

## A Novel approach to convert speech to Text and Vice-Versa and Translate from English to Arabic Language

Ameera Al-Rehili<sup>1</sup>, Dalal Al-Juhani<sup>2</sup>, Maha Al-Maimani<sup>3</sup> and Munir Ahmed<sup>4</sup>  
College of Computer Science and Engineering, Taibah University, Al-Madinah Al-Munwarah, Saudi Arabia.  
[mahmed@taibahu.edu.sa](mailto:mahmed@taibahu.edu.sa)

### ABSTRACT

This paper shows benefits, analysis, design, and testing of a desktop application which is able to translate English text to Arabic text, pronounce English and/or Arabic text, and recognize the English speech to convert it then into a corresponding English text, to help users to complete their tasks easily especially those with special needs.

**Keywords:** Text to Speech, Speech to Text, English to Arabic Translation, Arabic voice, Speech Recognition, TTS, SR, Arabic TTS, English TTS

### 1. INTRODUCTION

In the current era, technology evolved in a few past years to keep pace with the requirements of the people and facilitate their works. Today, computer technology used by everyone, young, old people, and even those with visual disabilities.

New programs have been successful through the development of pronunciation of words, as text and speech recognition techniques to provide an integrated solution for the blind and visually impaired, making them not always associated with a person to read or write for them, and made them able to deal with the computer easily. From these programs, speech to text and text to speech converters, but what about an application that has the ability to combine these functions with other attributes just like translation that is what we aim to.

### 2. LITERATURE REVIEW

In this part we listed some of the known programs that are related to the idea of our project. We offer them according to their performance for each function our project is interested at.

#### 2.1 Convert Text to Speech

A Text-to-Speech (TTS) system converts normal language text into speech. one example of such a system is Verbose Text to Speech.

#### 2.1.1 Verbose Text to Speech

Verbose is a text-to-speech application. It is what is often referred to as a system text reader. It uses Microsoft Sam, a built-in speech system in Windows, for reading out loud. The program reads text from different sources. First, there is a text field where you can write or paste text from other applications. If you write anything there and click on "Read aloud", a playback window will come up and you will hear Sam reading your text, along with a visual representation of the sound levels. Additionally, Verbose can read text from text files, such as .txt or even .doc files [1]. Figure 1 shows the main window of Verbose Text to Speech.

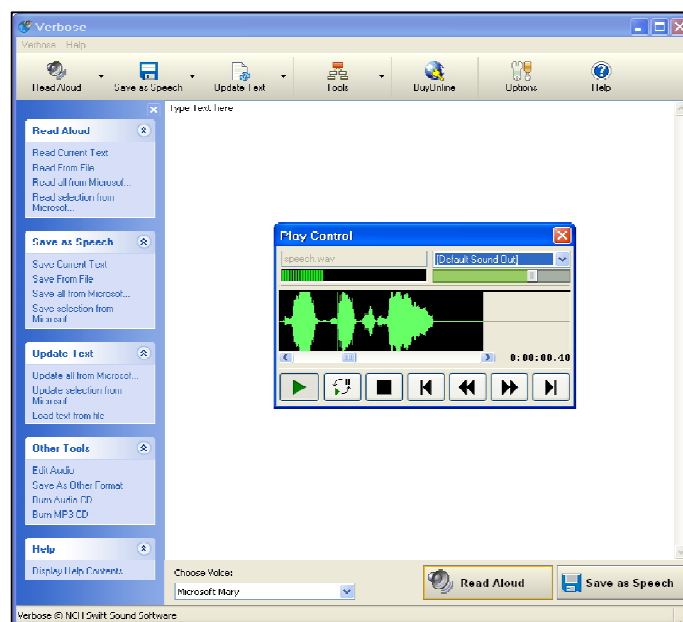


Figure 1: Verbose Text to Speech - Main Window

#### 2.2 Convert audio to text

In this part we have collected most of the known programs that convert user's speech to text. One of them is Dragon Naturally Speaking

### 2.2.1 Dragon Naturally Speaking

Dragon Naturally Speaking is a software used to recognize natural speech and convert it to text. Figure 2 shows the main window of Dragon Naturally Speaking.

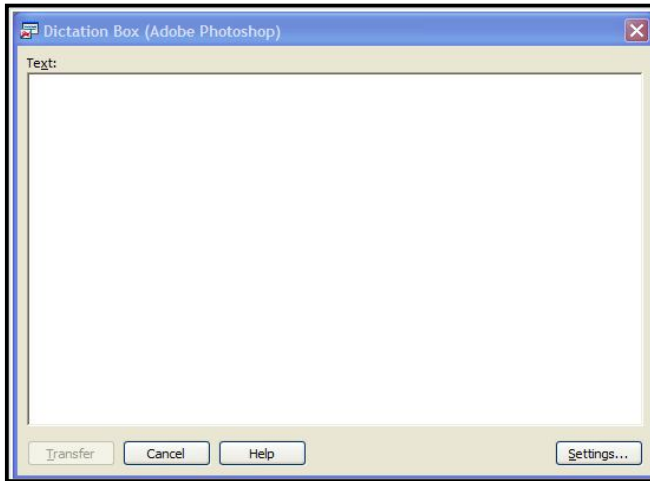


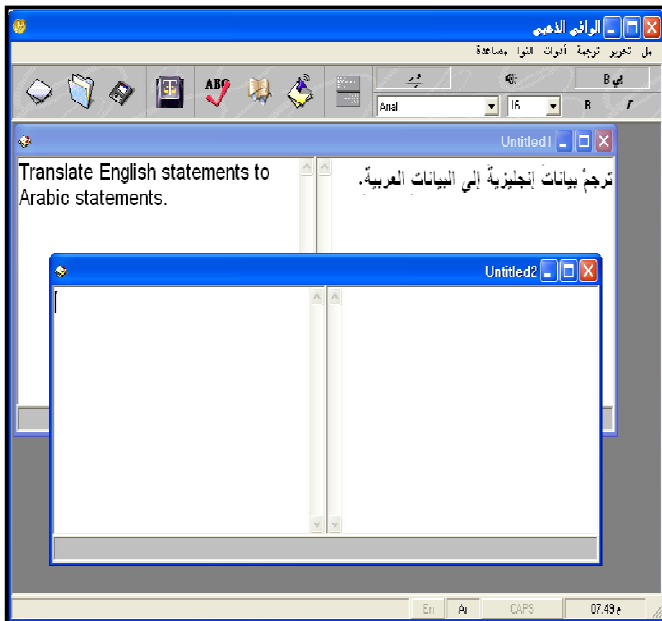
Figure 2: Dragon NaturallySpeaking : Dictation box [2]

### 2.3 Language Translation

The function of translation from English to Arabic Language will be provided in our system. In this part we presented the most well known desktop application that provides this function which is Golden Al-Wafi Translator.

#### 2.3.1 Golden Al-Wafi Translator

Golden Al-Wafi Translator 6.0 is English/Arabic translation software for advanced users and professional translators. Expanded and specialized dictionaries and multi-document translation makes this software suitable for advanced translation purposes [3]. Figure 3 shows how it looks.



## 3. THE PROPOSED SYSTEM

The aim of this system is to develop a desktop application that can combine between three functions, converting English speech to text (programming a computer to recognize natural human speech), translating the English text into Arabic text, and converting English or Arabic text to speech (programming a computer to read text aloud). The user can press a button labeled Read then the program will be able to read a selected text by the user. The user also can use each function separately in order to provide more benefits for different users.

### 3.1 Benefits of the proposed system

The proposed system will help users with special needs to use computers smoothly and easily without the need for any assistance. This system can be beneficial to a wide range of users in many different ways such as learning, while pronouncing English text will help the English language learners in learning English speech, and pronouncing Arabic text will help the non Arabs in learning Arabic Language. Converting the English speech to English text helps in entering text without the need to write it, this could be beneficial for people having problems with their hands, and for regular people it will reduce the time and effort associated with the writing process. This text can be translated to an Arabic corresponding text which in turn can be pronounced with an Arabic voice to help people trying to learn the Arabic language. In contrast a text can be written in English or Arabic to be pronounced for either non Arab or those people don't know the English language.

### 3.2 Requirements of the proposed system

There are two types of requirements; user requirements and system requirements. In the user requirements users want to convert English speech to text with a high quality performance. Users also want to convert English or Arabic text to speech, translate from English to Arabic Language and save or load either the text or the text pronouncing on/from their PCs. Users also want to deal with an easy to use program. As for the system requirements, the proposed system will need a PC and MS Windows environment.

Some of the tools that we have used to achieve our goals are the visual basic.net as a development environment for the application, Microsoft Access as a database to accomplish the part related to the translation from English to Arabic, and Microsoft Speech SDK5.1 for converting English text to speech and vice-versa.

## 4. ANALYZE THE PROPOSED SYSTEM

In this part, the analysis of the application will be presented. The analysis of the application includes: flowchart and data flow diagrams.

### 4.1 Program Flow Chart

The flowchart of our program (shown at Figure 4) describes the steps followed by the user from the beginning of launching the application till the end of using it.

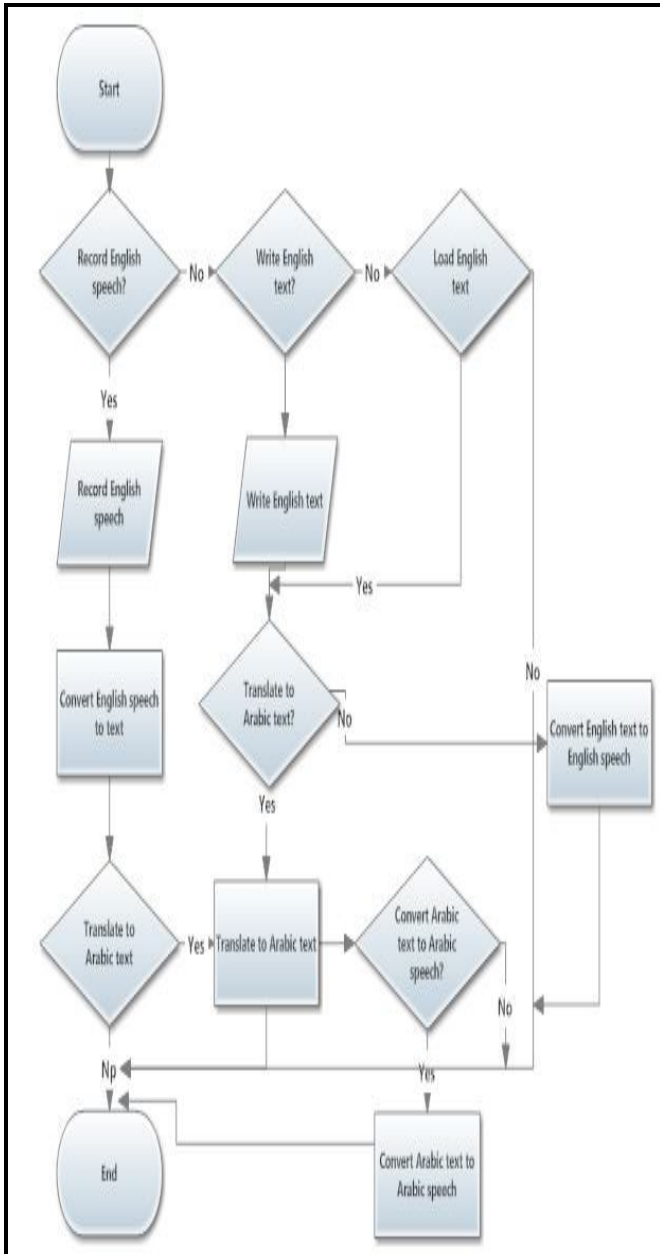


Figure 2: System Flow Chart

### 4.2. Context Diagram

The context data flow diagram is shown at Figure 5. The user can write, load, and save the English text, record English speech, and save the English pronouncing. While the translation and pronouncing system is able to pronounce

English and Arabic text, in addition to the display of the English to Arabic translation.

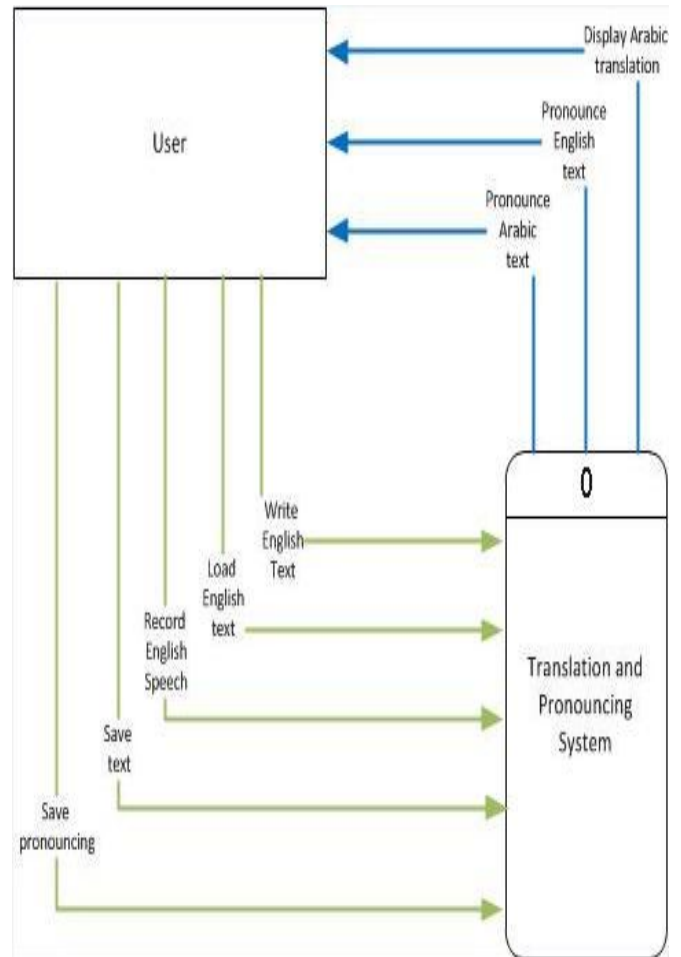


Figure 3: Level 0 Context Diagram

### 4.3. Data flow diagram

Data flow diagrams (DFDs) reveal relationships among and between the various components in a program or system. DFDs are an important technique for modeling a system’s high-level detail by showing how input data is transformed to output results through a sequence of functional transformations [4].

Figure 6 shows the level 1 data flow diagram. The user chooses to enter either an English text or an English speech. For the English speech, the system will use the SDK that recognizes the speech and translates it into text. The system then translates the English text using the English to Arabic dictionary database. Also, the application has the feature of converting English text into English speech using the SDK or Arabic text into Arabic speech using the phonemic database.

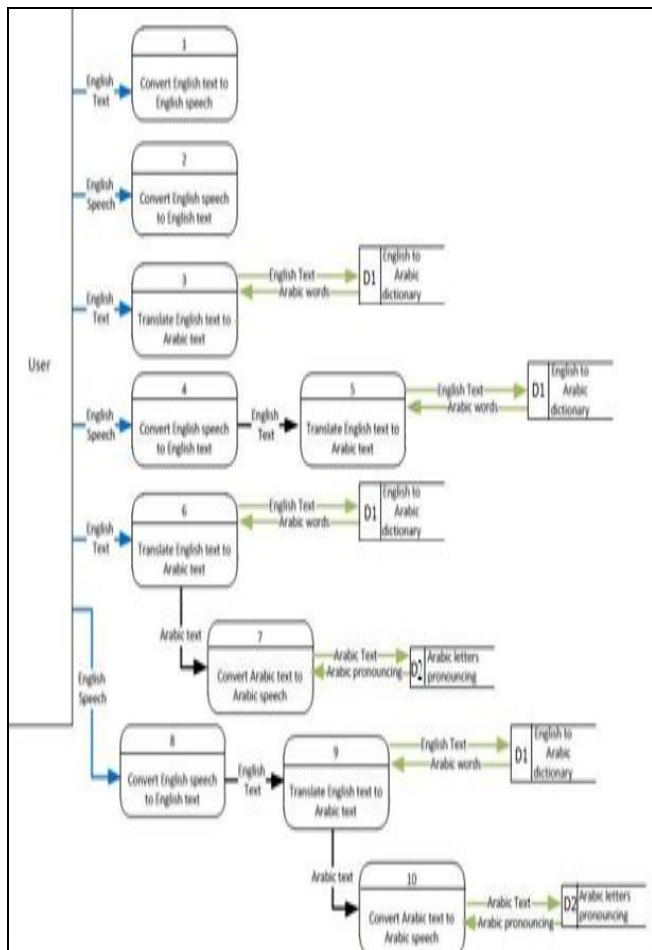


Figure 4: Level 1 Data Flow Diagram

In the proposed system the user can enter an English speech and convert it into English text, translate an English text to Arabic text, and pronounce an English text or an Arabic text. The proposed system achieve that by following one of two paths.

The first path:

1. The user launches the application.
2. The user presses the record button.
3. The user then enters the English text using his/her voice.
4. The application will convert the English speech into text.
5. The user presses the translation button.
6. The application displays the Arabic translation in the Arabic translation text area.
7. The user presses the reading button.
8. The application pronounces the Arabic text.

Second Alternative Path:

1. The user launches the application.
2. The user then enters the English text in the specified area by using the keyboard or loading a text file from his/her PC.
3. The user presses the reading button.
4. The application will pronounce the English text that have been entered by the user.

5. The user presses the translation button.
6. The application displays the Arabic translation in the Arabic translation text area.
7. The user presses the reading button.
8. The application pronounces the Arabic text.

#### 4.4 Translating the English text into Arabic text

The basic idea of the translation in our application depend on coding and database. Figure 7 will illustrate the idea.

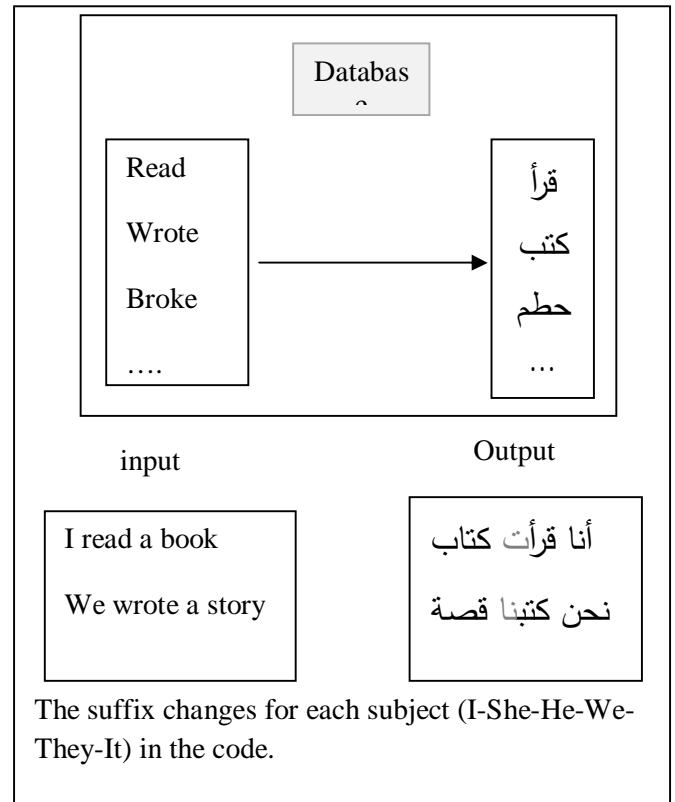


Figure 5: Past verbs translation method

#### 4.5 Pronouncing Arabic Text

Arabic Language is one of the most common languages around the world. There are more than 100 million native speakers of Arabic Language, as Arabic Language is the Holy Qur'an Language, the book of Islam.

To start converting the Arabic text to a corresponding Arabic speech, there are two important notions that need to be described: grapheme and phoneme. Graphemes are usually considered to be the smallest functional units of a written language. The most common definition is that graphemes are the corresponding units of phonemes in spoken language [5]. The general steps followed by the application to convert Arabic text to Arabic speech are showed in Figure 6:

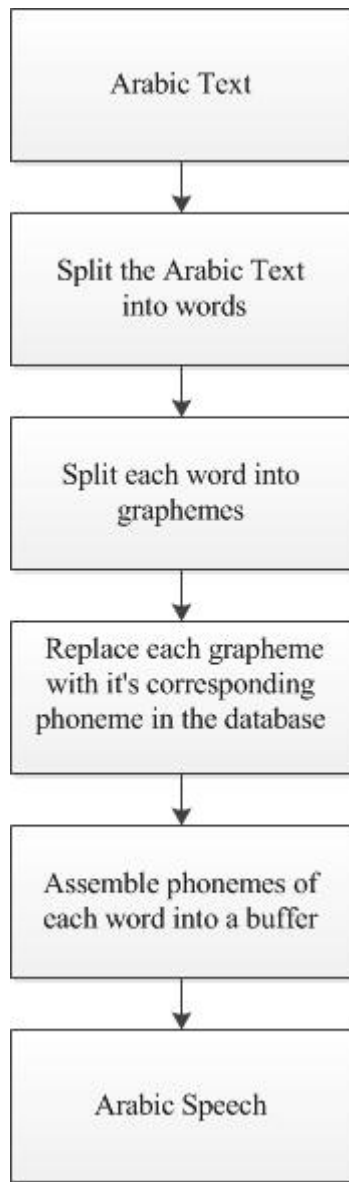


Figure 6: Steps to convert Arabic text to speech (adopted) [6]

In this part, our goal was to make the speech good and natural to listen as human speech, and giving each Arabic letter a corresponding sound according to the diacritic of that letter is not enough. So to achieve this goal we consider that Arabic Language consist of 28 letters, 25 are consonants and three letters are vowels (ي, و, ا), in addition to the pronunciation grammar rules which will be described in the next.

Arabic Language consists of many rules to convert grapheme to phoneme. Some of these rules are listed below:

1. If the definite article precedes a consonant which is pronounced roughly in the same articulation areas as Lam /l/ or behind the upper teeth (Shamsi letters),the

letter Lam is elided, and the next letter is geminated, e.g. [6]

الشَّمْسُ ← أَشْمَسُ

The Shamsi letters are:

ت, ث, د, ذ, ر, ز, س, ش, ص, ض, ط, ظ, ل, ن

2. If the definite article precedes a consonant which is not Shamsi (called Qamari letters), the letter Lam will be a consonant, e.g. [6]

والْقَمَرُ ← وَقَمَّر

The Qamari letters are:

أ, ب, ج, ح, خ, ع, غ, ف, ق, ك, م, و, هـ, ي

3. The Tanween diacritics (indicating a "un", "in", or "an" suffix) should be replaced by explicit suffixes, e.g. [6]

خَاتَمًا ← خَاتَمِ ان

4. If a letter is geminated, it is split into two consonants; the first one is a consonant while the second consonant is a vowel according to the sign of the germination (Tashdeed), e.g. [6]

فَلَّاحٌ ← فَلَ لَاح

5. In the Arabic language specific words must be stored in a database with full audio, e.g.

الله, الرحمن, بسم

## 5. DESIGN THE PROPOSED SYSTEM

In this part, the system architecture and the database design of the system will be presented.

### 5.1 System Architecture

Figure 9 shows the system architecture. It consists of a computer device at which the application is installed, the user of this application; and a local database which is used for the translation from English to Arabic and the conversion of Arabic text to Arabic speech.

In the computer device, the application will be installed. The application queries the English to Arabic database for the Arabic translation of words. The application then concatenates these words together and forms a complete Arabic statement. While using the Arabic words pronouncer, the application queries the database for the pronunciation of any Arabic letter and hence the Arabic words will be pronounced.

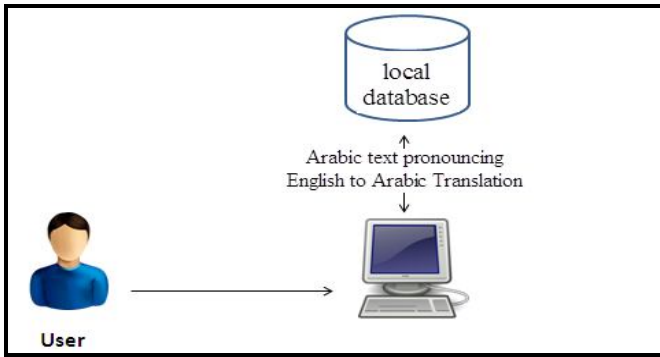


Figure 7: System architecture

## 5.2 Database Development

### 5.2.1 English to Arabic translation

The database used for translating English to Arabic text is created using Microsoft Access 2007. It contains of nine tables (see Figure 8).

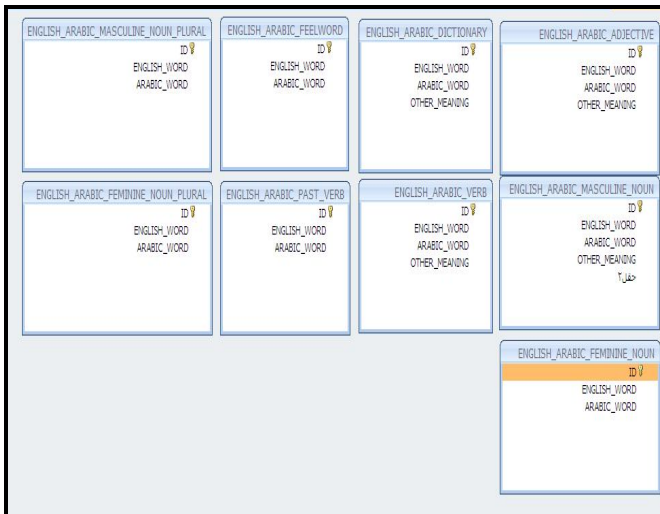


Figure 8: Translation database

### 5.2.2 Arabic text pronouncing

The phonemic files have been taken from an open source Arabic text to speech program which has been made by some students in King Abdulaziz City for Science and Technology (KACST). We have merged and edited those files using Sound Forge 8.0 to be suitable for our algorithm.

The audios have been named according to their order in the Arabic alphabets from 1 to 28, and categorized according to the available diacritics: Fatha, Kasra, Damma, Sukoon, Tanween fatha, Tanween damma, Tanween kasra, Mad fatha, Mad kasra, and Mad damma. so that there are 28 audios under each category. We have also added an audio file numbered "29" which indicates to the spaces between words.

## 6. SYSTEM TESTING

“Software testing is any activity aimed at evaluating an attribute or capability of a program or system and determining that it meets its required results.” [7].

We tested all the following interfaces as end users. In the main window of our application the user can choice to convert text to speech, convert speech to text, read the help manual or know some information about the application developers from the button "about us". If the user chooses to convert text to speech a window labeled Convert Text to Speech will appear, just like Figure 9.

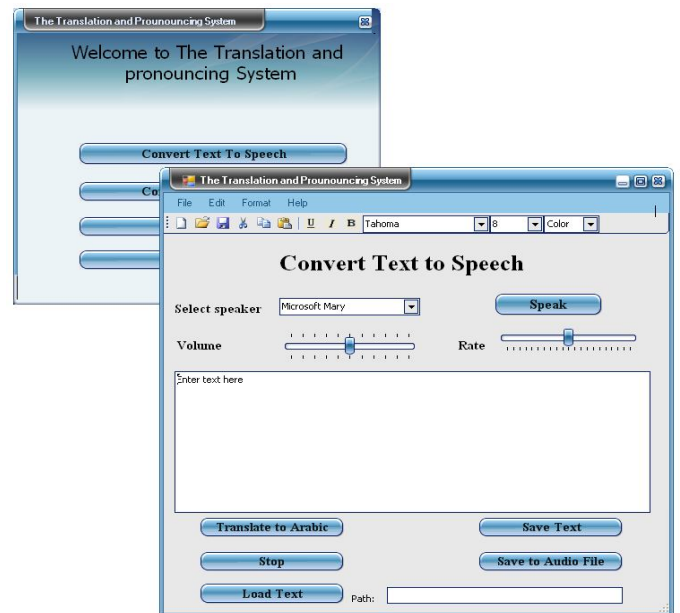


Figure 9: Main window testing

The user can save the speech; change the text color and font through the text editor which is shown at Figure 10.



Figure 10: Save to file audio testing

Searching for particular words and replacing them is allowed in our application. When the user click on the search button an interface just like Figure 11 will appear.

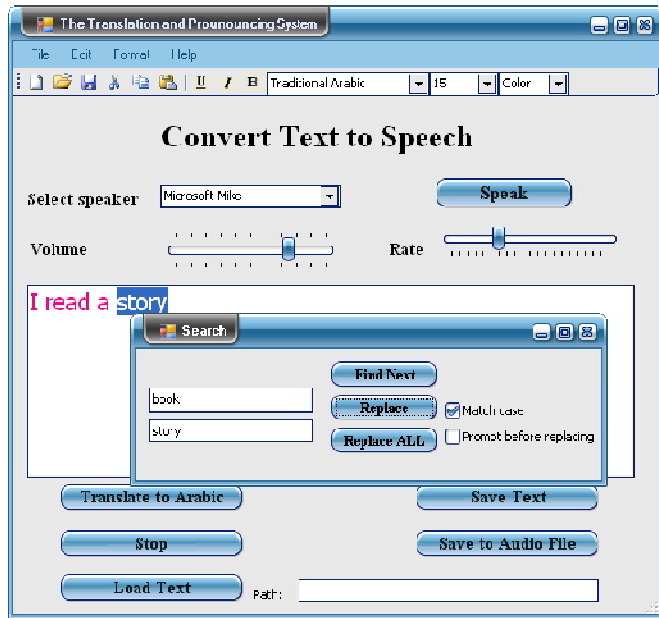


Figure 11: Search testing

When the user click on "Translate to Arabic" button the an interface just like Figure 12 will appear.

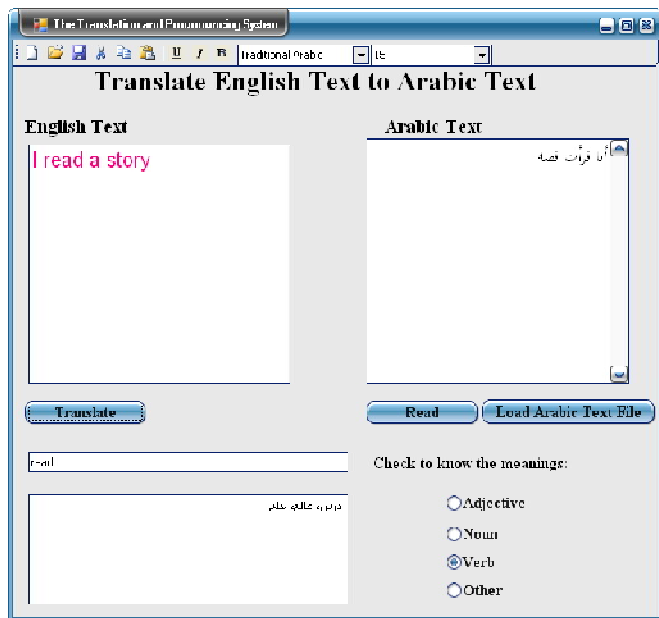


Figure 12: Translate English text to Arabic text testing

The user in the above interface can load an Arabic text which will be read by our application.

If the user chooses the second option in the main window which is convert speech to text an interface just like Figure

13 will appear, and he can convert his speech to text, translate or save that text.

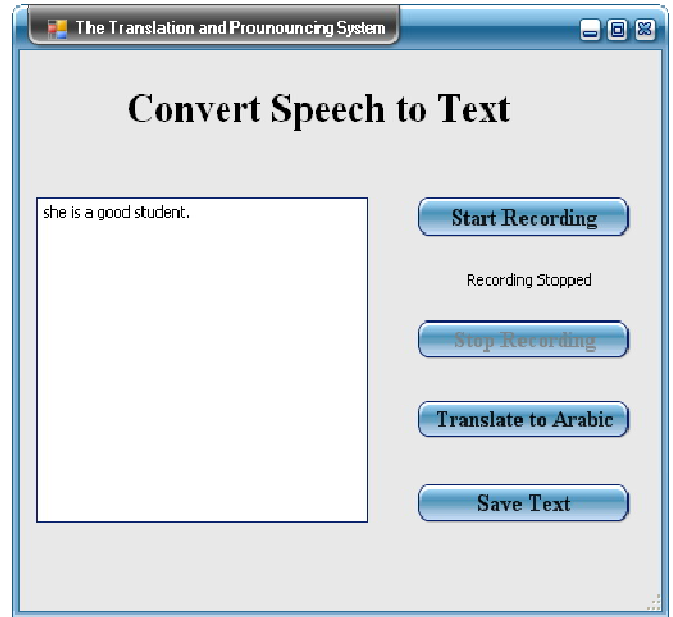


Figure 13: Convert speech to text testing

## 7. CONCLUSION AND FUTURE RESEARCH

The main idea of our project has been showed up at this paper, which is a desktop application that works primarily as a converter between text and speech, for the English language in both directions, and as a pronouncer for Arabic texts to help those trying to learn either English or Arabic. English to Arabic translator has been also embedded to make the application more beneficial to the learners.

In future, We intent to improve the application so that it will be able to translate from Arabic to English, as it's now limited to the translation of English text to Arabic text only. Another thing is about the Arabic text pronouncer as it can read Arabic text with diacritic in a better way, so we wish if a tool is made to diacritic letters automatically rather than the manual way. We also hope that we can build an Arabic speech recognizer and add it to the application.

## REFERENCES

1. J Fernández. (2011, August ) Verbose Text to Speech. [Online]. <http://verbose-text-to-speech.software.informer.com/>
2. K Vulisetti. (2011, August) Dragon NaturallySpeaking. [Online]. <http://dragon-naturallyspeaking.software.informer.com/>
3. Ismael Mireles. (2011, August) Software.Informer. [Online]. <http://golden-al-wafi-translator.software.informer.com/>

4. David C. Yen William S. Davis, **The Information System Consultant's Handbook: Systems Analysis and Design.**: CRC Press, CRC Press LLC, 1998.
5. R., Nottbusch, G., Will, U. Weingarten, Morphemes, **syllables and graphemes in written word production.** Berlin: TRENDS IN LINGUISTICS STUDIES AND MONOGRAPHS, 2004.
6. M. Elshafei Ahmed, **TOWARD AN ARABIC TEXT-TO-SPEECH SYSTEM.** Dhahran, Saudia Arabia: The Arabian Journal for Science and Engineering, 1991.
7. William C Hetzel, **The Complete Guide to Software Testing**, 2nd ed.: Wellesley, Mass. : QED Information Sciences, 1988.

### Author Biographies

**Ameera Al-Rehili** has completed her BSc in Computer Science at Taibah University, al-Madinah, KSA in June 2012. Her area of research interest is Artificial Intelligence, Translating and Communication Systems.

**Dalal Al-Juhani** has completed her BSc in Computer Science at Taibah University, al-Madinah, KSA in June 2012. Her area of research interest is Artificial Intelligence, Translating and Communication Systems.

**Maha Al-Maimani** has completed her BSc in Computer Science at Taibah University, al-Madinah, KSA in June 2012. Her area of research interest is Artificial Intelligence, Translating and Communication Systems.

**Professor Dr Munir Ahmed** is a professional member of the Institution of Engineering and Technology (MIET), United Kingdom (UK). He is completing his DProf - Doctor in Professional Studies (Computer Communications Engineering - Information Security) with Middlesex University, London, UK in September 2012. He has completed partly his EdD - Doctorate in Education (Information Communications Technology) from University of Greenwich, London, UK in 2006. He earned his PhD in Digital Communications Systems Engineering from London Institute of Technology, London, UK in 1997; his MSc in Information Systems Engineering – Computer Networking from South Bank University, London, UK in 1994 and BSc in Electrical Engineering – Electronics and Telecommunications from the University of AJK, Kashmir in 1990. He holds permanent positions as Professor of Computer Networks and Security Engineering, Chairman of Advisory Board and Director of Research at London College of Research, Reading, UK. Since September 2006, he works for Taibah University, Saudi Arabia as Professor of Computer Networks and Communications Engineering on contractual basis. He is a leader of

Security Engineering Research Group (SERG) - London, UK. He is also a reviewer of different international journals. He has extensive experience in the commercial sector and has held a variety of high-level positions in the industry, including Chief Executive Officer (CEO), Chief Operations Officer (COO), Training Director and Chief Network Architect in the UK. His current research activities aim to consolidate his skills and extensive commercial experience with various research areas in the field of Computer Networking and Communications Engineering. His particular areas of focus include Wireless Sensor Networks, Routing Protocols, and Information Security. Professor Ahmed has gone to author or co-author 6 books with leading international publishers in Germany and has had above 250 advanced research activities including papers and articles in international journals and conferences; technical manuals, workshops and presentations in industrial milieu.