No	saamtan	saallkhu	No	saamtan	saallkhu	No	saamtan	saallkhu	No	saamtan	saallkhu
No	baadhaakhin	baadhaakhun	Yes	baadhaakhun	baadhaakhun	Yes	baadhaakhun	baadhaakhun	Yes	baadhaakhun	baadhaakhun
Yes	baadhaakhin	baadhaakhin	Yes	baadhaakhin	baadhaakhin	No	ayn	baadhaakhin	No	baathaahu	baadhaakhin
No	baadhaakhun	baadhaakhanaa	No	baadhaakhun	baadhaakhanaa	No	baadhaakhun	baadhaakhanaa	No	baadhaakhun	baadhaakhanaa
No	min	dhid	No	min	dhid	No	ghaadh	dhid	No	dub	dhid
Yes	zzaand	zzaand	Yes	zzaand	zzaand	Yes	zzaand	zzaand	Yes	zzaand	zzaand
Yes	rraasaadaa	rraasaadaa	Yes	rraasaadaa	rraasaadaa	Yes	rraasaadaa	rraasaadaa	Yes	rraasaadaa	rraasaadaa
No	faadhin	kaadaam	No	faadhin	kaadaam	No	kaadhaaa	kaadaam	No	kaarrnan	kaadaam
No	thull	ttaawd	No	thull	ttaawd	No	daaaun	ttaawd	No	kaarrnun	ttaawd
Yes	daasaa	daasaa	No	dhaafaarr	daasaa	No	ghaasaallaa	daasaa	Yes	daasaa	daasaa
No	daama	daagll	No	ghaanaamu	daagll	No	daahrran	daagll	No	daama	daagll

	Recording 1			Recording 2			Recording 3			Recording 4	
Match	Recognised words	499 words	Match	Recognised words	499 words	Match	Recognised words	499 words	Match	Recognised words	499 words
Yes	daaahu	daaahu	No	daAho	daaahu	Yes	daaahu	daaahu	Yes	daaahu	daaahu
Yes	daaf	daaf	No	ghaath	daaf	No	dhaab	daaf	No	dub	daaf
Yes	daama	daama	No	ghaanaamu	daama	No	dhaanaa	daama	No	ghaanaamu	daama
No	ghill	dub	No	ghill	dub	Yes	dub	dub	Yes	dub	dub
Yes	diykk	diykk	Yes	diykk	diykk	Yes	diykk	diykk	Yes	diykk	diykk
Yes	naadaabaa	naadaabaa	Yes	naadaabaa	naadaabaa	Yes	naadaabaa	naadaabaa	Yes	naadaabaa	naadaabaa
No	sum	hudidaa	No	sum	hudidaa	No	wujjidaa	hudidaa	Yes	hudidaa	hudidaa
Yes	mudun	mudun	Yes	mudun	mudun	Yes	mudun	mudun	Yes	mudun	mudun
No	saajjaaa	saadaa	No	sum	saadaa	No	saajjaaa	saadaa	No	saabaa	saadaa
No	haathu	ahdu	No	aaawidu	ahdu	No	sum	ahdu	No	sum	ahdu
No	dhaaa	maahdi	No	naabu	maahdi	No	sum	maahdi	No	maahwu	maahdi
No	waattaan	waaadan	Yes	waaadan	waaadan	No	waadhaaa	waaadan	No	waattaan	waaadan
No	ewaajjun	waaadun	No	ewaajjun	waaadun	No	ewaajjun	waaadun	No	waaadan	waaadun
No	faadhin	waaadin	No	waattaan	waaadin	No	daaaun	waaadin	No	daaaun	waaadin
Yes	dhaakky	dhaakky	Yes	dhaakky	dhaakky	No	maahdi	dhaakky	No	yaakhti	dhaakky
No	daama	dhaamaa	No	daama	dhaamaa	No	daama	dhaamaa	No	daama	dhaamaa
No	Dhana	dhaallaa	No	Dhana	dhaallaa	No	dhaanaa	dhaallaa	No	naathaarraa	dhaallaa
No	min	faadh	Yes	faadh	faadh	No	sum	faadh	No	faan	faadh
Yes	kaadhaaa	kaadhaaa	Yes	kaadhaaa	kaadhaaa	No	kkaabaa	kaadhaaa	No	ghaadha	kaadhaaa
No	shaathaaf	shaadhaa	No	yin	shaadhaa	No	shaathaaf	shaadhaa	No	shaathaaf	shaadhaa
No	ghill	dhaawd	No	ghill	dhaawd	No	daaaun	dhaawd	No	dhaab	dhaawd
No	dhid	dhiib	No	min	dhiib	No	dhid	dhiib	No	dhid	dhiib
Yes	dhaab	dhaab	No	ghaadh	dhaab	Yes	dhaab	dhaab	No	ghaadha	dhaab
No	daaaun	dhull	No	mudun	dhull	No	daaaun	dhull	No	mUdhi	dhull

Recording 1			Recording 2			Recording 3			Recording 4		
Match	Recognised words	499 words	Match	Recognised words	499 words	Match	Recognised words	499 words	Match	Recognised words	499 words
No	rraahibaa	kkaadhibaa	No	sum	kkaadhibaa	No	naadaabaa	kkaadhibaa	Yes	kkaadhibaa	kkaadhibaa
No	kaataallaa	adhaarraa	No	faadhanaa	adhaarraa	No	baadaa	adhaarraa	No	ghaadaarraa	adhaarraa
Yes	aaadhunaa	aaadhunaa	Yes	aaadhunaa	aaadhunaa	Yes	aaadhunaa	aaadhunaa	Yes	aaadhunaa	aaadhunaa
Yes	shaahaadhaa	shaahaadhaa	Yes	shaahaadhaa	shaahaadhaa	Yes	shaahaadhaa	shaahaadhaa	Yes	shaahaadhaa	shaahaadhaa
No	minhu	mundhu	Yes	mundhu	mundhu	Yes	mundhu	mundhu	Yes	mundhu	mundhu
No	diykk	mUdhi	No	maahdi	mUdhi	No	maahdi	mUdhi	No	atyin	mUdhi
No	fillizzin	faadhanaa	No	faadh	faadhanaa	No	kaarrnan	faadhanaa	No	kaarrnun	faadhanaa
Yes	faadhun	faadhun	Yes	faadhun	faadhun	No	ttaabllan	faadhun	No	ttaabllan	faadhun
No	aaathin	faadhin	No	aaathin	faadhin	No	ttaabllin	faadhin	No	thaawbin	faadhin
No	dhaaba	thaahaarraa	No	naahaataa	thaahaarraa	No	ghaattaa	thaahaarraa	No	naahaataa	thaahaarraa
Yes	kaarrn	kaarrn	No	kaarrnin	kaarrn	No	ttaabllan	kaarrn	Yes	kaarrn	kaarrn
Yes	rraakkaallaa	rraakkaallaa	Yes	rraakkaallaa	rraakkaallaa	No	maakkaanaa	rraakkaallaa	Yes	rraakkaallaa	rraakkaallaa
No	daama	dhaarraa	No	faattaarraa	dhaarraa	No	ghaadaarraa	dhaarraa	No	dhaarraat	dhaarraa
No	maahwu	rraagw	No	minhu	rraagw	No	ttaabllan	rraagw	No	naathaarraa	rraagw
No	ayiyaa	ttaayrr	No	faaallaa	ttaayrr	No	sum	ttaayrr	No	kaarrnan	ttaayrr
No	sihrr	sirr	No	yin	sirr	No	sihrr	sirr	No	saadghu	sirr
No	faadh	rraad	Yes	rraad	rraad	No	sum	rraad	No	ghaadha	rraad
Yes	rrubaa	rrubaa	No	sum	rrubaa	Yes	rrubaa	rrubaa	No	ghaath	rrubaa
No	suhub	surrurr	No	fillizzin	surrurr	No	shibll	surrurr	No	saaerr	surrurr
No	haattaamaa	haarraamaa	No	allaaman	haarraamaa	No	haattaamaa	haarraamaa	Yes	haarraamaa	haarraamaa
No	saanaami	sirry	No	saanaami	sirry	No	faarraasi	sirry	No	saamtan	sirry
No	faadhin	faattaarraa	Yes	faattaarraa	faattaarraa	No	rraabaattaa	faattaarraa	No	haafaathaa	faattaarraa
No	sum	jjuhrri	No	jjaahun	jjuhrri	No	sum	jjuhrri	No	yaahillu	jjuhrri
No	thaaby	fikkrri	No	sum	fikkrri	No	dhaakky	fikkrri	No	faakkaa	fikkrri

	Recording 1			Recording 2			Recording 3			Recording 4	
Match	Recognised words	499 words	Match	Recognised words	499 words	Match	Recognised words	499 words	Match	Recognised words	499 words
No	daahrrin	daahrrun	No	daaaun	daahrrun	No	daaaun	daahrrun	Yes	daahrrun	daahrrun
No	daaaan	daahrrin	No	daaaan	daahrrin	No	daaaan	daahrrin	No	daaaan	daahrrin
No	daahrrun	daahrran	No	jjaahun	daahrran	No	daahrrun	daahrran	No	daahrrun	daahrran
Yes	zzaafaa	zzaafaa	Yes	zzaafaa	zzaafaa	Yes	zzaafaa	zzaafaa	Yes	zzaafaa	zzaafaa
Yes	zzaaamaa	zzaaamaa	No	daama	zzaaamaa	Yes	zzaaamaa	zzaaamaa	Yes	zzaaamaa	zzaaamaa
Yes	zzaakky	zzaakky	No	dhaakky	zzaakky	Yes	zzaakky	zzaakky	Yes	zzaakky	zzaakky
Yes	zzuhaall	zzuhaall	No	jjaahun	zzuhaall	Yes	zzuhaall	zzuhaall	Yes	zzuhaall	zzuhaall
No	zzaaamaa	zzaarraaa	No	daama	zzaarraaa	No	dhaaba	zzaarraaa	No	zzaakky	zzaarraaa
Yes	zzirr	zzirr	No	yin	zzirr	No	zzaaamaa	zzirr	No	zzaand	zzirr
Yes	rruzzik	rruzzik	No	sum	rruzzik	No	sum	rruzzik	No	baazzaaghaa	rruzzik
Yes	azzaafaa	azzaafaa	No	kaazzaahin	azzaafaa	No	khaazzaakaa	azzaafaa	Yes	azzaafaa	azzaafaa
Yes	jjuzzurr	jjuzzurr	Yes	jjuzzurr	jjuzzurr	No	sum	jjuzzurr	No	daasaa	jjuzzurr
No	faasin	faazzaa	No	faasin	faazzaa	Yes	faazzaa	faazzaa	Yes	faazzaa	faazzaa
Yes	jjaawzzu	jjaawzzu	Yes	jjaawzzu	jjaawzzu	No	sum	jjaawzzu	No	jjaazzaa	jjaawzzu
Yes	kkaanzzi	kkaanzzi	No	khaadukk	kkaanzzi	No	taathill	kkaanzzi	Yes	kkaanzzi	kkaanzzi
Yes	fillizzan	fillizzan	Yes	fillizzan	fillizzan	Yes	fillizzan	fillizzan	Yes	fillizzan	fillizzan
Yes	fillizzun	fillizzun	Yes	fillizzun	fillizzun	Yes	fillizzun	fillizzun	Yes	fillizzun	fillizzun
Yes	fillizzin	fillizzin	Yes	fillizzin	fillizzin	No	asaall	fillizzin	No	asaall	fillizzin
Yes	shaams	shaams	Yes	shaams	shaams	Yes	shaams	shaams	Yes	shaams	shaams
No	baasaattaa	ghaasaallaa	No	rraasaadaa	ghaasaallaa	No	baasaattaa	ghaasaallaa	Yes	ghaasaallaa	ghaasaallaa
Yes	saahw	saahw	Yes	saahw	saahw	Yes	saahw	saahw	No	saamtan	saahw
No	kkiys	kkys	No	kkiys	kkys	No	kkiys	kkys	No	saaaf	kkys
No	maawzz	dhirrs	No	daaf	dhirrs	No	daaf	dhirrs	No	daasaa	dhirrs
No	sirry	sum	No	saawghin	sum	No	shaahy	sum	Yes	sum	sum

Recording 1			Recording 2			Recording 3			Recording 4		
Match	Recognised words	499 words	Match	Recognised words	499 words	Match	Recognised words	499 words	Match	Recognised words	499 words
No	sillkkin	saakkaabaa	Yes	saakkaabaa	saakkaabaa	Yes	saakkaabaa	saakkaabaa	Yes	saakkaabaa	saakkaabaa
No	suhuf	sihrr	No	suhub	sihrr	Yes	sihrr	sihrr	No	saahw	sihrr
No	orrsin	rrusull	No	fillizzun	rrusull	Yes	rrusull	rrusull	No	nusirraa	rrusull
No	haasaan	asaall	No	haasaan	asaall	No	haasaan	asaall	No	haasaan	asaall
Yes	naasiyaa	naasiyaa	Yes	naasiyaa	naasiyaa	Yes	naasiyaa	naasiyaa	Yes	naasiyaa	naasiyaa
Yes	haabaasaa	haabaasaa	Yes	haabaasaa	haabaasaa	Yes	haabaasaa	haabaasaa	Yes	haabaasaa	haabaasaa
No	haasaan	haarraasaa	Yes	haarraasaa	haarraasaa	No	haasaan	haarraasaa	Yes	haarraasaa	haarraasaa
No	fillizzin	faarraasi	No	sum	faarraasi	No	rraathi	faarraasi	No	ghaasaallaa	faarraasi
No	faasin	orrsan	No	faasin	orrsan	No	fillizzin	orrsan	No	fillizzan	orrsan
No	haasaan	orrsun	No	haasaan	orrsun	No	rrusull	orrsun	No	rrusull	orrsun
No	faasin	orrsin	No	faasin	orrsin	No	fillizzin	orrsin	No	fillizzin	orrsin
No	shibll	shaadhw	No	shibll	shaadhw	No	shibll	shaadhw	Yes	shaadhw	shaadhw
Yes	shaas	shaas	Yes	shaas	shaas	Yes	shaas	shaas	Yes	shaas	shaas
Yes	shaathaaf	shaathaaf	Yes	shaathaaf	shaathaaf	No	shaadaat	shaathaaf	Yes	shaathaaf	shaathaaf
Yes	shaatt	shaatt	Yes	shaatt	shaatt	Yes	shaatt	shaatt	Yes	shaatt	shaatt
Yes	shugll	shugll	No	saawghin	shugll	No	shibll	shugll	No	shaadhaa	shugll
Yes	kaash	kaash	No	kaas	kaash	Yes	kaash	kaash	Yes	kaash	kaash
Yes	shaakk	shaakk	Yes	shaakk	shaakk	Yes	shaakk	shaakk	No	shaatt	shaakk
Yes	naashizz	naashizz	Yes	naashizz	naashizz	Yes	naashizz	naashizz	No	baasaattaa	naashizz
Yes	shaahy	shaahy	No	saamghi	shaahy	Yes	shaahy	shaahy	Yes	shaahy	shaahy
No	shaajjaa	shaajjaarr	No	shaajjaa	shaajjaarr	No	shaajjaa	shaajjaarr	No	shaadhaa	shaajjaarr
Yes	shibll	shibll	Yes	shibll	shibll	Yes	shibll	shibll	Yes	shibll	shibll
No	sum	shukkrr	No	shaakk	shukkrr	No	shaatt	shukkrr	No	shugll	shukkrr
Yes	waashm	waashm	No	waasy	waashm	Yes	waashm	waashm	No	rraashaa	waashm
Yes	rrushidaa	rrushidaa	Yes	rrushidaa	rrushidaa	Yes	rrushidaa	rrushidaa	Yes	rrushidaa	rrushidaa

	Recording 1			Recording 2			Recording 3			Recording 4	
Match	**Recognised words**	**499 words**	**Match**	**Recognised words**	**499 words**	**Match**	**Recognised words**	**499 words**	**Match**	**Recognised words**	**499 words**
Yes	aaashudu	aaashudu	Yes	aaashudu	aaashudu	Yes	aaashudu	aaashudu	Yes	aaashudu	aaashudu
Yes	rraashaa	rraashaa	No	rraasaa	rraashaa	Yes	rraashaa	rraashaa	Yes	rraashaa	rraashaa
Yes	rrimshu	rrimshu	No	rrusull	rrimshu	No	sum	rrimshu	Yes	rrimshu	rrimshu
Yes	rryshi	rryshi	Yes	rryshi	rryshi	No	sum	rryshi	No	aaashudu	rryshi
No	kkaabshun	kkaabshan	No	khaajjaall	kkaabshan	Yes	kkaabshan	kkaabshan	Yes	kkaabshan	kkaabshan
Yes	kkaabshun	kkaabshun	Yes	kkaabshun	kkaabshun	Yes	kkaabshun	kkaabshun	Yes	kkaabshun	kkaabshun
No	khaajjaall	kkaabshin	No	khaajjaall	kkaabshin	No	shaahy	kkaabshin	No	kkaabshan	kkaabshin
No	faasun	kaasaa	No	faasun	kaasaa	No	haasaan	kaasaa	No	kaas	kaasaa
No	saamghi	sum	No	faan	sum	No	shaahy	sum	No	saamghi	sum
No	faanaarr	saanaaa	No	faanaarr	saanaaa	No	faanaarr	saanaaa	Yes	saanaaa	saanaaa
Yes	saah	saah	No	saaaf	saah	Yes	saah	saah	Yes	saah	saah
Yes	waasy	waasy	No	waafy	waasy	Yes	waasy	waasy	No	waasaatti	waasy
No	sillkkin	suws	No	thullth	suws	No	shaams	suws	No	suhuf	suws
Yes	saayd	saayd	No	fyhi	saayd	Yes	saayd	saayd	No	saaaf	saayd
Yes	sihrr	sihrr	No	faan	sihrr	No	saaerr	sihrr	No	saahw	sihrr
No	baasaattaa	asaarraa	No	faasun	asaarraa	No	kaazzaah	asaarraa	No	kaasaa	asaarraa
Yes	nusirraa	nusirraa	Yes	nusirraa	nusirraa	No	naasiyaa	nusirraa	No	naasiyaa	nusirraa
Yes	yaasudu	yaasudu	No	naasaab	yaasudu	No	naasaab	yaasudu	Yes	yaasudu	yaasudu
No	waasia	rraasaa	Yes	rraasaa	rraasaa	Yes	rraasaa	rraasaa	Yes	rraasaa	rraasaa
No	waasy	kurrsi	Yes	kurrsi	kurrsi	No	waasy	kurrsi	No	fillizzan	kurrsi
Yes	faasun	faasun	Yes	faasun	faasun	No	haasaan	faasun	Yes	faasun	faasun
No	kazzaHon	faasaa	No	kazzaHon	faasaa	No	kazzaHon	faasaa	No	haasaan	faasaa
No	faSon	faasin	No	faSon	faasin	No	rraafaaa	faasin	No	rraafaaa	faasin
No	maallaaf	dhaaghaath	No	baahaak	dhaaghaath	No	dhaaba	dhaaghaath	No	daama	dhaaghaath
No	baattaallaa	waadhaaa	No	waaadun	waadhaaa	No	waattaan	waadhaaa	No	dhaaba	waadhaaa
No	daama	dhaanaa	No	maakkaanaa	dhaanaa	No	waanaa	dhaanaa	No	daaaun	dhaanaa

Recording 1			Recording 2			Recording 3			Recording 4		
Match	Recognised words	499 words	Match	Recognised words	499 words	Match	Recognised words	499 words	Match	Recognised words	499 words
No	daama	dhaallaa	No	daama	dhaallaa	No	daama	dhaallaa	No	daama	dhaallaa
No	diykk	dhyk	No	sum	dhyk	No	diykk	dhyk	No	sum	dhyk
No	daafirraa	dhaafaarr	No	waaathun	dhaafaarr	No	zzaafaa	dhaafaarr	No	maakkaanaa	dhaafaarr
No	waathaafaa	dhaarraabaa	No	waathaafaa	dhaarraabaa	No	ghaadha	dhaarraabaa	No	naadaabaa	dhaarraabaa
No	suhub	dhuha	Yes	dhuha	dhuha	Yes	dhuha	dhuha	Yes	dhuha	dhuha
No	ghill	dhidu	No	ghaaythu	dhidu	No	shibll	dhidu	No	naabu	dhidu
Yes	rraadhiyaa	rraadhiyaa	Yes	rraadhiyaa	rraadhiyaa	Yes	rraadhiyaa	rraadhiyaa	Yes	rraadhiyaa	rraadhiyaa
No	allaafun	adhud	No	allaaman	adhud	No	ttaabllin	adhud	No	aaadhunaa	adhud
No	thaabaattaa	faadhaallaa	No	faaallaa	faadhaallaa	No	waathaabaa	faadhaallaa	No	thaabaattaa	faadhaallaa
No	mundhu	maarraadhaa	No	naadaabaa	maarraadhaa	No	naadaabaa	maarraadhaa	No	naadaabaa	maarraadhaa
No	thaabaattaa	arraadhaa	No	haattaamaa	arraadhaa	No	baadaa	arraadhaa	No	dhaarraabaa	arraadhaa
No	thaaby	aaarrdhi	No	sum	aaarrdhi	No	thaaby	aaarrdhi	No	kaarrnun	aaarrdhi
No	thaawbun	kaarrdhan	No	saamtun	kaarrdhan	No	ttaabllun	kaarrdhan	No	kaarrnan	kaarrdhan
No	kaarrnun	kaarrdhun	No	kaarrnun	kaarrdhun	No	ttaabllan	kaarrdhun	No	kaarrnun	kaarrdhun
No	ttaabllin	kaarrdhin	No	saamtan	kaarrdhin	No	ttaabllin	kaarrdhin	No	kaarrnan	kaarrdhin
No	faadh	ttaak	No	faadh	ttaak	No	faadh	ttaak	No	ghaat	ttaak
No	maakkaanaa	haattaallaa	No	haattaamaa	haattaallaa	No	saakaattaa	haattaallaa	No	haattaamaa	haattaallaa
No	daama	ttaamaaa	No	daama	ttaamaaa	No	daama	ttaamaaa	No	kaarrnan	ttaamaaa
No	sum	ttib	No	faadh	ttib	No	sum	ttib	No	sum	ttib
No	ttaabllin	ttaabaaa	No	faaallaa	ttaabaaa	No	rraabaattaa	ttaabaaa	No	rraadhiyaa	ttaabaaa
Yes	waattaan	waattaan	Yes	waattaan	waattaan	Yes	waattaan	waattaan	No	waaathun	waattaan
No	baarrkin	rraattib	No	rraakkaaa	rraattib	No	rraakkaaa	rraattib	No	waaathan	rraattib
No	sum	ottuf	Yes	ottuf	ottuf	No	dhaaghaath	ottuf	No	thaakaaf	ottuf
No	taathill	kirrttu	No	minhu	kirrttu	No	taathill	kirrttu	No	minhu	kirrttu
Yes	waasaatti	waasaatti	Yes	waasaatti	waasaatti	Yes	waasaatti	waasaatti	Yes	waasaatti	waasaatti

Recording 1			Recording 2			Recording 3			Recording 4		
Match	Recognised words	499 words	Match	Recognised words	499 words	Match	Recognised words	499 words	Match	Recognised words	499 words
No	khaasaafaa	baasaattaa	Yes	baasaattaa	baasaattaa	Yes	baasaattaa	baasaattaa	Yes	baasaattaa	baasaattaa
Yes	nukaattan	nukaattan	Yes	nukaattan	nukaattan	Yes	nukaattan	nukaattan	Yes	nukaattan	nukaattan
Yes	nukaattun	nukaattun	Yes	nukaattun	nukaattun	Yes	nukaattun	nukaattun	Yes	nukaattun	nukaattun
Yes	nukaattin	nukaattin	Yes	nukaattin	nukaattin	Yes	nukaattin	nukaattin	Yes	nukaattin	nukaattin
No	faadh	thaahaarr	No	baahaak	thaahaarr	No	baathu	thaahaarr	No	naahaataa	thaahaarr
No	naabu	kkaathu	No	naabu	kkaathu	No	naabu	kkaathu	No	daama	kkaathu
Yes	waathaafaa	waathaafaa	Yes	waathaafaa	waathaafaa	Yes	waathaafaa	waathaafaa	Yes	waathaafaa	waathaafaa
No	naabu	thaarrf	No	ghaath	thaarrf	No	maallaaf	thaarrf	No	daama	thaarrf
No	daaf	thifrr	No	daasaa	thifrr	No	daaf	thifrr	No	daafirraa	thifrr
No	daaaan	thull	No	daaaun	thull	No	daaaun	thull	No	daaaan	thull
Yes	naathaarraa	naathaarraa	Yes	naathaarraa	naathaarraa	No	naadaabaa	naathaarraa	No	dhaaba	naathaarraa
No	naabu	naathufaa	No	naabu	naathufaa	Yes	naathufaa	naathufaa	Yes	naathufaa	naathufaa
No	kkaadhibaa	athimaa	No	rraadhiyaa	athimaa	No	rraadhiyaa	athimaa	No	ghaadha	athimaa
Yes	haafaathaa	haafaathaa	Yes	haafaathaa	haafaathaa	Yes	haafaathaa	haafaathaa	Yes	haafaathaa	haafaathaa
Yes	kaaythi	kaaythi	No	fyhi	kaaythi	No	aaawidu	kaaythi	Yes	kaaythi	kaaythi
No	thull	haathu	No	ayn	haathu	No	naabu	haathu	No	kaarrnun	haathu
No	waaathun	waaathan	No	waattaan	waaathan	No	waaadun	waaathan	No	waattaan	waaathan
Yes	waaathun	waaathun	Yes	waaathun	waaathun	No	daaaun	waaathun	No	daaaun	waaathun
Yes	waaathin	waaathin	No	waaadin	waaathin	No	daaaan	waaathin	No	daaaan	waaathin
No	faadhin	adhaall	No	faadhin	adhaall	No	faadhin	adhaall	No	ahdu	adhaall
No	sihrr	saaaf	No	ghaath	saaaf	Yes	saaaf	saaaf	No	thaakaaf	saaaf
Yes	attaash	attaash	Yes	attaash	attaash	No	haawaas	attaash	Yes	attaash	attaash
No	baakhaasaa	akks	No	haakkaa	akks	No	ttaaf	akks	No	kaas	akks
No	ottuf	akrr	No	haattaamaa	akrr	No	sum	akrr	Yes	akrr	akrr
No	fyhi	ayn	No	aaamill	ayn	No	haay	ayn	No	kaarrn	ayn

Recording 1			Recording 2			Recording 3			Recording 4		
Match	Recognised words	499 words	Match	Recognised words	499 words	Match	Recognised words	499 words	Match	Recognised words	499 words
Yes	ejjll	ejjll	Yes	ejjll	ejjll	No	sum	ejjll	No	faadhin	ejjll
No	sum	omrr	No	saamghi	omrr	No	aaathin	omrr	No	kaarrnun	omrr
No	sihrr	saaerr	No	saaaf	saaerr	No	saanaaa	saaerr	Yes	saaerr	saaerr
No	naasaab	naaasaa	No	naasaab	naaasaa	No	naasaab	naaasaa	No	daasaa	naaasaa
No	daama	naaomaa	Yes	naaomaa	naaomaa	Yes	naaomaa	naaomaa	No	dhaamaa	naaomaa
Yes	waasia	waasia	Yes	waasia	waasia	Yes	waasia	waasia	Yes	waasia	waasia
No	thulluthin	kaae	No	saaniyaa	kaae	No	zzaakky	kaae	No	saaaf	kaae
No	saahw	saao	No	saahw	saao	No	rrusull	saao	No	saahw	saao
No	waaadin	waarriaan	No	waarriain	waarriaan	No	waaadun	waarriaan	No	waaadun	waarriaan
Yes	waarriaun	waarriaun	Yes	waarriaun	waarriaun	Yes	waarriaun	waarriaun	Yes	waarriaun	waarriaun
Yes	waarriain	waarriain	Yes	waarriain	waarriain	No	waadhaaa	waarriain	No	waaadan	waarriain
Yes	ghaajjaarr	ghaajjaarr	No	ttaajjaan	ghaajjaarr	No	rraaghaad	ghaajjaarr	No	ghaadha	ghaajjaarr
No	taathill	ghaadha	No	naathaarraa	ghaadha	No	rraad	ghaadha	No	daama	ghaadha
No	naashizz	ghaashaa	No	rraashaa	ghaashaa	No	rraashaa	ghaashaa	Yes	ghaashaa	ghaashaa
No	dub	ghaadh	No	rraad	ghaadh	No	naabu	ghaadh	No	naabu	ghaadh
No	waasia	ghaafiyaa	No	waasia	ghaafiyaa	No	ayiyaa	ghaafiyaa	No	sum	ghaafiyaa
No	rraahufaa	ghaarraakaa	No	haarraakkaa	ghaarraakaa	No	rraabaattaa	ghaarraakaa	No	ghaattaa	ghaarraakaa
No	sum	ghaaythu	No	ghaanaamu	ghaaythu	No	aaajjidu	ghaaythu	Yes	ghaaythu	ghaaythu
No	waattaan	ghaattaa	No	faan	ghaattaa	No	khaath	ghaattaa	No	saah	ghaattaa
No	waathaafaa	ghaadaarraa	Yes	ghaadaarraa	ghaadaarraa	No	naadaabaa	ghaadaarraa	No	ghaadha	ghaadaarraa
No	orrsun	ghusn	No	faasun	ghusn	No	naashizz	ghusn	No	waaathan	ghusn
No	dhihni	ghill	No	min	ghill	No	min	ghill	No	min	ghill
No	saawghun	saaghurraa	No	baakhillaa	saaghurraa	No	saawghin	saaghurraa	No	saawghin	saaghurraa
Yes	rraaghaad	rraaghaad	Yes	rraaghaad	rraaghaad	Yes	rraaghaad	rraaghaad	No	naadaabaa	rraaghaad
No	waaathin	ttaaghiyaa	No	rraadhiyaa	ttaaghiyaa	No	rraadhiyaa	ttaaghiyaa	No	maakkunaa	ttaaghiyaa

	Recording 1			Recording 2			Recording 3			Recording 4	
Match	**Recognised words**	**499 words**	**Match**	**Recognised words**	**499 words**	**Match**	**Recognised words**	**499 words**	**Match**	**Recognised words**	**499 words**
No	maarraadhaa	maarraaghaa	No	maarraadhaa	maarraaghaa	No	maakkunaa	maarraaghaa	No	maarraadhaa	maarraaghaa
Yes	saadghu	saadghu	Yes	saadghu	saadghu	No	faadhanaa	saadghu	No	faadhanaa	saadghu
No	saamaakki	saamghi	No	saamtin	saamghi	No	saanaaa	saamghi	No	saamaakki	saamghi
No	saawghun	saawgan	No	thulluthun	saawgan	No	saawghun	saawgan	No	saawghun	saawgan
Yes	saawghun	saawghun	Yes	saawghun	saawghun	No	saamtun	saawghun	No	thaawbun	saawghun
No	saawgan	saawghin	No	thulluthin	saawghin	No	saadaa	saawghin	No	saamtan	saawghin
No	ath	haaf	Yes	haaf	haaf	No	ghaadh	haaf	No	aaathaa	haaf
Yes	waafy	waafy	Yes	waafy	waafy	Yes	waafy	waafy	No	waasaatti	waafy
Yes	maallaaf	maallaaf	Yes	maallaaf	maallaaf	Yes	maallaaf	maallaaf	Yes	maallaaf	maallaaf
Yes	faakkaa	faakkaa	No	tyn	faakkaa	Yes	faakkaa	faakkaa	No	khaath	faakkaa
Yes	faan	faan	Yes	faan	faan	Yes	faan	faan	Yes	faan	faan
Yes	fijjll	fijjll	Yes	fijjll	fijjll	No	tyn	fijjll	No	faadhin	fijjll
No	faan	furrn	No	waattaan	furrn	No	faan	furrn	No	daama	furrn
No	faadhaallaa	faaallaa	No	faadhaallaa	faaallaa	No	ttaahaanaa	faaallaa	No	ttaahaanaa	faaallaa
No	dhaafaarr	rraafaaa	No	rraasaa	rraafaaa	No	rraasaa	rraafaaa	Yes	rraafaaa	rraafaaa
No	sum	daafirraa	No	ghaasaallaa	daafirraa	No	dhaakky	daafirraa	No	naathufaa	daafirraa
Yes	afwu	afwu	Yes	afwu	afwu	Yes	afwu	afwu	No	kaarrdhun	afwu
Yes	shaarraafaa	shaarraafaa	Yes	shaarraafaa	shaarraafaa	Yes	shaarraafaa	shaarraafaa	Yes	shaarraafaa	shaarraafaa
Yes	ttaarraafi	ttaarraafi	Yes	ttaarraafi	ttaarraafi	No	rraathi	ttaarraafi	Yes	ttaarraafi	ttaarraafi
No	saamtan	khaallfu	No	saamtan	khaallfu	No	saamtan	khaallfu	No	ghaaythu	khaallfu
No	allaafin	allaafan	Yes	allaafan	allaafan	No	allaafin	allaafan	Yes	allaafan	allaafan
Yes	allaafun	allaafun	Yes	allaafun	allaafun	No	allaafan	allaafun	No	allaafan	allaafun
Yes	allaafin	allaafin	Yes	allaafin	allaafin	Yes	allaafin	allaafin	Yes	allaafin	allaafin
No	saarrjju	sujjuk	No	sum	sujjuk	No	shibll	sujjuk	Yes	sujjuk	sujjuk
No	taathill	kullw	No	hunuw	kullw	No	sum	kullw	No	hunuw	kullw

Recording 1			Recording 2			Recording 3			Recording 4		
Match	Recognised words	499 words	Match	Recognised words	499 words	Match	Recognised words	499 words	Match	Recognised words	499 words
Yes	daakaakkaa	daakaakkaa	No	faakkaa	daakaakkaa	No	faakkaa	daakaakkaa	No	aaathin	daakaakkaa
No	faadhin	kaallaam	No	kaarrnun	kaallaam	No	kaarrnan	kaallaam	No	kaarrnan	kaallaam
No	taathill	kidrr	No	hibrr	kidrr	No	hibrr	kidrr	No	daagllu	kidrr
No	waaathin	kudaa	No	yin	kudaa	No	faadh	kudaa	No	sum	kudaa
Yes	saakaattaa	saakaattaa	Yes	saakaattaa	saakaattaa	Yes	saakaattaa	saakaattaa	Yes	saakaattaa	saakaattaa
Yes	fukidaa	fukidaa	Yes	fukidaa	fukidaa	Yes	fukidaa	fukidaa	Yes	fukidaa	fukidaa
No	sum	thaakullaa	Yes	thaakullaa	thaakullaa	Yes	thaakullaa	thaakullaa	No	maakkunaa	thaakullaa
Yes	saabaakaa	saabaakaa	Yes	saabaakaa	saabaakaa	Yes	saabaakaa	saabaakaa	No	thaabaattaa	saabaakaa
No	thaabaattaa	abaakaa	No	baadhaakhun	abaakaa	No	rraabaattaa	abaakaa	No	waathaafaa	abaakaa
Yes	ghaasaaku	ghaasaaku	Yes	ghaasaaku	ghaasaaku	No	maasaahaa	ghaasaaku	Yes	ghaasaaku	ghaasaaku
No	baadhaakhun	baarrkan	No	baadhaakhun	baarrkan	No	min	baarrkan	No	ayn	baarrkan
No	lleyn	baarrkun	No	lleyn	baarrkun	No	baadhaakhin	baarrkun	No	sum	baarrkun
No	baadhaakhin	baarrkin	No	baadhaakhin	baarrkin	No	ayn	baarrkin	No	ayn	baarrkin
No	rraakkaallaa	rraakkaadhaa	Yes	rraakkaadhaa	rraakkaadhaa	No	baattaallaa	rraakkaadhaa	No	maakkaathaa	rraakkaadhaa
Yes	jjaarraakkaa	jjaarraakkaa	Yes	jjaarraakkaa	jjaarraakkaa	Yes	jjaarraakkaa	jjaarraakkaa	No	ath	jjaarraakkaa
No	ghill	kkaawa	No	dhaallaa	kkaawa	No	dhaallaa	kkaawa	No	kaarrnun	kkaawa
No	min	kkaahaan	Yes	kkaahaan	kkaahaan	Yes	kkaahaan	kkaahaan	Yes	kkaahaan	kkaahaan
No	tyn	kkaallb	No	khaawy	kkaallb	No	daaaan	kkaallb	No	khaadukk	kkaallb
Yes	kkiys	kkiys	Yes	kkiys	kkiys	Yes	kkiys	kkiys	No	saaaf	kkiys
Yes	kkuwa	kkuwa	No	thullaat	kkuwa	No	thullaat	kkuwa	Yes	kkuwa	kkuwa
Yes	rraakkibaa	rraakkibaa	Yes	rraakkibaa	rraakkibaa	No	fukidaa	rraakkibaa	Yes	rraakkibaa	rraakkibaa
No	rraakkaallaa	rraakkaaa	No	rraakkaallaa	rraakkaaa	No	haakkaa	rraakkaaa	No	faakkaa	rraakkaaa
Yes	maakkunaa	maakkunaa	Yes	maakkunaa	maakkunaa	Yes	maakkunaa	maakkunaa	Yes	maakkunaa	maakkunaa
Yes	haarraakkaa	haarraakkaa	Yes	haarraakkaa	haarraakkaa	Yes	haarraakkaa	haarraakkaa	Yes	haarraakkaa	haarraakkaa
No	sum	birraakku	No	dhiib	birraakku	No	sum	birraakku	No	minhu	birraakku

Recording 1			Recording 2			Recording 3			Recording 4		
Match	Recognised words	499 words	Match	Recognised words	499 words	Match	Recognised words	499 words	Match	Recognised words	499 words
Yes	saamaakki	saamaakki	Yes	saamaakki	saamaakki	Yes	saamaakki	saamaakki	Yes	saamaakki	saamaakki
No	sillkkin	sillkkan	No	saamtan	sillkkan	No	fillizzin	sillkkan	No	saamtan	sillkkan
Yes	sillkkun	sillkkun	No	saamtun	sillkkun	No	saamtun	sillkkun	No	saamtun	sillkkun
Yes	sillkkin	sillkkin	No	saamtan	sillkkin	No	saamtan	sillkkin	No	saamtan	sillkkin
Yes	llaayth	llaayth	Yes	llaayth	llaayth	No	sum	llaayth	Yes	llaayth	llaayth
Yes	lliyn	lliyn	Yes	lliyn	lliyn	No	atyin	lliyn	No	mudun	lliyn
No	daama	llumaat	Yes	llumaat	llumaat	No	daama	llumaat	No	daama	llumaat
No	taathill	olluw	No	taathill	olluw	No	waarriain	olluw	No	hunuw	olluw
No	ghaadaarraa	ghaallaaa	No	maakkaanaa	ghaallaaa	No	dhaanaa	ghaallaaa	No	ghaadaarraa	ghaallaaa
No	dhihni	jjaalliy	No	dhihni	jjaalliy	No	jjaady	jjaalliy	No	dhihni	jjaalliy
No	daaahu	daagllu	No	ghaanaamu	daagllu	No	ghaanaamu	daagllu	No	daaahu	daagllu
No	sum	amaalli	Yes	amaalli	amaalli	No	saanaaa	amaalli	No	saanaami	amaalli
No	faadhin	ttaabllan	No	sum	ttaabllan	No	ewaajjan	ttaabllan	Yes	ttaabllan	ttaabllan
No	min	ttaabllun	No	faadhin	ttaabllun	No	ttaabllan	ttaabllun	No	ewaajjun	ttaabllun
No	faadhin	ttaabllin	Yes	ttaabllin	ttaabllin	No	waaathin	ttaabllin	No	ttaabllan	ttaabllin
No	faan	haam	No	daama	haam	No	kaarrn	haam	No	sum	haam
No	ghill	yaawm	Yes	yaawm	yaawm	No	ghaallaaa	yaawm	Yes	yaawm	yaawm
Yes	maawzz	maawzz	Yes	maawzz	maawzz	Yes	maawzz	maawzz	No	maahwu	maawzz
Yes	min	min	Yes	min	min	Yes	min	min	Yes	min	min
No	kaarrnin	aaamill	No	kaarrnin	aaamill	No	kaarrnin	aaamill	No	kaarrnan	aaamill
Yes	amaallaa	amaallaa	Yes	amaallaa	amaallaa	No	khaamaanaa	amaallaa	Yes	amaallaa	amaallaa
No	min	numuw	Yes	numuw	numuw	Yes	numuw	numuw	Yes	numuw	numuw
No	daama	faahaamaa	Yes	faahaamaa	faahaamaa	No	thaawbanaa	faahaamaa	Yes	faahaamaa	faahaamaa
No	min	ghaanaamu	Yes	ghaanaamu	ghaanaamu	Yes	ghaanaamu	ghaanaamu	Yes	ghaanaamu	ghaanaamu

	Recording 1			Recording 2			Recording 3			Recording 4	
Match	Recognised words	499 words	Match	Recognised words	499 words	Match	Recognised words	499 words	Match	Recognised words	499 words
Yes	saanaami	saanaami	No	thiny	saanaami	No	saanaaa	saanaami	Yes	saanaami	saanaami
Yes	allaaman	allaaman	Yes	allaaman	allaaman	Yes	allaaman	allaaman	Yes	allaaman	allaaman
No	Aallamon	allaamun	No	Aallamon	allaamun	No	Aallamon	allaamun	No	Aallamon	allaamun
No	Aallaman	allaamin	No	Aallaman	allaamin	No	kaadhaaa	allaamin	No	kaadhaaa	allaamin
Yes	waanaa	waanaa	Yes	waanaa	waanaa	Yes	waanaa	waanaa	Yes	waanaa	waanaa
No	naahaataa	naahrr	No	naahaataa	naahrr	Yes	naahrr	naahrr	No	daama	naahrr
Yes	nibrr	nibrr	Yes	nibrr	nibrr	Yes	nibrr	nibrr	Yes	nibrr	nibrr
No	min	nuwrr	No	min	nuwrr	No	minhu	nuwrr	No	min	nuwrr
Yes	faanaarr	faanaarr	Yes	faanaarr	faanaarr	No	thaanaat	faanaarr	No	ghaanaamu	faanaarr
Yes	saaniyaa	saaniyaa	Yes	saaniyaa	saaniyaa	Yes	saaniyaa	saaniyaa	No	saanaaa	saaniyaa
No	aaamill	hunuw	No	aaamill	hunuw	No	kaarrnin	hunuw	No	aaamill	hunuw
Yes	maakkaanaa	maakkaanaa	Yes	maakkaanaa	maakkaanaa	Yes	maakkaanaa	maakkaanaa	No	kaallaam	maakkaanaa
No	maahwu	naahnu	No	maahwu	naahnu	No	maahwu	naahnu	No	ghill	naahnu
No	sum	dhihni	No	diykk	dhihni	No	ayn	dhihni	No	lliyn	dhihni
No	kaarrnin	kaarrnan	No	kaarrnun	kaarrnan	No	kaarrnin	kaarrnan	Yes	kaarrnan	kaarrnan
Yes	kaarrnun	kaarrnun	No	fadhon	kaarrnun	Yes	kaarrnun	kaarrnun	Yes	kaarrnun	kaarrnun
Yes	kaarrnin	kaarrnin	No	kaarrnan	kaarrnin	Yes	kaarrnin	kaarrnin	No	kaarrnan	kaarrnin
No	daama	ghaarraahu	No	ghaanaamu	ghaarraahu	No	maahwu	ghaarraahu	Yes	ghaarraahu	ghaarraahu
No	ghaanaamu	thaallaahu	No	faAalla	thaallaahu	No	ghaanaamu	thaallaahu	No	faaallaa	thaallaahu
No	hibrr	hirr	No	daahrrun	hirr	No	hibrr	hirr	No	yin	hirr
Yes	haawaas	haawaas	No	maallaaf	haawaas	No	maallaaf	haawaas	Yes	haawaas	haawaas
No	sum	huwid	No	witrr	huwid	No	fukidaa	huwid	No	sum	huwid
Yes	rraahibaa	rraahibaa	Yes	rraahibaa	rraahibaa	Yes	rraahibaa	rraahibaa	Yes	rraahibaa	rraahibaa
Yes	rraahufaa	rraahufaa	Yes	rraahufaa	rraahufaa	Yes	rraahufaa	rraahufaa	Yes	rraahufaa	rraahufaa

Recording 1			Recording 2			Recording 3			Recording 4		
Match	Recognised words	499 words	Match	Recognised words	499 words	Match	Recognised words	499 words	Match	Recognised words	499 words
No	kkaawa	kaahaarraa	Yes	kaahaarraa	kaahaarraa	No	ghaattaa	kaahaarraa	Yes	kaahaarraa	kaahaarraa
No	dhaallaa	naawaahaa	Yes	naawaahaa	naawaahaa	Yes	naawaahaa	naawaahaa	Yes	naawaahaa	naawaahaa
Yes	minhu	minhu	Yes	minhu	minhu	Yes	minhu	minhu	Yes	minhu	minhu
No	lliyn	fyhi	No	lliyn	fyhi	No	kkiys	fyhi	No	kaaythi	fyhi
No	jjaahun	jjaahan	No	jjaahun	jjaahan	No	dhaaa	jjaahan	Yes	jjaahan	jjaahan
No	jjaahon	jjaahun	No	jjaahon	jjaahun	No	daahrrun	jjaahun	Yes	jjaahun	jjaahun
Yes	jjaahin	jjaahin	No	jjaahan	jjaahin	No	daaaan	jjaahin	No	jjaahan	jjaahin
Yes	witrr	witrr	Yes	witrr	witrr	Yes	witrr	witrr	No	waaathan	witrr
No	waajjibaa	wujjidaa	No	waajjibaa	wujjidaa	Yes	wujjidaa	wujjidaa	No	waajjibaa	wujjidaa
No	sum	aaawidu	No	ayn	aaawidu	No	sum	aaawidu	No	allaafan	aaawidu
No	ghill	dhaawuw	No	ghill	dhaawuw	No	ghill	dhaawuw	Yes	dhaawuw	dhaawuw
Yes	maahwu	maahwu	Yes	maahwu	maahwu	Yes	maahwu	maahwu	Yes	maahwu	maahwu
Yes	llaahwi	llaahwi	Yes	llaahwi	llaahwi	Yes	llaahwi	llaahwi	Yes	llaahwi	llaahwi
Yes	saahwaa	saahwaa	Yes	saahwaa	saahwaa	Yes	saahwaa	saahwaa	Yes	saahwaa	saahwaa
No	jjaarrwun	jjaarrwan	No	jjaarrwun	jjaarrwan	No	jjaarrwun	jjaarrwan	No	daahrrun	jjaarrwan
No	jjaarrwin	jjaarrwun	No	daaaon	jjaarrwun	No	ttaabllan	jjaarrwun	No	daaaun	jjaarrwun
No	jjaalliy	jjaarrwin	No	jjaalliy	jjaarrwin	No	ghaallaaa	jjaarrwin	Yes	jjaarrwin	jjaarrwin
No	yin	yaad	Yes	yaad	yaad	No	ghaadh	yaad	Yes	yaad	yaad
Yes	yusrr	yusrr	Yes	yusrr	yusrr	Yes	yusrr	yusrr	Yes	yusrr	yusrr
Yes	yin	yin	Yes	yin	yin	No	daaaun	yin	No	daaaan	yin
Yes	saayaarraa	saayaarraa	Yes	saayaarraa	saayaarraa	No	saaniyaa	saayaarraa	No	saaniyaa	saayaarraa
Yes	ayiyaa	ayiyaa	No	sum	ayiyaa	No	haayuUa	ayiyaa	Yes	ayiyaa	ayiyaa
Yes	saawyi	saawyi	Yes	saawyi	saawyi	No	saanaami	saawyi	Yes	saawyi	saawyi
Yes	ttaayu	ttaayu	Yes	ttaayu	ttaayu	No	aaaw	ttaayu	Yes	ttaayu	ttaayu
Yes	haayuUa	haayuUa	Yes	haayuUa	haayuUa	Yes	haayuUa	haayuUa	No	aaathaa	haayuUa

Recording 1			Recording 2			Recording 3			Recording 4		
Match	**Recognised words**	**499 words**	**Match**	**Recognised words**	**499 words**	**Match**	**Recognised words**	**499 words**	**Match**	**Recognised words**	**499 words**
No	aaathin	atyan	No	atumaa	atyan	No	rraadhiyaa	atyan	No	kaarrnan	atyan
No	Aatyan	atyun	No	Aatyan	atyun	Yes	atyun	atyun	No	mudun	atyun
No	dhaakky	atyin	No	dhaakky	atyin	No	rraadhiyaa	atyin	No	allaafin	atyin
195	No. Recognised words		174	No. Recognised words		144	No. Recognised words		172	No. Recognised words	
39.078	Average		34.869	Average		28.857	Average		34.468	Average	
Average all = 34.318											

Table 1- Improved SLT table analysis for the four recordings.

Appendix M

Letter/Diacritic pair analysis

Letter/Diacritic pair analysis

Arabic letter	Name of letter	Alghamdi English letter	Overall		Fat ha 37.3%		Dhamma 33.4%		Kasra 34.5%	
			Total Number of words	Recog. rate (%)	Total Number of words	Recog. rate (%)	Total Number of words	Recog. rate (%)	Total Number of words	Recog. rate (%)
ش	sheen	sh	34	57.4%	17	58.4%	4	31.3%	5	90%
ن	noon	n	61	56.1%	24	56.3%	8	57.3%	3	75%
ز	zain	zz	28	52.7%	16	54.4%	3	75%	3	25%
ف	faa	f	57	44.3%	28	39.2%	4	50%	8	47.5%
ح	haa	h	43	44.2%	22	48.9%	7	44%	3	50%
س	seen	s	64	43.8%	32	52.5%	8	38.5%	9	25%
ط	ta	tt	36	43.1%	24	38.4%	3	8.3%	3	25%

Arabic letter	Name of letter	Alghamdi English letter	Overall		Fat ha 37.3%		Dhamma 33.4%		Kasra 34.5%	
			Total Number of words	Recog. rate (%)	Total Number of words	Recog. rate (%)	Total Number of words	Recog. rate (%)	Total Number of words	Recog. rate (%)
ك	kaaf	kk	43	40.7%	27	39.1%	3	58.3%	3	66.7%
م	meem	m	62	40.3%	20	44.9%	5	61.7%	5	60%
ب	baa	b	71	40.1%	38	43.4%	4	56.3%	4	12.5%
ه	haa	h	45	38.9%	16	31.3%	9	33.3%	3	25%
ي	yaa	y	50	36%	13	45.8%	4	62.5%	3	41.7%
و	waaw	w	71	35.9%	28	52.8%	4	43.8%	4	37.5%
ج	jeem	jj	51	34.8%	28	36%	9	20.8%	5	50%

Arabic letter	Name of letter	Alghamdi English letter	Overall		Fat ha 37.3%		Dhamma 33.4%		Kasra 34.5%	
			Total Number of words	Recog. rate (%)	Total Number of words	Recog. rate (%)	Total Number of words	Recog. rate (%)	Total Number of words	Recog. rate (%)
ص	saad	s	39	33.3%	23	32.2%	5	33.3%	3	33.3%
د	daal	d	58	30.9%	26	30.3%	10	47.5%	3	83.3%
ق	qaaf	k	51	29.4%	31	33.8%	5	25%	4	31.3%
ت	taa	t	31	29%	7	7.1%	3	8.3%	3	25%
ر	raa	rr	126	28.4%	58	42.8%	6	35.4%	6	50%
خ	khaa	k	28	27.7%	16	28.6%	3	16.7%	3	0%

Arabic letter	Name of letter	Alghamdi English letter	Overall		Fat ha 37.3%		Dhamma 33.4%		Kasra 34.5%	
			Total Number of words	Recog. rate (%)	Total Number of words	Recog. rate (%)	Total Number of words	Recog. rate (%)	Total Number of words	Recog. rate (%)
ل	laam	ll	77	27.3%	31	49.2%	8	28.1%	6	45.8%
ذ	thaal	dh	33	25.8%	16	26.6%	3	41.7%	4	0%
ث	thaa	th	34	25%	16	28.1%	7	3.6%	3	16.7%
ظ	tha	th	22	25%	22	27.8%	5	5%	5	5%
ع	ain	a	75	24%	44	25.4%	10	7.5%	6	20.8%
أ	alef	a	19	18.4%	12	27.1%	3	0%	3	8.3%

Arabic letter	Name of letter	Alghamdi English letter	Overall		Fat ha 37.3%		Dhamma 33.4%		Kasra 34.5%	
			Total Number of words	Recog. rate (%)	Total Number of words	Recog. rate (%)	Total Number of words	Recog. rate (%)	Total Number of words	Recog. rate (%)
غ	ghain	g	34	18.4%	21	26.4%	3	16.7%	3	0%
ض	dhad	dh	30	14.2%	16	18.4%	4	25%	5	15%

Table 1-Letter/Diacritic pair analysis

Appendix N

Further analysis of the odd pairs cases

Alef dhamma analysis

Misrecognised as	No. of recog. words	dhamma		Third letter		Second letter		First letter		
		R (%)	L	R (%)	L	R (%)	L	R (%)	L	
thullth + faadh + ghill + faazzaa	0 out of 4	33.4	u	29	taa ت	27.7	khaa خ	18.4	alef أ	أُخت aukht
badhakon + daama + dhaaba	0 out of 4	33.4	u	18.4	alef أ	28.4	raa ر	40.1	baa ب	بَرَأ baarraau
maahdi + diykk + atyin	0 out of 4	33.4	u	25.8	dhaal ذ	18.4	alef ؤ	40.3	meem م	مؤذِ mUdhi

Table 1- Alef dhamma analysis

Baa Kasra Analysis

Misrecognised as	No. of recog. words	kasra R (%)	L	Third letter R (%)	L	Second letter R (%)	L	First letter R (%)	L		
dhirrs + rryshi + ghaashaa + naashizz	0 out of 4	34.5	i	28.4	raa ر	57.4	ش sheen	40.1	baa ب	بِشر	bishrr
sum + rraadhiyaa	2 out of 4	34.5	i	27.3	laam ل	40.1	baa ب	34.8	jeem ج	جُبِلَ	jjubillaa
thiny + kaaythi + kkaanzzi	0 out of 4	34.5	i	40.1	baa ب	27.3	laam ل	29.4	qaaf ق	قَلبِ	kaallbi
dhiib + sum + minhu	0 out of 4	34.5	i	40.7	kaaf ك	28.4	raa ر	40.1	baa ب	بِرَكُ	birraakku

Table 2- Baa Kasra analysis

Taa 'fat ha' Analysis

Misrecognised as	No. of recog. words	fat ha		Third	letter	Second	letter	First	letter		
		R (%)	L	R (%)	L	R (%)	L	R (%)	L		
ttaabllin + sum + rraakkaaa + khaasaafaa	0 out of 4	37.3	aa	29	taa ت	40.1	baa ب	25	thaa ث	thaabaataa	ثَبَتَ
kaash + thaakhn + sum + saahw	0 out of 4	37.3	aa	29	taa ت	44.2	haa ح	29	taa ت	taaht	تَحت
Ttaabllin + thaawbun	0 out of 4	37.3	aa	27.3	laam ل	25	tha ظ	29	taa ت	taathill	تَظِل
dhaakky + rraakkaaa + sum	0 out of 4	37.3	aa	36	yaa ي	29.4	qaaf ق	29	taa ت	taaky	تَقي
saamtan + thull + dhaamaat + saamghi	0 out of 4	37.3	aa	28.4	raa ر	40.3	meem م	29	taa ت	taamrr	تَمر
maakkaanaa + haattaamaa + aaadhunaa	0 out of 4	37.3	aa	27.3	laam ل	29	taa ت	29.4	qaaf ق	kaataallaa	قَتَلَ
saamtan + maakkaathaa	2 out of 4	37.3	aa	29	taa ت	44.2	haa ح	56.1	noon ن	naahaataa	نَحَتَ

Table 3- Taa 'fat ha' analysis

Taa dhamma

Misrecognised as	No. of recog. words	dhamma		Third letter		Second letter		First letter			
		R (%)	L	R (%)	L	R (%)	L	R (%)	L		
thullaat + dhuUll + sum + ttaaf	0 out of 4	33.4	u	29	taa ت	35.9	waaw و	29	taa ت	twt	ثُوت
haattaamaa + aaadhunaa + sum	1 out of 4	33.4	u	40.3	meem م	29	taa ت	24	ain ع	atumaa	عَثُمَ
sum + sillkkan + thull	0 out of 4	33.4	u	29	taa ت	40.3	meem م	36	yaa ي	yumitu	يُمِتُ

Table 4- Taa dhamma analysis

Thaa dhamma analysis

Misrecognised as	No. of recog. words	dhamma		Third letter		Second letter		First letter			
		R (%)	L	R (%)	L	R (%)	L	R (%)	L		
thaarrf + saaaf + daaf + dhaanaa	0 out of 4	33.4	u	25	thaa ث	27.3	laam ل	25	thaa ث	thullth	ثُلث
faaallaa + waanaa	0 out of 4	33.4	u	29	taa ة	27.3	laam ل	25	thaa ث	thullaat	ثُلَّة
sum + taathill + dhaakhaahu	0 out of 4	33.4	u	35.9	waaw و	25	thaa ث	34.8	jeem ج	jjuthw	جُثْو
ghaaythu + baasaall	2 out of 4	33.4	u	-	-	25	thaa ث	40.1	baa ب	baathu	بَثُ
kaarrdhun + fillizzin + fillizzan + saamtan	0 out of 4	33.4	u	25	thaa ث	27.3	laam ل	25	thaa ث	thulluthin	ثُلُث
fillizzan + thulluthin + fillizzin + fillizzun	0 out of 4	33.4	u	25	thaa ث	27.3	laam ل	25	thaa ث	thulluthun	ثُلُث
haasaan + fillizzin + fillizzan + saamtan	0 out of 4	33.4	u	25	thaa ث	27.3	laam ل	25	thaa ث	thulluthan	ثُلُث

Table 5- Thaa dhamma analysis

Khaa kasra analysis

Misrecognised as	No. of recog. words	kasra		Third letter		Second letter		First letter			
		R (%)	L	R (%)	L	R (%)	L	R (%)	L		
hibrr + faadhanaa	0 out of 4	34.5	i	28.4	raa ر	30.9	daal د	27.7	khaa خ	khidrr	خِدر
naasiyaa + naathufaa + baashimaa	1 out of 4	34.5	i	27.3	laam ل	27.7	khaa خ	40.1	baa ب	baakhillaa	بَخِلَ
sum + dhaakky + rraathi + mUdhi	0 out of 4	34.5	i	-	-	27.7	khaa خ	40.3	meem م	mukhi	مُخ

Table 6- Khaa kasra analysis

Thaal kasra analysis

Misrecognised as	No. of recog. words	kasra		Third letter		Second letter		First letter			
		R (%)	L	R (%)	L	R (%)	L	R (%)	L		
min + dhid	0 out of 4	34.5	i	40.1	baa ب	18.4	alef ئ	25.8	thaal ذ	dhiib	ذِئب
sum + rraahibaa + naadaabaa	1 out of 4	34.5	i	40.1	baa ب	25.8	ذ thaal	40.7	kaaf ك	kkaadhibaa	كَذِبَ
maahdi + diykk + atyin	0 out of 4	34.5	i	25.8	thaal ذ	18.4	alef ؤ	40.3	meem م	mUdhi	مؤذِ
diykk + sum + ayn + lliyn	0 out of 4	34.5	i	56.1	ن noon	38.9	haa ه	25.8	thaal ذ	dhihni	ذِهن

Table 7- Thaal kasra analysis

Ta dhamma analysis

Misrecognised as	n of recog. words	dhamma		Third letter		Second letter		First letter			
		R (%)	L	R (%)	L	R (%)	L	R (%)	L		
faadh + ghaat	0 out of 4	33.4	u	-	-	29.4	qaaf ق	43.1	ta ط	ttaak	طُق
sum + dhaaghaath + thaakaaf	1 out of 4	33.4	u	44.3	faa ف	43.1	ta ط	24	ain ع	ottuf	عُطُف
minhu + taathill	0 out of 4	33.4	u	43.1	ta ط	28.4	raa ر	29.4	qaaf ق	kirrttu	قِرطُ

Table 8- Ta dhamma analysis

Tha dhamma analysis

Misrecognised as	n of recog. words	dhamma		Third letter		Second letter		First letter			
		R (%)	L	R (%)	L	R (%)	L	R (%)	L		
naabu + daama	0 out of 4	33.4	u	-	-	25	ظ tha	40.7	ك kaaf	kkaathu	كَظْ
daaaun + daaaan	0 out of 4	33.4	u	-	-	27.3	ل laam	25	ظ tha	thull	ظْل
naabu	2 out of 4	33.4	u	44.3	ف faa	25	ظ tha	56.1	ن noon	naathufaa	نَظْفَ
ayn + thull + naabu + kaarrnun	0 out of 4	33.4	u	-	-	25	ظ tha	44.2	ح haa	haathu	حَظْ
ghaanaamu + sum + aaajjidu	1 out of 4	33.4	u	25	ظ tha	36	ي yaa	18.4	غ ghain	ghaaythu	غَيظْ

Table 9- Tha dhamma anaylsis

Tha kasra analysis

Misrecognised as	n of recog. words	kasra R (%)	L	Third letter R (%)	L	Second letter R (%)	L	First letter R (%)	L		
ttaabllin + kkaathu	0 out of 4	34.5	i	56.1	ن noon	25	ظ tha	18.4	أ alef	aaathin	أظِن
ttaabllin + thaawbun	0 out of 4	34.5	i	27.3	ل laam	25	ظ tha	29	ت taa	taathill	تَظِل
daasaa + daaf + daafirraa	0 out of 4	34.5	i	28.4	ر raa	44.3	ف faa	25	ظ tha	thifrr	ظِفر
rraadhiyaa + kkaadhibaa + ghaadha	0 out of 4	34.5	i	40.3	م meem	25	ظ tha	24	ع ain	athimaa	عَظِمَ
fyhi + aaawidu	2 out of 4	34.5	i	25	ظ tha	36	ي yaa	29.4	ق qaaf	kaaythi	قَيظِ

Table 10- Tha kasra anaysis

Ain dhamma analysis

Misrecognised as	n of recog. words	dhamma		Third	letter	Second	letter	First	letter		
		R (%)	L	R (%)	L	R (%)	L	R (%)	L		
taathill + saamtan	0 out of 4	33.4	u	35.9	waaw و	29	taa ت	24	ain ع	otw	عُتو
fillizzin + waasia +sum	0 out of 4	33.4	u	28.4	raa ر	25	thaa ث	24	ain ع	othirraa	عُثِرَ
faasin + fillizzin + fillizzan	0 out of 4	33.4	u	43.8	seen س	28.4	raa ر	24	ain ع	orrsan	عُرسًا
haasaan + rrusull	0 out of 4	33.4	u	43.8	seen س	28.4	raa ر	24	ain ع	orrsun	عُرسِ
faasin + fillizzin	0 out of 4	33.4	u	43.8	seen س	28.4	raa ر	24	ain ع	orrsin	عُرسٌ
sum + dhaaghaath + thaakaaf	1 out of 4	33.4	u	44.3	faa ف	43.1	ta ط	24	ain ع	ottuf	عُطف
saamghi + sum + aaathin + kaarrnun	0 out of 4	33.4	u	28.4	raa ر	40.3	meem مُ	24	ain ع	omrr	عُمر
daama + dhaamaa	2 out of 4	33.4	u	40.3	meem مُ	24	ain ع	56.1	noon ن	naaomaa	نَعْمَ
saahw + rrusull	0 out of 4	33.4	u	24	ain ع	18.4	alef ا	33.3	saad ص	saao	صاعُ
taathill + waarriain + hunuw	0 out of 4	33.4	u	35.9	waaw و	27.3	laam ل	24	ain ع	olluw	عُلو

Table 11- Ain dhamma anaylsis

Ghain kasra analysis

Misrecognised as	n of recog. words	kasra R (%)	L	Third letter R (%)	L	Second letter R (%)	L	First letter R (%)	L		
min + dhihni	0 out of 4	34.5	i	-	-	27.3	laam ل	18.4	ghain غ	ghill	غِل
rraadhiyaa + waaathin + maakkunaa	0 out of 4	34.5	i	36	yaa ي	18.4	ghain غ	43.1	ta ط	ttaaghiyaa	طَغِي
saamtin + saamaakki + saanaaa + saamaakki	0 out of 4	34.5	i	18.4	غ ghain	40.3	meem م	33.3	ص saad	saamghi	صَمغ

Table 12- Ghain kasra analysis

Appendix O

Problematic letter/diacritic pair alternatives

Ara. l	Name of letter	diacritic	T	Table 1 English letter with diacritic		Alternatives															Table 2 Eng. l with diacritic
						1		2		3		4		5		6		7			
				Eng. L	rr%	Eng. l	rr%	Eng. l	rr%	Eng. l	rr%	Eng. l	rr%	Eng. l	rr%	Eng. l	rr%	Eng. l	rr%		
أ	alef	dhamma	3	au	0	aaou	0	ao	0	aoo	0	aou	0	aau	0	aao	0	aaoo	0	au
ب	baa	kasra	4	Bi	12.5	bbee	0	be	25	bie	0	bee	6.3	bbi	0	bbe	6.3	bbie	6.3	be
ت	taa	fat ha	7	taa	3.6	ttaa	0	ta	7.1	tta	3.6	-	-	-	-	-	-	-	-	ta
ت	taa	dhamma	3	Tu	8.3	ttou	0	to	0	too	0	tou	0	ttu	0	tto	0	ttoo	0	tu
ث	thaa	dhamma	7	thu	7.1	tthou	0	tho	10.7	thoo	0	thou	3.6	tthu	0	ttho	3.6	ttho	3.6	Tho
خ	khaa	kasra	3	Ki	8.3	kee	0	khi	25	khe	0	khie	16.7	khee	0	ke	8.3	kie	0	Khi
ذ	thaal	kasra	4	Dhi	6.3	dhee	0	thi	6.3	the	6.3	thie	0	thee	6.3	dhe	6.3	dhie	0	dhi

Ara. l	Name of letter	diacritic	T	Eng. L	rr%	1 Eng. l	rr %	2 Eng. l	rr %	3 Eng. l	rr %	4 Eng. l	rr %	5 Eng. l	rr %	6 Eng. l	rr %	7 Eng. l	rr %	Table 2 Eng. l with diacritic
ط	ta	dhamma	3	Ttu	8.3	ttou	0	tu	16.7	to	8.3	too	0	tou	0	tto	8.3	ttoo	8.3	tu
ظ	tha	dhamma	5	Thu	15	tthou	0	tho	20	thoo	0	thou	5	tthu	5	ttho	5	tthoo	0	Tho
ظ	tha	kasra	5	Thi	10	tthee	0	the	15	thie	0	thee	5	tthi	0	tthe	5	tthie	0	the
ع	ain	dhamma	10	O	7.5	aoo	0	oo	5	oou	2.5	ooo	0	ou	0	ao	2.5	aou	5	o
غ	ghain	Kasra	3	Gi	0	gee	0	ghi	16.7	ghe	8.3	ghie	8.3	ghee	8.3	ge	8.3	gie	8.3	ghi

Table 1 Problematic letter/diaciritc pair alternatives

* **Ara. L:** Arabic letter   ***T:** Total   ***Eng. L:** English letter   ***rr %** : recognition rate %

Appendix P

Improved LDPT recognition analysis

Improved LDPT recognition analysis

Recording 4			Recording 3			Recording 2			Recording 1		
499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match
dhaaa	tyn	No	dhaaa	tyn	No	dhaaa	tyn	No	dhaaa	tyn	No
aaathen	khaadukk	No	aaathen	ttaabllin	No	aaathen	ttaabllin	No	aaathen	ttaabllin	No
saaghaa	saawghun	No	saaghaa	saah	No	saaghaa	saah	No	saaghaa	faakkaa	No
edhaa	eewaajjan	No	edhaa	yin	No	edhaa	tathell	No	edhaa	sum	No
zzaarr	zzaafaa	No	zzaarr	zzaaamaa	No	zzaarr	zzaaamaa	No	zzaarr	saah	No
kaas	faazzaa	No	kaas	shaams	No	kaas	kaas	Yes	kaas	haaf	No
aaamaall	amaallaa	No	aaamaall	kaarrnin	No	aaamaall	kaarrnin	No	aaamaall	amaalli	No
jjaatha	daasaa	No	jjaatha	dhaaba	No	jjaatha	jjaatha	Yes	jjaatha	jjaatha	Yes
shaah	shaahaadhaa	No	shaah	shaah	Yes	shaah	shugll	No	shaah	shaarraath	No
ttaaf	zzaafaa	No	ttaaf	ghaat	No	ttaaf	thaarrf	No	ttaaf	haaf	No
haayaaaaa	ath	No	haayaaaaa	haayuUa	No	haayaaaaa	ayiyaa	No	haayaaaaa	saayaarraa	No
kkaaasu	khaas	No	kkaaasu	ghaat	No	kkaaasu	khaas	No	kkaaasu	khaath	No
aukht	faazzaa	No	aukht	ghill	No	aukht	faadh	No	aukht	thollth	No
baadaa	faadhanaa	No	baadaa	ghaadha	No	baadaa	baadaa	Yes	baadaa	dhaaghaath	No
aaaw	kaarrn	No	aaaw	kkaatho	No	aaaw	ghill	No	aaaw	sum	No
aaakkaallaa	maakkaanaa	No	aaakkaallaa	maakkaanaa	No	aaakkaallaa	haattaallaa	No	aaakkaallaa	aaakkaallaa	Yes
saaaaall	thaakullaa	No	saaaaall	thaawbanaa	No	saaaaall	saakaattaa	No	saaaaall	saakaattaa	No
dhuUll	daahrrun	No	dhuUll	daaaun	No	dhuUll	daaaun	No	dhuUll	sum	No
baaiisaa	daasaa	No	baaiisaa	ghaaytho	No	baaiisaa	ghaaytho	No	baaiisaa	sum	No
baarraau	dhaaba	No	baarraau	baadaa	No	baarraau	daama	No	baarraau	ghaadha	No

Recording 4			Recording 3			Recording 2			Recording 1		
499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match
swai	suhub	No	swai	saawyi	No	swai	sillkkin	No	swai	saawyi	No
daaaan	daaaan	Yes	daaaan	daahrran	No	daaaan	jjaahan	No	daaaan	daaaun	No
daaaun	sum	No	daaaun	daaaun	Yes	daaaun	jjaahin	No	daaaun	daaaun	Yes
daaain	daaaan	No	daaain	daaaan	No	daaain	daaaan	No	daaain	daaaan	No
thaaby	sum	No	thaaby	maahdi	No	thaaby	waafy	No	thaaby	mUdhi	No
dhaaba	dhaaba	Yes	dhaaba	dhaab	No	dhaaba	dub	No	dhaaba	naabu	No
baazzaaghaa	baaiisaa	No	baazzaaghaa	baaiisaa	No	baazzaaghaa	baaiisaa	No	baazzaaghaa	sum	No
baasaall	baasaall	Yes	baasaall	baasaattaa	No	baasaall	baasaall	Yes	baasaall	naasaab	No
baahaak	naahaata	No	baahaak	sum	No	baahaak	baahaak	Yes	baahaak	sum	No
khaabaatt	haabaasaa	No	khaabaatt	kaadhaaa	No	khaabaatt	khaabaatt	Yes	khaabaatt	khaabaatt	Yes
kkaabaa	dhaaba	No	kkaabaa	kkaabaa	Yes	kkaabaa	ghaadh	No	kkaabaa	kathu	No
dhaanb	dhaanb	Yes	dhaanb	dhaanb	Yes	dhaanb	daama	No	dhaanb	zzaand	No
baashimaa	baashimaa	Yes	baashimaa	baashimaa	Yes	baashimaa	baashimaa	Yes	baashimaa	sum	No
saabaa	saah	No	saabaa	saabaa	Yes	saabaa	subull	No	saabaa	faadhun	No
faarraabu	ghill	No	faarraabu	faarraabu	Yes	faarraabu	faarraabu	Yes	faarraabu	faarraabu	Yes
naasaab	naasaab	Yes	naasaab	naasaab	Yes	naasaab	naasaab	Yes	naasaab	naasaab	Yes
waajjibaa	waajjibaa	Yes	waajjibaa	waajjibaa	Yes	waajjibaa	waajjibaa	Yes	waajjibaa	waajjibaa	Yes
thaabaata	khaasaafaa	No	thaabaata	thaabaattaa	No	thaabaata	thaabaata	Yes	thaabaata	thaabaata	Yes
baattaallaa	ghaasaallaa	No	baattaallaa	maakkaanaa	No	baattaallaa	baattaallaa	Yes	baattaallaa	baattaallaa	Yes
beshrr	beshrr	Yes	beshrr	ghaashaa	No	beshrr	rryshi	No	beshrr	dhirrs	No
burrjj	burrjj	Yes	burrjj	daaaun	No	burrjj	min	No	burrjj	burrjj	Yes

Recording 4			Recording 3			Recording 2			Recording 1		
499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match
jjubellaa	sum	No	jjubellaa	ghaattaa	No	jjubellaa	jjubellaa	Yes	jjubellaa	sum	No
rraabaattaa	thaabaata	No	rraabaattaa	rraabaattaa	Yes	rraabaattaa	rraabaattaa	Yes	rraabaattaa	rraabaattaa	Yes
subull	sillkkun	No	subull	saamtun	No	subull	subull	Yes	subull	sillkkun	No
haallaabaa	haallaabaa	Yes	haallaabaa	haallaabaa	Yes	haallaabaa	haallaabaa	Yes	haallaabaa	haallaabaa	Yes
kaallbe	kkaanzzi	No	kaallbe	sum	No	kaallbe	sum	No	kaallbe	thiny	No
naabu	naabu	Yes	naabu	naabu	Yes	naabu	naabu	Yes	naabu	naabu	Yes
thaawbanaa	thaawbun	No	thaawbanaa	ttaabllin	No	thaawbanaa	tholluthun	No	thaawbanaa	tholluthun	No
thaawbun	kaarrnun	No	thaawbun	ttaabllan	No	thaawbun	thaawbin	No	thaawbun	thaawbin	No
thaawbin	thaawbin	Yes	thaawbin	ttaabllin	No	thaawbin	tholluthin	No	thaawbin	allaafin	No
taht	saahw	No	taht	sum	No	taht	thaakhn	No	taht	saahw	No
dhaamaat	dhaamaat	Yes	dhaamaat	dhaamaa	No	dhaamaat	dhaamaat	Yes	dhaamaat	dhaamaat	Yes
tathell	thaawbun	No	tathell	ttaabllin	No	tathell	ttaabllin	No	tathell	ttaabllin	No
saattaat	saakkaat	No	saattaat	saakkaat	No	saattaat	saattaat	Yes	saattaat	faakkaa	No
saakkaat	khaath	No	saakkaat	saattaat	No	saakkaat	sum	No	saakkaat	sum	No
dhaarraat	dhaamaa	No	dhaarraat	dhaarraat	Yes	dhaarraat	dhaarraat	Yes	dhaarraat	dhaarraat	Yes
haazzaat	faazzaa	No	haazzaat	haazzaat	Yes	haazzaat	haazzaat	Yes	haazzaat	haazzaat	Yes
shaadaat	shaathaaf	No	shaadaat	shaadaat	Yes	shaadaat	shaajjaa	No	shaadaat	shaajjaa	No
thaanaat	ghaadha	No	thaanaat	thaanaat	Yes	thaanaat	thaanaat	Yes	thaanaat	thaanaat	Yes
jjaafaat	dhaafaarr	No	jjaafaat	jjaafaat	Yes	jjaafaat	akks	No	jjaafaat	jjaafaat	Yes
otw	tathell	No	otw	saamtan	No	otw	baatho	No	otw	adhaall	No
ghaat	faanaarr	No	ghaat	ghaat	Yes	ghaat	sum	No	ghaat	sum	No
taky	sum	No	taky	rraakkaaa	No	taky	dhaakky	No	taky	dhaakky	No

Recording 4			Recording 3			Recording 2			Recording 1		
499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match
tamrr	saamghi	No	tamrr	dhaamaat	No	tamrr	daama	No	tamrr	saamtan	No
tyn	kaarrn	No	tyn	kkiys	No	tyn	tyn	Yes	tyn	lliyn	No
twt	taht	No	twt	sum	No	twt	dhuUll	No	twt	kaarrn	No
kaatallaa	maakkunaa	No	kaatallaa	haattaamaa	No	kaatallaa	haattaamaa	No	kaatallaa	maakkaanaa	No
sutirraa	sutirraa	Yes	sutirraa	sutirraa	Yes	sutirraa	sillkkin	No	sutirraa	sutirraa	Yes
atumaa	aaadhunaa	No	atumaa	aaadhunaa	No	atumaa	haattaamaa	No	atumaa	sum	No
yumitu	tholl	No	yumitu	sum	No	yumitu	sum	No	yumitu	mudun	No
yaakhti	kaaythe	No	yaakhti	dhaakky	No	yaakhti	fijjll	No	yaakhti	dhaakky	No
naahaata	naahaata	Yes	naahaata	naahaata	Yes	naahaata	naahaata	Yes	naahaata	jjaahaathaa	No
saamtun	saamtun	Yes	saamtun	saamtan	No	saamtun	saamtun	Yes	saamtun	saamtun	Yes
saamtan	saamtan	Yes	saamtan	saamtun	No	saamtan	saamtan	Yes	saamtan	saamtan	Yes
saamtin	saamtan	No	saamtin	saamtan	No	saamtin	saamtin	Yes	saamtin	saamtan	No
thollth	haadaath	No	thollth	daaf	No	thollth	saaaf	No	thollth	thaarrf	No
thaakaaf	thaakaaf	Yes	thaakaaf	thaakaaf	Yes	thaakaaf	thaakaaf	Yes	thaakaaf	dhaaghaath	No
maakkaathaa	baakhaasaa	No	maakkaathaa	maakkaanaa	No	maakkaathaa	dhaakhaahu	No	maakkaathaa	maakkaathaa	Yes
ghaath	ghaath	Yes	ghaath	rraad	No	ghaath	ghaath	Yes	ghaath	ghaath	Yes
haadaath	haadaath	Yes	haadaath	haadaath	Yes	haadaath	haadaath	Yes	haadaath	haadaath	Yes
shaarraath	shaah	No	shaarraath	shaathaaf	No	shaarraath	shaarraath	Yes	shaarraath	shaadaat	No
ath	ath	Yes	ath	haaf	No	ath	haaf	No	ath	haf	No
thaawy	saaaf	No	thaawy	saadaa	No	thaawy	saawyi	No	thaawy	saawyi	No
thaakhn	kkaahaan	No	thaakhn	dhaaa	No	thaakhn	thaakhn	Yes	thaakhn	faadhun	No
baathaahu	dhaakhaahu	No	baathaahu	baasaall	No	baathaahu	tathell	No	baathaahu	ghaasaallaa	No

Recording 4			Recording 3			Recording 2			Recording 1		
499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match
thaabaattaa	thaabaattaa	Yes	thaabaattaa	thaabaata	No	thaabaattaa	thaabaata	No	thaabaattaa	thaabaata	No
thaajjaa	shaatt	No	thaajjaa	thaajjaa	Yes	thaajjaa	thaajjaa	Yes	thaajjaa	thaajjaa	Yes
thiny	sum	No	thiny	dhihni	No	thiny	tyn	No	thiny	saaniyaa	No
thollaat	faanaarr	No	thollaat	waanaa	No	thollaat	faaallaa	No	thollaat	faaallaa	No
waathaabaa	waathaabaa	Yes	waathaabaa	waathaabaa	Yes	waathaabaa	waathaabaa	Yes	waathaabaa	waathaabaa	Yes
othirraa	baakhillaa	No	othirraa	waasia	No	othirraa	waasia	No	othirraa	hujjub	No
jjuthw	ghaaytho	No	jjuthw	tathell	No	jjuthw	tathell	No	jjuthw	sum	No
aaathaa	kaasaa	No	aaathaa	haakkaa	No	aaathaa	haasaan	No	aaathaa	haakkaa	No
rraathi	llaayth	No	rraathi	waafy	No	rraathi	rraathi	Yes	rraathi	rraathi	Yes
baatho	dhaafaarr	No	baatho	ghaaytho	No	baatho	baasaall	No	baatho	baasaall	No
tholluthin	fillizzan	No	tholluthin	fillizzan	No	tholluthin	fillizzin	No	tholluthin	kaarrdhun	No
tholluthun	fillizzun	No	tholluthun	fillizzin	No	tholluthun	tholluthin	No	tholluthun	fillizzan	No
tholluthan	saamtan	No	tholluthan	fillizzun	No	tholluthan	fillizzin	No	tholluthan	kaarrdhun	No
llujjaajj	llujjaajj	Yes	llujjaajj	llujjaajj	Yes	llujjaajj	llujjaajj	Yes	llujjaajj	llujjaajj	Yes
jjaarraakkaa	daakaakkaa	No	jjaarraakkaa	jjaarraakkaa	Yes	jjaarraakkaa	jjaarraakkaa	Yes	jjaarraakkaa	jjaarraakkaa	Yes
dhaajjaa	naathofaa	No	dhaajjaa	dhaajjaa	Yes	dhaajjaa	baattaallaa	No	dhaajjaa	sum	No
jjaas	jjaas	Yes	jjaas	ghaath	No	jjaas	jjaas	Yes	jjaas	jjaas	Yes
khaajjaall	faadhanaa	No	khaajjaall	faadhin	No	khaajjaall	ttaajjaan	No	khaajjaall	sum	No
jjaahaathaa	naahaata	No	jjaahaathaa	dhaakhaahu	No	jjaahaathaa	jjaahaathaa	Yes	jjaahaathaa	jjaahun	No
ttaajjaan	faadhanaa	No	ttaajjaan	faadhin	No	ttaajjaan	ttaajjaan	Yes	ttaajjaan	ttaajjaan	Yes
shaajjaa	daaaan	No	shaajjaa	shaajjaa	Yes	shaajjaa	shaajjaa	Yes	shaajjaa	shaajjaa	Yes

Recording 4			Recording 3			Recording 2			Recording 1		
499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match
ajjaazzaa	ajjaazzaa	Yes	ajjaazzaa	faadhanaa	No	ajjaazzaa	ajjaazzaa	Yes	ajjaazzaa	ajjaazzaa	Yes
saajjaaa	thaajjaa	No	saajjaaa	thaajjaa	No	saajjaaa	shaajjaarr	No	saajjaaa	thaajjaa	No
jjudhm	sum	No	jjudhm	sum	No	jjudhm	jjuthw	No	jjudhm	jjuhd	No
jjaady	dhaakky	No	jjaady	sum	No	jjaady	jjaady	Yes	jjaady	dhaakky	No
jjaazzaa	jjaazzaa	Yes	jjaazzaa	khaazzaakaa	No	jjaazzaa	jjuzzurr	No	jjaazzaa	jjaazzaa	Yes
haajjaa	kkaabshan	No	haajjaa	thaajjaa	No	haajjaa	tyn	No	haajjaa	haajjaa	Yes
jjaawk	sum	No	jjaawk	ghill	No	jjaawk	jjaawk	Yes	jjaawk	jjaawk	Yes
jjaamaall	daaaun	No	jjaamaall	kaarrnin	No	jjaamaall	sum	No	jjaamaall	jjaamaall	Yes
jjuhd	ghill	No	jjuhd	dub	No	jjuhd	jjuhd	Yes	jjuhd	jjaafaat	No
jjidu	naabu	No	jjidu	ghaadh	No	jjidu	ghill	No	jjidu	sum	No
waajjaadaa	naadaabaa	No	waajjaadaa	waajjaadaa	Yes	waajjaadaa	wujjidaa	No	waajjaadaa	waajjibaa	No
aaajjidu	aaajjidu	Yes	aaajjidu	aaajjidu	Yes	aaajjidu	aaajjidu	Yes	aaajjidu	aaajjidu	Yes
hujjub	haadaath	No	hujjub	sum	No	hujjub	hujjub	Yes	hujjub	sum	No
daarraajjaa	ghaadha	No	daarraajjaa	dhaajjaa	No	daarraajjaa	dhaajjaa	No	daarraajjaa	ghaajjaarr	No
saarrju	saamtan	No	saarrju	saamtan	No	saarrju	saarrju	Yes	saarrju	khaajjaall	No
waahaajji	waahaajji	Yes	waahaajji	waahaajji	Yes	waahaajji	waahaajji	Yes	waahaajji	waahaajji	Yes
eewaajjan	eewaajjan	Yes	eewaajjan	eewaajjan	Yes	eewaajjan	eewaajjin	No	eewaajjan	eewaajjan	Yes
eewaajjun	eewaajjun	Yes	eewaajjun	aaadhunaa	No	eewaajjun	eewaajjun	Yes	eewaajjun	eewaajjun	Yes
eewaajjin	sum	No	eewaajjin	eewaajjin	Yes	eewaajjin	eewaajjin	Yes	eewaajjin	eewaajjan	No
haadhaarraa	haattaamaa	No	haadhaarraa	haabaasaa	No	haadhaarraa	thaabaata	No	haadhaarraa	haamaallaa	No
kaazzaah	kaazzaahun	No	kaazzaah	kaazzaah	Yes	kaazzaah	kaazzaahun	No	kaazzaah	haazzaat	No
suhuf	suhuf	Yes	suhuf	suhub	No	suhuf	suhuf	Yes	suhuf	suhuf	Yes

Recording 4			Recording 3			Recording 2			Recording 1		
499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match
haadhw	khaadukk	No	haadhw	haatho	No	haadhw	haatho	No	haadhw	haatho	No
haattaamaa	daama	No	haattaamaa	daama	No	haattaamaa	haattaamaa	Yes	haattaamaa	haattaamaa	Yes
haasaan	haasaan	Yes	haasaan	haasaan	Yes	haasaan	haasaan	Yes	haasaan	haasaan	Yes
haakkaa	khaath	No	haakkaa	faadh	No	haakkaa	tyn	No	haakkaa	haakkaa	Yes
haallaahu	haallaahu	Yes	haallaahu	haayaaaaa	No	haallaahu	ghaanaamu	No	haallaahu	haallaahu	Yes
haay	haatho	No	haay	haay	Yes	haay	fyhi	No	haay	thiny	No
haamaallaa	khaamaanaa	No	haamaallaa	khaamaanaa	No	haamaallaa	faadhanaa	No	haamaallaa	haamaallaa	Yes
hibrr	hibrr	Yes	hibrr	hibrr	Yes	hibrr	hibrr	Yes	hibrr	hibrr	Yes
husn	haasaan	No	husn	faasin	No	husn	husn	Yes	husn	faasin	No
ttaahaanaa	faahaamaa	No	ttaahaanaa	ttaahaanaa	Yes	ttaahaanaa	ttaahaanaa	Yes	ttaahaanaa	kaahaarraa	No
suhub	suhub	Yes	suhub	suhub	Yes	suhub	suhub	Yes	suhub	suhub	Yes
yaahillu	faahaamaa	No	yaahillu	yaahillu	Yes	yaahillu	yaahillu	Yes	yaahillu	sum	No
maasaahaa	maasaahaa	Yes	maasaahaa	maasaahaa	Yes	maasaahaa	maasaahaa	Yes	maasaahaa	maasaahaa	Yes
faarraahi	kkaahaan	No	faarraahi	shaahy	No	faarraahi	ttaarraafi	No	faarraahi	sum	No
maarraahu	naahaata	No	maarraahu	maarraahu	Yes	maarraahu	maarraahu	Yes	maarraahu	maarraahu	Yes
kaazzaahanaa	kaazzaahun	No	kaazzaahanaa	kaazzaahun	No	kaazzaahanaa	kaazzaahun	No	kaazzaahanaa	kaazzaahun	No
kaazzaahin	waadhaaa	No	kaazzaahin	waadhaaa	No	kaazzaahin	kaazzaahin	Yes	kaazzaahin	waadhaaa	No
kaazzaahun	kaazzaahun	Yes	kaazzaahun	kaazzaahun	Yes	kaazzaahun	kaazzaahun	Yes	kaazzaahun	kaazzaahun	Yes
dhaakhaahu	dhaakhaahu	Yes	dhaakhaahu	dhaakhaahu	Yes	dhaakhaahu	dhaakhaahu	Yes	dhaakhaahu	dhaakhaahu	Yes
khaadukk	ath	No	khaadukk	haadaath	No	khaadukk	faadh	No	khaadukk	faadh	No

Recording 4			Recording 3			Recording 2			Recording 1		
499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match
khaath	kaaythe	No	khaath	faadh	No	khaath	haaf	No	khaath	haaf	No
khaashaaa	kaazzaah	No	khaashaaa	rraashaa	No	khaashaaa	khaashaaa	Yes	khaashaaa	sum	No
khaasaa	kaasaa	No	khaasaa	faasaa	No	khaasaa	faasun	No	khaasaa	faasun	No
dhaakhaarraa	daakaakkaa	No	dhaakhaarraa	naahaata	No	dhaakhaarraa	daafirraa	No	dhaakhaarraa	naahaata	No
khaazzaakaa	khaazzaakaa	Yes	khaazzaakaa	khaazzaakaa	Yes	khaazzaakaa	khaasaafaa	No	khaazzaakaa	kaazzaahun	No
khaasaafaa	khaasaafaa	Yes	khaasaafaa	khaasaafaa	Yes	khaasaafaa	khaasaafaa	Yes	khaasaafaa	khaasaafaa	Yes
khaamaanaa	khaamaanaa	Yes	khaamaanaa	dhaamaa	No	khaamaanaa	faadhanaa	No	khaamaanaa	haamaallaa	No
khaawy	kaadhaaa	No	khaawy	saadaa	No	khaawy	saawyi	No	khaawy	saawyi	No
khaas	khaas	Yes	khaas	haaf	No	khaas	khaas	Yes	khaas	haaf	No
khidrr	faadhanaa	No	khidrr	hibrr	No	khidrr	hibrr	No	khidrr	hibrr	No
khums	shaams	No	khums	shaams	No	khums	shaams	No	khums	khums	Yes
baakhaasaa	baakhaasaa	Yes	baakhaasaa	baakhaasaa	Yes	baakhaasaa	baakhaasaa	Yes	baakhaasaa	maakkaathaa	No
baakhillaa	baakhillaa	Yes	baakhillaa	baashimaa	No	baakhillaa	ghaaytho	No	baakhillaa	naasiyaa	No
rraakhusaa	rraakhusaa	Yes	rraakhusaa	rraakhusaa	Yes	rraakhusaa	rraakhusaa	Yes	rraakhusaa	rraahufaa	No
saarraakhaa	saabaakaa	No	saarraakhaa	shaarraafaa	No	saarraakhaa	saawghun	No	saarraakhaa	saarraakhaa	Yes
mukhi	mUdhi	No	mukhi	rraathi	No	mukhi	dhaakky	No	mukhi	sum	No
saallkhu	saamtan	No	saallkhu	saamtun	No	saallkhu	saamtun	No	saallkhu	saamtan	No
baadhaakhun	baadhaakhun	Yes	baadhaakhun	baadhaakhun	Yes	baadhaakhun	baadhaakhin	No	baadhaakhun	baadhaakhun	Yes
baadhaakhin	baathaahu	No	baadhaakhin	baadhaakhin	Yes	baadhaakhin	baadhaakhin	Yes	baadhaakhin	baadhaakhin	Yes
baadhaakhanaa	baadhaakhun	No	baadhaakhanaa	baadhaakhun	No	baadhaakhanaa	baadhaakhun	No	baadhaakhanaa	baadhaakhun	No

Recording 4			Recording 3			Recording 2			Recording 1		
499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match
dhid	dub	No	dhid	ghaadh	No	dhid	dhid	Yes	dhid	min	No
zzaand	zzaand	Yes	zzaand	zzaand	Yes	zzaand	zzaand	Yes	zzaand	zzaand	Yes
rraasaadaa	rraasaadaa	Yes	rraasaadaa	rraasaadaa	Yes	rraasaadaa	rraasaadaa	Yes	rraasaadaa	rraasaadaa	Yes
kaadaam	kaarrnan	No	kaadaam	kaadhaaa	No	kaadaam	faadhin	No	kaadaam	faadhin	No
ttaawd	kaarrnun	No	ttaawd	daaaun	No	ttaawd	tholl	No	ttaawd	tholl	No
daasaa	daasaa	Yes	daasaa	daasaa	Yes	daasaa	daasaa	Yes	daasaa	dhaafaarr	No
daagll	daama	No	daagll	daaaan	No	daagll	daama	No	daagll	ghaanaamu	No
daaahu	daaahu	Yes	daaahu	daaahu	Yes	daaahu	daaahu	Yes	daaahu	daaahu	Yes
daaf	ghaath	No	daaf	dhaab	No	daaf	ghaath	No	daaf	ghaath	No
daama	dhaamaat	No	daama	dhaanaa	No	daama	daama	Yes	daama	ghaanaamu	No
dub	dub	Yes	dub	dub	Yes	dub	ghill	No	dub	ghill	No
diykk	diykk	Yes	diykk	diykk	Yes	diykk	diykk	Yes	diykk	diykk	Yes
naadaabaa	naadaabaa	Yes	naadaabaa	naadaabaa	Yes	naadaabaa	naadaabaa	Yes	naadaabaa	naadaabaa	Yes
hudidaa	wujjidaa	No	hudidaa	wujjidaa	No	hudidaa	mudun	No	hudidaa	atyin	No
mudun	mudun	Yes	mudun	mudun	Yes	mudun	mudun	Yes	mudun	mudun	Yes
saadaa	saabaa	No	saadaa	saakkaabaa	No	saadaa	saajjaaa	No	saadaa	sum	No
ahdu	sum	No	ahdu	sum	No	ahdu	tathell	No	ahdu	aaawidu	No
maahdi	maahwu	No	maahdi	sum	No	maahdi	rraathi	No	maahdi	maahwu	No
waaadan	waaadun	No	waaadan	waadhaaa	No	waaadan	faadhin	No	waaadan	waaadan	Yes
waaadun	waaadan	No	waaadun	waadhaaa	No	waaadun	waadhaaa	No	waaadun	waaadun	Yes

Recording 4			Recording 3			Recording 2			Recording 1		
499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match
waaadin	waajjaadaa	No	waaadin	waajjaadaa	No	waaadin	waaadin	Yes	waaadin	waattaan	No
dhaakky	yaakhti	No	dhaakky	maahdi	No	dhaakky	dhaakky	Yes	dhaakky	dhaakky	Yes
dhaamaa	daama	No	dhaamaa	daama	No	dhaamaa	daama	No	dhaamaa	daama	No
dhaallaa	maakkunaa	No	dhaallaa	dhaanaa	No	dhaallaa	naathaarraa	No	dhaallaa	naathaarraa	No
faadh	faan	No	faadh	daaaan	No	faadh	min	No	faadh	faadh	Yes
kaadhaaa	ghaadha	No	kaadhaaa	kaadhaaa	Yes	kaadhaaa	baadaa	No	kaadhaaa	kkaabaa	No
shaadhaa	shaathaaf	No	shaadhaa	shaathaaf	No	shaadhaa	shaathaaf	No	shaadhaa	shaathaaf	No
dhaawd	dhaab	No	dhaawd	daaaun	No	dhaawd	ghill	No	dhaawd	ghill	No
dhiib	tyn	No	dhiib	daagll	No	dhiib	daagll	No	dhiib	min	No
dhaab	dhaaba	No	dhaab	dhaab	Yes	dhaab	dhaab	Yes	dhaab	ghaadh	No
dhull	daama	No	dhull	daaaun	No	dhull	daaaun	No	dhull	mudun	No
kkaadhibaa	sum	No	kkaadhibaa	ghaattaa	No	kkaadhibaa	sum	No	kkaadhibaa	khaadukk	No
adhaarraa	kaarrnan	No	adhaarraa	ghaadaarraa	No	adhaarraa	faadhin	No	adhaarraa	haattaallaa	No
aaadhunaa	aaadhunaa	Yes	aaadhunaa	aaadhunaa	Yes	aaadhunaa	aaadhunaa	Yes	aaadhunaa	aaadhunaa	Yes
shaahaadhaa	shaahaadhaa	Yes	shaahaadhaa	shaahaadhaa	Yes	shaahaadhaa	shaahaadhaa	Yes	shaahaadhaa	shaahaadhaa	Yes
mundhu	mundhu	Yes	mundhu	mundhu	Yes	mundhu	minhu	No	mundhu	mundhu	Yes
mUdhi	llaayth	No	mUdhi	maahdi	No	mUdhi	sum	No	mUdhi	maahdi	No
faadhanaa	kaarrnun	No	faadhanaa	kaarrnan	No	faadhanaa	fillizzin	No	faadhanaa	faadh	No
faadhun	ttaabllan	No	faadhun	ttaabllan	No	faadhun	faadhun	Yes	faadhun	faadhun	Yes
faadhin	kaarrnan	No	faadhin	ttaabllin	No	faadhin	faadhin	Yes	faadhin	faadhin	Yes

Recording 4			Recording 3			Recording 2			Recording 1		
499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match
thaahaarraa	naahaata	No	thaahaarraa	ghaattaa	No	thaahaarraa	dhaaba	No	thaahaarraa	naahaata	No
kaarrn	kaarrn	Yes	kaarrn	ttaabllan	No	kaarrn	kaarrnin	No	kaarrn	saamtan	No
rraakkaallaa	rraakkaallaa	Yes	rraakkaallaa	maakkaanaa	No	rraakkaallaa	rraakkaallaa	Yes	rraakkaallaa	rraakkaallaa	Yes
dhaarraa	dhaarraat	No	dhaarraa	ghaadaarraa	No	dhaarraa	daama	No	dhaarraa	naathaarraa	No
rraagw	ghaadaarraa	No	rraagw	ttaabllan	No	rraagw	maahwu	No	rraagw	baathaahu	No
ttaayrr	sum	No	ttaayrr	sum	No	ttaayrr	ayiyaa	No	ttaayrr	faaallaa	No
sirr	saadghu	No	sirr	sihrr	No	sirr	sihrr	No	sirr	sihrr	No
rraad	ghaadha	No	rraad	faadh	No	rraad	faadh	No	rraad	rraad	Yes
rrubaa	ghaath	No	rrubaa	rrubaa	Yes	rrubaa	rrubaa	Yes	rrubaa	sum	No
surrurr	shaah	No	surrurr	shibll	No	surrurr	suhub	No	surrurr	fillizzin	No
haarraamaa	haarraamaa	Yes	haarraamaa	haattaamaa	No	haarraamaa	haattaamaa	No	haarraamaa	haattaamaa	No
sirry	saamtan	No	sirry	faarraasi	No	sirry	saanaami	No	sirry	saanaami	No
faattaarraa	haafaathaa	No	faattaarraa	rraabaattaa	No	faattaarraa	thaabaata	No	faattaarraa	saakaattaa	No
jjuhrri	yaahillu	No	jjuhrri	sum	No	jjuhrri	sum	No	jjuhrri	jjaahun	No
fikkrri	faakkaa	No	fikkrri	taky	No	fikkrri	sum	No	fikkrri	taky	No
daahrrun	daahrrun	Yes	daahrrun	daaaun	No	daahrrun	daahrrin	No	daahrrun	daaaun	No
daahrrin	daaaan	No	daahrrin	daaaan	No	daahrrin	daaaan	No	daahrrin	daaaan	No
daahrran	daahrrun	No	daahrran	daahrrun	No	daahrran	daahrrun	No	daahrran	jjaahun	No
zzaafaa	zzaafaa	Yes	zzaafaa	zzaafaa	Yes	zzaafaa	zzaafaa	Yes	zzaafaa	zzaafaa	Yes
zzaaamaa	dhaamaa	No	zzaaamaa	zzaaamaa	Yes	zzaaamaa	zzaaamaa	Yes	zzaaamaa	daama	No

Recording 4			Recording 3			Recording 2			Recording 1		
499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match
zzaakky	sum	No	zzaakky	zzaakky	Yes	zzaakky	zzaakky	Yes	zzaakky	dhaakky	No
zzuhaall	dhaakhaahu	No	zzuhaall	zzuhaall	Yes	zzuhaall	zzuhaall	Yes	zzuhaall	zzuhaall	Yes
zzaarraaa	zzaakky	No	zzaarraaa	dhaaba	No	zzaarraaa	zzaaamaa	No	zzaarraaa	daama	No
zzirr	zzaand	No	zzirr	zzaaamaa	No	zzirr	zzirr	Yes	zzirr	dhid	No
rruzzik	baazzaaghaa	No	rruzzik	sum	No	rruzzik	rruzzik	Yes	rruzzik	sum	No
azzaafaa	azzaafaa	Yes	azzaafaa	sum	No	azzaafaa	azzaafaa	Yes	azzaafaa	kaazzaahun	No
jjuzzurr	daasaa	No	jjuzzurr	sum	No	jjuzzurr	jjuzzurr	Yes	jjuzzurr	jjuzzurr	Yes
faazzaa	faazzaa	Yes	faazzaa	faazzaa	Yes	faazzaa	faasin	No	faazzaa	faasin	No
jjaawzzu	jjaazzaa	No	jjaawzzu	jjaawzzu	Yes	jjaawzzu	jjaawzzu	Yes	jjaawzzu	jjaawzzu	Yes
kkaanzzi	sum	No	kkaanzzi	tathell	No	kkaanzzi	kkaanzzi	Yes	kkaanzzi	sum	No
fillizzan	fillizzan	Yes	fillizzan	fillizzin	No	fillizzan	fillizzan	Yes	fillizzan	fillizzan	Yes
fillizzun	fillizzan	No	fillizzun	fillizzan	No	fillizzun	fillizzun	Yes	fillizzun	fillizzun	Yes
fillizzin	asaall	No	fillizzin	asaall	No	fillizzin	fillizzin	Yes	fillizzin	fillizzin	Yes
shaams	shaams	Yes	shaams	shaams	Yes	shaams	shaams	Yes	shaams	shaams	Yes
ghaasaallaa	ghaasaallaa	Yes	ghaasaallaa	baasaattaa	No	ghaasaallaa	baasaattaa	No	ghaasaallaa	rraasaadaa	No
saahw	saamtan	No	saahw	saahw	Yes	saahw	saahw	Yes	saahw	saahw	Yes
kkys	saaaf	No	kkys	kkiys	No	kkys	kkiys	No	kkys	kkiys	No
dhirrs	daasaa	No	dhirrs	dhaaghaath	No	dhirrs	maawzz	No	dhirrs	daaf	No
sum	sum	Yes	sum	shaahy	No	sum	saawghin	No	sum	faan	No
saakkaabaa	saakkaabaa	Yes	saakkaabaa	saakkaabaa	Yes	saakkaabaa	sillkkin	No	saakkaabaa	saakkaabaa	Yes
sihrr	saahw	No	sihrr	sihrr	Yes	sihrr	suhuf	No	sihrr	suhub	No

Recording 4			Recording 3			Recording 2			Recording 1		
499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match
rrusull	nusirraa	No	rrusull	rrusull	Yes	rrusull	orrsin	No	rrusull	fillizzun	No
asaall	haasaan	No	asaall	haasaan	No	asaall	haasaan	No	asaall	haasaan	No
naasiyaa	naasiyaa	Yes	naasiyaa	naasiyaa	Yes	naasiyaa	naasiyaa	Yes	naasiyaa	naasiyaa	Yes
haabaasaa	haabaasaa	Yes	haabaasaa	haabaasaa	Yes	haabaasaa	haabaasaa	Yes	haabaasaa	haabaasaa	Yes
haarraasaa	haarraasaa	Yes	haarraasaa	haasaan	No	haarraasaa	haasaan	No	haarraasaa	haarraasaa	Yes
faarraasi	ghaasaallaa	No	faarraasi	rraathi	No	faarraasi	fillizzin	No	faarraasi	ttaarraafi	No
orrsan	fillizzan	No	orrsan	orrsan	Yes	orrsan	haasaan	No	orrsan	haasaan	No
orrsun	rrusull	No	orrsun	rrusull	No	orrsun	rrusull	No	orrsun	rrusull	No
orrsin	fillizzin	No	orrsin	fillizzin	No	orrsin	fillizzin	No	orrsin	fillizzin	No
shaadhw	shaadhw	Yes	shaadhw	shibll	No	shaadhw	shibll	No	shaadhw	shibll	No
shaas	shaas	Yes	shaas	shaas	Yes	shaas	shaas	Yes	shaas	shaas	Yes
shaathaaf	shaathaaf	Yes	shaathaaf	shaadaat	No	shaathaaf	shaathaaf	Yes	shaathaaf	shaathaaf	Yes
shaatt	shaatt	Yes	shaatt	shaatt	Yes	shaatt	shaatt	Yes	shaatt	sum	No
shugll	shaadhaa	No	shugll	shibll	No	shugll	shugll	Yes	shugll	shugll	Yes
kaash	kaash	Yes	kaash	kaash	Yes	kaash	kaash	Yes	kaash	kaas	No
shaakk	shaatt	No	shaakk	shaakk	Yes	shaakk	shaakk	Yes	shaakk	shaakk	Yes
naashizz	baasaattaa	No	naashizz	naashizz	Yes	naashizz	naashizz	Yes	naashizz	naashizz	Yes
shaahy	shaahy	Yes	shaahy	shaahy	Yes	shaahy	shaahy	Yes	shaahy	saamghi	No
shaajjaarr	shaadhaa	No	shaajjaarr	shaajjaa	No	shaajjaarr	shaajjaa	No	shaajjaarr	shaajjaa	No
shibll	shibll	Yes	shibll	shibll	Yes	shibll	shibll	Yes	shibll	shibll	Yes

Recording 4			Recording 3			Recording 2			Recording 1		
499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match
shukkrr	shugll	No	shukkrr	shaatt	No	shukkrr	sum	No	shukkrr	shaakk	No
waashm	rraashaa	No	waashm	waashm	Yes	waashm	waashm	Yes	waashm	waasy	No
rrushidaa	rrushidaa	Yes	rrushidaa	rrushidaa	Yes	rrushidaa	rrushidaa	Yes	rrushidaa	rrushidaa	Yes
aaashudu	aaashudu	Yes	aaashudu	aaashudu	Yes	aaashudu	aaashudu	Yes	aaashudu	aaashudu	Yes
rraashaa	rraashaa	Yes	rraashaa	rraashaa	Yes	rraashaa	rraashaa	Yes	rraashaa	rraasaa	No
rrimshu	rrimshu	Yes	rrimshu	rrimshu	Yes	rrimshu	rrimshu	Yes	rrimshu	rrusull	No
rryshi	sum	No	rryshi	sum	No	rryshi	rryshi	Yes	rryshi	rryshi	Yes
kkaabshan	kkaabshan	Yes	kkaabshan	kkaabshan	Yes	kkaabshan	haam	No	kkaabshan	khaajjaall	No
kkaabshun	kkaabshun	Yes	kkaabshun	kkaabshun	Yes	kkaabshun	kkaabshun	Yes	kkaabshun	kabshin	No
kkaabshin	kkaabshan	No	kkaabshin	kkaabshin	Yes	kkaabshin	kkaabshan	No	kkaabshin	kkaabshan	No
kaasaa	kaas	No	kaasaa	kaasaa	Yes	kaasaa	faasun	No	kaasaa	faasun	No
sum	saamtan	No	sum	shaahy	No	sum	saamghi	No	sum	faan	No
saanaaa	saanaaa	Yes	saanaaa	faanaarr	No	saanaaa	saanaaa	Yes	saanaaa	faanaarr	No
saah	saah	Yes	saah	saah	Yes	saah	saah	Yes	saah	saaaf	No
waasy	waasaatti	No	waasy	waasy	Yes	waasy	waasy	Yes	waasy	waafy	No
suws	suhuf	No	suws	shaams	No	suws	sillkkin	No	suws	thollth	No
saayd	saamtin	No	saayd	saayd	Yes	saayd	saayd	Yes	saayd	fyhi	No
sihrr	saahw	No	sihrr	saaeerr	No	sihrr	sihrr	Yes	sihrr	saadghu	No
asaarraa	kaasaa	No	asaarraa	kaazzaah	No	asaarraa	baasaattaa	No	asaarraa	faasun	No
nusirraa	naasiyaa	No	nusirraa	naasiyaa	No	nusirraa	nusirraa	Yes	nusirraa	nusirraa	Yes
yaasudu	yaasudu	Yes	yaasudu	naasaab	No	yaasudu	yaasudu	Yes	yaasudu	yaasudu	Yes

Recording 4			Recording 3			Recording 2			Recording 1		
499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match
rraasaa	rraasaa	Yes	rraasaa	rraasaa	Yes	rraasaa	rraasaa	Yes	rraasaa	rraasaa	Yes
kurrsi	kurrsi	Yes	kurrsi	waasy	No	kurrsi	waasy	No	kurrsi	kurrsi	Yes
faasun	faasun	Yes	faasun	haasaan	No	faasun	faasun	Yes	faasun	faasun	Yes
faasaa	haasaan	No	faasaa	haasaan	No	faasaa	haasaan	No	faasaa	haasaan	No
faasin	rraasaadaa	No	faasin	rraasaadaa	No	faasin	rraafaaa	No	faasin	rraafaaa	No
dhaaghaath	daama	No	dhaaghaath	ghaadha	No	dhaaghaath	maallaaf	No	dhaaghaath	baahaak	No
waadhaaa	dhaaba	No	waadhaaa	waattaan	No	waadhaaa	waathaafaa	No	waadhaaa	waaadun	No
dhaanaa	daaaun	No	dhaanaa	waanaa	No	dhaanaa	daama	No	dhaanaa	maakkaanaa	No
dhaallaa	daama	No	dhaallaa	naadaabaa	No	dhaallaa	daama	No	dhaallaa	daama	No
dhyk	sum	No	dhyk	diykk	No	dhyk	diykk	No	dhyk	sum	No
dhaafaarr	naahaata	No	dhaafaarr	baasaattaa	No	dhaafaarr	daafirraa	No	dhaafaarr	waaathun	No
dhaarraabaa	naadaabaa	No	dhaarraabaa	ghaadha	No	dhaarraabaa	naahaata	No	dhaarraabaa	naadaabaa	No
dhuha	dhuha	Yes	dhuha	dhuha	Yes	dhuha	suhub	No	dhuha	waaathan	No
dhidu	naabu	No	dhidu	shibll	No	dhidu	ghill	No	dhidu	naabu	No
rraadhiyaa	rraadhiyaa	Yes	rraadhiyaa	rraadhiyaa	Yes	rraadhiyaa	rraadhiyaa	Yes	rraadhiyaa	rraadhiyaa	Yes
adhud	aaadhunaa	No	adhud	ttaabllin	No	adhud	ttaabllin	No	adhud	allaaman	No
faadhaallaa	thaabaata	No	faadhaallaa	thaabaata	No	faadhaallaa	thaabaata	No	faadhaallaa	baattaallaa	No
maarraadhaa	naadaabaa	No	maarraadhaa	naadaabaa	No	maarraadhaa	mundhu	No	maarraadhaa	naadaabaa	No
arraadhaa	dhaarraabaa	No	arraadhaa	baadaa	No	arraadhaa	thaabaata	No	arraadhaa	haattaamaa	No
aaarrdhi	kaarrnun	No	aaarrdhi	thaaby	No	aaarrdhi	thaaby	No	aaarrdhi	sum	No

Recording 4			Recording 3			Recording 2			Recording 1		
499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match
kaarrdhan	kaarrnan	No	kaarrdhan	ttaabllun	No	kaarrdhan	kaarrnun	No	kaarrdhan	kaarrnun	No
kaarrdhun	kaarrnun	No	kaarrdhun	ttaabllan	No	kaarrdhun	thaawbun	No	kaarrdhun	alamun	No
kaarrdhin	kaarrnan	No	kaarrdhin	ttaabllin	No	kaarrdhin	ttaabllin	No	kaarrdhin	saamtun	No
ttuk	ghaat	No	ttuk	faadh	No	ttuk	faadh	No	ttuk	faadh	No
haattaallaa	haattaamaa	No	haattaallaa	thaabaata	No	haattaallaa	thaabaata	No	haattaallaa	haattaallaa	Yes
ttaamaaa	kaarrnan	No	ttaamaaa	daama	No	ttaamaaa	daama	No	ttaamaaa	daama	No
ttib	sum	No	ttib	sum	No	ttib	sum	No	ttib	faadh	No
ttaabaaa	sum	No	ttaabaaa	rraabaattaa	No	ttaabaaa	ttaabllin	No	ttaabaaa	faaallaa	No
waattaan	waaathun	No	waattaan	waattaan	Yes	waattaan	waattaan	Yes	waattaan	waattaan	Yes
rraattib	waaathan	No	rraattib	rraakkaaa	No	rraattib	baarrkin	No	rraattib	rraattib	Yes
otuf	sum	No	otuf	dhaaghaath	No	otuf	otuf	Yes	otuf	sum	No
kirrtu	minhu	No	kirrtu	sum	No	kirrtu	tathell	No	kirrtu	minhu	No
waasaatti	waasaatti	Yes	waasaatti	waasaatti	Yes	waasaatti	waasaatti	Yes	waasaatti	waasaatti	Yes
baasaattaa	baasaattaa	Yes	baasaattaa	baasaattaa	Yes	baasaattaa	khaasaafaa	No	baasaattaa	baasaattaa	Yes
nukaattan	nukaattan	Yes	nukaattan	nukaattan	Yes	nukaattan	nukaattan	Yes	nukaattan	nukaattan	Yes
nukaattun	nukaattin	No	nukaattun	nukaattun	Yes	nukaattun	nukaattun	Yes	nukaattun	nukaattun	Yes
nukaattin	nukaattin	Yes	nukaattin	nukaattin	Yes	nukaattin	nukaattin	Yes	nukaattin	nukaattin	Yes
thaahaarr	naahaata	No	thaahaarr	baatho	No	thaahaarr	faadh	No	thaahaarr	naahaata	No
kkaatho	dhaaba	No	kkaatho	ttaabllan	No	kkaatho	naabu	No	kkaatho	naabu	No
waathaafaa	waathaafaa	Yes	waathaafaa	waathaafaa	Yes	waathaafaa	waathaafaa	Yes	waathaafaa	waathaafaa	Yes
thaarrf	ghaath	No	thaarrf	maallaaf	No	thaarrf	baatho	No	thaarrf	ghaath	No

Recording 4			Recording 3			Recording 2			Recording 1		
499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match
thefrr	daafirraa	No	thefrr	daaf	No	thefrr	daaf	No	thefrr	sum	No
tholl	daama	No	tholl	daaaun	No	tholl	daaaan	No	tholl	daaaun	No
naathaarraa	naahaata	No	naathaarraa	naadaabaa	No	naathaarraa	naathaarraa	Yes	naathaarraa	naathaarraa	Yes
naathofaa	naathofaa	Yes	naathofaa	naathofaa	Yes	naathofaa	naathofaa	Yes	naathofaa	naathofaa	Yes
athemaa	ghaadha	No	athemaa	rraadhiyaa	No	athemaa	kkaadhibaa	No	athemaa	sum	No
haafaathaa	haafaathaa	Yes	haafaathaa	haafaathaa	Yes	haafaathaa	haafaathaa	Yes	haafaathaa	haafaathaa	Yes
kaaythe	aaamill	No	kaaythe	aaawidu	No	kaaythe	sum	No	kaaythe	sum	No
haatho	kaarrnun	No	haatho	faadh	No	haatho	naabu	No	haatho	faadh	No
waaathan	waaadun	No	waaathan	waaadun	No	waaathan	waaathun	No	waaathan	waattaan	No
waaathun	waattaan	No	waaathun	waaathun	Yes	waaathun	waaathun	Yes	waaathun	waaathun	Yes
waaathin	daaaan	No	waaathin	daaaan	No	waaathin	waaathin	Yes	waaathin	waaadin	No
adhaall	ahdu	No	adhaall	kaadhaaa	No	adhaall	ahdu	No	adhaall	faadhin	No
saaaf	thaakaaf	No	saaaf	saah	No	saaaf	sihrr	No	saaaf	ghaath	No
attaash	attaash	Yes	attaash	haawaas	No	attaash	attaash	Yes	attaash	attaash	Yes
akks	kaas	No	akks	kkiys	No	akks	baakhaasaa	No	akks	haakkaa	No
akrr	akrr	Yes	akrr	sum	No	akrr	saattaat	No	akrr	haattaamaa	No
ayn	kaarrn	No	ayn	haay	No	ayn	fyhi	No	ayn	aaaw	No
eejjll	sum	No	eejjll	sum	No	eejjll	eejjll	Yes	eejjll	eejjll	Yes
omrr	allaaman	No	omrr	aaamill	No	omrr	aaaw	No	omrr	haamaallaa	No
saaeerr	saaeerr	Yes	saaeerr	saaeerr	Yes	saaeerr	sihrr	No	saaeerr	saaeerr	Yes

Recording 4			Recording 3			Recording 2			Recording 1		
499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match
naaasaa	daasaa	No	naaasaa	naasaab	No	naaasaa	naasaab	No	naaasaa	naathofaa	No
naaomaa	dhaamaa	No	naaomaa	naaomaa	Yes	naaomaa	naaomaa	Yes	naaomaa	naaomaa	Yes
waasia	waasia	Yes	waasia	waasia	Yes	waasia	waasia	Yes	waasia	waasia	Yes
kaaee	kaaythe	No	kaaee	zzaakky	No	kaaee	kaaythe	No	kaaee	saaniyaa	No
saao	saadghu	No	saao	saahw	No	saao	saahw	No	saao	saahw	No
waarriaan	waaadun	No	waarriaan	waaadun	No	waarriaan	waaadin	No	waarriaan	waarriain	No
waarriaun	waarriaun	Yes	waarriaun	waarriaun	Yes	waarriaun	waarriaun	Yes	waarriaun	waarriaun	Yes
waarriain	waaadan	No	waarriain	waadhaaa	No	waarriain	waarriain	Yes	waarriain	waarriain	Yes
ghaajjaarr	ghaadha	No	ghaajjaarr	rraaghaad	No	ghaajjaarr	dhaajjaa	No	ghaajjaarr	ghaajjaarr	Yes
ghaadha	daama	No	ghaadha	naathaarraa	No	ghaadha	tathell	No	ghaadha	naathaarraa	No
ghaashaa	naashizz	No	ghaashaa	rraashaa	No	ghaashaa	naashizz	No	ghaashaa	rraashaa	No
ghaadhu	naabu	No	ghaadhu	naabu	No	ghaadhu	rraad	No	ghaadhu	rraad	No
ghaafiyaa	sum	No	ghaafiyaa	rraadhiyaa	No	ghaafiyaa	waasia	No	ghaafiyaa	sum	No
ghaarraakaa	ghaattaa	No	ghaarraakaa	rraabaattaa	No	ghaarraakaa	naathofaa	No	ghaarraakaa	haarraakkaa	No
ghaaytho	waaadin	No	ghaaytho	aaajjidu	No	ghaaytho	sum	No	ghaaytho	ghaanaamu	No
ghaattaa	saah	No	ghaattaa	khaath	No	ghaattaa	naathofaa	No	ghaattaa	zzaarr	No
ghaadaarraa	ghaadha	No	ghaadaarraa	naadaabaa	No	ghaadaarraa	waathaafaa	No	ghaadaarraa	naadaabaa	No
ghusn	waaathan	No	ghusn	naashizz	No	ghusn	orrsin	No	ghusn	faasun	No
ghill	min	No	ghill	min	No	ghill	dhihni	No	ghill	min	No
saaghurraa	saawghin	No	saaghurraa	saawghin	No	saaghurraa	saawghun	No	saaghurraa	faadhanaa	No

Recording 4			Recording 3			Recording 2			Recording 1		
499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match
rraaghaad	naadaabaa	No	rraaghaad	rraaghaad	Yes	rraaghaad	rraaghaad	Yes	rraaghaad	rraaghaad	Yes
ttaaghiyaa	maakkunaa	No	ttaaghiyaa	rraadhiyaa	No	ttaaghiyaa	waaathin	No	ttaaghiyaa	rraadhiyaa	No
maarraaghaa	maarraadhaa	No	maarraaghaa	maakkaanaa	No	maarraaghaa	maarraahu	No	maarraaghaa	maarraadhaa	No
saadghu	faadhanaa	No	saadghu	faadhanaa	No	saadghu	saadghu	Yes	saadghu	saadghu	Yes
saamghi	saamtan	No	saamghi	saanaaa	No	saamghi	saamtin	No	saamghi	saamghi	Yes
saawgan	saawghun	No	saawgan	saawghun	No	saawgan	saawghun	No	saawgan	saawghun	No
saawghun	thaawbun	No	saawghun	saamtun	No	saawghun	ayn	No	saawghun	ayn	No
saawghin	saamtan	No	saawghin	saadaa	No	saawghin	saawgan	No	saawghin	tholluthin	No
haaf	ghaath	No	haaf	haaf	Yes	haaf	ath	No	haaf	haaf	Yes
waafy	waasaatti	No	waafy	waafy	Yes	waafy	waafy	Yes	waafy	waafy	Yes
maallaaf	maallaaf	Yes	maallaaf	maallaaf	Yes	maallaaf	maallaaf	Yes	maallaaf	maallaaf	Yes
faakkaa	sum	No	faakkaa	daaaun	No	faakkaa	faakkaa	Yes	faakkaa	khaath	No
faan	zzaand	No	faan	faan	Yes	faan	faan	Yes	faan	faan	Yes
fijjll	faadhin	No	fijjll	sillkkin	No	fijjll	fijjll	Yes	fijjll	fijjll	Yes
furrn	daama	No	furrn	faan	No	furrn	faadhin	No	furrn	waattaan	No
faaallaa	faaallaa	Yes	faaallaa	ttaahaanaa	No	faaallaa	daama	No	faaallaa	faadhaallaa	No
rraafaaa	rraafaaa	Yes	rraafaaa	rraasaa	No	rraafaaa	dhaafaarr	No	rraafaaa	rraasaa	No
daafirraa	daasaa	No	daafirraa	dhaakky	No	daafirraa	sum	No	daafirraa	ghaasaallaa	No
afwu	kaarrdhun	No	afwu	afwu	Yes	afwu	afwu	Yes	afwu	afwu	Yes
shaarraafaa	shaarraafaa	Yes	shaarraafaa	shaarraafaa	Yes	shaarraafaa	shaarraafaa	Yes	shaarraafaa	shaarraafaa	Yes

Recording 4			Recording 3			Recording 2			Recording 1		
499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match
ttaarraafi	ttaarraafi	Yes	ttaarraafi	rraathi	No	ttaarraafi	ttaarraafi	Yes	ttaarraafi	ttaarraafi	Yes
khaallfu	haarraakkaa	No	khaallfu	saamtan	No	khaallfu	ghaaytho	No	khaallfu	saamtan	No
allaafan	allaafan	Yes	allaafan	haasaan	No	allaafan	allaafin	No	allaafan	allaafan	Yes
allaafun	allaafan	No	allaafun	allaafan	No	allaafun	allaafun	Yes	allaafun	allaafun	Yes
allaafin	allaafin	Yes	allaafin	allaafin	Yes	allaafin	allaafin	Yes	allaafin	allaafin	Yes
sujjuk	sujjuk	Yes	sujjuk	shibll	No	sujjuk	saarrju	No	sujjuk	sum	No
kullw	hunuw	No	kullw	sum	No	kullw	aaamill	No	kullw	hunuw	No
daakaakkaa	daakaakkaa	Yes	daakaakkaa	haakkaa	No	daakaakkaa	daakaakkaa	Yes	daakaakkaa	haakkaa	No
kaallaam	kaarrnan	No	kaallaam	kaarrnan	No	kaallaam	faadhin	No	kaallaam	kaarrnun	No
kidrr	daagllu	No	kidrr	hibrr	No	kidrr	tathell	No	kidrr	hibrr	No
kudaa	sum	No	kudaa	kudaa	Yes	kudaa	waaathin	No	kudaa	yin	No
saakaattaa	saakaattaa	Yes	saakaattaa	saakaattaa	Yes	saakaattaa	saakaattaa	Yes	saakaattaa	saakaattaa	Yes
fukidaa	fukidaa	Yes	fukidaa	fukidaa	Yes	fukidaa	fukidaa	Yes	fukidaa	fukidaa	Yes
thaakullaa	maakkunaa	No	thaakullaa	saakaattaa	No	thaakullaa	thaakullaa	Yes	thaakullaa	thaakullaa	Yes
saabaakaa	thaabaata	No	saabaakaa	saabaakaa	Yes	saabaakaa	saabaakaa	Yes	saabaakaa	saabaakaa	Yes
abaakaa	rraabaattaa	No	abaakaa	rraabaattaa	No	abaakaa	thaabaata	No	abaakaa	sum	No
ghaasaaku	ghaasaaku	Yes	ghaasaaku	ghaasaaku	Yes	ghaasaaku	ghaasaaku	Yes	ghaasaaku	ghaasaaku	Yes
baarrkan	naathofaa	No	baarrkan	daaaan	No	baarrkan	baadhaakhun	No	baarrkan	baadhaakhun	No
baarrkun	waaathun	No	baarrkun	baadhaakhin	No	baarrkun	sum	No	baarrkun	sum	No
baarrkin	ayn	No	baarrkin	ayn	No	baarrkin	baadhaakhin	No	baarrkin	baadhaakhin	No

Recording 4			Recording 3			Recording 2			Recording 1		
499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match
rraakkaadhaa	maakkaathaa	No	rraakkaadhaa	rraakkaallaa	No	rraakkaadhaa	rraakkaallaa	No	rraakkaadhaa	rraakkaadhaa	Yes
jjaarraakkaa	ath	No	jjaarraakkaa	jjaarraakkaa	Yes	jjaarraakkaa	jjaarraakkaa	Yes	jjaarraakkaa	jjaarraakkaa	Yes
kkaawa	kaarrnun	No	kkaawa	dhaallaa	No	kkaawa	ghill	No	kkaawa	dhaallaa	No
kkaahaan	kkaahaan	Yes	kkaahaan	kkaahaan	Yes	kkaahaan	kkaahaan	Yes	kkaahaan	kkaahaan	Yes
kkaallb	khaadukk	No	kkaallb	daaaan	No	kkaallb	tyn	No	kkaallb	khaawy	No
kkiys	saaaf	No	kkiys	kkiys	Yes	kkiys	kkiys	Yes	kkiys	kkiys	Yes
kkuwa	kkuwa	Yes	kkuwa	sum	No	kkuwa	kkuwa	Yes	kkuwa	allaafin	No
rraakkibaa	rraakkibaa	Yes	rraakkibaa	fukidaa	No	rraakkibaa	rraakkibaa	Yes	rraakkibaa	rraakkibaa	Yes
rraakkaaa	rraakkaallaa	No	rraakkaaa	rraakkaallaa	No	rraakkaaa	rraakkaallaa	No	rraakkaaa	rraakkaallaa	No
maakkunaa	maakkunaa	Yes	maakkunaa	maakkunaa	Yes	maakkunaa	maakkunaa	Yes	maakkunaa	maakkunaa	Yes
haarraakkaa	haarraakkaa	Yes	haarraakkaa	haarraakkaa	Yes	haarraakkaa	haarraakkaa	Yes	haarraakkaa	haarraakkaa	Yes
berraakku	berraakku	Yes	berraakku	ghaasaaku	No	berraakku	berraakku	Yes	berraakku	daaahu	No
saamaakki	saamaakki	Yes	saamaakki	saamaakki	Yes	saamaakki	saamaakki	Yes	saamaakki	saamaakki	Yes
sillkkan	saamtan	No	sillkkan	saamtun	No	sillkkan	sillkkin	No	sillkkan	saamtan	No
sillkkun	saamtun	No	sillkkun	sillkkun	Yes	sillkkun	sillkkun	Yes	sillkkun	saamtun	No
sillkkin	saamtan	No	sillkkin	saamtan	No	sillkkin	sillkkin	Yes	sillkkin	saamtan	No
llaayth	llaayth	Yes	llaayth	mUdhi	No	llaayth	llaayth	Yes	llaayth	llaayth	Yes
lliyn	mudun	No	lliyn	atyin	No	lliyn	lliyn	Yes	lliyn	lliyn	Yes
llumaat	daama	No	llumaat	naaomaa	No	llumaat	numuw	No	llumaat	llumaat	Yes
olluw	hunuw	No	olluw	olluw	Yes	olluw	aaamill	No	olluw	aaamill	No

Recording 4			Recording 3			Recording 2			Recording 1		
499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match
ghaallaaa	ghaadaarraa	No	ghaallaaa	dhaanaa	No	ghaallaaa	baatho	No	ghaallaaa	maakkaanaa	No
jjaalliy	dhihni	No	jjaalliy	jjaady	No	jjaalliy	dhihni	No	jjaalliy	dhihni	No
daagllu	daaahu	No	daagllu	ghaanaamu	No	daagllu	ghaanaamu	No	daagllu	ghaanaamu	No
amaalli	sum	No	amaalli	saanaaa	No	amaalli	amaalli	Yes	amaalli	amaalli	Yes
ttaabllan	ttaabllan	Yes	ttaabllan	waattaan	No	ttaabllan	faadhin	No	ttaabllan	faadhin	No
ttaabllun	faadhun	No	ttaabllun	ttaabllan	No	ttaabllun	faadhun	No	ttaabllun	waaadan	No
ttaabllin	faadhin	No	ttaabllin	waaathin	No	ttaabllin	faadhin	No	ttaabllin	thaawbin	No
haam	sum	No	haam	kaarrn	No	haam	faan	No	haam	daama	No
yaawm	yaawm	Yes	yaawm	ghaallaaa	No	yaawm	ghill	No	yaawm	yaawm	Yes
maawzz	maallaaf	No	maawzz	maawzz	Yes	maawzz	maawzz	Yes	maawzz	maawzz	Yes
min	min	Yes	min	min	Yes	min	min	Yes	min	min	Yes
aaamill	kaarrnan	No	aaamill	kaarrnin	No	aaamill	kaarrnin	No	aaamill	kaarrnin	No
amaallaa	amaallaa	Yes	amaallaa	khaamaanaa	No	amaallaa	amaallaa	Yes	amaallaa	haamaallaa	No
numuw	numuw	Yes	numuw	numuw	Yes	numuw	numuw	Yes	numuw	numuw	Yes
faahaamaa	faahaamaa	Yes	faahaamaa	ttaahaanaa	No	faahaamaa	daama	No	faahaamaa	daama	No
ghaanaamu	ghaanaamu	Yes	ghaanaamu	ghaanaamu	Yes	ghaanaamu	min	No	ghaanaamu	ghaanaamu	Yes
saanaami	saanaami	Yes	saanaami	saanaaa	No	saanaami	saanaami	Yes	saanaami	thiny	No
allaaman	allaaman	Yes	allaaman	sum	No	allaaman	allaaman	Yes	allaaman	allaaman	Yes
allaamun	allaaman	No	allaamun	allaaman	No	allaamun	Aallamon	No	allaamun	Aallamon	No
allaamin	kaadhaaa	No	allaamin	kaadhaaa	No	allaamin	alamin	No	allaamin	alamin	No

Recording 4			Recording 3			Recording 2			Recording 1		
499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match
waanaa	waanaa	Yes	waanaa	waanaa	Yes	waanaa	waanaa	Yes	waanaa	waanaa	Yes
naahrr	naahaata	No	naahrr	naahrr	Yes	naahrr	naahrr	Yes	naahrr	naahaata	No
nibrr	nibrr	Yes	nibrr	nibrr	Yes	nibrr	nibrr	Yes	nibrr	nibrr	Yes
nuwrr	sum	No	nuwrr	min	No	nuwrr	min	No	nuwrr	sum	No
faanaarr	ghaanaamu	No	faanaarr	saanaaa	No	faanaarr	faanaarr	Yes	faanaarr	faanaarr	Yes
saaniyaa	saanaaa	No	saaniyaa	saaniyaa	Yes	saaniyaa	saaniyaa	Yes	saaniyaa	saaeerr	No
hunuw	aaamill	No	hunuw	kaarrnin	No	hunuw	aaamill	No	hunuw	aaamill	No
maakkaanaa	kaallaam	No	maakkaanaa	maakkaanaa	Yes	maakkaanaa	maakkaanaa	Yes	maakkaanaa	maakkaanaa	Yes
naahnu	daaahu	No	naahnu	maahwu	No	naahnu	maahwu	No	naahnu	maahwu	No
dhihni	lliyn	No	dhihni	ayn	No	dhihni	minhu	No	dhihni	diykk	No
kaarrnan	kaarrnan	Yes	kaarrnan	kaarrnin	No	kaarrnan	kaarrnin	No	kaarrnan	kaarrnun	No
kaarrnun	kaarrnun	Yes	kaarrnun	kaarrnun	Yes	kaarrnun	kaarrnun	Yes	kaarrnun	kaarrnun	Yes
kaarrnin	kaarrnan	No	kaarrnin	kaarrnin	Yes	kaarrnin	kaarrnin	Yes	kaarrnin	kaarrnan	No
ghaarraahu	ghaarraahu	Yes	ghaarraahu	maahwu	No	ghaarraahu	maarraahu	No	ghaarraahu	ghaanaamu	No
thaallaahu	faaallaa	No	thaallaahu	ghaanaamu	No	thaallaahu	ghaanaamu	No	thaallaahu	faaallaa	No
hirr	yin	No	hirr	hibrr	No	hirr	hibrr	No	hirr	daahrrin	No
haawaas	haaf	No	haawaas	maallaaf	No	haawaas	haawaas	Yes	haawaas	maallaaf	No
huwid	aaawidu	No	huwid	sum	No	huwid	aaawidu	No	huwid	witrr	No
rraahibaa	rraahibaa	Yes	rraahibaa	rraahibaa	Yes	rraahibaa	rraahibaa	Yes	rraahibaa	rraahibaa	Yes
rraahufaa	rraahufaa	Yes	rraahufaa	rraahufaa	Yes	rraahufaa	rraahufaa	Yes	rraahufaa	rraahufaa	Yes
kaahaarraa	kaahaarraa	Yes	kaahaarraa	ghaattaa	No	kaahaarraa	kkaawa	No	kaahaarraa	kaahaarraa	Yes

Recording 4			Recording 3			Recording 2			Recording 1		
499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match
naawaahaa	naawaahaa	Yes	naawaahaa	naawaahaa	Yes	naawaahaa	dhaallaa	No	naawaahaa	naawaahaa	Yes
minhu	minhu	Yes	minhu	minhu	Yes	minhu	minhu	Yes	minhu	minhu	Yes
fyhi	llaayth	No	fyhi	kkiys	No	fyhi	lliyn	No	fyhi	diykk	No
jjaahan	jjaahan	Yes	jjaahan	dhaaa	No	jjaahan	jjaahun	No	jjaahan	jjaahun	No
jjaahun	jjaahun	Yes	jjaahun	daahrrun	No	jjaahun	jjaahun	Yes	jjaahun	jjaahun	Yes
jjaahin	jjaahan	No	jjaahin	daaaan	No	jjaahin	jjaahin	Yes	jjaahin	jjaahan	No
witrr	waaathan	No	witrr	witrr	Yes	witrr	witrr	Yes	witrr	witrr	Yes
wujjidaa	waajjibaa	No	wujjidaa	wujjidaa	Yes	wujjidaa	waajjibaa	No	wujjidaa	waajjibaa	No
aaawidu	sum	No	aaawidu	haam	No	aaawidu	aaawidu	Yes	aaawidu	sum	No
dhaawuw	dhaawuw	Yes	dhaawuw	ghill	No	dhaawuw	ghill	No	dhaawuw	ghill	No
maahwu	maahwu	Yes	maahwu	maahwu	Yes	maahwu	maahwu	Yes	maahwu	maahwu	Yes
llaahwi	llaahwi	Yes	llaahwi	llaahwi	Yes	llaahwi	llaahwi	Yes	llaahwi	llaahwi	Yes
saahwaa	saahwaa	Yes	saahwaa	saahwaa	Yes	saahwaa	saahwaa	Yes	saahwaa	saahwaa	Yes
jjaarrwan	daahrrun	No	jjaarrwan	jjaarrwun	No	jjaarrwan	jjaarrwun	No	jjaarrwan	jjaarrwun	No
jjaarrwun	daaaun	No	jjaarrwun	jjaarrwin	No	jjaarrwun	jjaarrwin	No	jjaarrwun	jarwun	No
jjaarrwin	jjaarrwin	Yes	jjaarrwin	ghaallaaa	No	jjaarrwin	jjaalliy	No	jjaarrwin	jjaalliy	No
yaad	yaad	Yes	yaad	ghaadh	No	yaad	yin	No	yaad	yaad	Yes
yusrr	yusrr	Yes	yusrr	yusrr	Yes	yusrr	yusrr	Yes	yusrr	yusrr	Yes
yin	daaaan	No	yin	daaaun	No	yin	yin	Yes	yin	yin	Yes
saayaarraa	saaniyaa	No	saayaarraa	saaniyaa	No	saayaarraa	saayaarraa	Yes	saayaarraa	saayaarraa	Yes
ayiyaa	ayiyaa	Yes	ayiyaa	haayuUa	No	ayiyaa	ayiyaa	Yes	ayiyaa	sum	No

Recording 4			Recording 3			Recording 2			Recording 1		
499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match	499 words	Recognised words	Match
saawyi	thaawy	No	saawyi	saanaami	No	saawyi	saawyi	Yes	saawyi	saawyi	Yes
ttaayu	ttaayu	Yes	ttaayu	aaaw	No	ttaayu	ttaayu	Yes	ttaayu	ttaayu	Yes
haayuUa	aaathaa	No	haayuUa	haayuUa	Yes	haayuUa	haayuUa	Yes	haayuUa	haayuUa	Yes
atyan	kaarrnan	No	atyan	rraadhiyaa	No	atyan	rraadhiyaa	No	atyan	atumaa	No
atyun	mudun	No	atyun	atyun	Yes	atyun	mudun	No	atyun	mudun	No
atyin	allaafin	No	atyin	rraadhiyaa	No	atyin	baarrkin	No	atyin	ghafiya	No
	Total	156		Total	158		Total	213		Total	183
	Average	31.2		Average	31.6		Average	42.6		Average	36.6

Table 1- Improved LDPT analysis for the four recordings

Appendix Q

Diacritem analysis

Diacritem analysis

Arabic letter	Name of letter	English letter	Fat ha							
			T overall	R.r overall (%)	T start	R.r start (%)	T midd.	R.r midd. (%)	T end	R.r end (%)
أ	alef	a	12	29.2	10	35	1	0	1	0
ب	baa	b	38	43.4	18	29.2	8	45.8	12	62.5
ت	taa	t	7	17.9	4	0	1	0	2	62.5
ث	thaa	th	16	32.8	12	33.3	2	37.5	2	12.5
ج	jeem	j	28	38.4	15	40	8	35	5	40
ح	haa	h	22	53.4	16	46.9	5	65	1	100
خ	khaa	kh	16	31.3	12	25	3	41.7	1	25
د	daal	d	26	34.6	12	27.1	5	43.9	9	41.7
ذ	thaal	th	16	29.7	8	25	6	37.5	2	50
ر	raa	r	58	38.4	18	61.1	24	29.7	16	17.2
ز	zain	z	16	50	5	60	8	36.5	3	58.3
س	seen	s	32	56.3	18	48.6	9	71.1	5	45
ش	sheen	sh	17	55.9	14	60.7	1	25	2	37.5
ص	saad	s	23	38	16	32.8	4	47.9	3	58.3
ض	dhad	dh	16	15.6	10	22.5	3	0	3	8.3
ط	ta	t	24	35.4	14	21.4	6	37.5	4	68.8
ظ	tha	th	22	38.9	4	0	3	75	2	62.5
ع	ain	a	44	23.9	30	25.8	4	20	10	20
غ	ghain	gh	21	20.2	16	21.9	2	25	3	0
ف	faa	f	28	46.4	17	27.9	4	62.5	7	92.9
ق	qaaf	q	31	37.9	20	26.3	7	59.2	4	31.3
ك	kaaf	k	27	35.2	12	31.3	9	44.4	6	45.8

Arabic letter	Name of letter	English letter	Fat ha							
			T overall	R.r overall (%)	T start	R.r start (%)	T midd.	R.r midd. (%)	T end	R.r end (%)
ل	**laam**	**l**	31	35.5	2	87.5	13	38.6	16	25
م	**meem**	**m**	20	44.2	11	59.1	10	38.6	9	33.3
ن	**noon**	**n**	24	64.6	12	66.7	5	57.9	7	64.3
ه	**haa**	**h**	16	40.6	8	43.4	7	32.1	1	75
و	**waaw**	**w**	28	60.7	22	59.1	5	70	1	100
ي	**yaa**	**y**	13	44.2	5	45	2	29.2	6	50
َ	**Fat ha**	**aa**	665	39.8	-	-	-	-	-	-
َ	**Fat ha**	**a**	7	17.9	-	-	-	-	-	-

Table 1- 'Fat ha' diacritem analysis

Dhamma Diacritem Analysis

Arabic letter	Name of letter	English letter	Dhamma							
			T overall	R.r overall (%)	T start	R.r start (%)	T midd.	R.r midd. (%)	T end	R.r end (%)
أ	alef	a	3	0	1	0	1	0	1	0
ب	baa	b	4	62.5	1	50	1	25	2	87.5
ت	taa	t	3	0	1	0	1	0	1	0
ث	thaa	th	7	0	5	0	1	0	1	0
ج	jeem	j	9	19.4	6	16.7	2	25	1	25
ح	haa	h	7	42.9	3	16.7	2	87.5	1	75
خ	khaa	kh	3	33.3	1	25	1	75	1	0
د	daal	d	10	30	1	100	1	100	7	28.6
ذ	thaal	th	3	58.3	1	0	1	100	1	75
ر	raa	r	6	37.5	4	50	1	0	1	50
ز	zain	z	3	75	1	75	1	50	1	75
س	seen	s	8	34.4	6	41.7	1	25	1	0
ش	sheen	sh	4	56.3	2	25	1	100	1	75
ص	saad	s	5	45	3	25	1	75	1	75
ض	dhad	dh	4	12.5	2	25	1	0	1	0
ط	ta	t	3	8.3	1	0	1	25	1	0
ظ	tha	th	5	20	1	0	1	100	3	0
ع	ain	a	10	15	8	9.4	1	75	1	0
غ	ghain	gh	3	16.7	1	0	1	0	1	50
ف	faa	f	4	43.8	2	50	1	75	1	0

Arabic letter	Name of letter	English letter	Dhamma							
			T overall	R.r overall (%)	T start	R.r start (%)	T midd.	R.r midd. (%)	T end	R.r end (%)
ق	qaaf	q	5	45	3	25	1	50	1	100
ك	kaaf	k	3	66.7	1	50	1	100	1	50
ل	laam	l	8	25	2	62.5	4	6.3	2	25
م	meem	m	5	70	3	58.3	1	100	1	75
ن	noon	n	8	53.3	6	70.8	1	0	1	0
ه	haa	h	9	52.8	1	0	1	100	7	53.6
و	waaw	w	4	56.3	1	25	1	25	2	87.5
ي	yaa	y	4	62.5	2	50	1	75	1	75
ُ	dhamma	u	133	39.3	-	-	-	-	-	-
ُ	dhamma	o	17	10	-	-	-	-	-	-

Table 2- Dhamma diacritem analysis

Kasra diacritem analysis

Arabic letter	Name of letter	English letter	Kasra							
			T overall	R.r overall (%)	T start	R.r start (%)	T midd.	R.r midd. (%)	T end	R.r end (%)
أ	alef	a	3	0	1	0	1	0	1	0
ب	baa	b	4	25	2	37.5	1	25	1	0
ت	taa	t	3	33.3	1	25	1	75	1	0
ث	thaa	th	3	16.7	1	0	1	0	1	50
ج	jeem	j	5	65	1	0	3	75	1	100
ح	haa	h	3	50	1	100	1	50	1	0
خ	khaa	kh	3	8.3	1	0	1	25	1	0
د	daal	d	3	75	1	100	1	100	1	0
ذ	thaal	th	4	0	2	0	1	0	1	0
ر	raa	r	6	37.5	1	75	3	50	2	0
ز	zain	z	3	25	1	25	1	25	1	25
س	seen	s	9	33.3	6	16.7	1	100	1	0
ش	sheen	sh	5	80	1	100	3	83.3	1	50
ص	saad	s	3	41.7	1	25	1	50	1	50
ض	dhad	dh	5	25	3	8.3	1	100	1	0
ط	ta	t	3	41.7	1	0	1	25	1	100
ظ	tha	th	5	0	1	0	3	0	1	0
ع	ain	a	6	54.2	4	62.5	1	75	1	0
غ	ghain	gh	3	8.3	1	0	1	0	1	25

Arabic letter	Name of letter	English letter	Kasra							
			T overall	R.r overall (%)	T start	R.r start (%)	T midd.	R.r midd. (%)	T end	R.r end (%)
ف	faa	f	8	37.5	5	45	2	0	1	75
ق	qaaf	q	4	37.5	2	0	1	100	1	50
ك	kaaf	k	3	83.3	1	75	1	75	1	100
ل	laam	l	6	45.8	1	50	4	43.8	1	50
م	meem	m	5	50	2	100	2	0	1	50
ن	noon	n	3	50	1	100	1	50	1	0
ه	haa	h	3	33.3	1	0	1	100	1	0
و	waaw	w	4	50	1	75	2	12.5	1	100
ي	yaa	y	3	50	1	50	1	50	1	50
ِ	kasra	i	109	39.7	-	-	-	-	-	-
ِ	kasra	e	9	17.5	-	-	-	-	-	-

Table 3- Kasra diacritem analysis

T overall= Test overall
R.r overall= Recognition rate overall
T start, T midd., T end= Test start, Test middle, Test end
R.r start, R.r. midd., R.r. end= Recogition rate start

Appendix R

Diacritem alternatives

Ara. l	Name of letter	position	T	Table3 English letter with diacritic		'Fat ha' alternatives															Table 4 English letter with diacritic
						1		2		3		4		5		6		7			
				Eng. l	R %	Eng. l	r %	Eng. l	r %	Eng. l	r %	Eng. l	r %	Eng. l	r %	Eng. l	r %	Eng. l	r %	
أ	alef	Middle	1	aaa	0	aa	25	aaaa	0	a	0	-	-	-	-	-	-	-	-	aa
أ	alef	End	1	aaa	0	aa	0	aaaa	0	a	0	-	-	-	-	-	-	-	-	aaa
ت	taa	Start	4	ta	0	taa	6.3	tta	0	ttaa	0	-	-	-	-	-	-	-	-	ta
ت	taa	Middle	1	ta	0	taa	0	tta	6.3	ttaa	0	-	-	-	-	-	-	-	-	tta
ض	dhad	Middle	3	dhaa	0	dha	8.3	ddha	0	ddhaa	0	-	-	-	-	-	-	-	-	dha
ظ	tha	Middle	3	thaa	0	tha	8.3	ttha	0	tthaa	0	-	-	-	-	-	-	-	-	thaa
غ	ghain	End	3	ghaa	0	gha	0	ga	0	gaa	8.3	-	-	-	-	-	-	-	-	gaa

Table 1- 'fat ha' diacritem alternatives

Ara.l	Name of letter	position	T	Table3 English letter with diacritic		Alternatives														Table 4 English letter with diacritic
						1		2		3		4		5		6		7		
				Eng. l	R %	Eng. l	r %	Eng. l	r %	Eng. l	r %	Eng. l	r %	Eng. l	r %	Eng. l	r %	Eng. l	r %	
خ	khaa	End	1	khu	0	kho	0	khoo	25	khou	0	ku	0	ko	0	koo	0	kou	0	khoo
ذ	thal	Start	1	dhu	0	dho	0	dhoo	0	dhou	0	thu	25	tho	0	thoo	0	thou	0	thu
ر	raa	Middle	1	ru	0	ro	0	roo	0	rou	0	rru	0	rro	0	rroo	25	rrou	0	rro
س	seen	End	1	su	0	so	0	soo	0	sou	0	ssu	0	sso	0	ssoo	0	ssou	0	su
ض	dhad	Middle	1	dhu	0	dho	0	dhoo	0	dhou	0	ddhu	0	ddho	0	ddhoo	0	ddhou	0	dhu
ض	dhad	End	1	dhu	0	dho	25	dhoo	0	dhou	0	ddhu	0	ddho	0	ddhoo	0	ddhou	0	dho
ط	ta	Start	1	ttu	0	tto	25	ttoo	0	ttou	0	tu	0	to	0	too	0	tou	0	tto
ط	ta	End	1	ttu	0	tto	0	ttoo	0	ttou	25	tu	0	to	0	too	0	tou	0	ttou
ظ	tha	Start	1	tho	0	thu	50	thoo	25	thou	0	tthu	0	ttho	0	tthoo	0	tthou	0	thu
ظ	tha	End	1	tho	0	thu	25	thoo	0	thou	0	tthu	0	ttho	0	tthoo	0	tthou	0	thu
ع	ain	End	1	o	0	ao	0	oo	0	oou	0	ooo	0	ou	0	au	25	aou	0	au
غ	ghain	Start	1	ghu	0	gho	0	ghoo	0	ghou	0	gu	0	go	0	goo	0	gou	0	ghu
غ	ghain	Middle	1	ghu	0	gho	0	ghoo	0	ghou	0	gu	25	go	25	goo	0	gou	0	gu
ف	faa	End	1	fu	0	fo	0	foo	0	fou	0	ffu	0	ffo	0	ffoo	0	ffou	0	fu

Ara.l	Name of letter	position	T	Eng. l	R %	Eng. l	r %	Eng. l	r %	Eng. l	r %	Eng. l	r %	Eng. l	r %	Eng. l	r %	Eng. l	r %	Table 4 English letter with diacritic
ل	laam	Middle	4	llu	6.3	llo	0	lloo	0	llou	0	lu	6.3	lo	6.3	loo	0	lou	0	lu
ن	noon	Middle	1	nu	0	no	25	noo	0	nou	0	nnu	0	nno	0	nnoo	0	nnou	0	no
ن	noon	End	1	nu	0	no	0	noo	0	nou	0	nnu	0	nno	0	nnoo	0	nnou	0	nu
ه	haa	Start	1	hu	0	ho	25	hoo	0	hou	0	hhu	25	hho	25	hhoo	0	hhou	0	ho

Table 2- Dhamma diacritem alternatives

Ara. l	Name of letter	position	T	Table3 English letter with diacritic		Alternatives															Table 4 English letter with diacritic
						1		2		3		4		5		6		7			
				Eng. l	R %	Eng. l	r %	Eng. l	r %	Eng. l	r %	Eng. l	r %	Eng. l	r %	Eng. l	r %	Eng. l	r %	
ب	baa	End	1	be	0	bi	25	bee	0	bie	0	bbi	0	bbe	0	bbee	0	bbie	0	bi
ت	taa	End	1	ti	0	te	0	tee	0	tie	0	tti	0	tte	0	ttee	0	ttie	0	ti
ث	thaa	Start	1	thi	0	the	25	thee	0	thie	0	tthi	0	tthe	0	tthee	0	tthie	0	the
ث	thaa	Middle	1	thi	0	the	0	thee	0	thie	0	tthi	0	tthe	0	tthee	0	tthie	0	thi
ج	jeem	Start	1	jji	0	jje	0	jjee	0	jjie	0	jji	0	jje	25	jjee	25	jjie	0	jje
ح	haa	End	1	hi	0	he	0	hee	0	hie	0	hhi	0	hhe	0	hhee	0	hhie	0	hi
خ	khaa	Start	1	khi	0	khe	0	khee	0	khie	0	ki	0	ke	0	kee	0	kie	0	khi
خ	khaa	End	1	khi	0	khe	0	khee	0	khie	0	ki	0	ke	0	kee	0	kie	0	khi
د	daal	End	1	di	0	de	0	dee	0	die	0	ddi	25	dde	25	ddee	0	ddie	0	ddi
ر	raa	End	2	rri	0	rre	0	rree	0	rrie	0	ri	12.5	re	0	rree	0	rrie	0	ri
س	seen	End	1	si	0	se	25	see	0	sie	0	ssi	25	sse	25	ssee	0	ssie	0	ssi
ض	dhad	Start	3	dhi	8.3	dhe	0	dhee	0	dhie	0	ddhi	0	ddhe	0	ddhee	0	ddhie	0	dhi
ض	dhad	End	1	dhi	0	dhe	25	dhee	0	dhie	0	ddhi	0	ddhe	0	ddhee	0	ddhie	0	dhe

Ara. l	Name of letter	position	T	Table3 English letter with diacritic		Alternatives														Table 4 English letter with diacritic
						1		2		3		4		5		6		7		
				Eng. l	R %	Eng. l	r %	Eng. l	r %	Eng. l	r %	Eng. l	r %	Eng. l	r %	Eng. l	r %	Eng. l	r %	
ط	ta	Start	1	tti	0	tte	0	ttee	0	ttie	0	ti	0	te	0	tee	0	tie	0	tti
ع	ain	End	1	ee	0	e	0	eee	0	i	0	ie	0	ai	25	ae	0	ii	0	ai
غ	ghain	Start	1	ghi	0	ghe	0	ghee	0	ghie	0	gi	0	ge	0	gee	0	gie	0	ghi
غ	ghain	Middle	1	ghi	0	ghe	0	ghee	0	ghie	0	gi	0	ge	0	gee	0	gie	0	ghe
ف	faa	Middle	2	fi	0	fe	0	fee	0	fie	0	ffi	0	ffe	0	ffee	0	ffie	0	fi
ق	qaaf	Start	2	ki	0	ke	0	kee	0	kie	0	kki	12.5	kke	0	kkee	0	kkie	0	kki
م	meem	Middle	2	mi	0	me	12.5	mee	0	mie	0	mmi	0	mme	0	mmee	0	mmie	0	me
ن	noon	End	1	ni	0	ne	0	nee	0	nie	0	nni	0	nne	0	nnee	0	nnie	0	Ni
ه	haa	Start	1	hi	0	he	25	hee	0	hie	0	hhi	0	hhe	0	hhee	0	hhie	0	hi
ه	haa	End	1	hi	0	he	0	hee	0	hie	0	hhi	25	hhe	0	hhee	0	hhie	0	hhi
و	waaw	Middle	2	wi	12.5	we	12.5	wee	0	wie	0	wwi	0	wwe	0	wwee	0	wwie	0	wi

Table 3- kasra diacritem alternatives

Ara. l= Arabic letter   T=Total number of words   Eng. l= English letter   R%=Recognition rate

Appendix S

The transliteration comparison survey

The transliteration comparison survey

A short article was chosen from the newspaper and transliterated using both Alghamdi's transliteration system and the improved transliteration table (SLT). Then it was presented to friends, family and students at the University of Bahrain, and they were also asked to fill in a short survey and 50 people took part.

زوج ..للبيع!
ضاقت أمريكية ذرعاً بتعلق زوجها الشديد بألعاب الفيديو, فأقدمت على سبيل المزاح على عرضه للبيع على موقع الكتروني. و قالت ان زوجها أمضى وقتا طويلا على ألعاب الفيديو فقررت عرضه للبيع. و أشارت الى أنها تلقت عروضاً لشراء زوجها بعد ساعات من نشر عبارة "زوج للبيع لمن يتقدم بالسعر الأعلى".

Please read both transliterations and answer the following questions:

Transliteration 1
Alghamdi's transliteration
Dhaqt amrykyh tharAn btAlq zwjha alshadyd balAab alfydyw, faqdmt Ala sbyl almzah Ala Ardhh llbyA Ala mawqA alktrwny. W qalt an zawjha amdha wqta twyla ala alAab alfydyw fqrrt Ardhh llbyA. W ashart ala anha tlqt Arwdha lshraa zawjha bAd saAat mn nshr Abart "zwj llbyA lmn ytqdm balsAr alaAla".

Transliteration 2
SLT table
Dhakt amrrykkyh dharran btallk zzawjjjha allshadyd ballaab allfydyw, fakdmt ala sbyll allmzzah alla ardhh llllbya ala mawka allkktrrwny. W kallt an zzawjjha amdha wakta ttawyllan alla allab allfydyw fkrrrrt arrdhh llllbya. W asharrt alla anha tllkt arrwdha llshrraa zzawjjha bad saaat mn nshrr abarrt "zzwjj llllbya llmn ytkdm ballsarr allaala".

Please circle one of the numbers for each question.

Question 1
Transliteration 1 is

1 Very Easy to read	2 Easy	3 OK	4 Challenging	5 Difficult

Question 2

Transliteration 2

1 Very Easy to read	2 Easy	3 OK	4 Challenging	5 Difficult

Question 3

I am able to read transliteration 1 without referring to the Arabic writing.

1 Strongly disagree	2 Disagree	3 Neither agree or disagree	4 Agree	5 Strongly agree

---

Question 4

I am able to read transliteration 2 without referring to the Arabic writing.

1 Strongly disagree	2 Disagree	3 Neither agree or disagree	4 Agree	5 Strongly agree

---

(Preference)
Please circle one of the numbers for each question.

Question 1
Which transliteration would you prefer to use?

Transliteration 1	Transliteration 2

You like the above transliteration because:

The results are as follows:
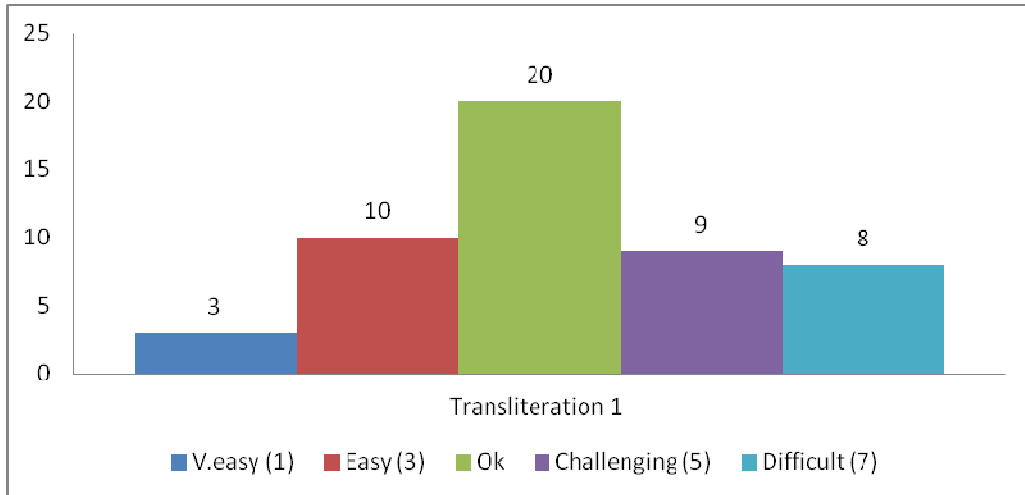
Figure 17- Transliteration 1 chart

20 people found reading transliteration 1 Ok while 10 thought that it was easy, 3 mentioned that it was very easy. And 9 found it challenging while only 8 thought it was difficult.
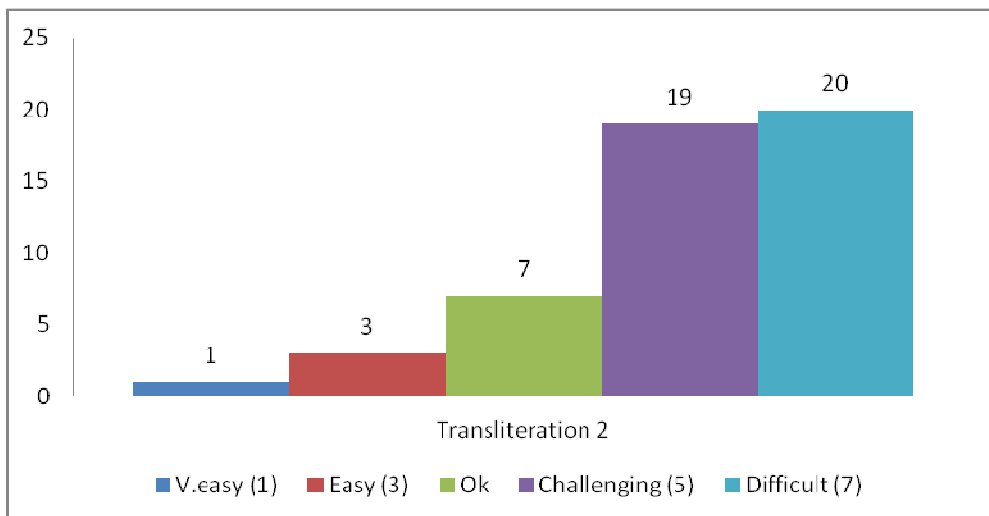
Figure18-Trasnliteration 2 chart

20 thought that transliteration 2 was difficult to read, 19 found it challenging, 7 mentioned that it was ok, and 3 said that it was easy while only 1 thought that it was very easy.
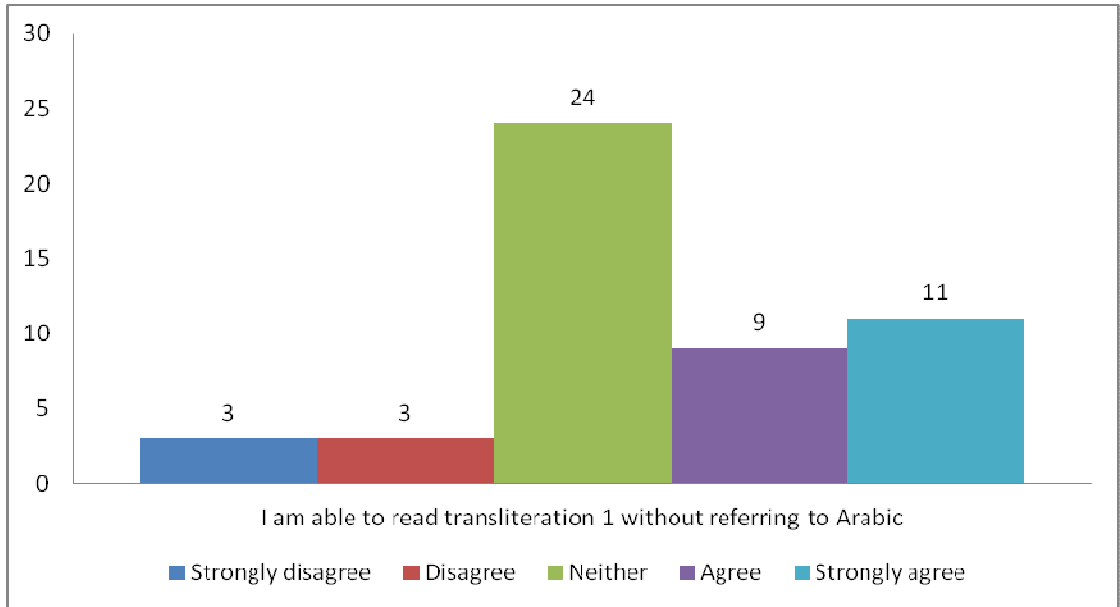
Figure 19- Ability to read transliteration 1 chart
11 strongly agreed that they were able to read transliteration 1 without referring to Arabic, 9 agreed, 24 neither agreed nor disagreed,   and 3 disagreed and 3 strongly agreed.
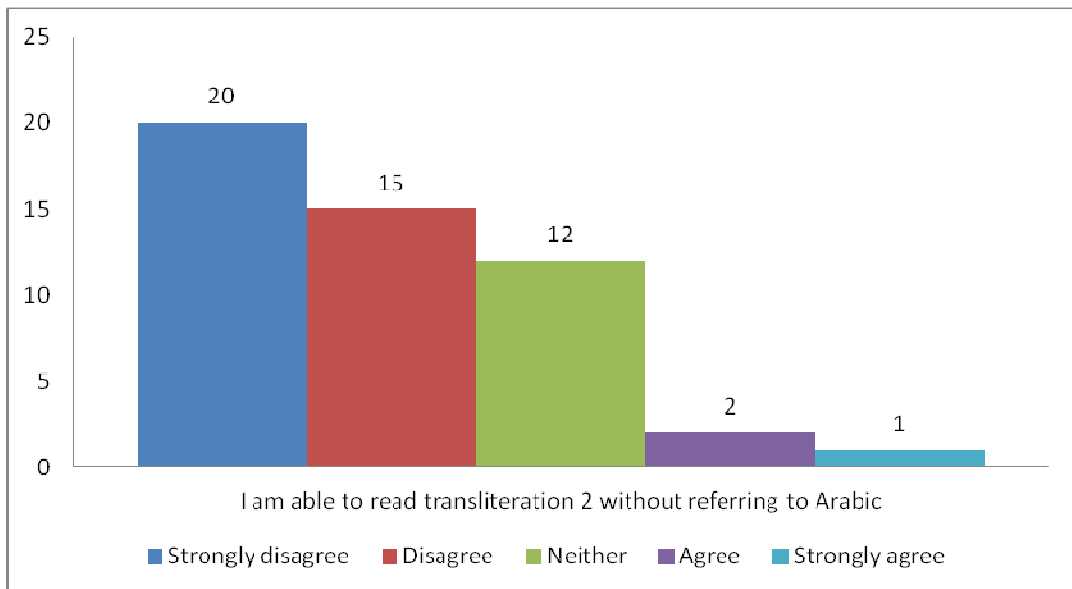
Figure 20- Ability to read transliteration 2 chart
20 strongly disagreed that they were able to read transliteration 2 without referring to Arabic, 15 disagreed, 12 neither agreed nor disagreed,   2 agreed and only 1 strongly agreed.
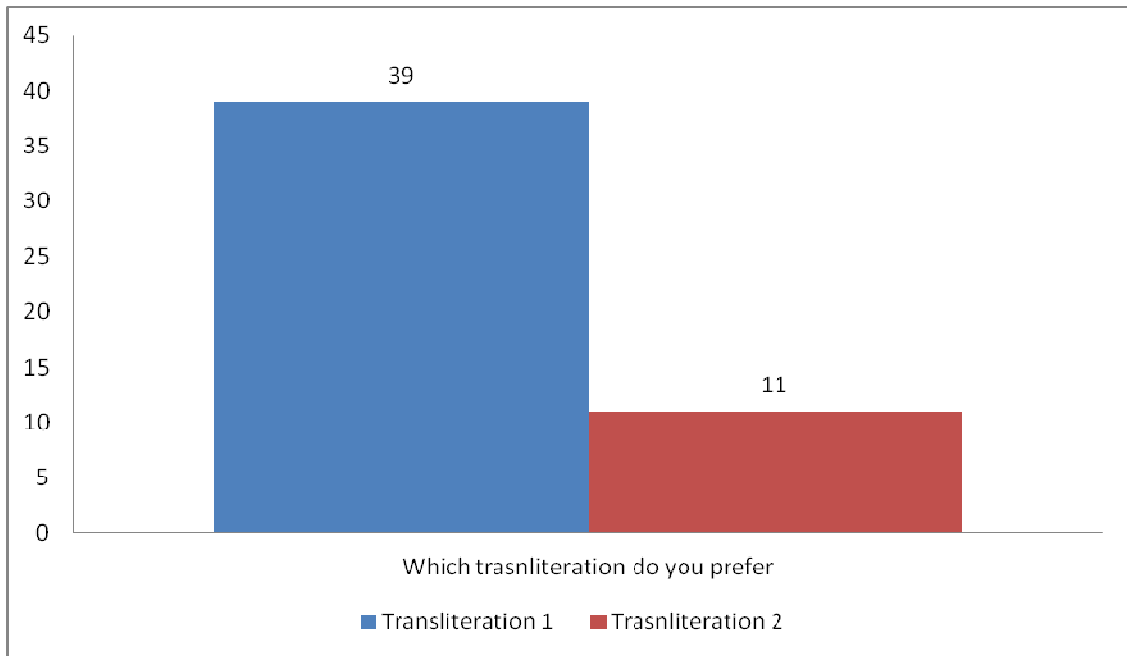
Figure 21- Transliteration 1 and 2 comparison chart

When asked to choose the transliteration they preferred, 39 thought that transliteration 1 was better, whilst 11 chose transliteration 2.

The reason for choosing transliteration 1 as mentioned in the survey is that it contained plain simple letters, while the second contained some doubled letters which made it more difficult to read.

Transliteration 2 is specially made for words in a list and altered to match the recognition of the speech recognition. On the other hand Alghamdi's transliteration was specially made for big chunks of texts. That is why it is easier to read.

Appendix T

....

New list of (kha) words

New list of 'kha' words

	Group 1 Arabic	Group1 English	Group 2 Arabic	Group 2 English
13	أخت	aukht	دُخر	thukhr
30	خبط	khaabaatt	خَصَم	khaSam
71	يخت	yaakhti	صَخر	Sakhri
84	ثخن	thaakhn	فَخر	fakhr
103	خجل	khaajjaall	خَطِر	khaTar
147	ضخه	Dhakhahu	صَخَبُ	Sakhab
148	خدك	khaadukk	خَطُك	khaTuk
149	خث	khaath	خَر	khar
150	خشع	khaashaaa	خَدَعَ	khadaaa
151	خص	khaasaa	خَضَ	khaDh
152	ذخر	thakhira	سَخِرَ	sakhira
153	خزق	khazaqa	خَفَقَ	khafaqa
154	خسف	kasafa	خَلَفَ	khalafa
155	خمن	khamana	خَفَفَ	khafafa
156	خوي	khaawy	خَدِي	khady
157	خس	khaas	خَل	khal
158	خدر	khidrr	خِصر	khiSr
159	خمس	khums	خُدر	khudr
160	بخس	bakhasa	دَخَلَ	dakhala
161	بخل	baakhillaa	نَخِرَ	nakhira
162	رخص	rakhusa	صَخُبَ	Sakhuba
163	صرخ	sarakha	شَمَخَ	shamakha
164	مخ	mukhi	دُخ	dukh
165	سلخ	saallkhu	شَرخُ	sharkhu
166	بذخٌ	bathakhun	وَسَخٌ	wasakhun
167	بذخ	bathakhin	وَسَخِ	wasakhin
168	بذخاً	bathakhan	وَسَخاً	wasakhan
393	خلف	khaallfu	خَفَرُ	khafru

Table 1- New list of 'kha' words

Appendix U

Alghamdi and improved DT table comparison
Accuracy evaluation by the two experts

Appendix U: Original evaluation results by the two experts

	Alghamdi	T	R1	R2	R3	R4	Improved table 2	T	R1	R2	R3	R4
1	dhaaa	70	71	70	72	70	Dhaaa	70	71	70	70	70
2	aathin	52	55	50	45	50	Aaathen	54	60	60	60	60
3	saagha	50	45	45	45	45	Saaghaa	55	60	60	60	55
4	ethaa	70	70	70	70	70	Edhaa	65	70	71	70	70
5	zaar	82	80	70	75	80	Zzaarr	80	85	85	85	85
6	qaas	71	70	71	70	70	Kaas	65	65	70	65	70
7	aamal	91	80	81	80	80	Aaamaall	90	90	90	90	90
8	jatha	82	75	75	80	80	jjaatha	85	80	85	85	85
9	shaah	65	65	65	65	65	shaah	65	65	65	65	65
10	taaf	81	70	69	70	70	ttaaf	80	85	85	85	85
11	hayaaa	60	60	50	50	50	haayaaaaa	50	50	50	50	50
12	kaas	46	40	40	40	40	kkaaasu	50	50	50	50	50
13	aukht	60	50	50	55	55	aukht	60	50	50	55	55
14	baada	76	60	60	55	55	baadaa	80	80	85	85	80
15	aaw	92	91	90	90	90	aaaw	90	91	90	89	90
16	aakala	92	80	80	80	85	aaakkaallaa	90	90	90	92	90
17	saaal	50	40	40	45	40	saaaaall	45	45	50	45	50
18	dhuUl	51	50	50	50	50	dhuUll	46	45	50	50	50
19	baiisa	56	60	60	55	60	baaiisaa	55	55	55	55	55
20	baraa	46	30	30	30	30	baarraau	60	60	60	60	60
21	swai	44	40	40	40	40	swai	44	40	42	40	40
22	daaaan	45	40	40	45	40	daaaan	45	40	40	45	41
23	daaaun	45	40	40	40	40	daaaun	45	40	40	41	40
24	daaain	45	45	45	45	45	daaain	45	45	44	45	45
25	thaby	64	62	60	60	60	thaaby	70	72	70	70	70
26	dhaba	44	30	40	30	30	dhaaba	50	45	45	45	50
27	bazagha	62	72	70	70	70	baazzaaghaa	70	70	70	70	71
28	basal	63	62	60	50	60	baasaall	62	70	73	70	70
29	bahaq	52	50	50	50	50	baahaak	50	55	55	50	50
30	khabat	46	40	40	40	40	khaabaatt	50	50	50	50	50
31	kaba	30	30	25	30	30	kkaabaa	40	55	55	55	55
32	thanb	62	62	70	62	70	dhaanb	60	60	62	62	62
33	bashima	82	80	80	80	80	baashimaa	80	80	80	80	80
34	saba	62	60	65	65	60	saabaa	65	65	65	65	65
35	farabu	84	75	80	80	80	faarraabu	90	90	90	90	90

	Alghamdi	T	R1	R2	R3	R4	Improved table 2	T	R1	R2	R3	R4
36	nasab	83	80	78	80	80	naasaab	85	80	85	85	80
37	wajiba	91	90	90	90	90	waajjibaa	90	90	90	90	90
38	thabata	82	80	80	80	80	thaabaata	80	80	80	80	85
39	batala	62	50	50	50	50	baattaallaa	55	60	60	60	60
40	bishr	82	80	85	85	85	beshrr	85	85	85	85	85
41	burj	90	90	90	90	90	burrjj	90	90	90	90	90
42	jubila	82	80	80	80	80	jjubellaa	85	85	85	85	85
43	rabata	62	50	50	50	50	rraabaattaa	65	70	70	70	70
44	subul	82	70	70	70	70	subull	80	80	80	80	80
45	halaba	70	62	62	65	65	haallaabaa	65	65	70	65	65
46	qalbi	62	70	70	70	70	kaallbe	60	65	65	65	62
47	naabu	80	80	80	80	80	naabu	80	80	80	80	80
48	thawban	62	70	70	75	75	thaawban	65	65	65	70	65
49	thawbun	62	70	70	70	70	thaawbun	65	65	70	65	70
50	thawbin	62	75	70	70	70	thaawbin	65	70	65	70	70
51	taht	70	70	70	60	60	taht	70	70	70	60	60
52	dhamat	62	70	70	60	60	dhaamaat	65	70	70	70	70
53	tathil	52	50	55	50	50	tathell	55	60	55	55	55
54	satat	42	40	40	40	40	saattaat	45	50	50	50	50
55	sakat	46	40	42	40	42	saakkaat	50	45	45	45	45
56	tharat	42	42	40	40	40	dhaarraat	40	40	40	40	40
57	hazat	58	50	50	60	60	haazzaat	65	70	65	65	70
58	shadat	62	55	55	50	50	shaadaat	65	65	65	70	65
59	thanat	60	70	70	70	70	thaanaat	55	60	55	55	60
60	jafat	81	75	75	75	75	jjaafaat	80	80	80	80	80
61	otw	20	35	35	30	35	otw	20	35	35	30	35
62	ghat	50	55	50	50	50	ghaat	55	60	60	60	60
63	taqy	55	60	60	60	60	taky	50	55	55	50	50
64	tamr	72	80	80	80	80	tamrr	75	75	80	75	75
65	tyn	65	60	55	55	55	tyn	65	60	55	55	55
66	twt	65	50	50	50	50	twt	65	50	50	50	50
67	qatala	66	70	70	70	65	kaatallaa	60	60	60	65	60
68	sutira	72	80	80	80	80	sutirraa	75	85	80	80	80
69	atuma	51	40	40	40	40	atumaa	55	55	55	60	55
70	yumitu	72	70	70	70	70	yumitu	72	70	70	70	70
71	yakhti	42	42	42	42	40	yaakhti	45	45	50	50	50
72	nahata	50	50	50	50	50	naahaata	45	42	42	42	42
73	samtun	42	40	42	42	40	saamtun	45	45	45	42	45
74	samtan	42	42	45	42	42	saamtan	45	45	42	45	42
75	samtin	42	42	42	45	42	saamtin	45	45	42	45	42

	Alghamdi	T	R1	R2	R3	R4	Improved table 2	T	R1	R2	R3	R4
76	thulth	62	70	70	70	70	thollth	60	75	75	75	70
77	thaqaf	42	45	45	50	50	thaakaaf	40	40	40	40	40
78	makatha	65	70	70	70	65	maakkaathaa	60	55	55	60	60
79	ghath	46	55	55	55	55	ghaath	50	50	50	55	50
80	hadath	44	35	35	40	40	haadaath	42	50	50	50	50
81	sharath	60	60	60	60	60	shaarraath	55	55	55	55	55
82	ath	50	50	42	42	42	ath	50	50	42	42	42
83	thawy	51	45	50	50	50	thaawy	55	55	55	55	55
84	thakhn	46	50	50	45	45	thaakhn	50	50	50	50	50
85	bathahu	61	60	55	60	55	baathaahu	65	65	65	70	70
86	thabata	80	80	80	80	80	thaabaattaa	80	85	80	85	80
87	thaja	81	80	80	80	80	thaajjaa	85	90	90	90	90
88	thiny	80	75	75	75	75	thiny	80	75	75	75	75
89	thulat	33	40	35	40	40	thollaat	35	40	40	40	40
90	wathaba	80	70	80	80	70	waathaabaa	80	80	85	85	80
91	othira	42	30	30	30	30	othirraa	40	40	40	50	50
92	juthw	31	40	40	40	35	jjuthw	35	45	45	35	35
93	aaatha	40	35	35	35	40	aaathaa	42	40	40	40	42
94	rathi	71	70	70	70	70	rraathi	75	75	80	80	80
95	bathu	72	80	80	80	80	baatho	75	80	75	80	80
96	thuluthin	66	65	60	65	65	tholluthin	65	70	70	70	65
97	thuluthun	65	65	60	60	60	tholluthun	65	65	70	70	65
98	thuluthan	65	60	65	65	60	tholluthan	65	70	65	65	70
99	lujaj	90	90	90	90	90	llujjaajj	85	85	85	85	85
100	jaraka	80	80	80	80	80	jjaarraakkaa	75	75	75	75	75
101	dhaja	30	20	20	20	20	dhaajjaa	35	40	40	40	35
102	jas	26	30	25	30	25	jjaas	25	30	30	30	30
103	khajal	42	50	50	50	50	khaajjaall	40	50	50	50	45
104	jahatha	40	30	30	30	30	jjaahaathaa	35	35	35	35	35
105	tajan	44	40	40	45	45	ttaajjaan	50	50	50	50	50
106	shaja	81	80	80	80	80	shaajjaa	80	85	85	85	85
107	ajaza	62	50	50	50	50	ajjaazzaa	70	70	70	70	70
108	sajaa	61	60	55	55	55	saajjaaa	70	70	75	75	70
109	juthm	50	60	65	65	65	jjudhm	40	45	40	45	45
110	jady	90	80	82	80	80	jjaady	90	90	90	90	90
111	jaza	72	60	60	63	60	jjaazzaa	75	75	70	75	75
112	haja	71	65	65	65	70	haajjaa	75	75	70	70	70
113	jawq	70	55	55	60	60	jjaawk	65	65	65	65	65
114	jamal	89	80	85	85	85	jjaamaall	85	90	90	90	90
115	juhd	77	75	80	75	75	jjuhd	75	80	75	80	75

	Alghamdi	T	R1	R2	R3	R4	Improved table 2	T	R1	R2	R3	R4
116	jidu	78	75	80	75	70	jjidu	80	80	80	80	85
117	wajada	88	85	85	85	85	waajjaadaa	90	90	90	90	90
118	aajidu	86	85	85	85	85	aaajjidu	90	90	90	90	90
119	hujub	64	62	60	60	60	hujjub	70	75	75	75	75
120	daraja	82	80	81	80	80	daarraajjaa	80	80	80	80	80
121	sarju	81	80	80	80	80	saarrju	85	80	80	80	80
122	wahaji	71	75	75	70	70	waahaajji	70	70	70	70	70
123	ewajan	55	50	55	50	55	eewaajjan	50	55	55	55	60
124	ewajun	55	50	55	45	40	eewaajjun	50	55	50	50	50
125	ewajin	55	55	50	45	55	eewaajjin	50	50	55	50	55
126	hadhara	44	42	42	40	42	haadhaarraa	50	55	50	50	55
127	qazah	43	40	40	45	40	kaazzaah	50	55	55	55	55
128	suhuf	45	45	45	45	45	suhuf	45	45	45	45	45
129	hathw	51	50	50	51	50	haadhw	55	55	60	55	55
130	hatama	55	50	50	55	55	haattaamaa	50	50	50	50	50
131	hasan	55	55	55	50	50	haasaan	60	60	60	55	60
132	haka	55	45	45	45	50	haakkaa	60	55	55	55	55
133	halahu	42	35	35	35	40	haallaahu	40	40	40	40	40
134	hay	56	55	55	55	55	haay	60	65	65	65	65
135	hamala	71	70	72	70	70	haamaallaa	75	75	77	75	75
136	hibr	70	75	70	70	75	hibrr	70	75	75	75	75
137	husn	70	60	65	65	65	husn	70	60	65	65	65
138	tahana	51	45	45	45	50	ttaahaanaa	55	60	60	60	60
139	suhub	55	70	70	70	70	suhub	55	70	70	70	70
140	yahilu	55	55	55	55	55	yaahillu	60	60	60	60	60
141	masaha	55	55	50	50	55	maasaahaa	50	55	55	55	55
142	farahi	43	40	42	50	42	faarraahi	50	50	50	50	50
143	marahu	51	50	50	50	50	maarraahu	55	60	60	60	60
144	qazahan	70	70	60	60	60	kaazzaahan	65	62	65	65	65
145	qazahin	70	70	60	70	70	kaazzaahin	65	62	62	62	62
146	qazahun	70	70	70	60	60	kaazzaahun	65	65	62	65	62
147	dhakhahu	34	20	20	15	15	dhaakhaahu	40	40	45	45	45
148	khaduk	42	30	30	30	35	khaadukk	40	45	45	45	45
149	khath	45	50	50	50	50	khaath	50	50	50	50	55
150	khashaa	26	25	25	25	25	khaashaaa	30	30	35	30	30
151	khasa	21	20	25	25	20	khaasaa	30	30	30	30	30
152	thakhara	32	25	25	30	30	dhaakhaarraa	30	40	35	40	30
153	khazaqa	25	30	20	20	25	khaazzaakaa	20	30	30	30	30
154	khasafa	50	50	50	50	50	khaasaafaa	55	55	55	55	55
155	khamana	63	60	60	60	60	khaamaanaa	60	55	55	55	55

	Alghamdi	T	R1	R2	R3	R4	Improved table 2	T	R1	R2	R3	R4
156	khawy	61	50	50	50	50	khaawy	65	65	60	65	60
157	khas	58	55	50	60	50	khaas	60	62	62	62	62
158	khidr	61	65	65	60	60	khidrr	60	62	62	62	62
159	khums	60	60	60		60	khums	60	62	62	62	62
160	bakhasa	42	42	42	42	42	baakhaasaa	45	45	50	50	50
161	bakhila	60	65	65	65	65	baakhillaa	60	65	65	65	65
162	rakhusa	44	40	40	40	40	rraakhusaa	42	40	40	40	40
163	sarakha	42	42	42	45	45	saarraakhaa	45	50	50	50	50
164	mukhi	62	70	70	65	65	mukhi	62	70	70	65	65
165	salkhu	43	50	50	50	50	saallkhu	50	55	55	55	55
166	bathakhun	31	20	20	20	25	baadhaakhun	35	35	35	35	35
167	bathakhin	30	25	20	20	20	baadhaakhin	35	40	40	35	35
168	bathakhan	30	25	20	20	20	baadhaakhan	35	40	35	40	35
169	dhid	20	15	15	15	15	dhid	20	15	15	15	15
170	zand	92	85	85	85	85	zzaand	90	90	90	90	90
171	rasada	65	60	60	60	65	rraasaadaa	70	70	75	70	75
172	qadam	60	55	60	55	60	kaadaam	55	60	60	60	60
173	tawd	44	40	40	40	40	ttaawd	50	55	55	55	55
174	dasa	75	80	80	80	80	daasaa	70	70	70	70	70
175	dagl	60	50	50	50	50	daagll	60	55	55	55	55
176	daahu	22	20	20	20	20	daaahu	25	25	25	25	25
177	daf	80	80	80	80	80	daaf	80	80	80	80	80
178	dama	46	35	35	35	35	daama	50	45	50	45	50
179	dub	90	90	90	90	90	dub	90	90	90	90	90
180	diyk	80	80	80	80	80	diykk	80	85	85	85	85
181	nadaba	91	90	85	90	90	naadaabaa	90	90	90	90	90
182	hudida	66	50	50	50	50	hudidaa	70	80	80	80	80
183	mudun	90	85	85	85	90	mudun	90	85	85	85	90
184	sada	80	65	65	65	65	saadaa	75	70	70	70	70
185	ahdu	50	40	40	40	40	ahdu	50	40	40	40	40
186	mahdi	88	85	85	85	85	maahdi	90	90	90	90	90
187	waadan	64	45	45	50	45	waaadan	70	75	70	70	70
188	waadun	64	45	50	50	45	waaadun	70	70	70	75	70
189	waadin	65	45	50	50	50	waaadin	70	75	70	70	75
190	thaky	70	70	65	70	70	dhaakky	75	75	75	75	75
191	thama	70	70	70	70	65	dhaamaa	70	75	75	75	75
192	thala	70	65	70	70	70	dhaallaa	70	70	70	70	75
193	fath	80	75	75	75	75	faadh	75	75	75	70	75
194	qathaa	61	60	60	60	60	kaadhaaa	65	60	65	60	60
195	shatha	72	70	70	70	70	shaadhaa	75	75	80	80	80

	Alghamdi	T	R1	R2	R3	R4	Improved table 2	T	R1	R2	R3	R4
196	thawd	71	65	65	65	65	dhaawd	70	80	80	80	80
197	thiib	30	20	20	20	20	dhiib	25	30	30	30	30
198	thaab	61	55	55	55	60	dhaab	60	60	55	60	55
199	thul	80	80	80	80	80	dhull	75	75	75	80	80
200	kathiba	82	80	80	80	75	kkaadhibaa	80	85	80	85	80
201	athara	45	40	35	35	40	adhaarraa	50	55	50	55	55
202	aathuna	62	62	62	60	62	aaadhunaa	65	70	70	70	70
203	shahatha	42	42	42	42	42	shaahaadhaa	50	55	55	55	55
204	munthu	70	70	70	70	70	mundhu	65	62	65	60	65
205	mUthi	55	50	50	50	55	mUdhi	55	55	55	55	55
206	fathan	60	75	75	75	75	faadhan	55	55	60	55	55
207	fathun	60	70	75	75	75	faadhun	55	55	55	55	60
208	fathin	60	75	75	75	75	faadhin	55	55	55	60	60
209	thahara	65	65	65	60	65	thaahaarraa	65	65	60	60	60
210	qarn	67	65	65	65	65	kaarrn	65	60	60	60	60
211	rakala	88	80	85	85	85	rraakkaallaa	85	90	90	90	90
212	dhara	50	50	50	50	50	dhaarraa	55	55	55	55	55
213	ragw	35	30	30	30	35	rraagw	40	45	45	45	45
214	tayr	61	55	55	60	60	ttaayrr	60	60	60	65	60
215	sir	80	80	80	80	80	sirr	85	85	90	85	90
216	rad	90	90	90	90	90	rraad	90	90	90	90	90
217	ruba	90	80	80	80	80	rrubaa	90	90	90	90	90
218	surur	82	70	70	70	70	surrurr	85	90	90	90	90
219	harama	67	50	50	50	50	haarraamaa	70	75	75	75	75
220	siry	80	80	80	80	80	sirry	80	80	80	85	85
221	fatara	61	55	60	60	60	faattaarraa	60	60	55	55	55
222	juhri	62	55	55	50	50	jjuhrri	65	65	60	65	65
223	fikri	83	80	80	80	80	fikkrri	85	90	90	85	90
224	dahrun	84	80	80	80	80	daahrrun	85	90	85	85	85
225	dahrin	84	80	80	80	80	daahrrin	85	85	90	90	85
226	dahran	84	80	80	80	80	daahrran	85	90	85	90	90
227	zafa	80	80	80	80	80	zzaafaa	80	85	80	85	80
228	zaama	50	45	45	45	45	zzaaamaa	42	45	45	45	45
229	zaky	80	80	80	80	80	zzaakky	85	90	90	90	90
230	zuhal	65	55	55	55	55	zzuhaall	65	75	75	75	75
231	zaraa	75	55	60	60	60	zzaarraaa	80	80	80	80	80
232	zir	90	90	90	90	90	zzirr	90	90	90	90	90
233	ruziq	65	62	62	62	62	rruzzik	60	65	60	60	60
234	azafa	62	55	55	55	55	azzaafaa	70	70	70	70	70
235	juzur	80	80	85	80	80	jjuzzurr	80	80	80	85	80

	Alghamdi	T	R1	R2	R3	R4	Improved table 2	T	R1	R2	R3	R4
236	faaza	85	80	85	85	85	faazzaa	85	90	90	90	90
237	jawzu	85	85	80	80	85	jjaawzzu	90	90	90	90	90
238	kanzi	80	85	80	80	80	kkaanzzi	80	80	80	80	80
239	filizan	80	75	80	80	75	fillizzan	85	85	85	85	85
240	filizun	80	80	80	75	80	fillizzun	85	85	85	85	85
241	filizin	80	80	80	80	80	fillizzin	85	85	85	85	85
242	shams	80	80	80	75	80	shaams	80	80	80	80	80
243	ghasala	65	65	60	65	65	ghaasaallaa	60	60	60	60	60
244	sahw	42	40	40	40	40	saahw	45	40	45	40	45
245	kys	45	45	40	45	40	kkys	45	50	50	50	50
246	dhirs	45	50	50	50	50	dhirrs	50	50	50	50	55
247	sum	90	90	90	90	90	sum	90	90	90	90	90
248	sakaba	85	85	85	85	85	saakkaabaa	85	90	90	90	85
249	sihr	42	40	40	40	40	sihrr	45	50	50	50	50
250	rusul	80	80	80	80	80	rrusull	80	80	85	80	85
251	asal	60	50	50	50	50	asaall	65	65	70	65	70
252	nasiya	75	80	80	80	80	naasiyaa	80	80	85	90	90
253	habasa	65	50	50	50	50	haabaasaa	70	70	70	70	70
254	harasa	65	50	50	50	50	haarraasaa	70	70	75	70	70
255	farasi	85	80	80	85	85	faarraasi	85	85	85	85	85
256	orsan	62	55	55	60	60	orrsan	65	65	70	65	70
257	orsun	62	55	55	60	55	orrsun	65	70	70	70	70
258	orsin	62	62	60	55	60	orrsin	65	62	62	62	62
259	shathw	65	60	60	60	60	shaadhw	60	60	60	60	60
260	shas	60	62	62	62	60	shaas	70	65	65	65	65
261	shathaf	60	60	62	62	60	shaathaaf	65	65	70	65	70
262	shat	60	60	62	62	60	shaatt	60	60	70	70	70
263	shugl	60	60	60	60	62	shugll	65	70	65	70	65
264	qash	65	62	65	62	62	kaash	60	65	65	65	65
265	shak	80	70	80	80	70	shaakk	80	80	85	85	85
266	nashiz	80	80	80	80	80	naashizz	75	75	70	70	75
267	shahy	75	75	75	75	75	shaahy	75	75	75	80	80
268	shajar	85	85	80	85	85	shaajjaarr	90	90	90	90	90
269	shibl	85	80	85	80	85	shibll	90	90	90	90	90
270	shukr	85	80	85	85	85	shukkrr	90	90	90	90	90
271	washm	85	85	85	85	80	waashm	85	85	80	85	85
272	rushida	85	85	80	85	85	rrushidaa	85	80	85	85	85
273	aashudu	80	85	85	85	85	aaashudu	85	85	90	85	90
274	rasha	85	85	80	85	80	rraashaa	85	85	90	85	85
275	rimshu	85	80	80	80	80	rrimshu	90	90	90	90	90

	Alghamdi	T	R1	R2	R3	R4	Improved table 2	T	R1	R2	R3	R4
276	ryshi	75	65	65	62	60	rryshi	80	75	80	80	80
277	kabshan	80	75	75	80	80	kkaabshan	80	80	80	80	85
278	kabshun	80	80	80	80	75	kkaabshun	80	85	80	85	80
279	kabshin	80	80	75	80	80	kkaabshin	80	80	80	80	80
280	qasa	30	30	30	25	25	kaasaa	25	25	25	25	25
281	sum	50	42	40	42	42	sum	50	42	40	42	42
282	sanaa	35	35	35	35	35	saanaaa	40	45	45	45	45
283	sah	50	55	55	50	50	saah	55	60	60	60	60
284	wasy	50	40	40	40	40	waasy	55	55	60	55	60
285	suws	35	30	30	30	30	suws	35	30	30	30	30
286	sayd	60	55	55	55	55	saayd	65	70	70	70	70
287	sihr	30	25	25	25	25	sihrr	30	30	30	30	30
288	asara	20	20	20	25	20	asaarraa	30	40	40	40	40
289	nusira	62	62	62	62	62	nusirraa	60	62	62	60	60
290	yasudu	60	60	60	60	60	yaasudu	60	62	60	62	60
291	rasa	50	50	50	50	55	rraasaa	55	50	50	50	50
292	qursi	42	30	30	30	30	kurrsi	40	40	40	40	40
293	fasun	50	55	50	50	50	faasun	55	60	55	55	55
294	fasa	50	50	50	55	55	faasaa	55	55	55	55	55
295	fasin	50	50	50	55	50	faasin	55	55	55	60	55
296	dhaghath	20	15	15	15	15	dhaaghaath	20	20	20	20	20
297	wadhaa	15	15	15	15	15	waadhaaa	20	15	20	20	20
298	dhana	50	42	45	42	50	dhaanaa	50	50	50	45	50
299	dhala	50	50	45	50	45	dhaallaa	55	55	55	55	50
300	dhyq	20	20	20	20	20	dhyk	20	15	20	15	20
301	dhafar	50	50	50	45	50	dhaafaarr	50	50	50	55	50
302	dharaba	50	45	50	50	50	dhaarraabaa	55	60	60	60	60
303	dhuha	15	15	15	15	15	dhuha	15	15	15	15	15
304	dhidu	30	25	20	25	20	dhidu	30	25	20	25	20
305	radhiya	30	30	35	35	35	rraadhiyaa	30	35	35	53	35
306	adhud	15	15	15	15	20	adhud	15	15	15	15	20
307	fadhala	30	30	30	30	30	faadhaallaa	35	40	40	40	40
308	maradha	30	30	25	25	30	maarraadhaa	30	30	35	30	30
309	aradha	15	20	15	20	15	arraadhaa	20	20	20	15	20
310	aardhi	25	25	25	25	25	aaarrdhi	25	25	20	25	25
311	qardhan	15	20	15	15	15	kaarrdhan	20	20	20	20	15
312	qardhun	15	20	20	20	15	kaarrdhun	20	15	15	20	20
313	qardhin	15	15	20	15	20	kaarrdhin	20	20	20	20	15
314	taq	15	15	15	15	15	ttuk	20	20	20	20	20
315	hatala	60	55	55	50	55	haattaallaa	55	55	55	55	55

	Alghamdi	T	R1	R2	R3	R4	Improved table 2	T	R1	R2	R3	R4
316	tamaa	15	15	15	15	15	ttaamaaa	20	20	25	20	25
317	tib	45	50	50	50	50	ttib	50	50	55	50	50
318	tabaa	15	15	15	15	15	ttaabaaa	20	20	20	20	20
319	watan	50	50	50	50	50	waattaan	55	55	50	55	55
320	ratib	60	60	55	60	55	rraattib	70	70	75	70	75
321	otuf	40	50	50	55	50	otuf	40	50	50	50	50
322	qirtu	15	15	15	15	15	kirrtu	20	25	20	25	20
323	wasati	30	30	30	30	25	waasaatti	35	40	40	40	40
324	basata	30	30	25	30	30	baasaattaa	30	30	30	30	30
325	nuqatan	15	15	15	15	15	nukaattan	15	20	15	20	15
326	nuqatun	15	15	20	15	20	nukaattun	15	15	20	20	15
327	nuqatin	15	15	15	15	15	nukaattin	15	20	20	20	20
328	thahar	25	20	20	20	20	thaahaarr	30	30	30	30	35
329	kathu	50	55	50	55	50	kkaatho	45	45	45	40	45
330	wathafa	50	60	60	60	60	waathaafaa	45	55	55	55	55
331	tharf	60	60	55	60	60	thaarrf	60	65	65	65	60
332	thifr	60	60	60	55	60	thefrr	65	65	70	65	70
333	thul	60	50	50	60	50	tholl	65	65	70	70	65
334	nathara	60	60	60	50	60	naathaarraa	65	65	70	70	70
335	nathufa	50	55	50	60	50	naathofaa	55	55	60	60	60
336	athima	25	30	30	30	30	athemaa	30	30	40	40	40
337	hafatha	25	25	25	30	25	haafaathaa	25	25	25	25	25
338	qaythi	15	15	15	15	15	kaaythe	20	20	20	20	20
339	hathu	30	45	45	40	45	haatho	30	35	30	30	30
340	waathan	15	15	15	15	15	waaathan	20	20	20	20	25
341	waathun	15	15	15	15	15	waaathun	20	20	20	25	20
342	waathin	15	15	15	15	15	waaathin	20	20	20	25	20
343	athal	40	40	40	40	40	adhaall	35	40	35	35	35
344	saaf	30	30	30	30	30	saaaf	35	35	35	35	35
345	atash	15	15	15	15	15	attaash	20	20	20	20	20
346	aks	15	15	15	15	15	akks	15	20	20	20	20
347	aqr	15	15	15	15	15	akrr	15	20	20	20	20
348	ayn	40	50	50	50	50	ayn	40	50	50	50	50
349	ejl	50	60	60	60	60	eejjll	60	65	60	65	65
350	omr	50	50	50	50	50	omrr	50	55	55	55	55
351	saer	25	30	30	25	25	saaeerr	30	40	40	40	40
352	naasa	30	30	30	30	25	naaasaa	30	30	30	35	30
353	naoma	30	30	30	30	30	naaomaa	25	25	25	25	30
354	wasia	25	25	25	25	25	waasia	30	30	30	30	30
355	qaae	20	15	15	15	15	kaaee	25	30	30	30	30

	Alghamdi	T	R1	R2	R3	R4	Improved table 2	T	R1	R2	R3	R4
356	saao	30	30	30	30	30	saao	30	30	30	30	30
357	wariaan	15	15	15	15	15	waarriaan	25	25	25	25	25
358	wariaun	15	15	15	15	15	waarriaun	25	25	25	20	25
359	wariain	15	15	15	15	15	waarriain	25	25	25	25	25
360	ghajar	50	50	55	50	50	ghaajjaarr	55	60	55	55	60
361	ghatha	15	15	15	20	15	ghaadha	20	20	20	20	20
362	ghasha	50	60	60	50	50	ghaashaa	50	50	50	55	50
363	ghadh	15	15	15	15	15	ghaadhu	15	15	20	15	20
364	ghafiya	40	40	40	45	40	ghaafiyaa	45	45	45	45	45
365	gharaqa	15	15	15	15	15	ghaarraakaa	15	15	15	20	15
366	ghaythu	15	15	20	15	15	ghaaytho	20	20	20	20	15
367	ghata	15	15	15	20	20	ghaattaa	25	25	25	20	25
368	ghadara	50	55	55	55	55	ghaadaarraa	45	45	45	45	50
369	ghusn	15	20	20	20	20	ghusn	15	20	20	20	20
370	ghil	62	62	60	62	62	ghill	65	70	70	70	70
371	saghura	30	25	30	25	25	saaghurraa	30	30	35	30	30
372	raghad	42	35	40	42	42	rraaghaad	42	42	42	42	45
373	taghiya	15	15	15	15	15	ttaaghiyaa	20	25	25	25	25
374	maragha	64	60	60	60	60	maarraaghaa	70	65	65	65	65
375	sadghu	16	15	15	15	15	saadghu	20	20	20	15	20
376	samghi	15	15	15	15	15	saamghi	20	15	20	20	20
377	sawgan	17	15	15	15	15	saawgan	20	15	20	20	15
378	sawghun	15	15	15	15	15	saawghun	20	20	15	20	20
379	sawghin	15	15	15	15	15	saawghin	20	20	20	20	15
380	haf	80	80	85	85	80	haaf	85	85	85	85	85
381	wafy	75	60	60	60	60	waafy	80	85	85	85	85
382	malaf	85	70	70	70	75	maallaaf	85	85	85	85	85
383	faka	85	80	80	80	80	faakkaa	85	85	85	85	85
384	fan	85	80	80	80	80	faan	85	90	90	90	90
385	fijl	80	80	80	80	80	fijjll	85	90	90	90	90
386	furn	85	85	85	85	85	furrn	85	85	90	85	85
387	faala	50	42	42	42	42	faaallaa	60	55	55	55	55
388	rafaa	50	42	42	42	42	rraafaaa	45	45	42	45	45
389	dafira	80	70	70	75	75	daafirraa	70	75	70	70	75
390	afwu	60	62	62	62	65	afwu	60	62	62	62	65
391	sharafa	85	80	80	80	80	shaarraafaa	85	90	90	92	90
392	tarafi	60	60	50	60	55	ttaarraafi	65	65	70	70	70
393	khalfu	50	50	50	50	50	khaallfu	45	45	50	45	45
394	alafan	42	40	42	50	45	allaafan	45	45	45	50	45
395	alafun	43	45	4	45	45	allaafun	42	45	50	50	50

	Alghamdi	T	R1	R2	R3	R4	Improved table 2	T	R1	R2	R3	R4
396	alafin	42	42	45	45	45	allaafin	45	45	50	50	45
397	sujuq	62	50	50	51	50	sujjuk	60	60	60	60	60
398	qulw	54	50	50	50	50	kullw	55	55	55	56	55
399	daqaka	60	62	62	62	62	daakaakkaa	60	60	65	60	60
400	qalam	60	60	60	60	60	kaallaam	70	72	70	70	70
401	qidr	60	60	62	60	60	kidrr	65	65	66	65	65
402	quda	61	60	60	60	60	kudaa	60	65	65	65	60
403	saqata	20	16	15	15	14	saakaattaa	20	20	20	15	15
404	fuqida	62	50	50	50	50	fukidaa	65	65	70	65	70
405	thaqula	50	45	45	42	42	thaakullaa	55	60	60	55	60
406	sabaqa	60	50	50	50	50	saabaakaa	60	60	60	60	60
407	abaqa	20	20	20	20	20	abaakaa	15	20	20	20	20
408	ghasaqu	15	15	15	15	15	ghaasaaku	15	20	15	20	15
409	barqan	50	42	42	42	42	baarrkan	42	50	50	50	50
410	barqun	51	42	45	45	45	baarrkun	42	50	50	50	50
411	barqin	50	45	45	45	45	baarrkin	42	50	50	50	50
412	rakadha	50	50	50	50	50	rraakkaadhaa	50	50	55	50	50
413	jaraka	80	80	80	80	80	jjaarraakkaa	85	85	85	85	85
414	kawa	80	70	80	70	80	kkaawa	85	90	90	90	90
415	kahan	80	80	80	80	80	kkaahaan	80	75	75	75	80
416	kalb	90	85	85	85	85	kkaallb	90	90	90	90	90
417	kiys	80	80	80	80	80	kkiys	85	85	85	85	85
418	kuwa	70	62	62	62	62	kkuwa	65	65	65	65	65
419	rakiba	80	80	80	80	80	rraakkibaa	80	80	85	80	85
420	rakaa	42	40	40	40	40	rraakkaaa	45	45	50	45	45
421	makuna	80	80	80	80	80	maakkunaa	85	85	85	85	85
422	haraka	50	42	42	40	40	haarraakkaa	55	55	50	55	55
423	biraku	75	75	75	75	75	berraakku	75	70	70	70	70
424	samaki	80	80	80	80	80	saamaakki	85	90	85	85	85
425	silkan	80	80	80	80	80	sillkkan	75	80	80	85	80
426	silkun	80	80	80	80	80	sillkkun	75	80	85	80	80
427	silkin	80	80	80	80	80	sillkkin	75	80	80	80	85
428	layth	75	75	75	75	75	llaayth	80	80	80	80	80
429	liyn	65	65	65	65	65	lliyn	70	70	70	70	70
430	lumat	65	65	65	65	65	llumaat	70	70	70	70	70
431	oluw	42	40	35	35	40	olluw	45	45	45	45	45
432	ghalaa	40	35	35	35	35	ghaallaaa	45	45	45	45	45
433	jaliy	70	70	70	70	70	jjaalliy	75	75	75	75	75
434	daglu	40	40	40	35	40	daagllu	45	40	40	45	40
435	amali	70	60	70	60	70	amaalli	70	65	65	65	65

	Alghamdi	T	R1	R2	R3	R4	Improved table 2	T	R1	R2	R3	R4
436	tablan	40	35	40	40	40	ttaabllan	50	45	50	50	50
437	tablun	40	40	40	35	35	ttaabllun	50	50	45	50	50
438	tablin	40	40	40	35	40	ttaabllin	50	45	50	50	45
439	ham	90	90	90	90	85	haam	90	90	90	90	90
440	yawm	80	80	75	75	75	yaawm	85	85	85	85	80
441	mawz	80	80	80	80	80	maawzz	80	80	85	80	80
442	min	90	90	90	90	90	min	90	90	90	90	90
443	aamil	42	42	40	40	35	aaamill	40	40	40	40	40
444	amala	42	35	35	40	35	amaallaa	40	42	42	42	42
445	numuw	80	75	75	80	80	numuw	80	75	75	80	80
446	fahama	80	80	80	80	80	faahaamaa	80	80	85	80	80
447	ghanamu	50	40	40	40	40	ghaanaamu	55	60	60	60	60
448	sanami	50	40	40	40	40	saanaami	50	45	45	45	45
449	alaman	42	35	35	42	42	allaaman	45	45	45	45	45
450	alamun	42	42	35	42	35	allaamun	42	45	45	45	45
451	alamin	42	35	42	35	35	allaamin	45	45	45	45	45
452	wana	62	55	55	55	55	waanaa	70	70	70	70	70
453	nahr	75	75	75	75	75	naahrr	70	75	75	70	75
454	nibr	81	75	75	75	75	nibrr	85	85	85	85	85
455	nuwr	65	65	65	65	65	nuwrr	65	70	70	70	7
456	fanar	85	80	80	80	80	faanaarr	85	85	85	85	85
457	saniya	60	65	65	70	70	saaniyaa	60	65	65	60	65
458	hunuw	40	35	35	42	42	hunuw	40	35	35	42	42
459	makana	80	80	80	80	80	maakkaanaa	80	80	80	80	80
460	nahnu	42	42	40	42	42	naahnu	50	45	45	45	45
461	thihni	60	65	65	70	70	dhihni	62	60	60	60	60
462	qarnan	62	60	62	60	62	kaarrnan	65	65	65	65	65
463	qarnun	62	62	62	60	60	kaarrnun	65	70	70	70	70
464	qarnin	62	62	60	62	62	kaarrnin	65	65	65	65	65
465	gharahu	42	42	42	42	42	ghaarraahu	45	50	50	50	50
466	thalahu	53	50	50	50	50	thaallaahu	55	55	55	55	55
467	hir	90	90	90	90	90	hirr	90	90	90	90	90
468	hawas	81	85	80	80	85	haawaas	80	80	80	85	80
469	huwid	75	75	76	75	75	huwid	75	75	75	75	75
470	rahiba	76	70	75	70	75	rraahibaa	75	75	80	75	75
471	rahufa	68	65	65	65	65	rraahufaa	70	70	75	70	75
472	qahara	68	55	55	55	55	kaahaarraa	70	70	75	75	75
473	nawaha	67	70	69	70	72	naawaahaa	65	65	65	65	65
474	minhu	75	75	75	75	75	minhu	75	75	75	75	75
475	fyhi	60	50	50	50	50	fyhi	60	50	50	50	50

	Alghamdi	T	R1	R2	R3	R4	Improved table 2	T	R1	R2	R3	R4
476	jaahan	80	80	80	80	80	jjaahan	80	85	85	85	85
477	jaahun	80	80	80	85	85	jjaahun	80	85	85	85	85
478	jaahin	81	80	80	80	85	jjaahin	80	85	85	85	85
479	witr	79	80	80	80	80	witrr	80	75	75	75	75
480	wujida	77	75	70	70	75	wujjidaa	85	90	90	90	90
481	aawidu	67	65	65	65	65	aaawidu	70	75	75	75	70
482	thawuw	52	50	50	50	50	dhaawuw	55	55	60	55	60
483	mahwu	41	40	40	40	40	maahwu	42	42	42	42	42
484	lahwi	72	75	75	75	75	llaahwi	70	75	75	75	75
485	sahwa	40	40	40	40	40	saahwaa	40	45	40	45	40
486	jarwan	86	80	80	80	80	jjaarrwan	85	85	85	85	90
487	jarwun	85	85	80	80	80	jjaarrwun	85	85	85	90	85
488	jarwin	86	85	80	80	80	jjaarrwin	85	90	85	85	90
489	yad	90	90	90	90	90	yaad	90	90	90	90	90
490	yusr	87	85	85	85	85	yusrr	90	90	90	90	90
491	yin	95	95	95	95	95	yin	95	95	95	95	95
492	sayara	75	75	80	75	75	saayaarraa	75	70	70	70	70
493	ayiya	30	24	20	20	20	ayiyaa	40	45	45	45	40
494	sawyi	60	70	75	75	75	saawyi	60	60	60	65	60
495	tayu	35	30	31	30	30	ttaayu	40	40	40	40	40
496	hayuUa	35	30	30	30	30	haayuUa	35	30	35	35	35
497	atyan	30	25	25	25	25	atyan	30	25	25	25	25
498	atyun	30	30	30	25	30	atyun	30	30	30	25	30
499	atyin	30	25	25	30	25	atyin	30	25	25	30	25
	Total	28458	27443	27374	27444	27457	Total	29159	29726	29847	29865	29773
	Average	57.03	54.996	54.858	54.998	55.024	Average	58.43	59.57	59.81	59.85	59.67

Table 1 Alghamdi and improved DT table comparison Accuracy evaluation by expert 1
T= Total
R1, R2, R3, R4= Recording 1, Recording 2, Recording 3, Recording 4.

	Alghamdi	T	R1	R2	R3	R4	Improved table 2	T	R1	R2	R3	R4
1	dhaaa	60	64	61	65	69	Dhaaa	60	64	61	65	69
2	aathin	69	70	72	70	68	Aaathen	86	83	87	84	83
3	saagha	70	65	65	68	72	Saaghaa	90	87	93	92	88
4	ethaa	88	70	80	88	86	Edhaa	72	68	75	78	76
5	zaar	95	88	90	88	87	Zzaarr	78	80	86	88	86
6	qaas	69	72	77	70	70	Kaas	66	75	79	74	71
7	aamal	86	80	88	85	88	Aaamaall	73	77	80	76	80
8	jatha	68	77	80	80	78	jjaatha	70	80	82	77	76
9	shaah	76	65	79	72	69	shaah	76	65	79	72	69
10	taaf	68	62	67	69	69	ttaaf	75	72	78	75	77
11	hayaaa	49	40	60	55	57	haayaaaa	51	49	58	57	53
12	kaas	69	59	70	70	67	kkaaasu	68	60	65	64	64
13	aukht	70	68	69	77	72	aukht	70	68	69	77	72
14	baada	82	85	78	84	79	baadaa	84	86	88	85	78
15	aaw	92	79	88	90	90	aaaw	88	89	92	90	87
16	aakala	80	78	80	80	78	aaakkaallaa	78	78	81	80	76
17	saaal	72	69	74	67	76	saaaaall	62	65	72	71	71
18	dhuUl	53	50	58	54	46	dhuUll	52	47	57	56	52
19	baiisa	52	55	67	63	51	baaiisaa	61	58	70	68	64
20	baraa	56	60	54	63	61	baarraau	53	54	60	61	56
21	swai	54	50	60	64	54	swai	54	50	60	64	54
22	daaaan	31	22	36	34	31	daaaan	31	22	36	34	31
23	daaaun	32	40	45	38	30	daaaun	32	40	45	38	30
24	daaain	31	36	35	32	36	daaain	31	36	35	32	36
25	thaby	57	55	56	55	52	thaaby	66	58	68	64	63
26	dhaba	43	40	48	41	44	dhaaba	54	53	66	63	61
27	bazagha	52	50	49	52	55	baazzaaghaa	50	46	53	52	52
28	basal	34	30	30	32	27	baasaall	41	36	47	39	42
29	bahaq	43	40	47	41	43	baahaak	45	38	46	46	45
30	khabat	37	35	38	35	40	khaabaatt	50	50	57	54	52
31	kaba	32	34	32	28	27	kkaabaa	42	35	37	34	40
32	thanb	67	62	64	63	58	dhaanb	64	65	60	62	63
33	bashima	74	70	80	88	87	baashimaa	72	67	82	81	82
34	saba	26	28	27	22	25	saabaa	33	30	30	30	31
35	farabu	67	65	56	65	72	faarraabu	60	60	62	63	60

	Alghamdi	T	R1	R2	R3	R4	Improved table 2	T	R1	R2	R3	R4
36	nasab	68	67	74	72	76	naasaab	73	71	74	69	76
37	wajiba	84	80	84	89	84	waajjibaa	79	77	86	82	79
38	thabata	81	80	90	87	84	thaabaata	84	86	89	91	90
39	batala	67	65	69	74	64	baattaallaa	63	64	65	68	71
40	bishr	79	77	68	74	79	beshrr	85	90	94	88	89
41	burj	87	80	89	94	93	burrjj	91	95	95	95	95
42	jubila	80	80	86	83	87	jjubellaa	83	81	88	82	84
43	rabata	57	54	58	53	56	rraabaattaa	60	52	61	58	59
44	subul	72	68	62	72	67	subull	77	82	86	79	83
45	halaba	51	44	56	43	47	haallaabaa	48	45	49	53	52
46	qalbi	35	40	45	43	42	kaallbe	37	45	47	45	46
47	naabu	86	88	70	78	84	naabu	86	88	70	78	84
48	thawban	84	86	85	88	88	thaawban	88	89	90	89	91
49	thawbun	76	80	80	80	80	thaawbun	87	87	93	90	90
50	thawbin	79	80	72	74	74	thaawbin	85	86	94	94	90
51	taht	45	44	50	50	54	taht	45	44	50	50	54
52	dhamat	26	26	27	22	26	dhaamaat	32	36	40	32	37
53	tathil	29	20	27	16	22	tathell	30	19	26	28	25
54	satat	31	30	36	23	27	saattaat	39	34	42	37	30
55	sakat	30	26	36	36	34	saakkaat	41	37	46	42	37
56	tharat	42	40	42	37	39	dhaarraat	38	35	40	38	37
57	hazat	54	50	58	53	58	haazzaat	63	60	67	63	62
58	shadat	48	44	40	44	43	shaadaat	58	55	59	62	58
59	thanat	61	56	63	68	57	thaanaat	62	54	65	67	60
60	jafat	69	63	70	67	68	jjaafaat	68	70	66	64	69
61	otw	11	8	12	13	10	otw	11	8	12	3	10
62	ghat	23	22	16	27	24	ghaat	27	20	35	31	28
63	taqy	56	49	57	56	53	taky	52	47	52	48	54
64	tamr	72	65	72	73	68	tamrr	77	70	78	74	72
65	tyn	80	76	90	85	90	tyn	80	76	90	85	90
66	twt	65	60	69	70	67	twt	65	60	69	70	67
67	qatala	72	72	63	68	63	kaatallaa	70	68	70	67	65
68	sutira	89	81	80	81	81	sutirraa	91	88	89	91	88
69	atuma	67	66	62	59	54	atumaa	68	65	67	60	55
70	yumitu	78	73	77	67	70	yumitu	78	73	77	67	70
71	yakhti	47	45	49	54	49	yaakhti	50	48	55	53	47
72	nahata	39	30	36	32	36	naahaata	43	36	46	42	38
73	samtun	45	41	45	43	46	saamtun	46	41	46	44	45
74	samtan	46	41	42	42	46	saamtan	47	40	45	43	45
75	samtin	46	42	50	48	43	saamtin	47	41	50	50	44

	Alghamdi	T	R1	R2	R3	R4	Improved table 2	T	R1	R2	R3	R4
76	thulth	65	66	63	68	61	thollth	71	78	75	73	74
77	thaqaf	62	60	53	58	56	thaakaaf	57	56	58	56	53
78	makatha	59	50	52	47	45	maakkaathaa	57	51	53	52	50
79	ghath	20	16	14	21	23	ghaath	26	20	22	21	24
80	hadath	28	27	27	27	27	haadaath	32	28	32	30	26
81	sharath	41	43	47	44	45	shaarraath	43	43	46	47	47
82	ath	12	12	15	15	14	ath	12	12	15	15	14
83	thawy	32	30	35	28	25	thaawy	34	29	35	32	32
84	thakhn	21	18	16	18	17	thaakhn	27	20	22	26	26
85	bathahu	64	58	65	63	57	baathaahu	69	66	74	70	64
86	thabata	80	79	77	76	78	thaabaattaa	76	75	80	80	76
87	thaja	68	66	69	69	67	thaajjaa	75	80	82	85	85
88	thiny	41	36	50	46	44	thiny	41	36	50	46	44
89	thulat	21	20	25	17	19	thollaat	26	23	30	31	26
90	wathaba	81	79	75	83	78	waathaabaa	83	84	87	83	82
91	othira	65	60	54	57	63	othirraa	66	60	64	62	61
92	juthw	32	41	32	43	38	jjuthw	32	33	40	42	40
93	aaatha	52	48	57	53	47	aaathaa	52	47	56	54	53
94	rathi	31	29	30	32	32	rraathi	35	37	42	39	39
95	bathu	72	71	73	79	67	baatho	78	79	82	84	75
96	thuluthin	82	72	75	78	77	tholluthin	85	75	83	81	85
97	thuluthun	80	78	78	75	72	tholluthun	85	76	84	75	83
98	thuluthan	83	77	74	72	76	tholluthan	85	80	89	86	83
99	lujaj	93	85	85	89	86	llujjaajj	90	84	87	90	90
100	jaraka	83	82	84	85	84	jjaarraakkaa	90	90	94	93	91
101	dhaja	31	30	39	36	38	dhaajjaa	40	36	43	40	39
102	jas	12	12	12	12	12	jjaas	14	14	12	13	12
103	khajal	73	72	73	68	67	khaajjaall	68	66	69	72	70
104	jahatha	82	78	84	75	77	jjaahaathaa	84	77	79	80	83
105	tajan	42	44	46	38	41	ttaajjaan	36	38	39	43	40
106	shaja	52	50	47	53	51	shaajjaa	63	60	68	66	63
107	ajaza	62	62	64	60	58	ajjaazzaa	59	60	64	62	57
108	sajaa	84	85	83	89	80	saajjaaa	82	79	85	85	83
109	juthm	53	52	45	48	50	jjudhm	47	48	46	47	47
110	jady	73	70	67	66	67	jjaady	75	71	77	73	73
111	jaza	82	80	74	79	72	jjaazzaa	88	85	88	83	90
112	haja	71	67	70	71	68	haajjaa	80	75	78	82	80
113	jawq	76	63	65	72	67	jjaawk	75	68	71	63	73
114	jamal	80	80	83	78	84	jjaamaall	80	76	79	80	82
115	juhd	77	74	68	74	76	jjuhd	77	75	78	77	78

	Alghamdi	T	R1	R2	R3	R4	Improved table 2	T	R1	R2	R3	R4
116	jidu	82	82	86	84	85	jjidu	81	79	83	86	84
117	wajada	93	84	86	85	87	waajjaadaa	90	90	86	90	91
118	aajidu	71	68	76	73	68	aaajjidu	67	67	66	73	72
119	hujub	62	58	67	54	53	hujjub	63	60	68	67	63
120	daraja	62	62	69	70	65	daarraajjaa	59	57	66	63	68
121	sarju	76	74	76	78	69	saarrjju	82	88	86	84	80
122	wahaji	78	74	78	75	76	waahaajji	74	69	78	76	73
123	ewajan	77	64	80	77	73	eewaajjan	75	70	83	82	77
124	ewajun	78	78	76	75	75	eewaajjun	73	76	73	77	74
125	ewajin	79	74	76	78	77	eewaajjin	73	72	77	75	72
126	hadhara	43	50	43	44	41	haadhaarraa	40	39	50	45	47
127	qazah	67	64	73	67	62	kaazzaah	64	62	73	67	62
128	suhuf	62	63	69	73	68	suhuf	62	69	78	74	71
129	hathw	52	60	53	46	54	haadhw	50	55	52	56	56
130	hatama	30	29	26	35	34	haattaamaa	40	43	48	47	44
131	hasan	79	65	76	77	81	haasaan	84	88	89	86	86
132	haka	63	54	54	55	50	haakkaa	78	70	80	74	73
133	halahu	44	49	42	43	46	haallaahu	60	63	68	64	62
134	hay	61	57	59	63	58	haay	62	55	63	62	66
135	hamala	68	63	64	64	62	haamaallaa	63	57	66	62	59
136	hibr	67	66	69	62	59	hibrr	70	68	74	73	69
137	husn	54	63	54	55	55	husn	54	63	54	55	55
138	tahana	43	50	52	57	58	ttaahaanaa	56	52	60	57	56
139	suhub	71	68	77	60	78	suhub	71	68	77	60	78
140	yahilu	69	65	64	59	57	yaahillu	78	70	76	77	72
141	masaha	58	49	58	53	51	maasaahaa	60	53	57	60	59
142	farahi	59	54	57	48	52	faarraahi	55	52	55	53	57
143	marahu	60	66	63	68	63	maarraahu	53	50	57	58	57
144	qazahan	76	69	72	67	68	kaazzaahan	67	70	78	74	69
145	qazahin	75	70	68	63	74	kaazzaahin	67	68	70	70	66
146	qazahun	69	70	59	64	62	kaazzaahun	67	69	70	75	71
147	dhakhahu	10	10	6	7	10	dhaakhaahu	20	15	18	22	16
148	khaduk	50	47	45	48	44	khaadukk	51	46	48	48	43
149	khath	47	46	45	52	51	khaath	53	51	52	54	57
150	khashaa	5	5	8	9	6	khaashaaa	10	12	12	14	13
151	khasa	21	16	24	18	21	khaasaa	22	20	23	25	21
152	thakhara	18	17	22	28	23	dhaakhaarraa	15	14	20	21	20
153	khazaqa	19	17	26	18	22	khaazzaakaa	18	21	25	22	21
154	khasafa	16	18	24	19	20	khaasaafaa	16	20	23	23	21
155	khamana	13	12	10	11	8	khaamaanaa	20	22	21	23	22

	Alghamdi	T	R1	R2	R3	R4	Improved table 2	T	R1	R2	R3	R4
156	khawy	10	6	9	6	7	khaawy	18	20	20	20	20
157	khas	56	45	53	47	43	khaas	59	63	60	62	65
158	khidr	57	57	48	53	57	khidrr	60	58	63	68	64
159	khums	65	59	67	73	69	khums	65	59	67	73	69
160	bakhasa	43	40	44	43	41	baakhaasaa	49	50	55	53	49
161	bakhila	32	31	36	31	26	baakhillaa	28	27	29	30	30
162	rakhusa	26	23	21	22	19	rraakhusaa	27	22	28	27	27
163	sarakha	22	17	18	16	20	saarraakhaa	18	20	20	17	17
164	mukhi	42	45	54	50	50	mukhi	42	45	54	50	50
165	salkhu	43	40	38	42	37	saallkhu	47	50	52	51	49
166	bathakhun	40	37	36	37	37	baadhaakhun	38	31	39	40	39
167	bathakhin	42	40	46	43	42	baadhaakhin	39	35	42	41	40
168	bathakhan	41	40	46	39	41	baadhaakhan	39	45	39	42	43
169	dhid	12	14	18	22	23	dhid	12	14	18	22	23
170	zand	86	78	75	82	81	zzaand	86	86	84	89	84
171	rasada	42	40	40	42	37	rraasaadaa	43	41	40	42	42
172	qadam	61	60	58	60	60	kaadaam	60	63	62	59	63
173	tawd	39	38	32	34	34	ttaawd	45	47	45	46	42
174	dasa	75	67	71	68	66	daasaa	80	82	80	86	82
175	dagl	42	40	35	38	36	daagll	45	43	41	45	42
176	daahu	31	33	37	31	34	daaahu	34	31	29	36	34
177	daf	75	70	67	63	75	daaf	75	69	77	74	76
178	dama	31	27	25	37	36	daama	32	27	29	34	31
179	dub	80	77	90	83	84	dub	80	77	90	83	84
180	diyk	88	76	86	76	79	diykk	89	90	85	90	91
181	nadaba	80	76	65	72	76	naadaabaa	84	82	79	85	82
182	hudida	76	69	76	60	79	hudidaa	80	75	73	85	81
183	mudun	68	64	65	70	78	mudun	68	64	65	70	78
184	sada	56	55	53	46	58	saadaa	57	52	58	57	57
185	ahdu	50	50	42	46	43	ahdu	50	50	42	46	43
186	mahdi	80	73	78	76	76	maahdi	84	78	85	84	81
187	waadan	60	58	54	55	54	waaadan	67	63	67	71	69
188	waadun	59	50	53	52	56	waaadun	66	66	70	68	67
189	waadin	61	60	53	57	56	waaadin	66	64	71	69	65
190	thaky	31	33	32	35	34	dhaakky	27	25	31	28	29
191	thama	24	24	20	23	21	dhaamaa	25	20	25	24	23
192	thala	63	60	53	59	61	dhaallaa	60	60	61	61	61
193	fath	21	16	18	17	17	faadh	24	21	20	21	22
194	qathaa	52	45	53	53	48	kaadhaaa	50	47	52	50	52
195	shatha	78	70	69	71	70	shaadhaa	75	68	72	71	71

	Alghamdi	T	R1	R2	R3	R4	Improved table 2	T	R1	R2	R3	R4
196	thawd	73	64	70	71	68	dhaawd	70	71	69	72	75
197	thiib	13	10	9	9	9	dhiib	11	10	9	9	11
198	thaab	60	56	64	67	65	dhaab	58	60	62	61	59
199	thul	45	49	42	45	44	dhull	47	51	56	49	48
200	kathiba	87	70	81	72	79	kkaadhibaa	86	88	92	91	93
201	athara	65	60	62	58	57	adhaarraa	66	61	65	66	64
202	aathuna	42	33	41	42	44	aaadhunaa	46	47	43	42	45
203	shahatha	54	43	47	47	50	shaahaadhaa	55	52	51	52	50
204	munthu	67	63	62	64	67	mundhu	69	66	68	69	67
205	mUthi	23	20	25	23	27	mUdhi	20	19	24	21	22
206	fathan	47	45	47	37	38	faadhan	44	39	40	42	42
207	fathun	46	45	42	38	40	faadhun	44	40	45	42	42
208	fathin	47	43	43	46	45	faadhin	44	42	45	43	45
209	thahara	61	57	58	54	56	thaahaarraa	59	52	56	58	57
210	qarn	65	50	54	51	57	kaarrn	67	67	64	63	64
211	rakala	80	70	74	79	75	rraakkaallaa	78	73	78	75	74
212	dhara	59	54	56	58	53	dhaarraa	63	58	68	64	62
213	ragw	30	22	21	22	23	rraagw	32	28	34	32	33
214	tayr	78	77	70	80	77	ttaayrr	81	79	87	83	82
215	sir	89	86	83	86	85	sirr	90	90	95	93	92
216	rad	87	79	75	78	81	rraad	88	82	85	86	83
217	ruba	59	50	47	47	52	rrubaa	61	58	62	62	58
218	surur	54	55	52	57	56	surrurr	51	50	57	56	49
219	harama	76	68	72	63	64	haarraamaa	73	69	76	73	72
220	siry	74	74	72	69	67	sirry	78	80	79	82	81
221	fatara	32	22	27	32	32	faattaarraa	30	34	31	29	32
222	juhri	39	30	39	28	33	jjuhrri	35	29	35	32	34
223	fikri	59	53	51	47	50	fikkrri	54	51	53	49	53
224	dahrun	68	61	66	63	64	daahrrun	63	62	58	62	63
225	dahrin	66	66	60	68	63	daahrrin	63	62	61	59	62
226	dahran	65	54	64	61	58	daahrran	64	63	61	60	60
227	zafa	79	68	77	72	75	zzaafaa	80	77	79	78	77
228	zaama	46	43	45	42	42	zzaaamaa	50	51	50	53	49
229	zaky	78	76	64	69	70	zzaakky	76	79	81	80	78
230	zuhal	79	78	74	69	74	zzuhaall	81	81	78	82	83
231	zaraa	76	68	72	73	69	zzaarraaa	77	73	76	79	74
232	zir	89	86	88	83	84	zzirr	82	79	85	83	83
233	ruziq	77	76	70	71	77	rruzzik	73	69	77	73	74
234	azafa	79	73	68	71	69	azzaafaa	75	75	73	73	72
235	juzur	88	82	83	83	86	jjuzzurr	84	79	82	81	80

	Alghamdi	T	R1	R2	R3	R4	Improved table 2	T	R1	R2	R3	R4
236	faaza	90	83	86	88	84	faazzaa	87	85	90	86	84
237	jawzu	89	82	90	89	83	jjaawzzu	87	83	85	87	91
238	kanzi	83	76	78	76	77	kkaanzzi	85	84	90	85	89
239	filizan	78	64	67	65	66	fillizzan	80	79	84	78	83
240	filizun	76	62	68	64	65	fillizzun	80	80	87	82	82
241	filizin	75	67	67	65	64	fillizzin	80	83	86	79	80
242	shams	70	73	74	74	77	shaams	74	79	78	80	75
243	ghasala	64	62	67	65	68	ghaasaallaa	63	66	63	68	66
244	sahw	46	56	49	50	50	saahw	52	55	53	50	50
245	kys	56	51	52	53	49	kkys	50	50	55	53	51
246	dhirs	58	51	54	54	52	dhirrs	57	55	56	54	55
247	sum	77	76	84	88	74	sum	80	76	84	88	74
248	sakaba	88	78	82	79	88	saakkaabaa	85	80	85	84	80
249	sihr	46	42	48	50	46	sihrr	47	50	50	51	50
250	rusul	57	49	60	58	55	rrusull	59	57	59	60	59
251	asal	36	29	38	32	31	asaall	38	40	41	42	40
252	nasiya	55	53	69	62	58	naasiyaa	60	55	66	64	59
253	habasa	51	57	52	48	50	haabaasaa	55	60	59	66	61
254	harasa	51	48	49	54	48	haarraasaa	49	48	52	51	49
255	farasi	86	83	82	75	76	faarraasi	83	85	82	89	87
256	orsan	48	43	46	39	38	orrsan	51	60	57	58	54
257	orsun	50	52	46	45	45	orrsun	53	57	61	60	57
258	orsin	49	46	47	47	47	orrsin	53	53	58	55	52
259	shathw	54	62	48	46	49	shaadhw	56	60	63	64	61
260	shas	32	21	23	23	21	shaas	38	40	42	40	38
261	shathaf	22	27	20	23	22	shaathaaf	28	27	30	31	25
262	shat	20	21	20	17	17	shaatt	27	27	36	31	34
263	shugl	17	19	15	13	17	shugll	20	21	27	24	21
264	qash	22	26	22	21	19	kaash	22	18	25	22	17
265	shak	60	65	58	53	53	shaakk	61	67	64	68	59
266	nashiz	77	67	70	72	68	naashizz	78	74	78	79	77
267	shahy	59	53	49	51	54	shaahy	63	67	63	59	58
268	shajar	78	77	67	66	69	shaajjaarr	72	70	75	73	72
269	shibl	84	83	78	80	79	shibll	85	89	79	84	85
270	shukr	83	81	86	84	80	shukkrr	84	84	84	81	87
271	washm	68	67	64	60	63	waashm	70	69	74	78	77
272	rushida	64	68	70	73	72	rrushidaa	63	62	77	74	68
273	aashudu	54	52	53	58	52	aaashudu	59	58	56	51	56
274	rasha	50	46	40	41	48	rraashaa	55	57	62	61	59
275	rimshu	53	50	52	53	52	rrimshu	54	58	53	60	60

	Alghamdi	T	R1	R2	R3	R4	Improved table 2	T	R1	R2	R3	R4
276	ryshi	56	49	50	50	50	rryshi	57	53	58	54	60
277	kabshan	70	71	72	73	76	kkaabshan	71	75	77	75	74
278	kabshun	68	67	70	64	63	kkaabshun	70	69	73	72	68
279	kabshin	70	66	72	75	68	kkaabshin	70	72	73	68	73
280	qasa	40	43	46	45	42	kaasaa	38	40	42	45	43
281	sum	50	60	65	54	57	sum	50	60	65	54	57
282	sanaa	31	32	33	32	27	saanaaa	34	36	42	32	37
283	sah	34	28	27	30	26	saah	47	48	52	51	50
284	wasy	24	23	21	16	18	waasy	30	34	37	32	35
285	suws	14	11	15	15	14	suws	14	11	15	15	14
286	sayd	21	15	18	16	18	saayd	22	23	21	23	21
287	sihr	27	19	18	22	26	sihrr	28	26	37	35	31
288	asara	18	21	12	15	10	asaarraa	17	16	19	20	21
289	nusira	35	32	28	29	33	nusirraa	38	38	40	32	34
290	yasudu	26	24	29	19	25	yaasudu	30	31	28	29	29
291	rasa	23	18	18	20	16	rraasaa	27	25	30	26	28
292	qursi	10	12	13	10	9	kurrsi	21	17	18	22	21
293	fasun	14	17	18	18	17	faasun	21	20	22	21	22
294	fasa	15	12	13	14	14	faasaa	21	18	20	22	19
295	fasin	15	14	10	13	12	faasin	21	19	22	20	22
296	dhaghath	10	11	9	9	9	dhaaghaath	14	17	16	11	10
297	wadhaa	14	12	14	14	11	waadhaaa	14	16	14	16	15
298	dhana	20	19	22	21	21	dhaanaa	23	19	26	26	24
299	dhala	25	18	20	21	22	dhaallaa	27	23	27	24	24
300	dhyq	17	17	18	20	19	dhyk	19	21	25	23	24
301	dhafar	19	17	11	15	14	dhaafaarr	22	26	22	21	20
302	dharaba	25	26	28	31	33	dhaarraabaa	27	24	21	25	24
303	dhuha	14	20	20	20	20	dhuha	14	20	20	20	20
304	dhidu	10	7	8	10	10	dhidu	10	7	8	10	10
305	radhiya	27	26	33	31	28	rraadhiyaa	30	26	28	30	31
306	adhud	14	20	22	19	17	adhud	14	20	22	19	17
307	fadhala	21	31	21	27	26	faadhaallaa	23	24	26	24	21
308	maradha	34	33	28	26	29	maarraadhaa	28	27	31	29	31
309	aradha	20	25	24	25	27	arraadhaa	17	20	22	21	24
310	aardhi	23	23	21	22	26	aaarrdhi	30	31	37	34	34
311	qardhan	25	19	24	26	25	kaarrdhan	28	25	27	29	27
312	qardhun	25	19	25	16	19	kaarrdhun	29	29	29	31	26
313	qardhin	25	20	20	23	21	kaarrdhin	29	26	28	29	29
314	taq	34	35	32	36	32	ttuk	37	40	45	43	41
315	hatala	27	26	28	27	27	haattaallaa	26	32	36	33	34

	Alghamdi	T	R1	R2	R3	R4	Improved table 2	T	R1	R2	R3	R4
316	tamaa	16	13	18	16	15	ttaamaaa	20	24	21	27	25
317	tib	37	33	38	35	34	ttib	42	43	42	40	45
318	tabaa	19	24	23	25	26	ttaabaaa	22	28	25	25	21
319	watan	40	40	46	41	43	waattaan	45	44	47	47	48
320	ratib	35	31	38	35	37	rraattib	41	43	42	41	47
321	otuf	34	23	33	27	28	otuf	34	23	33	27	28
322	qirtu	18	27	18	25	27	kirrtu	21	27	23	21	26
323	wasati	40	41	44	45	42	waasaatti	52	57	57	53	52
324	basata	53	49	55	52	53	baasaattaa	60	61	58	68	65
325	nuqatan	22	15	19	16	17	nukaattan	31	30	31	28	31
326	nuqatun	21	19	26	24	22	nukaattun	31	28	27	33	29
327	nuqatin	21	18	25	23	25	nukaattin	32	31	28	29	30
328	thahar	20	18	22	19	23	thaahaarr	21	25	28	22	22
329	kathu	13	18	26	23	25	kkaatho	15	14	15	16	14
330	wathafa	34	32	38	34	36	waathaafaa	37	35	38	37	40
331	tharf	25	29	35	28	31	thaarrf	25	30	30	33	30
332	thifr	28	19	26	21	20	thefrr	31	34	38	32	36
333	thul	35	36	36	34	36	tholl	37	38	39	42	39
334	nathara	16	16	19	18	17	naathaarraa	14	17	15	14	18
335	nathufa	13	15	20	18	18	naathofaa	17	19	20	20	20
336	athima	12	15	15	14	15	athemaa	16	17	14	15	16
337	hafatha	7	10	11	11	10	haafaathaa	15	12	17	15	15
338	qaythi	9	11	13	12	14	kaaythe	17	20	23	21	19
339	hathu	7	11	14	15	13	haatho	14	16	18	15	18
340	waathan	11	14	16	14	16	waaathan	20	21	28	21	24
341	waathun	12	9	12	11	11	waaathun	20	25	24	23	24
342	waathin	12	10	11	9	8	waaathin	20	26	20	26	21
343	athal	25	17	15	18	17	adhaall	29	28	27	30	27
344	saaf	27	19	22	21	18	saaaf	28	27	28	27	29
345	atash	32	26	30	27	27	attaash	40	41	44	41	45
346	aks	27	31	29	29	30	akks	28	33	37	35	37
347	aqr	22	18	21	15	19	akrr	27	36	35	35	31
348	ayn	28	27	28	22	22	ayn	28	27	28	22	22
349	ejl	33	28	31	29	30	eejjll	36	37	38	34	36
350	omr	37	27	31	31	29	omrr	41	44	47	42	45
351	saer	18	18	18	18	18	saaeerr	23	27	27	24	26
352	naasa	12	15	11	14	12	naaasaa	21	18	19	22	21
353	naoma	13	11	10	10	11	naaomaa	18	20	21	17	17
354	wasia	12	11	15	14	16	waasia	17	19	17	15	16
355	qaae	11	14	17	16	14	kaaee	12	16	17	19	19

	Alghamdi	T	R1	R2	R3	R4	Improved table 2	T	R1	R2	R3	R4
356	saao	11	15	13	17	16	saao	11	15	13	17	16
357	wariaan	15	14	19	15	17	waarriaan	20	22	23	24	21
358	wariaun	15	14	17	13	12	waarriaun	20	23	20	23	23
359	wariain	15	14	19	18	18	waarriain	20	21	23	23	21
360	ghajar	27	26	27	25	27	ghaajjaarr	28	21	27	26	26
361	ghatha	14	13	18	16	17	ghaadha	21	22	23	21	24
362	ghasha	34	27	33	30	28	ghaashaa	38	34	33	31	34
363	ghadh	16	18	20	21	19	ghaadhu	22	24	22	23	21
364	ghafiya	23	22	23	22	22	ghaafiyaa	24	23	28	25	21
365	gharaqa	15	20	23	21	22	ghaarraakaa	20	20	21	23	26
366	ghaythu	17	14	22	20	17	ghaaytho	23	26	22	21	26
367	ghata	14	16	17	14	17	ghaattaa	22	25	23	21	24
368	ghadara	42	38	45	42	39	ghaadaarraa	45	43	50	46	48
369	ghusn	23	40	35	40	26	ghusn	23	40	35	40	26
370	ghil	47	35	42	38	39	ghill	48	49	55	52	53
371	saghura	50	45	52	48	43	saaghurraa	49	55	53	52	51
372	raghad	64	66	68	66	69	rraaghaad	69	73	74	71	69
373	taghiya	21	23	26	23	24	ttaaghiyaa	23	31	27	29	30
374	maragha	77	72	80	74	72	maarraaghaa	80	82	80	85	84
375	sadghu	22	27	27	30	28	saadghu	27	30	32	27	28
376	samghi	67	59	63	63	60	saamghi	70	68	69	73	74
377	sawgan	70	72	78	74	73	saawgan	75	75	73	75	78
378	sawghun	68	70	73	70	71	saawghun	74	74	78	73	76
379	sawghin	69	75	79	76	75	saawghin	74	78	78	76	78
380	haf	80	84	86	82	85	haaf	87	89	86	90	93
381	wafy	58	64	64	62	58	waafy	68	67	69	72	71
382	malaf	59	58	52	53	57	maallaaf	64	64	67	68	65
383	faka	65	59	66	68	63	faakkaa	69	70	75	74	78
384	fan	77	66	74	70	72	faan	84	87	89	86	89
385	fijl	79	80	79	82	78	fijjll	83	84	82	85	85
386	furn	88	74	83	81	80	furrn	92	93	96	95	94
387	faala	21	30	30	30	30	faaallaa	32	32	31	29	28
388	rafaa	23	31	32	32	31	rraafaaa	32	33	32	37	35
389	dafira	41	33	40	37	35	daafirraa	48	47	49	48	50
390	afwu	80	78	72	74	73	afwu	80	78	72	74	73
391	sharafa	68	56	65	66	63	shaarraafaa	70	72	74	68	69
392	tarafi	43	50	49	43	47	ttaarraafi	40	45	43	41	39
393	khalfu	49	48	52	48	49	khaallfu	47	46	49	50	45
394	alafan	31	33	35	32	33	allaafan	35	37	39	35	34
395	alafun	31	27	34	31	27	allaafun	35	38	36	36	38

	Alghamdi	T	R1	R2	R3	R4	Improved table 2	T	R1	R2	R3	R4
396	alafin	32	37	39	38	35	allaafin	34	34	32	34	36
397	sujuq	34	34	40	38	36	sujjuk	38	40	41	45	43
398	qulw	21	21	22	21	20	kullw	24	27	25	25	28
399	daqaka	31	25	31	24	28	daakaakkaa	34	36	38	29	30
400	qalam	74	67	73	72	68	kaallaam	77	78	75	79	81
401	qidr	54	53	55	53	53	kidrr	56	57	53	52	56
402	quda	34	39	39	38	37	kudaa	37	40	41	37	39
403	saqata	74	69	73	68	70	saakaattaa	73	73	75	73	76
404	fuqida	83	77	83	71	79	fukidaa	85	87	85	86	87
405	thaqula	63	60	67	62	64	thaakullaa	66	68	67	63	67
406	sabaqa	63	64	68	61	61	saabaakaa	63	65	63	63	64
407	abaqa	57	55	53	58	56	abaakaa	59	60	57	58	56
408	ghasaqu	53	48	56	54	55	ghaasaaku	58	60	62	57	59
409	barqan	72	69	73	70	70	baarrkan	78	79	78	85	82
410	barqun	69	65	74	68	67	baarrkun	77	79	82	82	80
411	barqin	68	66	67	68	67	baarrkin	78	80	82	78	79
412	rakadha	36	32	37	32	29	rraakkaadhaa	38	37	39	42	35
413	jaraka	73	71	70	72	66	jjaarraakkaa	76	74	74	79	75
414	kawa	64	63	68	59	62	kkaawa	65	67	65	63	68
415	kahan	58	59	60	53	58	kkaahaan	59	56	58	59	54
416	kalb	94	89	90	91	89	kkaallb	96	97	95	97	95
417	kiys	85	80	86	83	81	kkiys	86	83	84	89	88
418	kuwa	72	73	77	74	76	kkuwa	78	75	76	78	74
419	rakiba	84	75	80	82	78	rraakkibaa	81	84	87	89	86
420	rakaa	32	37	37	34	32	rraakkaaa	40	38	39	42	45
421	makuna	82	78	82	78	83	maakkunaa	84	79	83	81	83
422	haraka	72	77	80	79	74	haarraakkaa	70	69	67	70	74
423	biraku	82	79	80	74	79	berraakku	80	84	86	85	88
424	samaki	90	85	92	89	86	saamaakki	92	90	94	93	92
425	silkan	69	70	68	67	69	sillkkan	72	73	71	68	73
426	silkun	75	79	77	73	78	sillkkun	72	72	68	69	74
427	silkin	80	74	80	81	80	sillkkin	73	77	75	75	73
428	layth	73	67	78	64	67	llaayth	75	69	72	78	78
429	liyn	42	38	45	37	39	lliyn	41	44	42	39	45
430	lumat	63	60	68	64	64	llumaat	60	60	67	65	63
431	oluw	54	48	57	51	49	olluw	52	57	54	52	52
432	ghalaa	21	27	28	25	25	ghaallaaa	25	20	23	21	20
433	jaliy	15	20	25	24	21	jjaalliy	18	21	21	24	23
434	daglu	18	20	30	28	26	daagllu	22	30	26	24	26
435	amali	29	24	38	34	29	amaalli	30	32	32	31	29

	Alghamdi	T	R1	R2	R3	R4	Improved table 2	T	R1	R2	R3	R4
436	tablan	63	71	69	68	69	ttaabllan	64	69	72	70	70
437	tablun	65	67	72	68	71	ttaabllun	67	72	74	69	72
438	tablin	70	68	73	74	72	ttaabllin	73	72	70	75	73
439	ham	79	75	84	79	82	haam	80	82	79	85	84
440	yawm	76	70	79	78	73	yaawm	78	80	82	81	82
441	mawz	80	81	84	82	79	maawzz	81	84	86	87	86
442	min	83	90	90	90	90	min	83	90	90	90	90
443	aamil	30	31	38	35	32	aaamill	38	45	43	42	40
444	amala	39	36	33	32	38	amaallaa	46	47	50	46	47
445	numuw	79	83	88	84	79	numuw	79	83	88	84	79
446	fahama	81	87	85	85	85	faahaamaa	77	85	87	87	84
447	ghanamu	62	65	70	68	64	ghaanaamu	65	68	69	67	68
448	sanami	39	44	45	43	38	saanaami	40	47	45	42	46
449	alaman	42	37	47	42	45	allaaman	45	47	50	53	53
450	alamun	44	52	42	46	48	allaamun	46	50	52	51	49
451	alamin	48	39	47	45	42	allaamin	45	47	49	52	50
452	wana	54	47	53	51	47	waanaa	56	51	57	56	54
453	nahr	71	67	73	70	68	naahrr	73	74	75	78	75
454	nibr	65	63	66	62	60	nibrr	67	64	69	68	70
455	nuwr	32	34	32	28	29	nuwrr	40	42	41	37	45
456	fanar	41	41	47	42	40	faanaarr	45	46	46	53	52
457	saniya	46	45	49	43	46	saaniyaa	47	46	49	48	50
458	hunuw	21	14	30	27	25	hunuw	21	14	30	27	25
459	makana	46	43	50	47	43	maakkaanaa	49	50	52	48	51
460	nahnu	39	35	46	42	37	naahnu	42	41	39	38	40
461	thihni	50	63	60	59	58	dhihni	54	59	63	58	62
462	qarnan	53	57	55	52	52	kaarrnan	58	60	64	59	63
463	qarnun	54	53	56	54	56	kaarrnun	57	60	60	63	60
464	qarnin	43	51	48	49	52	kaarrnin	57	62	61	60	59
465	gharahu	30	34	36	32	32	ghaarraahu	42	42	44	46	42
466	thalahu	70	62	73	69	66	thaallaahu	75	75	75	70	76
467	hir	61	58	63	58	56	hirr	69	70	71	67	69
468	hawas	76	63	76	71	67	haawaas	74	70	76	72	71
469	huwid	32	38	36	31	29	huwid	32	38	36	31	29
470	rahiba	58	46	56	49	52	rraahibaa	63	67	64	68	65
471	rahufa	56	53	58	56	48	rraahufaa	53	57	54	54	52
472	qahara	61	58	64	68	59	kaahaarraa	67	66	63	67	69
473	nawaha	67	72	73	64	68	naawaahaa	66	64	68	68	70
474	minhu	86	77	85	76	80	minhu	86	77	85	76	80
475	fyhi	30	34	34	35	37	fyhi	30	34	34	35	37

	Alghamdi	T	R1	R2	R3	R4	Improved table 2	T	R1	R2	R3	R4
476	jaahan	70	83	79	78	79	jjaahan	73	74	74	73	76
477	jaahun	78	84	85	82	78	jjaahun	73	74	76	72	72
478	jaahin	79	82	88	84	81	jjaahin	73	72	75	76	74
479	witr	89	84	84	82	85	witrr	92	93	93	91	90
480	wujida	80	78	83	79	85	wujjidaa	87	86	89	87	91
481	aawidu	40	47	47	43	45	aaawidu	50	51	49	52	52
482	thawuw	55	50	56	49	52	dhaawuw	62	60	59	63	62
483	mahwu	75	64	70	68	65	maahwu	80	77	74	78	74
484	lahwi	75	68	84	78	76	llaahwi	73	73	74	73	75
485	sahwa	75	82	79	85	81	saahwaa	73	73	75	74	74
486	jarwan	65	59	67	61	57	jjaarrwan	78	80	75	78	77
487	jarwun	62	68	65	62	61	jjaarrwun	76	76	75	77	76
488	jarwin	68	73	74	74	72	jjaarrwin	78	75	78	78	80
489	yad	90	79	90	83	87	yaad	92	85	87	90	91
490	yusr	79	79	80	83	78	yusrr	75	76	78	78	75
491	yin	90	86	95	93	91	yin	90	86	95	93	91
492	sayara	79	68	75	80	74	saayaarraa	75	77	75	74	75
493	ayiya	76	78	80	75	76	ayiyaa	80	79	82	81	80
494	sawyi	75	64	76	72	70	saawyi	67	70	66	69	68
495	tayu	71	67	74	77	73	ttaayu	75	80	75	77	78
496	hayuUa	71	66	70	68	64	haayuUa	70	65	68	70	69
497	atyan	57	54	57	53	56	atyan	57	54	57	53	56
498	atyun	59	57	59	56	58	atyun	59	57	59	56	58
499	atyin	72	75	72	72	71	atyin	72	75	72	72	71
	Total	25946	24869	26018	25457	25280	Total	26964	26808	28052	27727	27518
	Average	52	49.84	52.14	51.02	50.66	Average	54.04	53.72	56.22	55.57	55.15

Table 2 Alghamdi and improved DT table comparison Accuracy evaluation by expert 2

T= Total

R1, R2, R3, R4= Recording 1, Recording 2, Recording 3, Recording 4.

Appendix V

Published Papers

The First International Symposium on Computers and Arabic Language 2007

(ISCAL-07)

Arabic speech recognition using English based engines

Ghadeer Khalil, Graham Tranfield and Tony Allen

School of Computing and Informatics,
Nottingham Trent University, Clifton Lane, Nottingham, UK

NG8 11NS

This paper reports on research that is designed to evaluate the use of commercially available, English based speech engines, to recognise limited Arabic vocabularies. Although it is recognised that speech engines that are designed specifically for Arabic would have better recognition rates, using this approach would enable mixed language systems to be built, which is a typical requirement for medical applications in the Arabic world where much of the technical language is English but names of patients and other information is in Arabic. As a first step, an application has been built to recognise code words for the letters of the Arabic alphabet and it has been evaluated on 30 Arabic speakers. At present there are limits on its accuracy, and strategies that can improve performance are discussed.

1. Introduction

In recent years, the use of speech and natural language interface technologies have shown great promise for significantly improving the usability of many mobile computer based applications. Examples are use by the police (*Cohen, 2005)* and by medical staff (*Baumgart, 2005) and (*Moffett, *2003*). There are many reasons for this new focus but according to (Holmes, 2001), one of the main reasons is the recent introduction of reasonably effective speaker independent speech recognition technologies. Voice is a natural interface that the majority of people are capable of using without any technical training. The creation of speaker-independent, speech-enabled interface systems for mobile applications, are thus likely to be of increasing benefit to users.

There are several commercially available voice recognition systems such as Dragon Naturally speaking (Nuance, 2006) and IBM ViaVoice (IBM, 2006). The majority have been developed for the English language although there are several speech engines that have been developed for other languages (Peissner, 2002). Systems can additionally be categorised as those that are speaker dependent or independent and those that deal with limited vocabularies as opposed to those that try to recognise the whole of the specified language.

Arabic is one of the Semitic languages and is an important language in literature and religion. It is spoken by almost 250 Million people of which roughly 195 million are first language speakers and 55 million are second language speakers (Lewis, 2003). As a consequence, a number of Arabic speech recognition systems have been developed. The Sakr company have developed a system to recognise limited Arabic vocabularies for telephony applications (Sakhr, 2006) and other Arabic speech engines have been developed by IBM (IBM Research, 2006) and Aculab (Aculab, 2006).

However many applications in the Arabic world are in fact mixed language applications. For example, in hospital applications, a doctor may want to record Arabic names of patients, but use English words for the names of drugs. The creation of applications using two, or more speech engine would be difficult to implement on mobile devices so the idea of using English speech engines to recognize both English and Arabic words in one application is very attractive.

This research therefore investigates the feasibility of this approach by exploring the effectiveness of using English based speech engines to recognise Arabic words.

2. Methodology

The first stage has been to develop an application that can recognise the names of the Arabic letters of the alphabet in order to allow Arabic words to be spelt out. The application has been developed in Microsoft Visual Basic and uses the Microsoft Speech SDK 5.1 to create an interface to the Microsoft English (U.S.) V6.1 Recognizer speech recognition engine.

It was programmed to recognise a series of words that are used to represent the letters in a similar way to that used in the Civil Aviation Organisation code to identify letters of the English alphabet (Alpha, Bravo …..Zulu).

In fact, no such similar code exists for Arabic letters (except for a names code that was used by the Iraqi Army which was not available to the authors), so it was first necessary to create a code by choosing words that would be familiar to Arabic speakers, but that would be sufficiently different from one another to be easily distinguished by the application. Work was then undertaken to identify the best English spelling to represent the phonetic structure of these Arabic words. Finally the recognition rate of the application was evaluated with a variety of Arabic speakers with a wide range of characteristics such as age, sex and regional origin.

The development was undertaken in the following steps:

1. A web-based survey was used to collect the words that people most commonly associated with the letters of the Arabic alphabet, and up to 3 most chosen code-words candidates for each letter of the alphabet (suggested by the participants) were chosen.
2. Different English spellings for each of the chosen Arabic words were tried, in order to find the spelling that sounded closest to correct Arabic pronunciation.
3. The best word for each letter based on recognition rates by the ten users was identified.
4. The initial list of words was refined by finding words within this vocabulary that were most frequently confused by the application.
5. The application was then tested on a sample population of Arabic speakers

2.1. Lexicon construction

Table 1
Initial Code
Word Selection

Word	No. of People	Word	No. of People	Word	No. of people
أ Arnab* Asad	59 34	ز Zahraa Zarafa Zak kaah*	74 16 1	ق Galam Galb Gassi*	44 41 3
ب Batta Boostan* Baab	42 28 21	س Samaaka Samak Sakan*	56 34 5	ك Kalb Korrssay*	75 21
ت Toofah* Toot Tem sah	74 9 3	ش Shams* Shabaka	48 22	ل Laimoon Lail Lee bas*	63 23 5
ث Thaalab Thoor Thoom*	43 21 8	ص Sagor Soorah* Sadeeq	72 18 3	م Maawz Madrasa*	45 19
ج Jamal Jazar Jowz*	74 21 3	ض Dhifdaaa Dha baaab Dhameer*	90 2 2	ن Nasr* Naml	51 40
ح Hemar Ham mama* Hessan	44 14 6	ط Taawela Taa era Teen*	62 27 9	هـ Hood hood* Herra	85 7
خ Khaa roof Khawkh Kho soof*	78 7 2	ظ The laam* Tharf	47 42	و Wa rdda Wadi* Wet waat	77 9 2
د Dob Deek*	52 39	ع Ayn Asal Aali*	84 10 4	ي Yas meen* Yad Yam mama	83 10 2
ذ Thora The a bab *	63 2	غ Gazal* Ghoraab	65 30		
ر Roomaaan Reeesh*	47 21	ف Feeel Fanoos*	87 2		

2.2. Initial Word Selection

The initial selection of words was made by publishing a web-based survey. Friends, family and first year computing students at Al Ahlia University in Bahrain were invited to fill in the questionnaire and 100 people took part.

Table 1 shows the 3 most frequently suggested words for each letter of the alphabet as chosen by the participants. Also included in the table are the approximate number of people who chose each of the words (unsuitable words and blank spaces were omitted from the results).

2.3. Choosing the Spellings

The next stage was to find the most effective spelling for each of the words that were being considered. Being Arabic words, there is no 'correct' English spelling to be used in the vocabulary for the speech recognition application. For example the Arabic word جوز, can be spelt in many different ways such as jaws, jooz, or jows.

A variety of potential spellings were considered and these were then typed into a text to speech program called Free Natural Reader. The spelling was then chosen on the basis of how close each spelling sounded compared to normal Arabic pronunciation.

2.4. Selection of Words

The next phase was then to select the most appropriate word to use for each of the letters of the alphabet. Ten Arabic speaking students living in Nottingham (5 males and 5 females) were used in the study.

The list of words shown in table 1 was presented to the subjects. Each person was asked to read each word clearly using a microphone and the recognition rates were calculated. After that, the same table was presented to the same subjects and they were asked to conduct the same test again and the recognition rates were calculated again.

Word accuracy recognition or word accuracy percentage rates were defined using the formula:

$$\text{Word Accuracy} = \frac{\text{Number of words correctly recognized}}{\text{Total number of words tested}} \times 100$$

The results of the second attempts only are shown in Table 2.

Word	Accuracy rate %	Word	Accuracy rate %
Arnab*	90	Dhifdaaa	0
Asad	50	Dha baaab	0
		Dhameer*	20
Batta	20	Taawela	20
Boostan*	90	Taa era	0
Baab	80	Teen*	90
Toofah*	80	The laam*	50
Toot	50	Tharf	10
Tem sah	60		
Thaalab	10	Ayn*	50
Thoor	30	Asal	40
Thoom*	100	Aali	90
Jamal	40	Gazal*	90
Jazar	70	Ghoraab	70
Jowz*	100		
Hemar	10	Feeel	70
Ham mama*	90	Fanoos*	90
Hessan	10		
Khaa roof	20	Galam	30
Khawkh	10	Galb	40
Kho soof*	90	Gassi*	100
Dob	50	Kalb	50
Deek*	100	Korrssay*	90
Thora	20	Laimoon	50
The a bab*	60	Lail	30
		Lee bas*	80
Roomaaan	20	Maawz	30
Reeesh*	100	Madrasa*	80
Zahraa	20	Nasr*	80
Zarafa	10	Naml	40
Zak kaah*	80		
Samaaka	50	Hood hood*	90
Samak	50	Herra	30
Sakan*	80		
Shams*	90	Wa rdda	20
Shabaka	50	Wadi*	90
		Wet waat	10
Sagor	30	Yas meen*	100
Soorah*	90	Yad	40
Sadeeq	40	Yam mama	80

Table 2
Recognition rates
for candidate words

2.5 Refining the selection

From tables 1 & 2 it can be seen that if the most popular words in each section were to be selected as the lexicon then the average recognition rate would only be 46%. However, if the lexicon were to be chosen based on the words with the best recognition rate (indicated by *) in each section then the average recognition rate would be 85%. The following set therefore shows the set of chosen words.

Table 3 Set of Chosen words	أ Arnab	ح Ham mama	ز Zak kaah	ط Teen	ق Gassi	هـ Hood hood
	ب Boostan	خ Kho soof	س Sakan	ظ The laam	ك Korrssay	و Wadi
	ت Toofah	د Deek	ش Shams	ع Aali	ل Lee bas	ي Yas meen
	ث Thoom	ذ The a bab	ص Soorah	غ Gazal	م Madrasa	
	ج Jowz	ر Reeesh	ض Dhameer	ف Fanoos	ن Nasr	

3. Evaluation

This vocabulary was then tested more systematically on a range of different Arabic speakers. Of the thirty subjects, 16 were females and 14 males. They included a marketing specialist, 23 students (4 school students and 19 university students), two managers, and 4 teachers participated in the study. None of the participants had used a Speech Recognition application before.

The 28 chosen words were presented to the subjects and each person was asked to read each word clearly using a microphone. The recognition rates were calculated.

After a few minutes the same table was presented to the same subjects and they were asked to conduct the same test again and the recognition rates were calculated again.

The purpose of conducting the same test twice is that some of the subjects sounded shy first time, and by doing it for the second time they overcame their shyness and in effect became trained in using the application. The results of the second attempts only have been used in this study, and first attempts were considered as practice.

The subjects are from different parts of the Arab region, 26 spoke Gulf Arabic, 3 Egyptians and one spoke Lebanese. Gulf Arabs were from 3 different countries, 24 Bahrainis, one Qatari and a Saudi. The subjects are from 4 different age groups 10-15, 15-20, 20-25 and over 25.

Table 4
Evaluation results.
F/A = first attempt
S/A = second attempt

	Occupation	Sex	Region	Nationality	Age	Accuracy Rate F/A	Accuracy Rate S/A
1	Student	F	Gulf	Bahraini	20-25	93	97
2	Student	M	Gulf	Bahraini	20-25	80	100
3	Student	F	Gulf	Bahraini	10-15	40	63
4	Student	F	Gulf	Bahraini	10-15	73	80
5	Student	F	Gulf	Bahraini	10-15	43	43
6	Student	F	Gulf	Bahraini	20-25	40	80
7	Student	F	Gulf	Bahraini	10-15	37	63
8	Teacher	F	Gulf	Bahraini	over 25	90	93
9	Student	F	Gulf	Bahraini	20-25	80	90
10	Student	M	Gulf	Bahraini	20-25	63	77
11	Student	M	Gulf	Bahraini	20-25	73	87
12	Teacher	M	North African	Egyptian	over 25	60	73
13	Teacher	M	North African	Egyptian	over 25	60	60
14	Teacher	M	North African	Egyptian	over 25	73	80
15	Marketing Specialist	F	Gulf	Bahraini	20-25	73	83
16	Student	F	Gulf	Bahraini	15-20	30	53
17	Student	F	Gulf	Bahraini	15-20	53	73
18	Student	M	Gulf	Qatari	15-20	63	80
19	Student	F	Gulf	Saudi	20-25	73	83
20	Student	F	Gulf	Bahraini	20-25	63	73
21	Student	F	Gulf	Bahraini	20-25	53	60
22	Manager	M	Gulf	Bahraini	over 25	80	87
23	Student	F	Gulf	Bahraini	15-20	80	87
24	Student	F	Gulf	Bahraini	15-20	53	73
25	Student	M	Gulf	Bahraini	20-25	63	80
26	Student	M	Gulf	Bahraini	20-25	80	80
27	Student	M	Gulf	Bahraini	20-25	80	87
28	Student	M	Gulf	Bahraini	20-25	80	83
29	Student	M	Gulf	Bahraini	over 25	87	90
30	Manager	M	Levantine	Lebanese	over 25	90	97

The overall results gave an average recognition rate of 79% with males achieving a slightly higher result (83%) than females (76%).

Figure 1.
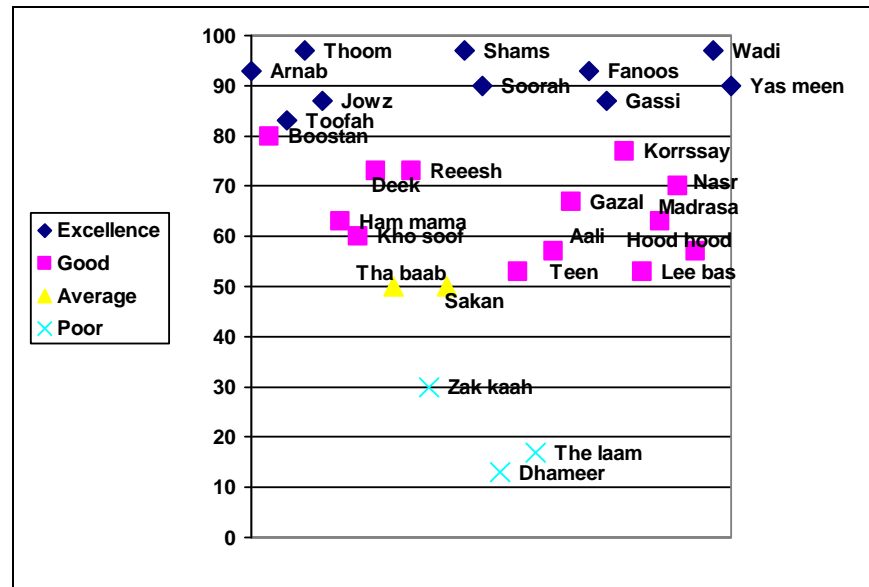The Accuracy
Rates of Words
Recognition

Figure 1 shows that the words Dhameer, The Laam and Zak Kaah had very poor recognition rates which was usually due to the application failing to make a match at all when these words were spoken. In other cases words were wrongly identified as shown in Table 5

Table 5
Misrecognition of
words

Word	Misrecognised as	
Kho soof	Gazal (27%)	
Deek	Teen (39)	
Aaali	Gassi (50%)	Wadi (45%)
Korrssay	Gassi (50%)	
Teen	Reeesh (33%)	

A final attempt is conducted to change some words to obtain better recognition rates.

So (Kho soof) is changed to (khoorfa kaan), (The a bab) is changed to (Thee kkraa), (Gassi) to (Ghaa noon), (Gazal) to (Ghaanna), (Wadi) to (Waseela) and (Thoom) to (Thamer) , Dhameer spelling is also changed to Dhameeer, so that the pronunciation is closer to Arabic.

The following table shows the finalized set of chosen words

Table 6
Final set of
Words

أ Arnab	ح Ham mama	ز Zak kaah	ط Teen	ق Ghaa noon	هـ Hood hood
ب Boostan	خ Khoorfa kaan	س Sakan	ظ The laam	ك Korrssay	و Waseela
ت Toofah	د Deek	ش Shams	ع Aaali	ل Lee bas	ي Yas meen
ث Thamer	ذ Thee kkraa	ص Soorah	غ Ghaanna	م Madrasa	
ج Jowz	ر Reeesh	ض Dhameeer	ف Fanoos	ن Nasr	

This vocabulary was then tested on a range of different Arabic speakers (the subjects are a sub-set of the 30 speakers used in the main experiment) in two different environments, a quiet and a noisy environment (the experiment was performed in a room with only one other individual present (the experimenter). Background noise was a factor, the air conditioner and other computers were on during the experiment. Of the twenty subjects, 10 were females and 10 males.

Then recognition rates were calculated.

Table 7
Recognition
rates

Environment	Males Accuracy rates	Females A/R	Average
Noisy	90	93	91.5
Quiet	91.1	93.4	92.25

The accuracy results in a noisy environment gave an average recognition rate of 91.5% vaguely lower than the rate in a quiet environment 92.2%.

The overall accuracy results have improved significantly and a final average recognition rate of 91.87% was achieved.

4. Conclusion & Discussion

The results show that it is possible to create an application for the purpose of recognizing the letters of the Arabic alphabet using a standard English speech recognition engine.

However, despite taking care to choose an appropriate set of words, the recognition rates are still limited. In addition, although some care has been taken to get a range of Arabic speakers, they mainly came from or lived in Bahrain; as did the author of the paper who was responsible for selecting the words and spellings for the application. It is likely that the recognition rates would be even lower for the full Arabic speaking population.

Future work will therefore concentrate on how this recognition rate can be improved. Three strategies are being considered. The first could be to select different words that can be more easily distinguished by the application, although this approach may well have been exhausted by the efforts made so far.

Secondly it may be possible to process the results of the speech recognition engine in a more intelligent way. At present the application simply uses the nearest match to a word in the vocabulary as defined by the speech

engine. However, it is possible to get confidence scores related to the degree of matching between the incoming sound and each of the words in the vocabulary. Static and intelligent thresholding (Chase, 1997) of these n-best confidence scores could be used as a means of generating more reliable recognition rates. Automatic error recovery methods such as re-speak with elimination (Murray, 1993) based on these confidence thresholds could also be investigated.

Finally, there may be opportunities to use information about the context in which the letters are being spelt in order to enhance recognition rates. N-grams (Chen, 1999) and predictive texting techniques (Dunlop, 2000) adapted to Arabic will be investigated as will the use of adaptive language modelling (Rosenfeld, 1994).

Eventually, the effects and constraints of porting the application onto a PDA will be considered with particular emphasis on achieving an optimal balance between memory footprint, processing speed and recognition rate (Lewis, 2003).

References

Aculab. (2006), "Prosody with Nuance speech recognition products "http://www.aculab.com/products/product_summary/nuance.htm.

Baumgart, D. (2005), Personal digital assistants in health care: experienced clinicians in the palm of your hand?, *Lancet, October,* 2005. Vol.366, 1210–22.

Chase L. (1997), Word and acoustic confidence annotation for large vocabulary speech recognition, In Proc. Of the 5th European Conference on Speech Communications and Technology, Rhodes, Greece, pp. 815-818.

Chen, S. F. and. Goodman. J. (1999), "An empirical study of smoothing techniques for language modeling", Comput. Speech Lang., vol. 13, pp. 359-394.

Cohen ,M. (2005), " The Voice Response Translator: A Valuable Police Tool" , NIJ JOURNAL, Issue 252.

Dunlop, M. D. & Crossan, A. (2000), Predictive Text Entry for Mobile Phones, Personal Technologies 4(2), pp. 134-143.

Holmes, J.N. Holmes, W.J. (2001) "Speech Synthesis and Recognition", Taylor & Francis.

IBM. (2006),Embedded ViaVoice Multiplatform Edition, http://www306.ibm.com/software/pervasive/embedded_viavoice_multiplatform/.

IBM Research. (2006), "IBM Text-to-Speech Research", http://www.research.ibm.com/tts/.

Lewis, J. R., Commarford, P. M. (2003), "Developing a voice-spelling alphabet for PDAs", IBM SYSTEMS JOURNAL. VOL 42, NO 4.

Microsoft. (2006), "Speech SDK 5.1", http://www.microsoft.com/downloads/details.aspx?FamilyId=5E86EC97-40A7-453F-B0EE-6583171B4530&displaylang=en#Overview

Moffett, S., Menon*, A.. (2003), "* Preparing doctors for bedside computing", *Lancet, July,* 2003. Vol.362.

Murray, A. C., Frankish, C. R., & Jones, D. M. (1993), Data-entry by voice: facilitating correction of misrecognitions, Interactive Speech Technology, Human Factors Issues in the Application of Speech Input/Output to Computers (Eds. Christopher Barber and Janet M Noyes) Taylor & Francis, London, pp. 137-144 ISBN 0-7484-0127X.

Nuance. (2006), "Dragon NaturallySpeaking", http://www.nuance.com/naturallyspeaking/.

Peissner, M. (2002), What the Relationship between Correct Recognition Rates and Usability Measures Can Tell Us about the Quality of a Speech Application. In Proceedings of 6th International Scientific Conference on Work With Display Units, Berchtesgaden, Germany, Page 296-298.

Rosenfeld, R. (1994), Adaptive statistical language modeling, PhD Dissertation, Carnegie-Mellon University, Pittsburgh, PA.

Sakhr.(2006),"Sakhr ASR",http://www.sakhr.com/Sakhr_e/Products/ASR.htm?Index=2&Main=Products&Sub=ASR