BORNO: A BENGALI VOICE COMMAND BASED INTERACTION SYSTEM

BY

Raven Mark Quiah

ID: 151-15-5059

AND

Soma Akter

ID: 151-15-5345

This Thesis Presented in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science in Computer Science and Engineering

Supervised By

Mr. Naziour Rahaman
Lecturer
Department of CSE
Daffodil International University

Co-Supervised By

Mr. Ahmed Al Marouf
Lecturer
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY
DHAKA, BANGLADESH

November, 2018

APPROVAL

This Thesis **" Borno: A Bengali Voice Command Based Interaction System**", submitted by **Raven Mark Quiah** and **Soma Akter** to the Department of Computer Science and Engineering, Daffodil International University, has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 20th November 2018.

BOARD OF EXAMINERS

_____

(Name) **Chairman**

Designation

Department of CSE

Faculty of Science & Information Technology

Daffodil International University

_____

(Name) **Internal Examiner**

Designation

Department of CSE

Faculty of Science & Information Technology

Daffodil International University

_____

(Name) **External Examiner**

Designation

Department of -------Jahangirnagar University

DECLARATION

We hereby declare that, this project has been done by us under the supervision of **Mr. Naziour Rahaman, Lecturer, Department of CSE** Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by:

_____

Mr. Naziour Rahaman
Lecturer
Department of CSE
Daffodil International University

Co- Supervised by:

_____

Mr. Ahmed Al Marouf
Lecturer
Department of CSE
Daffodil International University

Submitted by:

_____

Raven Mark Quiah
ID: -151-15-5059
Department of CSE
Daffodil International University

_____

Soma Akter
ID: -151-15-5345
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First we express our heartiest thanks and gratefulness to Almighty for His divine blessings that make us possible to complete the final year thesis successfully.

We are really very much grateful to **Mr. Naziour Rahaman**, **Lecturer**, Department of CSE Daffodil International University, Dhaka. He was so helpful in the whole bunch of time. Whenever we needed his help no matter what he was always there for us. His motivation, inspiration, and direction helped us to fulfill our thesis work very effectively. His endless patience ,scholarly guidance, continual encouragement , constant and energetic supervision, constructive criticism , valuable advice ,reading many inferior draft and correcting them at all stage have made it possible to complete this project.

We would also like to thank our co-supervisor **Mr. Ahmed Al Marouf, Lecturer,** Department of Computer Science and Engineering, Daffodil International University, Dhaka. He showed us many ideas on this human-computer interaction. He inspired and motivated us that we would be able to come up with this idea and would be able to make this work happen.

We would also like to express our heartiest gratitude to **Prof Dr. Syed Akhter Hossain**, Head of Computer Science and Engineering for all his appreciation, adoration and motivation.

We are also very much thankful to our family, siblings, and friends who were always there to criticize our work in a manner to improve that all the way long. So thanking all of them from the core of our heart.

ABSTRACT

We are introducing you “Borno”, a Bengali Voice Command Based Interaction System for the personal computer. The system uses the Speech Application Program Interface (SAPI) to recognize the speech. We have developed our own grammar in order to map the Bengali phonemes in English representation. We have also developed an algorithm to reduce the confusion of the machine in recognizing the same type of phonetics in the language. The algorithm will also help to understand the Bengali command with more accuracy. Another thing of this system is that we have used Microsoft speech synthesizer to synthesize the speech in Bengali also. It helps to implement human-computer interaction in a wide range. This feature will help the disabled people especially the blind people who are unable to read the text on the computer. This system can be widely used among the Bengali speaker around the world. This system can also be used in rural areas of Bangladesh. Most of the people of Bangladesh are illiterate. That’s why they face challenges while operating a computer in English. In order to use a computer and get the facilities of modern technology, they have to get the training. So, we think instead of using traditional mouse and keyboard if they are able to operate the computer with their own voice command in their native language, it will help them a lot to get the facilities of the computer. So, we design this system to help them in this regard. We achieve 90% of accuracy in the noise-free environment and 76% of accuracy in the noisy environment with the help of 20 individual candidates both male and female. Our goal is to raise the accuracy of the system in the future.

TABLE OF CONTENTS

CONTENTS	PAGE
Board of examiners	I
Declaration	Ii
Acknowledgements	Iii
Abstract	Iv
CHAPTER	
CHAPTER 1: INTRODUCTION	**1-5**
1.1 Introduction	1
1.2 Motivation	2
1.3 Rationale of the Study	3
1.4 Research Questions	4
1.5 Expected Output	5
CHAPTER 2: BACKGROUND	**6-8**
2.1 Introduction	6
2.2 Related Works	6
2.3 Research Summary	7
2.4 Scope of the Problem	7
2.5 Challenges	8
CHAPTER 3: RESEARCH METHODOLOGY	**9-19**
3.1 Introduction	9
3.2 Research Subject and Instrumentation	9-10

3.3 Data Collection Procedure 10-12

3.4 Statistical Analysis 12-13

3.5 Implementation Requirements 13-19

CHAPTER 4: EXPERIMENTAL RESULTS AND DISCUSSION 20-21

4.1 Introduction 20

4.2 Experimental Results 20

4.3 Descriptive Analysis 21

4.4 Summary 21

CHPATER 5: SUMMARY, CONCLUSION, RECOMMENDATION, AND IMPLICATION FOR FUTURE RESEARCH 22-24

5.1 Summary of the Study 22

5.2 Conclusions 22

5.3 Recommendations 23

5.4 Implication for Further Study 23-24

REFERENCES 25

LIST OF FIGURES

Figure 3.1 Namespaces used by the system 10

Figure 3.2 Graphical view of the recognition accuracy of the system 13

Figure 3.3 Block diagram of the voice recognition based interaction system 18

Figure 3.4 User interface of Bengali voice command based interaction system 19

LIST OF TABLES

Table I: Bengali voice command accuracy rate 11-12

Table II: Bengali phonemes to English representation grammar list 14-16

CHAPTER 1
INTRODUCTION

1.1 Introduction

In this current world, Human-Computer interaction is a giant term on the basis of digitalization and technological advancement. Tech world was always in search of some amazing inventions regarding this. We all know how the invention of the computer simply changed the outlook of the world as well as it changed the way of our day to day life. From then, everyone wanted a revolution in the human and computer interaction. As the more the interaction will get improved, the more the tech world will get revolutionary advancement.

"Speech Recognition" is one of the fastest and useful ways of interacting with computers. From early 1970, "Speech Recognition" has become one of the most interesting searching areas in the sector of Natural Language Processing.

In this, any type of voice or speech gets recognized by the system and the system respectively responds to them. This speech or voice is considered as inputs and the response given by the system is considered as output.

After searching Human-Computer interaction related work we found out that there are many works on the English Language. And people were trying to develop the sector of Bengali Language. As much work was not done on Bengali Language so we wanted to work with our mother tongue.

In this system, we will work with "Bengali Voice Command". We wanted to establish a commanding system on Computer where various commands in Bengali will be given by the user as inputs and computer will respond to each of the given commands. And then the system will interact with the user by following the given instructions.

In Computer there are so many options like opening a folder, opening a file, closing the current window, opening a new window, minimizing tabs, browsing internets and so on. And we will do these works by giving the command in Bangla. For that purpose,

we have built our own Grammar in Bengali and have mapped them in such a way so that the computer gets to understand the commands through mapping and can interact in a respective way.

We wanted to build a system for the computer which is not available and will leave a strong impact on human-computer interaction. This system is much easier for the user to interact only with their voice or speech rather than using several mouse clicks and so on.

We are hoping that this Bengali Voice Command Based Interaction system will be even more helpful to the blind people to interact with computers. As it is considered as a rather impossible task for them to get introduced with such things. So this system will make their life much easier in many ways. As for this, what they only need is their voice which will be considered as commands and then their computer will work as per as their given commands. Thus the system will be a great use for them.

1.2 Motivation

There is a simple saying that for every single invention, there works motivation behind the happening. And ours is not different from this. As there were so many motivational things that motivated us to work on this particular project.

In this era, information works like power. The more one can process something within the most required minimum time, the more it's regarded as useful. And now-a-days the biggest tech giants are in search of such innovation every single minute. So we wanted to do something with Bangla Language regarding this. So very first what motivated us was working with our mother tongue. As we earlier mentioned that from the very beginning more and more people are working with the English language. And as a Bengali, we wanted to enrich ours one with new technological terms. And thought to have a try on this particular researching sector.

Then another thought that came to our mind is about the total number of population and the literacy percentage of our country. There was a time when we couldn't even afford those six basic human rights. So people were more interested in earning rather than

acquiring knowledge. And thus a huge amount of our total population couldn't afford education.

Education is like a blessing and there is no doubt about it. But in a country like us, it's way too difficult to ensure education for all. As a result, a large number of the total population couldn't finish their schooling. May be we'll have a hard time finding how many people in our country knows English and can speak in that. But there will no one for sure who can't speak in their mother tongue Bangla. So this one also motivated us to work with Bangla Speech Recognition. Where every single educated or uneducated people will be able to interact with computers through their voice command in Bangla only. This is a great thing to even think of getting happened.

Also, there was something which motivated us in a great way to work with this is those visually impaired people who can't see the light of the earth. So they can't even think of using any digital technological gadgets. So as a responsible citizen we wanted to something for the welfare of our countrymen and all those blind persons.

All these things literally motivate us to the core to work with such an amazing project and to build something new. If this motivation can be turned to our imaginative output then we can think of having a great impact in the digital world as well as in the vast area of Bengali Speech Recognition.

1.3 Rationale of the Study

We have already discussed the importance of speech recognition with a specific language. And also it's a known fact that among all the languages, English is the most used and worldwide language regarding this particular searching area. As we have chosen our mother tongue Bengali as our specific voice recognizable language so here we would like to uncover the importance of Bengali as a language.

Very first we would like to proudly say about the Language Movement which was taken place on 21st February of 1952. In the history of the world, we Bangladeshi are the only nation who fought for their mother tongue and also achieved victory. There were brave sons like Rafiq, Shafiq, Barkat, Zabbar and so on who sacrificed their valuable lives

only to speak in Bengali. That was like a tremendous victory for all of us. And may be that was the beginning of proving our worth to the world. And after all these sacrifice and dedication we got this as our mother tongue.

Approximately on an average, 205 Millions of people speaks in Bengali as their first language. They are known as native speakers. And according to a static, on an average 3.05% of world population speaks in Bengali. So no doubt that the total number is huge.

And Bengali is used as a first language in Bangladesh. This is our mother tongue. Also, Bengali is used in several parts of India. Among 22 scheduled languages in India, Bengali is the most widely used 2nd language after Hindi. West Bengal, Tripura, and Barak Valley use Bangla as their official language. Bengali is also used in different parts of Brahmaputra valley of Assam. Also few people of Bihar, Mizoram, Meghalaya, Odisha, and Jharkhand speaks in Bengali. So the user is not that much negligible anymore. Day by day the total number of people speaking in Bengali in increasing. And according to a statistic, Bangla is on the top seventh list among the languages of the world. It is counted as the seventh most spoken native language by population.

So when we look at all these achievements, we proudly say to the world that yes, we speak in Bengali and this is our mother tongue. So when we were thinking to work on Human-Computer interaction, we always wanted to work with our own mother tongue. We wanted to enrich this language more and more so that Bengali language also can compete in the history of speech recognition with all dignity.

1.4 Research Questions

While working on this particular area we faced some of the existing challenges and also some questions regarding this research was hitting hardly on our head whether we would be able to implement our idea or not. Some research related questions are given below:

- Is it possible for this system to cope up with the vast area of Bengali Language like other existing developed implemented systems in English language?

- Will we be able to create adoptable command based grammar for the implementation of the idea?
- Will SAPI have compatibility to recognize Bengali Language?
- Will the accuracy level for both noisy and noise-free environment be trust worthy?
- How much the system will come in the help of illiterate and disables people?

So these were some of the exemplary basic questions that we are facing in our initial stage of implementation.

1.5 Expected Output

- The user should be able to operate his/her personal computer with Bengali Voice Command.
- Illiterate people would be able to easily interact with the computer with the help of easy user interface and Bengali Voice Command.
- Visually impaired people can easily interact with the computer with voice command.
- Disable people can easily use this without using mouse or keyboard at all.
- End of using traditional mouse and keyboard.

CHAPTER 2

BACKGROUND

2.1 Introduction

"Borno" is a Bengali virtual assistant for personal computer users. A lot of works are done before in the Bengali language. Speech-to-text, text-to-speech, Bengali voice control robot etc. works are done before. But there are few works those were done considering the disabled peoples, especially for the blind people. So through this system, we have tried our best to contribute to this sector. Our system will help the disabled and illiterate people to use the computer more comfortably.

2.2 Related Works

As we said before that there are a lot of works are done in the Bengali language. The most relevant work to our research is Bangla Voice Controlled Robot for Rescue Operation in Noisy Environment [1]. They tried to implement Bangla voice-controlled robotic motion based on an efficient and unique algorithm for accurately detecting Bangla voice commands. They used Mel-Frequency Cepstrum Coefficients. Some related works are done in the Bengali language are SHRUTI-II: A Vernacular Speech Recognition System in Bengali and an Application for Visually Impaired Community [2]. They introduced SHINX3 based Bengali Automated Speech Recognition (ASR) system and they implemented it in sending the email for the visually impaired people. Another work is Prodorshok I: A Bengali Isolated Speech Dataset for Voice-Based Assistive Technologies [3]. They tried to create speaker independent voice-command driven automated speech recognition (ASR) based assistive technologies to help and to improve human-computer interaction (HCI).

None of the works are done in building the Bengali Virtual Assistant. So, we think, it will be a great contribution to our country in building the virtual assistant in the Bengali language. "Borno" is our first Bengali Virtual Assistant which can recognize any

individual's voice in a noise-free and noisy environment. We tried our best to raise the accuracy of the system. Hopefully, very soon we will reach our goals.

2.3 Research Summary

We are going to implement a Bengali Virtual Assistant for the personal computer to help the visually impaired people and illiterate people to operate their computer in their native language. To do so, we have used Microsoft Speech Recognition technology. In this system, we have mapped the Bengali phonemes with English phonemes and thus have tried to help in recognizing the Bengali speech by the machine. Our system not only recognizes voice but it can also be able to synthesize the speech in Bengali.

We have also implemented an algorithm to reduce the confusion of the machine to take command in a noisy environment. Here we don't try to be perfectionist by eliminating all the problems, but we have tried to reduce some major problems found in Speech Recognition Based Systems.

2.4 Scope of the Problem

There are lot of works done in Bengali Speech to Text using SAPI. But we couldn't find any system available that works for personal computer. So we have made our mind to work with such a system which will work for personal computers. And for this purpose, we have also used SAPI. This is actually known as Speech Application Program Interface which is an Application Program Interface and it is shortly known as API.

There were some scope of problems that was arising while implementing our ideas into reality. And among them most significant problem was to work with a noisy environment.

2.5 Challenges

We faced a lot of challenges while implementing the project. One of the most common challenges is avoiding the noise from the background. The system detected the similar command from the noise. Another challenge that we faced was the mapping of English phones with Bengali phones. But we are happy to say that we have succeeded in most of the cases and have overcome most of the challenges successfully. But, still, there is some existing problems with the noise.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

In order to control our personal computer with voice is a common trend in this era. This is like a blessing for the disabled people. To make the computer more efficient to use for the disabled people, in fact, especially for the blind people, we have tried to develop this voice command based human-computer interaction system which will enable the disabled people to use the computer more efficiently. Our system can recognize the Bangla speech and according to the given command, it can control the pc. It can also feedback the visited directory or file name in Bengali phoneme. That will help the blind people to understand better about the place that they are visiting.

3.2 Research Subject and Instrumentation

3.2.1 Hardware Required for Developing the System

1. Processor: Intel Core i3
2. Operating System: Windows 10
3. Programming Language: C#
4. Microphone: Windows Default Microphone

3.2.2 Software Required for Developing the System

We have used Microsoft Visual Studio 2012 to design both the interface and developing purpose. In order to recognize speech, we have used **Microsoft.Speech.dll** library file. This library is used both as speech synthesizer and speech recognition.

.

3.2.3 Installation Introduction

In order to implement the system, we have used Microsoft Visual Studio 2012. So, at first, we have to install the .NET environment. In this we have used the C# programming language to write the code.

3.2.4 Additional Library Installation

As we have used Microsoft's speech recognition technology, that's why we have to install additional libraries which provide both the speech synthesis and voice recognition. Here Microsoft **Speech.dll** library has been used to recognize the speech. So, we have to download and install the library.

3.2.5 Essential Namespaces

Namespaces are used to extract the ability of the library. Some important namespaces are ***System.Speech.Recognition***, ***System.Speech.Synthesis*** etc. **Figure-6.1** elaborates the other namespaces.

```
1  using System;
2  using System.Collections.Generic;
3  using System.ComponentModel;
4  using System.Data;
5  using System.Drawing;
6  using System.Linq;
7  using System.Text;
8  using System.Threading.Tasks;
9  using System.Windows.Forms;
10 using System.Speech.Recognition;
11 using System.Speech.Synthesis;
12 using System.Diagnostics;
13 using System.IO;
14 using System.Collections.Generic;
```

Figure 3.1: Namespaces used by the system

3.3 Data Collection Procedure

Very first we would like to let you know that to test the accuracy level of our system we have gone through two types to data collection ways.

We wanted to work with both the noisy and noise free environment. So we had collected data of 20 students in a classroom and that was full of noise. So that collected data sets are usually noisy data sets.

Then we have collected another 20 data sets in a noise free environment. And those data collections were eventually used in order to measure the accuracy level. And we would like to tell that the accuracy is pretty much worthy of being a successful system.

Table I shows the data from the individual speaker and the accuracy rate.

TABLE I: BENGALI VOICE COMMAND RCOGNITION ACCURACY RATE

Candidates	**Accuracy Rate In Noise-Free Environment**	**Accuracy Rate In Noisy Environment**
Speaker 1	90%	70%
Speaker 2	90%	90%
Speaker 3	100%	100%
Speaker 4	90%	80%
Speaker 5	100%	90%
Speaker 6	80%	70%
Speaker 7	90%	60%
Speaker 8	80%	90%
Speaker 9	90%	70%
Speaker 10	90%	70%
Speaker 11	80%	70%
Speaker 12	90%	70%
Speaker 13	90%	90%
Speaker 14	100%	100%

Speaker 15	90%	60%
Speaker 16	90%	70%
Speaker 17	80%	90%
Speaker 18	100%	80%
Speaker 19	90%	90%
Speaker 20	90%	70%
Mean	90%	79%

3.4 Statistical Analysis

We get the accuracy level like,

Mean Accuracy in Noise- Free Environment= 90%

Mean Accuracy in Noisy Environment= 79%

We have tested our accuracy level of the system on 10 commands. And on the basis of the system's response on those given commands, we have estimated that accuracy level.

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1}$$

Now by following the formula we get,

Standard Variance for Noise Free Environment= 42.10 and for noisy environment= 156.84.

Standard Deviation for Noise Free Environment= 6.48 and for noisy environment= 12.52.

Figure 3.2 shows the graphical view of the recognition accuracy of the system.

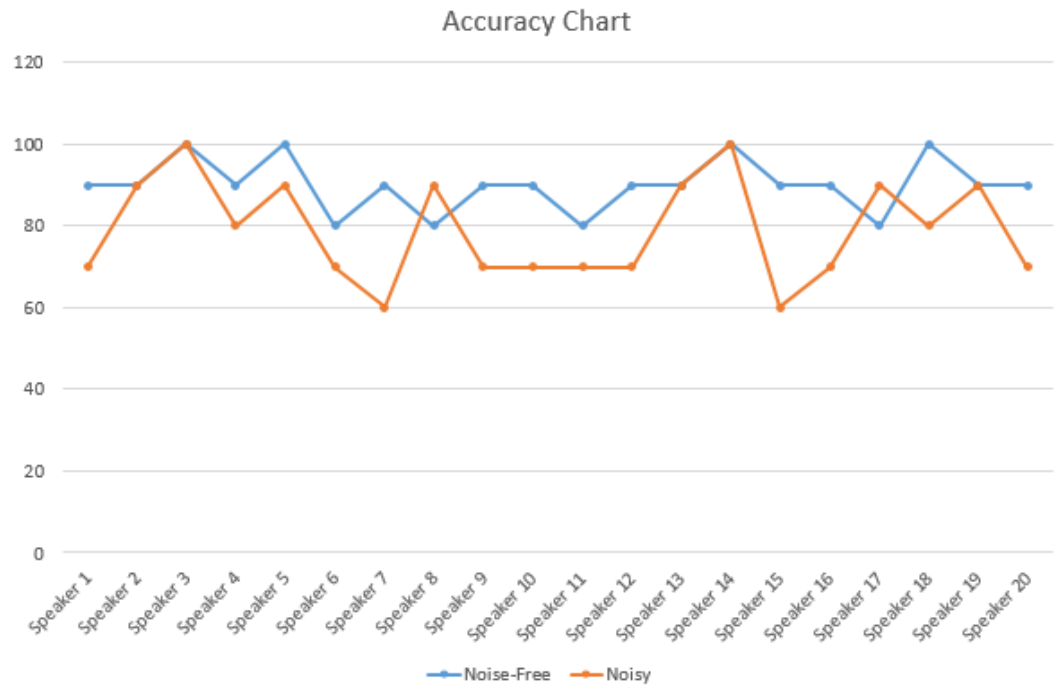

Figure 3.2 Graphical View of the Recognition Accuracy of the System.

3.5 Implementation Requirements

3.5.1 Bengali Phenomes To English Representation

In order to understand the Bangla Language by the machine, we have to create our own Grammar. As Bangla grammar is not that much available. Our main purpose is to find the Bengali Phones in the English Phonetics. That's how we mapped the English text to Bengali Text and then controlling the pc by those texts.

3.5.2 Creation Of Grammar

The grammar that we have built in order to make the machine to understand the Bengali language has taken a lot of time. Because in this, we have to try to get the clear phones according to Bengali phones. In this stage, we have found many difficulties in finding

the phones “চ” and “ছ”. Like the command “আছে”. If we map the English text “ACHEY” then the machine understand the phones as “একি”. This phone is so much distinct from our required phones. So, have to try another text to understand the phone “আছে” more accurately. If we add a consonant letter followed by a vowel in front of the “CH” then we get the more clear phones than the previous one. Finally, we have got the familiar phone in English of text “আছে” as “aye-chey”. In this way, we have created our other grammar that was compatible with the system in a systematic way. So while we were building our own grammar we were facing difficulties regarding these pronunciation issues. As everyone has their very own process to pronounce every single word. So to match that we had to come up with our own created grammars.

3.5.3 Grammar Lists

The other grammars that we have used in the system are being listed here:

TABLE II: BENGALI PHONEMES TO ENGLISH REPRESENTATION GRAMMAR LIST

Bengali Phonemes	**English Representation in MS Speech Recognition Engine**
উপরে উঠো	"Uporay otho"
সামনে আসো	"shamney asho"
লুকিয়ে থাকো	"lukiye thako"
ডেস্কটপ এ যাও	"desktop a jow"

নির্দেশনাগুলো দেখাও	"nirdeshona gulo dekhow"
ফাইল নির্বাচন কর	"File nirbachown koro"
নিচে নামো	"Nichay namo"
আমার কম্পিউটার এ কি আছে	"amar computer a ki ayechey"
চলন্ত অ্যাপ্লিকেশানগুলো দেখাও	"chawlonto application gulor numm bolo"
বর্ণ	"borno"
বন্ধ করো	"bondho koro"
তুমি কে	"tumi k"
C ড্রাইভ এ যাও	"C Drive a jow"
D ড্রাইভ এ যাও	"D Drive a jow"
E ড্রাইভ এ যাও	"E Drive a jow"
F ড্রাইভ এ যাও	"F Drive a jow"
ফোল্ডারগুলোর নাম বল	"folder gulor numm boolo"

পরিষ্কার করো	"porishkar koro"
ফোল্ডার দেখাও	"Folder dekhow"
ক্রমে যাও	"chrome a jow"
ফেসবুক এ যাও	"facebook a jow"
পিছনে যাও	"pichowney jow"
হোমে যাও	"home a jow"
ফাইলটি খোলো	"file ti kholo"
ফোল্ডারটি খোলো	"folder ti kholo"
ফোল্ডার নির্বাচন করো	"folder nirbachown koro"

3.5.4 Algorithm

Our main target was to reduce the noise recognition by the system because that is the main problem in speech recognition technology. This noise issue can be considered as a burning problem in this sector. In order to do so, we have developed our own algorithm in order to reduce similar speech pattern detection. Like "**বর্ণ**" and "**বন্ধ**". They sometimes confused the system whether they had been pronounced clearly or not. So, we have tried to reduce the sensibility of the microphone voice detection level, that's how the system can able to take the speech which is more accurately pronounced.

But, there is another problem arise here. Speaker has to give the command more clearly and slowly so that system can recognize the pattern more clearly and effectively.

We have also done another thing here to reduce the confusion of the system. Here we have initialized two different Voice Recognition Engines. We have set this in a way that when one engine will generate the initial command, the other existing engine will not engage. Like, when FirstSRE (Speech Recognition Engine) is recognizing, the SecondSRE is stopped. Then SecondSRE will start when the user gives command to navigate the hard disk of the computer. Then, FirstSRE will stop recognizing. Separating the two-part can reduce the available command in the system so that it can recognize less command and less confused. Thus we have developed the system.

Here is the algorithm that we have generated in order to reduce the confusion of the system.

1. INITIALIZE FIRST_SRE, SECOND_SRE;
2. IF (SRE_GRAMMAR = = NAVIGATION COMMAND)
   {

   EXECUTE;

   }
3. ELSE IF (SRE_GRAMMAR == HARD_DISK_COMMAND)
   {

   EXECUTE;

   STOP FIRST_SRE;

   START SECOND_SRE:

   }
4. ELSE {
   UNRECOGNIZED COMMAND;

   }

3.5.5 Working Procedure

After running the system user can able to command the system to do the work in the Bengali language. The system can provide the information with available hard disk memory locations. The user then will be able to command the system to go into any memory location. The system then gives the user feedback with the available file and directory name. The user then will be able to go into the directory or open the file with the voice command. Figure 3.3 shows the work flow of the Bengali speech recognition based interaction system.

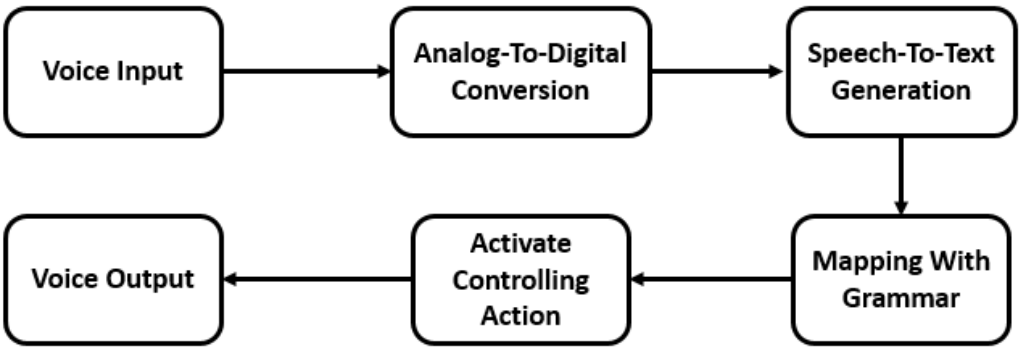

Figure 3.3: Block Diagram of the Voice Recognition Based Interaction System

3.5.6 Design and Development

We have tried to design the interface more convenient to the user so that illiterate people can easily understand the workflow. A Bengali user manual will also be provided so that they can command the system easily. The user can also minimize or maximize the interface with voice command. They can move anywhere in the computer with voice command. They can also navigate to the internet with voice command. The computer will feedback the selected folder/directory with a voice that will help them to better understand.

3.5.7 User Interface of the System

We have tried to give a more user-friendly look to the system interface. That's why we have made the interface simple enough to easily understand by the user. Figure-3.4 shows the initial user interface of the Bengali Voice Recognition System.

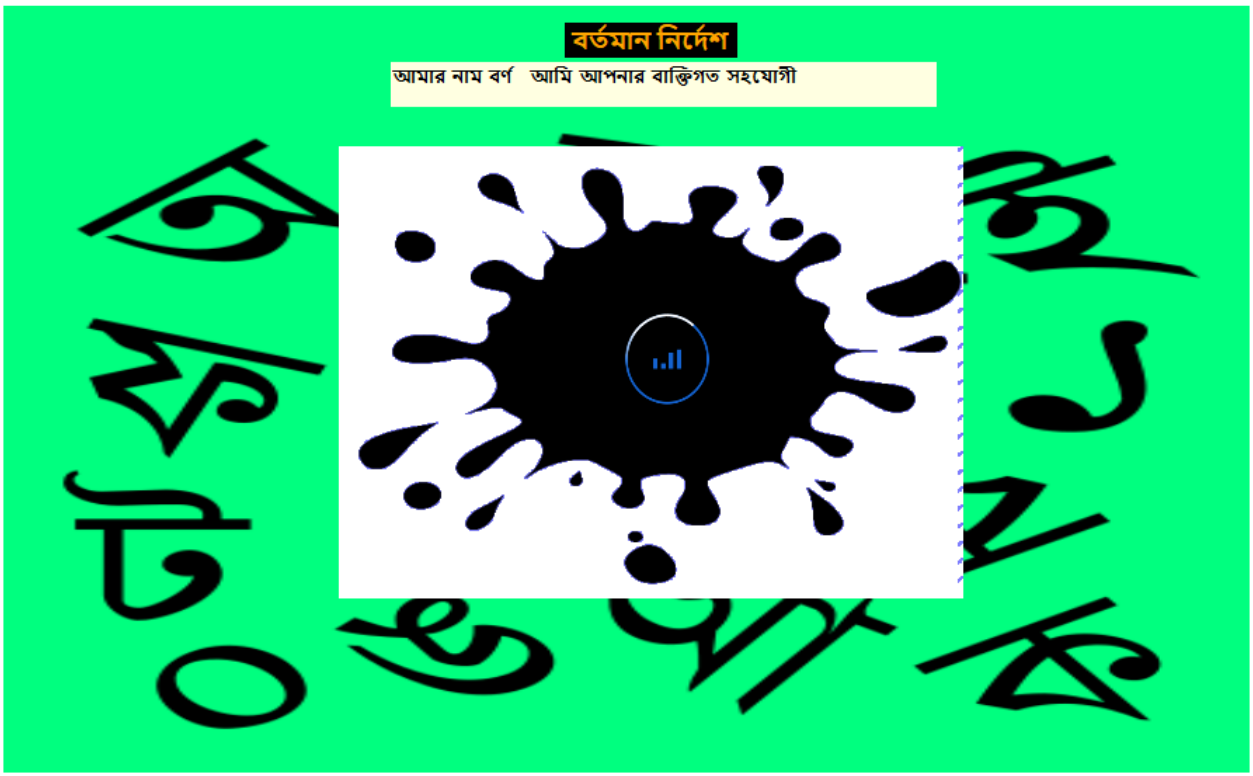

Figure 3.4: User Interface of Bengali Voice Command Based Interaction System

CHAPTER 4

EXPERIMENTAL RESULTS AND DISCUSSION

4.1 Introduction

The outcome of our project is quite satisfactory. Though there are some existimg problems in the system. And it's quite obvious to deal with a few problems in every system. The main and common problem of the voice recognition system is recognizing the similar phonemes and background noise. We can't fully eliminate these two problems but we have tried our best to reduce the problem through our algorithm. We can say that we are at a satisfactory level. In the future, we would try to develop more enhance algorithm to eliminate the problem.

4.2 Experimental Results

Here from the collected data sets we have measured a mean value by following the rule of,

Sum of Data Sets/ Total number of data sets.

Finally we get the accuracy level like,

Accuracy in Noise- Free Environment= 90%

Accuracy in Noisy Environment= 79%

On an average the accuracy level is quite good. Though the accuracy level of noisy environment is still not that high but we are working to get a high range.

4.3 Descriptive Analysis

As we have tried to implement Bengali phonemes through English phonemes, it's actually hard to map the two language phonemes. But, we are able to make the machine in term of recognizing the utterance of the Bengali language. In this case, we are 100% success. This system is much faster recognizing because the whole process would run in the random access memory (RAM). If we add additional files for speech mapping it will slow the system and it may irritate the user for waiting for the system to respond. At the end of the day, we can say that we are able to make the machine to recognize our beloved Bengali language. And that is our actual success regarding the research area and research purpose. As our main motto was to work with our mother tongue.

We have tested our system in both noisy and noise-free environment and our result is quite satisfactory. We have tested the accuracy of the system by testing both Male and Female testers. But we have found some problem in a noisy environment. As in a noisy environment the system was getting confused with the other noisy commands. So the accuracy level in a noise free environment is quite praiseworthy in terms of the accuracy level of the system in a noisy environment. So above all, this system can cope up with all the surrounding problems and we are working to solve the noisy issue.

4.4 Summary

In a nutshell, we can say that we have implemented our idea that we had dreamt of. We wanted to take Bengali voice command as input and then had dreamt of having such an output where given voice command will be followed and things will work out as per as those voice commands given only. And we are pretty much success in that attempt. Now any user can give their voice command in Bengali and the system will work according to that.

CHAPTER 5

SUMMARY, CONCLUSION, RECOMMENDATION, AND IMPLICATION FOR FUTURE RESEARCH

5.1 Summary of the Study

Though much works are done in Bengali voice recognition. But a few works are done in mapping the English phonemes with the Bengali phonemes. And actually that is our success. We will be able to map these two precious languages successfully. Bengali text to speech and speech to text projects are done before. There are a few works are done but none of them had introduced the Virtual Assistant in the Bengali language to operate a computer. There are some works done in voice control robot in Bengali, but not a personal computer. Though there are some limitations in our system, we will do further work on those limitations. And we will try to cope up with all the remaining issues and complications in future.

5.2 Conclusions

Bengali is the most precious and most beloved language for the Bengali people. Many memories are attached to this language. But our language is going dark due to lack of practice among the young generation. They even don't know many Bengali word meaning. This is a shame for our country. Doing a little contribution in this area satisfies us a lot. Again, this system will be a blessing for the disable people. This is another big success of our little work in the Bengali language.

5.3 Recommendations

As now-a-days the area of speech recognition and human-computer interaction is getting vast and larger every single passing day so concentrating on a particular area is a must. Any specific working idea on any specific language is a must to implement something better for the sake of invention something unique. So such recommendations regarding these would help.

We would recommend the system to the blind people and also to the illiterate folks. This system would come as a great help to these people to cope-up with the very fast pace of the current world.

Every single passing day the technology is inventing newer to newer things for the purpose of building a tech world. So such systems add a great contribution in their vision of creating a new technological era.

5.4 Implication for Further Study

In the future, we are very much into work with this project. We would like to improve it with some more unique and sophisticated ideas. We want to work with further implementations which would add much more definition to this existing one.

Here we have used speaker independence. Where system will work with any given voice or speech command in Bangla. But in future, we have thought to work with a particular voice for every particular system. As now-a-days people like to use their devices or gadgets way more personal. So in our future work, we would like to work with a specific voice recognition system where the system will be able to recognize the user. And the system will be more speaker dependent. For that, we have planned to work with wave frequency and to collect more and more data. We hope that very soon we would be able to implement the idea.

And another thing that we are planning to work with is having our own voice synthesizer and voice recognizer. Here we have used Microsoft's voice synthesis classes and voice recognizer to implement our idea. So in future, we are thinking to

build our own voice synthesizer and voice recognizer that will work with our own built Bangla Grammar.

In our research project, we have shown how any folder or file gets opened or closed by the given command of the user in Bangla. For our future work, we have also thought that by following the command the system will even respond to the user to read out any content of any file. And we believe that would be even much easier for the blind people to use the computer.

We hope that we would be able to implement all our ideas and would be able to execute all the plans to reality.

REFERENCES:

[1] Arnab Bhattacharjee, Md Zakaria Haider, Asir Intisar Khan, Dhiman Chowdhury, "Bangla Voice Controlled Robot for Rescue Operation in Noisy Environment", 2016 IEEE Region 10 Conference (TENCON),pp. 3284-3288,2016.

[2] Sandipan Mandal, Biswajit Das, Pabitra Mitra, "SHRUTI-II: A Vernacular Speech Recognition System in Bengali and an Application for Visually Impaired Community", 2010 IEEE Students Technology Symposium, pp.229-232, 3-4 April 2010.

[3] Mohi Reza, Warida Rashid, Moin Mostakim, "Prodorshok I: A Bengali Isolated Speech Dataset for Voice-Based Assistive Technologies", 5th IEEE R10 HTC, 2017.

[4] Sayem, A. "Speech analysis for alphabets in Bangla language: automatic speech recognition", International Journal of Engineering Research, vol. 3, issue no. 2, pp. 88-93, 1 February, 2014.

[5] Shikhor Kumer Roy, Abinash Kanti Ghosh, Aysa Siddika Asa, Md. Palash Uddin, Md. Rashedul Islam, and Masud Ibn Afjal, "Bengali consonants-voice to text conversion using machine learning tool", International Journal of Research in Computer Engineering and Electronics, vol. 6, issue no. 1, March 2017.

[6] Urmila Shrawankar, Dr. Vilas Thakare, "Speech User Interface for Computer Based Education System", 2010 International Conference on Signal and Image Processing, pp. 148-152, 2010.

[7] Prachi Khilari, Prof. Bhope V. P., "A REVIEW ON SPEECH TO TEXT CONVERSION METHODS", International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), vol. 4, issue 7, pp. 3067-3072, July 2015.

[8] Shaheena Sultana, M. A. H. Akhand, Prodip Kumer Das, M. M. Hafizur Rahman, "Bangla Speech-to-Text Conversion using SAPI", International Conference on Computer and Communication Engineering (ICCCE 2012), pp. 385-390, 3-5 July 2012.