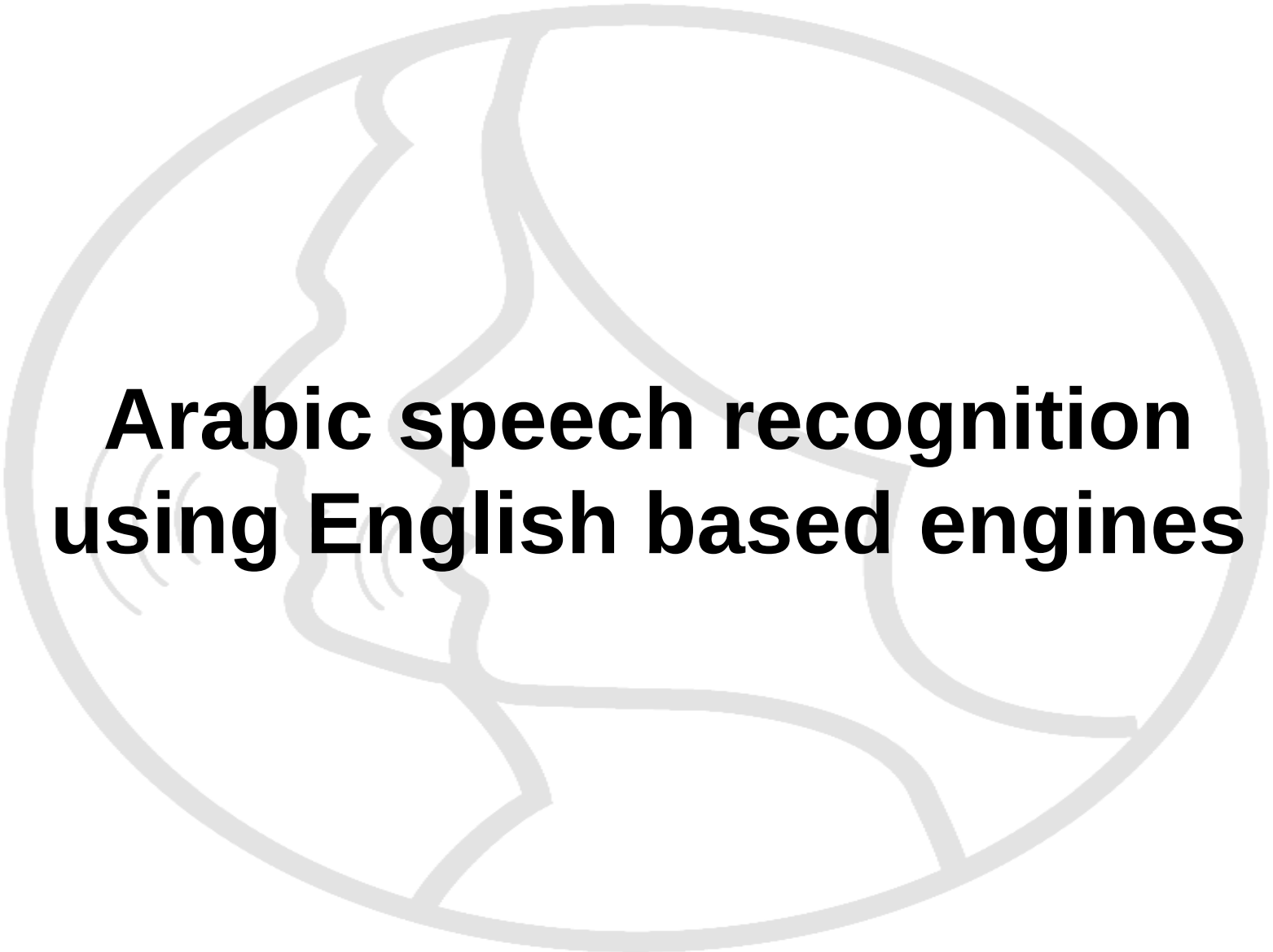
Ghadeer Ismail Khalil
ghadeer@itc.uob.bh



Arabic speech recognition using English based engines

- Introduction
- Emerging Technologies
- Aims and objectives
- Methodology
- Initial Word Selection
- Lexicon construction
- Choosing the Spellings
- Selection of Words
- Refining the selection
- Evaluation
- Conclusion & Discussion
- Future work

Introduction

Speech and natural language interface technologies can be used to improve the usability of many mobile computer based applications

e.g. police and medical staff

Emerging Technologies

Commercially available voice recognition systems

Dragon Naturally speaking

IBM ViaVoice

Arabic speech recognition systems

Sakr limited Arabic vocabularies recognition for telephony applications other Arabic speech engines have been developed by IBM and Aculab

Aims and objectives

Arabic is an important language in literature and religion. It is spoken by almost 250 Million people.

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. Normally, this would mean that a computer based application would need to use two speech recognition systems (Arabic & English).

This work aims to see if it is possible to create a mixed language application that uses only one (English) speech recognition engine to recognise both English and Arabic words.

Methodology

Developing 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.

Methodology

The development was undertaken in the following steps:

1- A web-based survey

2- Different English spellings for each of the chosen Arabic words were tried.

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.

5- The application was then tested on a sample population of Arabic speakers

Initial Word Selection

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.

Lexicon construction

Word	No. of People	Word	No. of People	Word	No. of People	Word	No. of People
أ Annab* Asad	59 34	ب Batta Boostan* Baab	42 28 21	ت Toofah* Toot Tem sah	74 9 3	ث Thaalab Thoor Thoom*	43 21 8
ج Jamal Jazar Jowz*	74 21 3	ح Hemar Ham mama* Hessan	44 14 6	خ Khaa roof Khawkh Kho soof*	78 7 2	د Dob Deek*	52 39
ذ Thora The a bab *	63 2	ر Roomaaan Reeesh*	47 21	ز Zahraa Zarafa Zak kaah*	74 16 1	س Samaaka Samak Sakan*	56 34 5
ش Shams* Shabaka	48 22	ص Sagor Soorah* Sadeeq	72 18 3	ض Dhifdaaa Dha baaab Dhameer*	90 2 2	ط Taawela Taa era Teen*	62 27 9

Table 1 Initial Code Word Selection

Word	No. of People	Word	No. of People	Word	No. of People	Word	No. of People
ظ The laam* Tharf	 47 42	ع Ayn Asal Aali*	 84 10 4	غ Gazal* Ghoraab	 65 30	ف Feeel Fanoos*	 87 2
ق Galam Galb Gassi*	 44 41 3	ك Kalb Korssay*	 75 21	ل Laimoon Lail Lee bas*	 63 23 5	م Maawz Madrasa*	 45 19
ن Nasr* Naml	 51 40	ھ Hood hood* Herra	 85 7	و Wa rdda Wadi* Wet waat	 77 9 2	ي Yas meen* Yad Yam mama	 83 10 2

Cont. Table 1 Initial Code Word Selection

Choosing the Spellings

A variety of potential spellings were considered and then typed into a text to speech program called Free Natural Reader.

Selection of Words

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

Ten Arabic speaking students living in Nottingham (5 males and 5 females) were used in the study.

Selection of Words

Word	Accuracy rate %	Word	Accuracy rate %	Word	Accuracy rate %	Word	Accuracy rate %
أ Arnab* Asad	 90 50	ب Batta Boostan* Baab	 20 90 80	ت Toofah* Toot Tem sah	 80 50 60	ث Thaalab Thoor Thoom*	 10 30 100
ج Jamal Jazar Jowz*	 40 70 100	ح Hemar Ham mama* Hessan	 10 90 10	خ Khaa roof Khawkh Kho soof*	 20 10 90	د Dob Deek*	 50 100
ذ Thora The a bab *	 20 60	ر Roomaaan Reeesh*	 20 100	ز Zahraa Zarafa Zak kaah*	 20 10 80	س Samaaka Samak Sakan*	 50 50 80
ش Shams* Shabaka	 90 50	ص Sagor Soorah* Sadeeq	 30 90 40	ض Dhifdaaa Dha baaab Dhameer*	 0 0 20	ط Taawela Taa era Teen*	 20 0 90

Table 2 Recognition rates for candidate words

Word	Accuracy rate %	Word	Accuracy rate %	Word	Accuracy rate %	Word	Accuracy rate %
ظ The laam* Tharf	 50 10	ع Ayn Asal Aali*	 50 40 90	غ Gazal* Ghoraab	 90 70	ف Feeel Fanoos*	 70 90
ق Galam Galb Gassi*	 30 40 100	ك Kalb Korssay*	 50 90	ل Laimoon Lail Lee bas*	 50 30 80	م Maawz Madrasa*	 30 80
ن Nasr* Naml	 80 40	هـ Hood hood* Herra	 90 30	و Wa rdda Wadi* Wet waat	 20 90 10	ي Yas meen* Yad Yam mama	 100 40 80

Cont. Table 2 Recognition rates for candidate words

Refining the selection

The set of chosen words :

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

Evaluation

This vocabulary was then tested on different Arabic speakers. Of the 30 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.

Evaluation

	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

Table 4 Evaluation results.

	Occupation	Sex	Region	Nationality	Age	Accuracy Rate F/A	Accuracy Rate S/A
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

Cont. Table 4 Evaluation results.

	Occupation	Sex	Region	Nationality	Age	Accuracy Rate F/A	Accuracy Rate S/A
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

Cont. Table 4 Evaluation results.

	Occupation	Sex	Region	Nationality	Age	Accuracy Rate F/A	Accuracy Rate S/A
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

Cont. Table 4 Evaluation results.

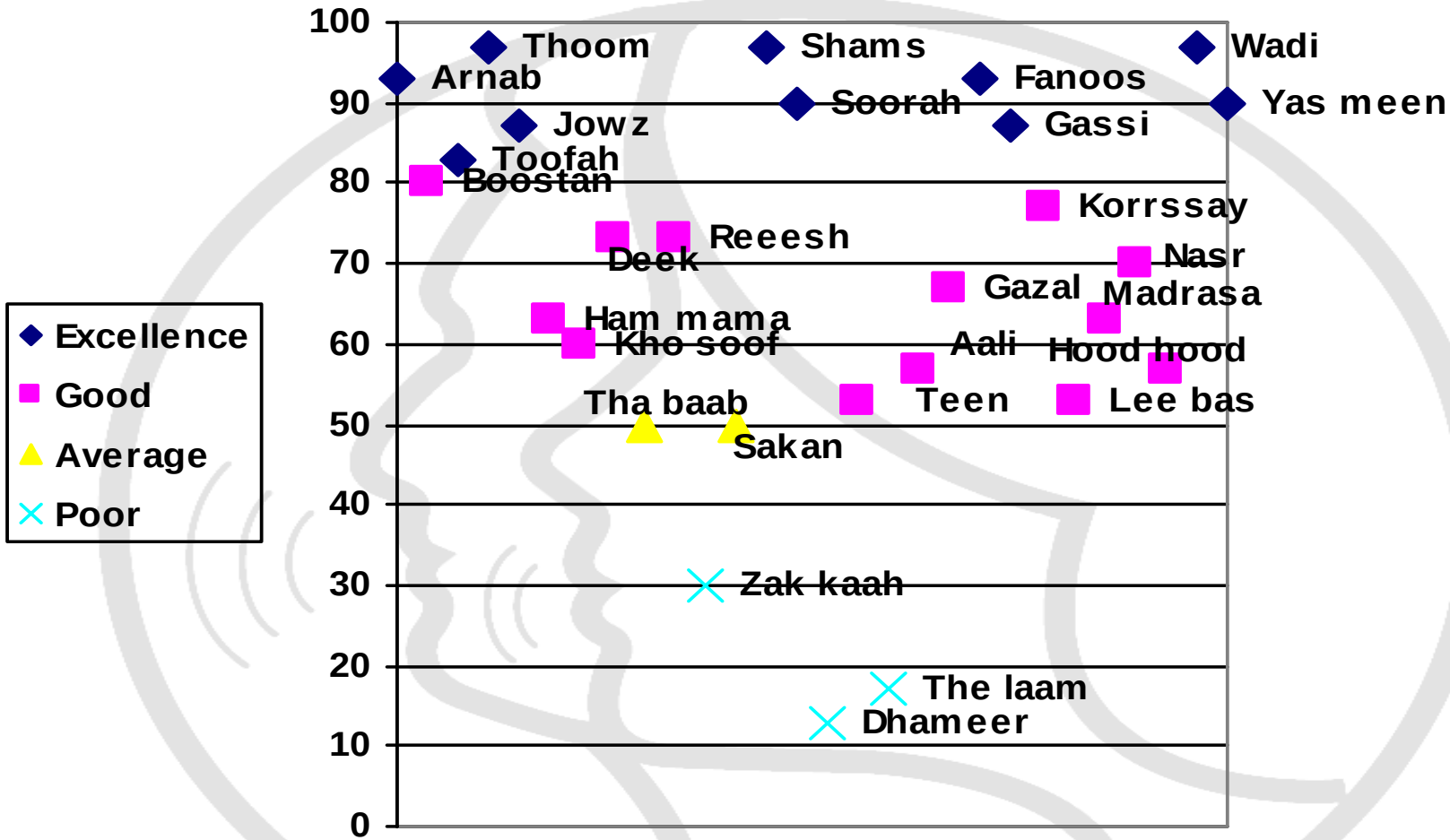


Figure 1.The Accuracy Rates of Words Recognition

Word	Misrecognised as	
Kho soof	Gazal (27%)	
Deek	Teen (39)	
Aaali	Gassi (50%)	Wadi (45%)
Korrssay	Gassi (50%)	
Teen	Reeesh (33%)	

Table 5 Misrecognition of words

A final attempt is conducted to change some words to obtain better recognition rates.

e.g. (Kho soof) is changed to (khoorfa kaan), (The a bab) is changed to (Thee kkraa), (Gassi) to (Ghaa noon), (Gazal) to (Ghaanna), etc..

أ Arnab	ح Ham mama	ز Zak kaah	ط Teen	ف 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	

Table 6 Final set of Words

Environment	Males Accuracy rates	Females A/R	Average
Noisy	90	93	91.5
Quiet	91.1	93.4	92.25

The overall accuracy results have improved significantly and a final average recognition rate of 91.87% was achieved.

Conclusion & Discussion

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.

It is likely that the recognition rates would be even lower for the full Arabic speaking population.

Future work

Improving recognition rates

1. Selecting different words that can be more easily distinguished by the application.

2. processing the results of the speech recognition engine in a more intelligent way.

(confidence scores related to the degree of matching between the incoming sound and each of the words in the vocabulary, Static and intelligent thresholding of these n-best confidence score, Automatic error recovery methods such as re-speak with elimination based on confidence thresholds)

3. using information about the context in which the letters are being spelt, N-grams and predictive texting techniques adapted to Arabic will be investigated as will the use of adaptive language modelling.

4. porting the application onto a PDA with particular emphasis on achieving an optimal balance between memory footprint, processing speed and recognition rate.

Thank you