

Be My Voice Application for People with Speech Difficulties

Amonah Ali

Dept. of Information Technology
King Abdulaziz University
Jeddah, KSA

Maryam Fallatah

Dept. of Information Technology
King Abdulaziz University
Jeddah, KSA

ABSTRACT

Speech Difficulties is the disability of producing speech sounds correctly or having problems with voice. The number of people that suffer from speech difficulties is about 7.5 million of different ages [1]. Speech difficulty is common and can happen out of nowhere or as a result of a medical problem. People with speech difficulties feel sad when others do not understand them or ask them to repeat what they said, sometimes it is a necessity to have someone by their side to help them in communicating. So, speech difficulty considered a major problem that has psychosocial, social and academic impact on people who have it. By using speech recognition and cloud computing to develop an android application that can translate incomprehensible speech into understandable speech and also may improve the pronunciation of words to help those people in communicating and their speech would be clear. Be my voice application will support android operating system and Arabic and English interfaces. It can help people with speech difficulties in communicating and to feel much better because it will allow people to understand them and all those people want is to be understood.

General Terms

Speech Recognition, Cloud Computing

Keywords

Speech Recognition, Cloud Computing, Speech difficulties, android phone application.

1. INTRODUCTION

Speech Difficulties is the disability of producing speech sounds correctly or fluently or having problems with voice. Some of its examples are difficulties pronouncing words, articulation disorders, and stuttering. People of all ages can have speech difficulty. It can happen out of nowhere or as a result of a medical problem [2]. The National Institutes of Health's wrote: "speech and language impairment may be any of several problems that make it difficult to communicate" [3]. In Addition, speech difficulty is common, and it has an impact on psychosocial, social and academic development [4]. Also, a questionnaire on the impact of speech disorders quality of life was conducted, and it stated: "Speech difficulties can result in difficulties in the communication process, lead to social isolation, and affect quality of life" [5], so speech difficulty considered a major problem that can prevent people who have it from communicating with others, affects them academically and also their health. All people with Speech difficulty wish to have a way for their voice and speech to be understandable and clear. By using the technology that is speech recognition in this proposed system that is meant to help those people in communicating, their voice would be clear and correct, and they will have the ability to describe their feelings and desires. Applications to

support them in communicating better have not been released yet, so they need an application to do so and has all the services that users with speech difficulty need to help them in communicating with others.

This paper is organized as follows: Section 2 describes related works. Then, the proposed system is explained in Section 3. Section 4 describes the testing results, Section 5 discussed lessons learned, and Section 6 explains the risks. Lastly, the conclusions are detailed in Section 7.

2. RELATED WORK

Many people suffer from speech difficulties around the world those people do not have a lot of application to support them specially an application that help them in communicating with others in a better way so, that's why the project team created Be My Voice. There is a number of applications that use the same technology of Be My Voice application in table 1.

Table 1. Comparison between Be My Voice and other similar applications

	Type	Idea	Limitation
Google Translate	Website and mobile application	Translate languages by typing, camera, and voice [6].	It need the internet for the translation to work and the voice has to be clear for the translator to translate probably.
QuickVoice Recorder	mobile and computer application	It allows recording voice memos, voice emails, meetings and ... etc [7].	Can not modify or edit the recorded voice.
Siri	MacOS System	Intelligent assistant embedded in IOS & MacOS depends on the principle of learning machine user requirements and suggestions solutions based on what learned [9].	The voice needs to be clear for Siri to understand the user requirements.

Shazam	PCs and mobile	Identifying music via microphone [8.]	Does not support arabic language platform.
	Type	Proposed Application	Advantages
Be My Voice	android application	It aims to give a voice to those people who have difficulties in speaking by translating their unclear speech into clear speech and understandable by people who are around them. Furthermore, it helps users in improving word pronunciation.	It recognizes individual speech, understands what they intend to say and converts it to a clear speech that matches their real voice so they can be heard and understood by the rest of the world.

3. BE MY VOICE (BMV)

This section describes the overall structure of the proposed system (Section 3.1). Then, it details the methodology for the system (Section 3.2). Lastly, Section 3.3 describes the developed phone application.

3.1 System Architecture

The proposed system consists of two steps:

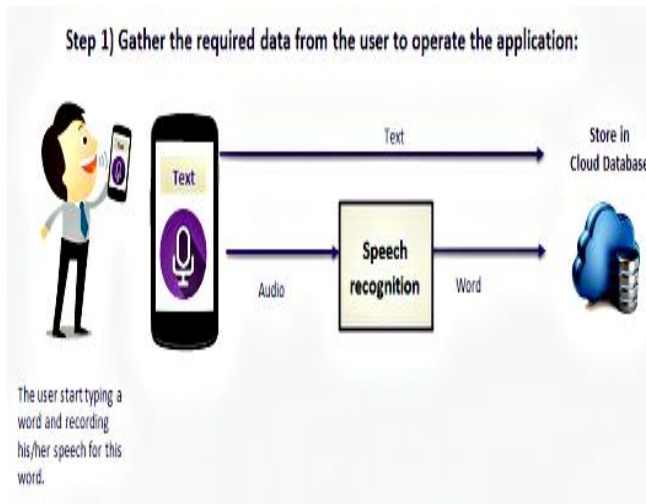


Fig 1: Be My Voice overall system step 1

In the first step the user is going to type the word and record his/her incomprehensible speech for the word. After that the text will be stored in cloud database also the result of the recorded audio after speech recognition operation.

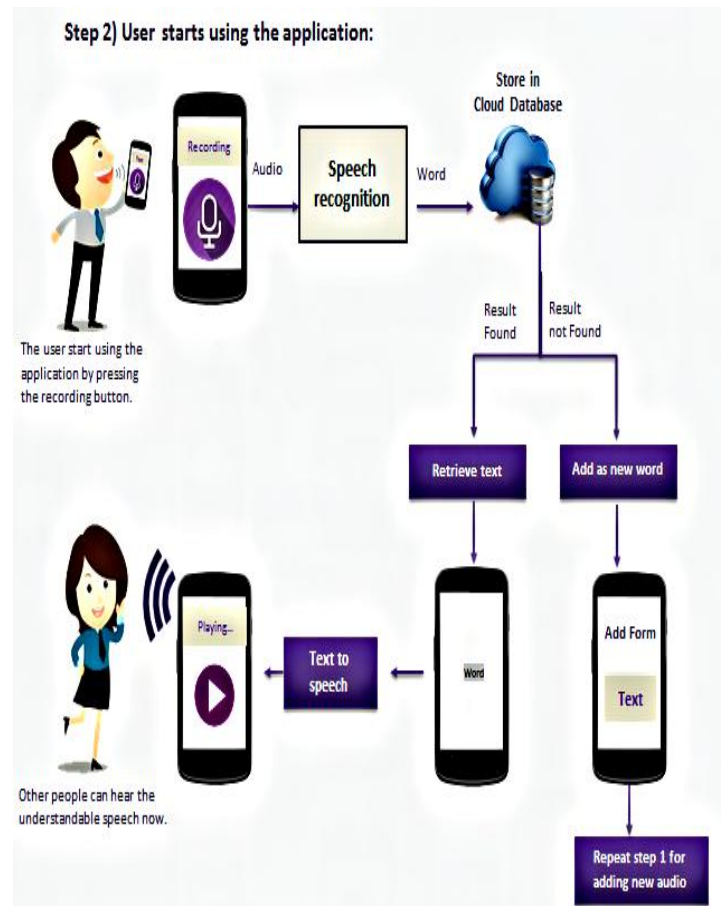


Fig 2: Be My Voice overall system step 2

In step 2 the user is starting to use the application by recording his/her incomprehensible speech then the speech recognition process will be applied on the recorded audio after that the result of the speech recognition will be searched in the cloud database. If the result is found, then the text against it will be retrieved and converted to speech and played and if not found then an option will be displayed for adding the word.

3.2 Methodology

The proposed system creates a dictionary with words for each user that will be stored in cloud database, for the dictionary to be created user must type a word and record it with her/his incomprehensible speech then the word will be saved in the user own dictionary with the result of the recorded audio after speech recognition operation. The result of this process is a dictionary that contain a number of words with the audios results of speech recognition. After the user's dictionary is created the user can use the application in communicating and others by recording the word that is incomprehensible then the application will apply the speech recognition process that is converting the incomprehensible recorded audio to text on the word then search the dictionary for the same text and if it is found then the application will convert what against the searched text to speech for the application to say it. If the text is not found in the dictionary, then the application will ask the user to add it.

3.3 Phone Application

Figure 3 shows the application dictionary that enable the user to build own dictionary with words and audios result after speech recognition

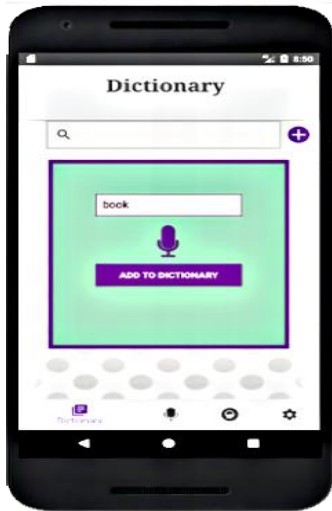


Fig 3: Users dictionary interface

Figure 4 shows the main screen of the application that enable the user to record his/her incomprehensible speech then the application translating it to understandable speech and say it.



Fig 4: Main screen interface

4. TESTING

The testing phase is an important phase to make sure the efficiency, quality and effectiveness is properly carried out. Three types of testing were applied on Be My Voice application: unit testing, system testing and usability testing. The testing has been done on people with speech difficulties starting from age 7 and up. The number of people with speech difficulties that participated are 10 people and the reason of the number being low is because there is a lot of people not accepting the speech difficulty as a problem or an obstruction to living properly. Participation was requested from a speech difficulty center in Jeddah and those who participated are required to be familiar with mobile applications.

4.1 Unit Testing

Unit testing is only successful when a piece of code or a function invokes another piece of code to make sure that the presumption is correct. If it ends up being wrong, then the unit test will fail [9]. Table 2 shows the add word test cases. To guarantee that the application is working properly functionality were tested in isolation from others: sign in, sign

up, forgot password, add word, delete word, play word, display dictionary, record speech and edit word.

Table 2. Add word test cases

Functionality	input	Expected result	Result
Add Word	Book	Successfully added	Pass
	No input entered	Fail addition	Pass
	1223	Fail addition	Pass
	Sleep	Successfully added	Pass

4.2 System Testing

The main point of system testing is to test the performance and functionality of a certain software versus its requirements specification [10]. The system must be tested as a whole that's why a number of test case scenarios has been made. Table 3 show a sample of different scenarios that used to test the system integration.

Table 3: Test case scenarios for system testing

Test case scenario
Sign up > add word and record audio in the dictionary
Sign in > home page > record audio > speak the word
Sign in > home page > record audio > display add new word form

4.3 Usability Testing

What usability testing interested with is determining how the system is easily used for people [11]. Even so the sample size is small the System Usability Scale is the best method for testing that's because it gives a strong and fast tool for measuring the usability on a small sample size also provide accurate results. System Usability Scale have 10 item questionnaires with five answer options for respondents that start from completely agree to completely disagree [12]. First computing the mean and standard deviations of the user answer. The participant was the actual user of the system. User 4 had an overall System Usability Scale rating that is below others. User 4, user 3, user 9, user 7 and user 6 found the system acceptable while user 1, user 5, user 8, user 2 and user 10 found the system excellent. The overall system average was acceptable (82) but the system needs improvement to be better. Table 4 summarize the test scores of each user and the overall test average.

Table 3: Usability testing results

Users	1	2	3	4	5	6	7	8	9	10
Score	90	87	70	66	92	77	79	89	74	97
Average	82									

5. LESSONS LEARNED

- Through the project the project team has gain knowledge on people with speech difficulties that helps in understanding them and know the right way to help them.
- Speech recognition is a new subject to the project team that they studied and learned to do the application in the right way.
- Team work is an important skill to have because when the team are understanding each other they sure will result in a great application.
- Collecting data is an important thing especially in application that encounters people themselves. The right data can help in resulting with the right application.

6. RISKS

- The internet connection could be Interrupted so, that way users cannot use the application and for Be My Voice application an internet connection is necessary for it to work.
- Users could forget the email or password so, to solve this the user can change the password by phone number.
- The accuracy of speech recognition that may result in converting a speech to a certain text to solve this problem using another signal processing algorithm that is more accurate.
- How the system is fast in retrieving the text from the database cloud along hundreds of words that maybe slow or not fast enough and to make this process better machine learning should be applied here for faster retrieving

7. CONCLUSION

Speech difficulties is a major problem and all people who has it wish to have a way for their speech to be understandable and clear and for them to communicate better with others. This problem can be solved by developing an application for translating their non-understandable speech to an understandable and clear one via using some technology like speech recognition. Be My Voice application can help them in many ways such as academically and socially also make them more confident. In the future the project team hope to improve Be My Voice application in many ways such as implementing machine learning in it to make the retrieving of words faster also to use another signal processing algorithm that is more accurate than the current one and to develop an IOS version.

8. ACKNOWLEDGMENTS

Apart from our efforts, the success of any project depends largely on the encouragement and guidelines of others. I take this opportunity to express my gratitude to the people who have been instrumental in making this project a successful one. I would like to show my greatest appreciation to the faculty members of King Abdullaziz University. I cannot thank them enough for the tremendous support and help. Without their encouragement and guidance, this project would still be a dream that could not come true. The guidance and

support received from all individuals who contributed and who are contributing to this project, was vital for this project to be successful. I am grateful for their constant support and help

9. REFERENCES

- [1] Nidcd.nih.gov. (2010). Statistics on Voice, Speech, and Language | NIDCD.[online] Available at: <https://www.nidcd.nih.gov/health/statistics/statistics-voicespeech-and-language> [Accessed 7 Jun. 2010].
- [2] Asha.org. (n.d.). Speech and Language Disorders and Diseases. [online] Available at: <http://www.asha.org/public/speech/disorders/>.
- [3] Medlineplus.gov. (n.d.). Speech and Communication Disorders: MedlinePlus. [online] Available at: <https://medlineplus.gov/speechandcommunicationdisorders.html> [Accessed 25 Oct. 2016].
- [4] aph.gov, (n.d.). Prevalence of different types of speech, language and communication disorders and speech pathology services in Australia. p.1.Tavel, P. 2007
- [5] Lúcio, G., Perilo, T., Vicente, L. and Friche, A. (2013). The impact of speech disorders quality of life: a questionnaire proposal. [online] Available at: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S2317-17822013000600610.
- [6] Google, (2006). About – Google Translate. [online] Translate.google.com. Available at: https://translate.google.com/about/intl/en_ALL/.
- [7] QuickVoice Recorder, and nFinity inc, (n.d.). QuickVoice® Recorder on the App Store. [online] App Store. Available at: <https://itunes.apple.com/sa/app/quickvoice-recorder/id284675296?mt=8> [Accessed 2016].
- [8] Shazam, (2015). Shazam – Windows Apps on Microsoft Store. [online] Microsoft Store. Available at: <https://www.microsoft.com/enus/store/p/shazam/9wzdncrfj0qq>.
- [9] O'BOYLE, B. (2015). What is Siri? Apple's personal voice assistant explained -Pocket-lint. [online] Pocket-lint.com. Available at: <http://www.pocketlint.com/news/112346-what-is-siri-apple-s-personal-voice-assistant-explained>.
- [10] OSHEROVE, R. (2014). The Art of Unit Testing. [online] Piazza-resources.s3.amazonaws.com.
- [11] Klaus, A. (2018). Testing – A Tutorial [online] Pdfs.semanticscholar.org. Available at: <https://pdfs.semanticscholar.org/presentation/0b0e/3763182eeb2f95893480ee0c0d77a3e894e4.pdf>.
- [12] Oladimeji, P. (2007). Levels of testing [online] Cs.swan.ac.uk. Available at: <http://www.cs.swan.ac.uk/~csmarkus/CS339/dissertation/OLadimejiP.pdf>.
- [13] McLellan, Sam, Andrew Muddimer, and S. Camille Peres. "The effect of experience on System Usability Scale ratings." *Journal of Usability Studies* 7.2 (2012): 56-67.